INSIGHTS IN HEALTH ECONOMICS: 2021

EDITED BY: Mihajlo Jakovljevic and Seiritsu Ogura
PUBLISHED IN: Frontiers in Public Health

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-872-1
DOI 10.3389/978-2-88976-872-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

INSIGHTS IN HEALTH ECONOMICS: 2021

Topic Editors:
Mihajlo Jakovljevic, Hosei University, Japan
Seiritsu Ogura, Hosei University, Japan

Citation: Jakovljevic, M., Ogura, S., eds. (2022). Insights in Health Economics: 2021. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-872-1

Table of Contents

frontiers | Frontiers in Public Health

TYPE Editorial
PUBLISHED 26 July 2022
DOI 10.3389/fpubh.2022.966741

OPEN ACCESS

EDITED AND REVIEWED BY
Chhabi Lal Ranabhat,
Eastern Kentucky University,
United States

*CORRESPONDENCE
Mihajlo Jakovljevic
sidartagothama@gmail.com

SPECIALTY SECTION
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

RECEIVED 11 June 2022
ACCEPTED 17 June 2022
PUBLISHED 26 July 2022

CITATION
Jakovljevic M and Ogura S (2022) Editorial: Insights in health economics: 2021. *Front. Public Health* 10:966741. doi: 10.3389/fpubh.2022.966741

Editorial: Insights in health economics: 2021

Mihajlo Jakovljevic[1,2,3]* and Seiritsu Ogura[4]

[1]Institute of Advanced Manufacturing Technologies, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia, [2]Institute of Comparative Economic Studies, Hosei University, Tokyo, Japan, [3]Department of Global Health Economics and Policy, University of Kragujevac, Kragujevac, Serbia, [4]Faculty of Economics, Hosei University, Tokyo, Japan

KEYWORDS

Health Economics, health financing, health systems, emerging markets, BRICS, health expenditure, global south, sustainability

Editorial on the Research Topic
Insights in Health Economics: 2021

In the third decade of the twenty first century, methodological applications are leading to the movement of the knowledge frontier in the field of Health Economics. This frontiers Research Topic aims to highlight the latest advancements in health economics at the forefront of science. This editorial initiative, of particular relevance, led by Specialty Chief Editors of the Health Economics section Prof. Mihajlo Jakovljevic and Prof. Seiritsu Ogura, was focused on new insights, novel developments, current challenges, latest discoveries, recent advances, and future perspectives in the field of Health Economics (1).

The Topic has succeeded to invite a total of 18 forward-looking contributions. Some of them were describing the state of the art, outlining recent developments and major accomplishments that have been achieved. Also, possible directions of further development were designated in some of these articles. An array of contributing institutions were encouraged to identify the greatest challenges in the inter-disciplinary developments at the cross-sections of health economics, policy, global, and public health, pandemic health emergencies, quality of life, and mainstream market economics.

There was a set of mutually interconnected European contributions. An Eastern European study focused on Health Technology Assessment framework in this vast region (Daubner-Bendes et al.). Clench-Aas and Holte elaborated on the crisis of political trust in Europe and the relationship between income and life satisfaction. This problem was observed on a scale of different levels including the national, community, and individual citizen level. On the other hand, a Scottish national study worked on an Analysis of multiple risky health behaviors and associated disease outcomes using Scottish linked hospitalization data (Olajide et al.).

A pandemic-focused trial has revealed bottle neck inefficiencies of overstretched health systems. A convenient example was the study pointing out the relationship to real gross domestic product per capita in 38 European countries during the first wave of the pandemic (Pardhan and Drydakis). Another piece of complementary evidence coming from Japan has shown that the pandemic has severely impacted physician visits in Japan (Kumagai). One more Japanese piece of work has dealt with financial literacy, financial

education, and smoking behavior (Watanapongvanich et al.). A single study of US origin focused on clustering and healthcare costs amongst the patients with multiple chronic conditions (Hajat et al.).

Chinese Universities encircle the largest set of studies that were successfully published in this Topic. Tang et al. have presented evidence of apparently puzzling connections between overweight/obesity and income-related inequality using Chinese labor force data. In another case a group led by Hao et al. has provided statistical evidence of the positive effect of the attempts of increased decentralization on the health outcome of communities, using DID analysis. Similarly, Xu and Lin published an elaborate econometric analysis of the effect of fiscal decentralization on the public health expenditure and public health, using continuous measures of fiscal decentralization. The fourth study is an interesting event-analysis of the food-industry firms that faced "crisis" from safety incidents, which examines the influence of their political connections and charitable donations, as well as the spill-over effects on their competitor firms (Xiang et al.). Particularly significant in an administrative system where health services almost never cross community boundaries is the study of spatial effect analysis of health spending patterns and trends in relationship to the health outputs in Chinese communities, establishing spill-over effects (Xu et al.). Finally, signaling the start of the expected and inescapable explosion of the demand for long-term care in the rapidly aging China, Tang et al. have provided a glimpse of the impact of long-term care insurance on the medical expenses and health status in Chinese society.

Vietnamese researchers did work on an interesting piece attempting to quantify the effects of various levels of dietary salt reduction on the prevalence of hypertension and the avoidable burden of stroke in Vietnam. These dynamic and complex consequences of salt intake in nutrition were observed through micro-simulation providing reliable assessment of the health and economic impacts (Aminde et al.). Another valuable Vietnamese contribution refers to a thorough exploration of complications of myocardial infarction after surgical treatment in Vietnam. This work contains an econometric part of the clinical equation, bringing valuable assessments of incremental cost, readmission risk, and duration of hospital admissions (Bui et al.). An Iranian study was particularly methodologically stringent adopting the framework of the systematic review and a meta-analysis. They managed to test the hypothesis of effectiveness of their national hospitals against the ongoing health sector evolution plan (Amini et al.).

Extensive Korean effort describes the market landscape and dynamics in relation to market exclusivity of the originator medicines in South Korea. This was conducted *via* adopting a retrospective cohort study design (Son). Last but not least, there was an ambitious piece of work originating from Sub-Saharan Africa. It was a comparative assessment of three national health systems namely the ones of Kenya, Tanzania, Uganda, and Zambia. This piece of research explores the causal relationship between donor commitments and disbursements for sexual and reproductive health aid (Kibira et al.).

The BRICS (Brazil, Russia, India, China, South Africa) Emerging markets continue to represent the engine of real economic growth worldwide. These nations shape the global demand for generic pharmaceuticals, medical goods, and services. Lock-downs caused by the pandemic have severely affected major supply chains and world trade routes. These changes alongside other upcoming challenges have decreased prospects for market recovery. The Global South nations, many of whom were represented in this topic, present a huge diversity in the historical legacy of their medical care financing and provision patterns. The burdens of premature mortality and absenteeism are multiplied by the prevalence and incidence of NCDs (2). Such health systems of LMICs countries will face difficult sustainability challenges due to long-term trends (3). An array of bottleneck vulnerabilities might be revealed once the entire health sectors are pushed to the limits of their resilience (4, 5). Current findings witness that the majority of global supply and demand is increasingly coming from the Asia—Pacific region. China, India, and South-East Asian ASEAN countries are most prominent representatives of this vast region (6).

Editors believe these valuable and diverse Topic contributions might open a new horizon of knowledge. Last but not least this is a unique opportunity to open the floor for a public debate on the Global South challenges from the perspective of academic health economics. The core aim of this special edition was to shed light on the progress made in the past decade in the Health Economics field (7). A diverse group of authors coming from academia, industry, governing authorities, and professional associations attempted to provide a thorough overview of the state of the art of the Health Economics field (8, 9). We hope that such an article Research Topic might inspire, inform, and provide direction and guidance to researchers in years to come.

Author contributions

MJ has prepared the manuscript draft, while SO has revised it for important intellectual content. Both authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Fuchs VR. *Who Shall Live?: Health, Economics and Social Choice*. Singapore: World Scientific (2011).

2. Jakovljevic M, Jakab M, Gerdtham U, McDaid D, Ogura S, Varavikova E, et al. Comparative financing analysis and political economy of noncommunicable diseases. *J Med Econ.* (2019) 22:722–7. doi: 10.1080/13696998.2019.1600523

3. Jakovljevic M, Liu Y, Cerda A, Simonyan M, Correia T, Mariita RM, et al. The Global South political economy of health financing and spending landscape–history and presence. *J Med Econ.* (2021) 24(Supp. 1):25–33. doi: 10.1080/13696998.2021.2007691

4. Jakovljevic M, Sugahara T, Timofeyev Y, Rancic N. Predictors of (in) efficiencies of healthcare expenditure among the leading Asian economies–comparison of OECD and Non-OECD Nations. *Risk Manage Healthc Policy*. (2020) 13:2261–80. doi: 10.2147/RMHP.S266386

5. Jakovljevic M, Timofeyev Y, Ranabhat C, Fernandes PO, Teixeira JP, Rancic N, et al. Real GDP growth rates and healthcare spending – comparison between the G7 and the EM7 countries. *Glob Health*. (2020) 16:64. doi: 10.1186/s12992-020-00590-3

6. Jakovljevic M, Lamnisos D, Westerman R, Chattu VK, and Cerda A. Future health spending forecast in leading emerging BRICS markets in 2030: health policy implications. *Health Res Policy Syst.* (2022) 20:23. doi: 10.1186/s12961-022-00822-5

7. Edwards RT, Charles JM, Lloyd-Williams H. Public health economics: a systematic review of guidance for the economic evaluation of public health interventions and discussion of key methodological issues. *BMC Publ Health*. (2013) 13:1001. doi: 10.1186/1471-2458-13-1001

8. Pitt C, Vassall A, Teerawattananon Y, Griffiths UK, Guinness L, Walker D, et al. Foreword: health economic evaluations in low-and middle-income countries: methodological issues and challenges for priority setting. *Health Econ.* (2016) 25(Suppl. 1):1–5. doi: 10.1002/hec.3319

9. Adeagbo CU, Rattanavipapong W, Guinness L, Teerawattananon Y. The development of the guide to economic analysis and research (GEAR) online resource for low-and middle-income countries' health economics practitioners: a commentary. *Value Health.* (2018) 21:569–72. doi: 10.1016/j.jval.2017.10.003

frontiers in Public Health

ORIGINAL RESEARCH
published: 08 January 2021
doi: 10.3389/fpubh.2020.612410

Quo Vadis HTA for Medical Devices in Central and Eastern Europe? Recommendations to Address Methodological Challenges

*Rita Daubner-Bendes[1], Sándor Kovács[1,2], Maciej Niewada[3], Mirjana Huic[4], Michael Drummond[5], Oriana Ciani[6,7], Carl Rudolf Blankart[8,9], Olena Mandrik[10], Aleksandra Torbica[6,11], John Yfantopoulos[12], Guenka Petrova[13], Malwina Holownia-Voloskova[3,14], Rod S. Taylor[7,15], Maiwenn Al[16], Oresta Piniazhko[17], László Lorenzovici[18,19], Rosanna Tarricone[6,11], Antal Zemplényi[1,20] and Zoltán Kaló[1,21]**

[1] Syreon Research Institute, Budapest, Hungary, [2] Centre for Health Technology Assessment, University of Pécs, Pécs, Hungary, [3] Department of Experimental and Clinical Pharmacology, Medical University of Warsaw, Warsaw, Poland, [4] HTA/EBM Consulting Centre, Zagreb, Croatia, [5] Centre for Health Economics, University of York, York, United Kingdom, [6] Centre for Research on Health and Social Care Management, SDA Bocconi School of Management, Milan, Italy, [7] Evidence Synthesis and Modelling for Health Improvement, College of Medicine and Health, Institute of Health Research, University of Exeter, Exeter, United Kingdom, [8] KPM Center for Public Management, University of Bern, Bern, Switzerland, [9] sitem-insel AG, Swiss Institute for Translational and Entrepreneurial Medicine, Bern, Switzerland, [10] School of Health and Related Research, Health Economics and Decision Science, The University of Sheffield, Sheffield, United Kingdom, [11] Department of Social and Political Science, Bocconi University, Milan, Italy, [12] School of Economics and Political Science, University of Athens, Athens, Greece, [13] Department of Social Pharmacy and Pharmacoeconomics, Faculty of Pharmacy, Medical University of Sofia, Sofia, Bulgaria, [14] Health Technology Assessment Department, State Budgetary Institution "Research Institute for Healthcare Organization and Medical Management of Moscow Healthcare Department", Moscow, Russia, [15] MRC/CSO Social and Public Health Sciences Unit and Robertson Centre for Biostatistics, Institute of Health and Well Being, University of Glasgow, Glasgow, United Kingdom, [16] Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, Netherlands, [17] HTA Department of State Expert Centre of Ministry of Health of Ukraine, Kyiv, Ukraine, [18] G. E. Palade University of Medicine, Pharmacy, Science and Technology, Tirgu Mures, Romania, [19] Syreon Research Romania, Tirgu Mures, Romania, [20] Division of Pharmacoeconomics, Faculty of Pharmacy, University of Pécs, Pécs, Hungary, [21] Centre for Health Technology Assessment, Semmelweis University, Budapest, Hungary

OPEN ACCESS

Edited by:
Mihajlo Jakovljevic, Hosei University, Japan

Reviewed by:
Tomas Tesar, Comenius University, Slovakia
Georgi Iskrov, Plovdiv Medical University, Bulgaria
Martin Dlouhy, University of Economics, Czechia

***Correspondence:**
Zoltán Kaló
kalo.zoltan@semmelweis-univ.hu

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *30 September 2020*
Accepted: *07 December 2020*
Published: *08 January 2021*

Citation:
Daubner-Bendes R, Kovács S, Niewada M, Huic M, Drummond M, Ciani O, Blankart CR, Mandrik O, Torbica A, Yfantopoulos J, Petrova G, Holownia-Voloskova M, Taylor RS, Al M, Piniazhko O, Lorenzovici L, Tarricone R, Zemplényi A and Kaló Z (2021) Quo Vadis HTA for Medical Devices in Central and Eastern Europe? Recommendations to Address Methodological Challenges. Front. Public Health 8:612410. doi: 10.3389/fpubh.2020.612410

Objectives: Methodological challenges in the evaluation of medical devices (MDs) may be different for early and late technology adopter countries, as well as the potential health technology assessment (HTA) solutions to tackle them. This study aims to provide guidance to Central and Eastern European (CEE) countries on how to address key challenges of HTA for MDs with special focus on the transferability of scientific evidence.

Methods: As part of the COMED Horizon 2020 project, a comprehensive list of issues related to MD HTA were identified based on a targeted literature review. Health technology assessment issues which pose a greater challenge or require different solutions in late technology adopter countries were selected. Draught recommendations to address these issues were developed and discussed in a focus group. The recommendations were then validated with a wider group of experts, including HTA and reimbursement decision makers from CEE countries in May and June 2020.

Results: A consolidated list of 11 recommendations were developed in 3 major areas: (1) clinical value assessment, focusing on the use of joint EU work, relying on real-world evidence, use of coverage with evidence development schemes, transferring

evidence from foreign countries and addressing the challenges of learning curve and centre effect; (2) economic value assessment, covering cost calculation of complex medical devices and transferability of economic evaluations of MDs; (3) HTA processes, related to the frequent product modifications and various indications of MDs.

Conclusions: Central and Eastern European countries with limited resources for conducting HTA, can benefit from HTA methods and evidence generated in early technology adopter countries. Considering the appropriate reuse of international HTA materials, late technology adopter countries can still implement HTA, even for MDs, which have a more limited evidence base compared with pharmaceuticals.

Keywords: medical device (MD), health technology assessement (HTA), methodological challenge, transferability, value assessment, Central-Eastern Europe (CEE), real world evidence (RWE)

INTRODUCTION

The use of HTA to improve the evidence base of health policy decisions has been increasing across Europe. In almost all countries where HTA has been introduced, it is mainly used to support public coverage decisions of new pharmaceuticals, partly because the availability of scientific evidence for medicines is quite substantial compared with other technologies. However, despite challenges (1), HTA has also been increasingly used to support coverage decision of MDs, which necessitates specific methodological guidelines for MDs (2).

Development and improvement of HTA methodologies is on the top agenda of the European Union (EU). The EU addressed the topic in several large scale projects financed through 7th Framework and the Horizon 2020 Research Programmes: the MedtechHTA project (3) was financed through 7th Framework Research Program, and the ongoing COMED project is financed through the Horizon 2020 Program. The COMED (Pushing the boundaries of Cost and Outcome analysis of Medical Technologies) project has multiple objectives. First, it aims to improve methods for economic evaluation for medical devices by addressing most relevant challenges in HTA of medical devices, second, to investigate health system performance through analysis of variation in access to medical technologies across different geographical areas; and finally, to strengthen the use of economic evaluation of MDs in policy making (4).

The COMED project put emphasis on extending the current knowledge on the transferability of HTA methodologies and reports for MDs especially across countries with different economic status. This is highly needed, as HTA implementation roadmaps of higher income Western European (WE) countries may not be applicable in lower income European countries, especially in Central and Eastern Europe for several reasons. The health status of population in Central and Eastern European (CEE) countries is significantly worse compared with WE countries (5, 6), which indicates greater demand for health technologies with substantial incremental health gain. On the other hand, the availability of public health care budgets to cover new technologies—including MDs—is more limited in CEE countries. Therefore, these countries incur a higher opportunity cost for inappropriate, not evidence-based policy decisions.

Central and Eastern European countries have more limited financial capacities for health technology assessment, furthermore, there is lack of sufficient human resources and expertise due to scarcity of post-graduate programs (7). Due to the lower market potential related to budget limitations, manufacturers tend to launch their new MDs later in CEE than in WE countries, which creates an opportunity for making use of accessible HTA reports prepared by influential HTA institutes in early adopter countries of new health technologies.

While the methodological challenges of WE and CEE countries in the evaluation of MDs may be similar, potential HTA solutions may be different due to the abovementioned reasons. This document focuses on lower income and/or small-size European countries which are late adopters of medical devices due to their limited market potential (8), with special focus on CEE EU member states. Our objectives are (1) to explore priority issues in conducting HTA for MDs, which are either specific to CEE countries, or are present in all EU countries but present greater challenges, or require different solutions in CEE countries; and (2) to provide guidance to CEE countries on how to address these challenges of MDs with special focus on the transferability of scientific evidence.

Given these objectives, this document focuses on HTA challenges that are more specific to medical devices as compared with pharmaceuticals. Also, HTA challenges that are equally relevant, and require similar solutions in early and later technology adopter countries, are not in our scope. Finally, it is not our intention to provide recommendations on how priority setting should be implemented at the national level, including which MDs should be selected for mandatory HTA prior to policy decisions and what should be the role of HTA in the pricing and reimbursement of MDs.

On the other hand, MDs usually subjected to HTA in the majority of countries are those which can deliver incremental effectiveness or safety to patients. Therefore, our recommendations may not be relevant to those decision support MDs (e.g., shared decision-making tools, digital platforms), which do not promise direct health gain.

MATERIALS AND METHODS

A targeted literature review was conducted in Scopus to identify a comprehensive list of issues of medical device HTA. A snowball method was also used to identify further relevant studies among the references of papers with full text review. The search strategy used with the date of the search and number of hits are summarised in **Supplementary Material 1**.

In a series of iterative brainstorming sessions, the results of the literature review were discussed with senior experts with experience in the HTA of MDs. Structured discussion was conducted about all issues where the experts had to judge which issues are relevant for European countries which are typically late adopters of MDs. The relevant issues were selected by considering certain requirements, such as non-redundancy, non-overlap and preference independence. Issues that were equally relevant to all European countries or do not require specific recommendations in CEE were excluded. The iterative discussion was continued until all participants agreed with the inclusion of the specific issue.

The iterative discussions were also used to propose draught recommendations on how to manage the identified issues. To provide sufficient details while preserving clarity, each recommendation had a short form and a detailed description.

The original plan was to refine and validate the draught list of recommendations in consecutive meetings with senior international experts. However, in the COVID-19 outbreak period we had to convert face-to-face meetings to virtual platforms. Participants to these virtual meetings were identified in an iterative process exploiting the professional networks of COMED partners. The main selection criteria were familiarity with the HTA ecosystem for MDs in late technology adopter countries with efforts for maintaining balanced geographical distribution of participants.

In April 2020 the list of issues and draught recommendations were presented to a small group of HTA experts with familiarity on national pricing and reimbursement processes of MDs in CEE countries in a virtual focus group meeting. After a thorough discussion, the descriptions of issues and recommendations were improved based on the consensus of participants.

The recommendations were validated with a wider group of experts, including HTA and reimbursement decision makers from CEE countries in May and June 2020. The validation process started with a webinar to present the issues and recommendations to a wider group of experts. Then, participants of the focus group meeting and the webinar in addition to COMED consortium members were asked to provide written feedback on draught recommendations. Finally, a virtual interactive meeting was organised to discuss the written feedback and facilitate consensus among participants.

Overall, 31 experts outside from the COMED consortium representing 14 CEE countries (Bulgaria, Croatia, Czech Republic, Greece, Hungary, Malta, Poland, Romania, Russia, Serbia, Slovakia, Slovenia, Turkey and Ukraine) contributed to the deliberative process.

RESULTS

Targeted Literature Review

A total of 563 records were identified by the literature search. After removal of duplicates and title-abstract screening, 20 articles were found to be eligible for this review. Integrating with the 8 extra articles identified by the snowball method, in total, 28 articles qualified for a full-text review, and 19 were considered eligible for the qualitative synthesis. The numbers of identified, screened and excluded papers are shown in the flowchart in **Supplementary Material 2**.

In the targeted literature review process, 33 issues were identified with potential relevance for medical device HTA in late adopter European countries. The 33 issues were merged and reduced to 9 important issues with special relevance to CEE countries after iterative brainstorming sessions (**Supplementary Material 3**). The 9 issues were grouped to clinical value assessment, economic value assessment and HTA process.

Challenges and Recommendations for HTA of MDs in Late Adopter Countries

Issue 1: Clinical Value Assessment

Challenge 1.1: Lower level of evidence for MDs

Evidence base is relatively limited for the majority of MDs compared with pharmaceuticals due to several reasons. Regulatory agencies do not mandate confirmatory efficacy and safety data on MDs from randomised clinical trials (RCTs) (1, 9–14). Hence, a substantial proportion of HTA reports rely on surrogate outcomes in estimating the health gain, often without proper validation (15). Although the recent EU regulation (2017/745 of the European Parliament) (16) made RCTs mandatory for Class IIb and III MDs, it is unforeseen whether the evidence base of MDs can be increased to the level of pharmaceutical, partly because blinding and proper randomisation cannot be implemented for all MDs (2). In several cases clinical trials for MDs are not adequately powered (i.e., sample size is too small) and/or the follow-up period is too short for performing HTA analysis (15, 17–19). Finally, MDs are often introduced into clinical practise quickly, especially in high income countries with great market potential, often even before the initiation of clinical trials (17). Then, when the MD is introduced in late adopter CEE countries, it is no longer feasible, or ethically justifiable to carry out RCTs (17).

Recommendation #1.1.a: Use relative effectiveness and safety assessment from joint EU work or use rigorous relative assessment from other jurisdictions. Relative effectiveness and safety assessments (RESAs) from the joint work of EU member states should be the starting point for CEE countries to judge the clinical effectiveness and safety of a MD (20). If this is not available, rigorous RESAs from other early adopter countries may be transferred to CEE countries. When relying on results from other countries, the absolute risk reduction should be adjusted based on the local baseline risk and uncertainty of the parameters should be explored with sensitivity analysis.

Iterative reassessment should be considered when better evidence becomes available.

Recommendation #1.1.b: Rely on real-world evidence, when evidence from explanatory randomised clinical trials is limited. If evidence from explanatory RCTs is limited, the best available evidence should be synthetized with special attention to real-world evidence (RWE) (21). Comparative observational studies, patient registries or claims databases are examples of real-world data (RWD) that can produce RWE on the effectiveness and safety of MDs. However, appropriate methods of bias-adjustment (e.g. counterfactual effect estimation) need to be applied (22). RWE from early adopter countries may be transferred to CEE countries.

Recommendation #1.1.c: Consider coverage with evidence development, when the scientific evidence from randomised clinical trials and real-world is premature. A possible solution for the decision-makers to deal with uncertainty is to apply coverage with evidence development (CED) schemes in a specified patient population tracked over a defined period of time under explicit requirement of further evidence generation, where the level or continuation of reimbursement is based on the clinical and economic outcomes achieved. Due to the administrative and financial burden of implementing CED schemes in countries with limited resources, they should be applied for MDs in areas with high unmet medical need and public health priority. CED schemes from early adopter countries may be transferred to CEE countries.

Challenge 1.2: Limited transferability of real-world data and real-world evidence from foreign countries

Even though RCTs may not reflect the effectiveness of MDs in the real-world, they still represent the gold standard of evidence generation for new technologies (23). However, more and more stakeholders are exploring new areas of evidence generation based on real-world data since rigour for its collection has improved and standardisation has been initiated (24). While RWD can increase the evidence base of health technologies, RWD cannot completely replace RCTs. For quality assessment purposes RWD and the derived RWE are virtually inseparable (25–27). Due to limited availability of efficacy data from explanatory RCTs and the methodological principles intrinsic to RCTs (28, 29), the relative weight of RWD in the evidence generation of MDs is greater compared with pharmaceuticals.

In CEE countries where MDs are introduced in later life cycle stages, HTA of MDs could be accelerated through adapting RWD to local jurisdictions. However, CEE countries have limited power to access the original RWD from other countries (10, 30). Therefore, transferability assessment cannot focus on RWD, but on RWE.

It is highly important to note, that RWE is more subject to local factors compared with protocol driven RCTs, partly because the benefit of MDs is dependent on capacity constraints for example in the follow-up care and heterogeneity of patient pathways. Consequently, transferability of RWE is more limited than transferability of evidence from RCTs.

Recommendation #1.2: Explore the feasibility of transferring RWE to late technology adopter countries in a stepwise approach. In order to overcome the uncertain quality and local relevance of RWD, a stepwise approach is recommended to explore the feasibility of transferring international RWE to CEE countries. The first step should be the systematic search and collection of potentially relevant RWEs, followed by an evaluation of the equivalence of technologies presented in the study of origin and local context. As a result, studies in which the technologies have no major differences should be channelled into further assessment. In the next step, the quality of evidence should be evaluated by using quality assessment tools, such as GRADE (31, 32), ROBINS-I (33) or the ISPOR initiative on RWE transparency (34) and include only studies with appropriate quality for quantitative synthesis (e.g., meta-analysis). As a last step to guarantee applicability of evidence in local jurisdiction, use of a transferability checklist [e.g., EUnetHTA Adaptation Toolkit (35)] should be considered to evaluate the variation in patient population, medical practise and health systems. In CEE countries with capacity constraints in the follow-up care and heterogeneous patient pathways, highly conservative effectiveness estimates should be applied. On the other hand, limited variability in the real-world outcomes across centres in the country of origin indicates less dependence on local factors, and consequently increased transferability of high-quality RWE with appropriate comparator can be assumed to late technology adopter countries.

Challenge 1.3: Limited transferability of surrogate endpoints

Surrogate endpoints (36) are intended to replace patient/clinical relevant final endpoints, providing the possibility of using indirect measurement in cases when direct measurement of long-term clinical effects is not feasible. However, reliability of surrogate outcomes in predicting effects on clinically meaningful outcomes has to be validated, i.e., the treatment effect on a surrogate endpoint needs to be predictive of the treatment effect on the final clinical outcomes (15, 37). Due to differences in populations and health care systems, surrogate outcome predictions from other jurisdictions may not be directly transferable to CEE countries without adjustment.

Recommendation #1.3: Reuse internationally validated surrogate endpoints with extensive sensitivity analyses. Reuse of internationally validated surrogate endpoints with extensive sensitivity analyses is highly recommended in countries with limited HTA resources. The use of surrogates should be limited to HTAs using the same intervention, class of technology and comparator as the validation. If modelled to determine cost-effectiveness, the uncertainty of the extrapolation of surrogates to final outcomes should be fully considered.

If the surrogate endpoint is endorsed by major regulatory agencies (e.g., FDA) or HTA bodies (e.g., NICE, HAS, IQWIG, CADTH) it should be accepted in the base case scenario. International predictions for long-term outcomes should be complemented with extensive sensitivity analysis by considering a wide range of scenarios based on local factors.

If major regulatory agencies or HTA bodies have not endorsed the surrogate outcome, consider the best available scientific evidence with special focus on surrogates validated by international clinical societies for local assessment with more conservative estimation of long-term benefits.

Challenge 1.4: Learning curve

The effectiveness of many MDs depends not only on the device itself, but also on how it is used (17, 18, 38). A learning curve is often associated with a MD, as user skills and training with the new technology can have important implications on clinical outcomes (1, 2, 13, 29, 39, 40). Central and Eastern European countries are generally late adopters, so evidence on learning curves from more developed health systems may be available and can be adapted for the assessment of a MD in a local context.

On the other hand, health care provision in CEE countries is traditionally less standardised (i.e., protocol driven) and due to resource constraints, "*ad hoc*" solutions are more considered. During the launch period of new MDs, the support of manufacturers in providing training or collecting outcomes data in patient registries is often less intensive in countries with limited market potential. These factors are expected to contribute to longer learning curves, which can result in inferior outcome initially due to incremental safety problems or treatment failures in CEE compared with more advanced health care systems.

Recommendation #1.4: In the introductory period of MDs consider inferior effectiveness and safety (1) based on learning curves from other countries (2) and by using Bayesian approach. Learning curve should be considered in the HTA of MDs. If local evidence is not available, the transferability of learning curve results from early adopter countries should be assessed. In more resource constrained health systems in CEE countries, differences in the length and shape of learning curves need to be considered, and sensitivity analysis should be used to assess different plausible scenarios. Another recommended method to adjust the expected learning curves is the use of expert opinion on the magnitude of inferior outcomes due to incremental safety problems or complication rates in CEE countries, channelled in the evaluation process using a Bayesian approach with informative priors. The estimations can be updated once local evidence is available. Shorter learning curves can be assumed, if tailor-made manufacturer support is available for the medical institution, for health care professionals, or for patients (e.g., in case of self-administered MDs, such as wearables) during the introductory period.

Challenge 1.5: Centre effect

Several external factors related to the ability of health care professionals and institutes can influence the clinical effectiveness of MDs. National health care systems are organised into different centres, which obviously differ from each other. The centre effect should be considered in a hierarchical order, as substantial variation in provider skills may reside at the individual health care professional, team or hospital level (13, 18, 41).

The institutional environment like the placement and size of the centre, has dual impact on outcomes delivered by the medical device. On one hand, the limited number of patients in smaller centres can limit the ability of individuals to acquire appropriate competencies, which extends the learning curve preventing the maintenance of required skills to achieve maximum performance. On the other hand, large volume and diversity of services can influence the ability of institutes to exploit maximum effectiveness of MDs. In this regard, economies of scale in large specialised centres accelerates the adoption of MDs (42) and allows institutes to employ more specialised operators and ancillary staff, which ultimately improves outcomes and reduces the average cost per case. In larger centres, economies of scope with broader portfolio of services may improve health outcomes in complicated cases (e.g., multimorbid patients) by ensuring access to specialty services.

Recommendation #1.5: Consider the relative effectiveness and safety of MDs in large volume centres with licenced health care professionals. Since scale and scope of services matter, primary judgement on the relative effectiveness and safety of MDs should focus on large volume centres with licenced health care professionals for a specific device. By introducing a threshold on a minimum number of patients to be treated at the institution and individual operator level, limiting the number of centres and by defining quality standards, the outcomes variation related to centre effect can be reduced. These quality assurance tools may be transferred from early adopter countries. Furthermore, by initiating a medical licencing system, the maintenance of required expertise can be guaranteed. In countries with decentralised health care system and small volume centres, inferior effectiveness should be assumed, especially where licenced health care professionals for the specific medical device are not available. Additionally, an HTA model grounded outcome-based risk sharing system can directly incentivize health outcome improvement.

Issue 2: Economic Value Assessment

Challenge 2.1: Cost calculation of complex medical devices for cost-effectiveness and budget impact analyses

The price transparency of complex MDs is limited due to several factors. At first manufacturers of MDs can collect income through different channels, and the prices for the equipment, consumables and maintenance depend on strategic considerations by manufacturers. At health care institutes the equipment represents fixed costs and booked among capital costs, while consumables and *ad hoc* maintenance expenses are variable costs and booked among operating costs. As the average cost is dependent on the economies of scale, it is not easy to calculate the actual cost per intervention of complex MDs with high upfront costs and unpredictable maintenance costs (43).

In addition, differences in local production functions among hospitals further increases the variability of actual cost, especially since resource utilisation within a hospital can also change over time due to organisational changes and experience of health care professionals. Finally, cost of necessary initial and continuous training of health care professionals should also be taken into account, when calculating the cost per case of complex MDs. On the other hand, the availability of public health care budgets

to cover new technologies—including MDs—is more limited in CEE countries, and due to less advanced control mechanism, prices of MDs in CEE countries may even be higher than in more affluent WE countries.

Recommendation #2.1: HTA for MDs should be considered primarily for national reimbursement decisions or centralised procurement by taking into account average expected payments, rather than actual costs. In countries with limited HTA resources, HTA for MDs should primarily focus on population level policies such as national reimbursement or centralised procurement decisions. To support the evidence base of national decisions, cost calculations should not depend on non-transparent price components, provider specific production functions and economies of scale. Consequently, instead of actual costs, the payment per case requested by health care providers or paid by health care payers (e.g., fees or charges) should be used in economic evaluations and budget impact analyses. If the focus of the national HTA assessment is the procedure completed by the MD rather than the MD itself, the payment for the procedure should be used, which covers costs of the MD, institutional cost and personnel cost as well.

Due to scarcity of HTA capacities at the local level, MD procurement decisions by local institutions may not be supported by full scope HTA reports. Still, decision-makers at local hospitals should be able to judge whether the payment per case by health care payers covers their expected local cost per case. This necessitates return on investment and financial sustainability calculation, which should be based on actual cost calculations by taking into account local production functions and economies of scale. On the other hand, if a CEE country still prefers hospital based HTA in addition to national assessments, each hospital should be able to calculate the actual cost of MDs per case from their own perspective, which may not be transferable to other hospitals.

Challenge 2.2: Limited transferability of economic evaluation of MDs

Being late adopters of MDs, CEE countries can potentially rely on HTA recommendations of early adopter countries. However, even if a given technology considered cost-effective in a WE country, it does not justify a positive recommendation in more resource constrained CEE countries, partly due to differences in the cost-effectiveness thresholds (44, 45). In addition, the variability in the incremental cost-effectiveness ratio of health technologies within different jurisdictions is documented in several publications (46–48). Therefore, transferring results and conclusions of an economic evaluation conducted in a higher income country along with the differences in patient population, comparator, patient pathways, outcomes, resource utilisation and unit costs and potential differences in the device itself represents several challenges (11, 49).

Recommendation #2.2: Adapt international economic models from early technology adopter countries after transferability assessment. In order to overcome the challenges of economic model adaptation from early adopter to CEE countries, a stepwise process is recommended. As a first step, a transferability of the model should be evaluated in particular to local relevance of the model concept and focusing on comparator, patient pathways and long-term outcome estimation. After adjusting the model to local circumstances, the use of country specific costs as a mandatory step is recommended. This should incorporate local patient pathways, average resource use in large centres and local unit cost. During the estimation of health outcome, as much local data as possible should be included to increase the relevance of the prediction. Discount rate, time-horizon of the assessment and other necessary input parameters should be adjusted according to the local methodological guideline. As a last step, a comprehensive sensitivity analysis is recommended to explore the effect of parameter uncertainty.

Issue 3: HTA Process

Challenge 3.1: Frequent product modifications and dynamic pricing

Unlike drugs, MDs often undergo several product modifications over time, which may have an impact on health gain, costs and patient experience (50). In late technology adopter CEE countries, often modified versions of MDs are introduced and have to be evaluated by HTA. In addition, prices of MDs more often change over time due to the market entry of new products (17, 39), iterative incremental developments over time (1, 2, 18, 39), and flexible procurement practises (17, 39). This raises the question of whether and when these small improvements have to be re-evaluated.

Recommendation #3.1: HTA should not be performed for a particular version of a MD, but for the group of devices with the same (or similar) characteristics. Full HTA is needed to support the initial national reimbursement decision or centralised procurement of a new MD. However, no HTA (or only fast track HTA) is necessary if (1) a different manufacturer introduces a similar MD with non-inferiority and no price increase; (2) the price of a reimbursed MD is reduced; (3) an improved MD emerges without price increase. A full HTA is still recommended, if total cost of the procedures with a modified MD is increased due to added value. A conservative approach should be followed when assessing the added clinical value of modified MDs.

Challenge 3.2: Diverse and numerous clinical indications of MDs

Several complex medical devices such as radiotherapy, robotic surgery, imaging diagnostics, can be used for heterogeneous patient populations in multiple indications. However, in countries with limited HTA resources it is not possible to conduct full scope HTA in all different indications.

Recommendation #3.2: Full scope HTA may not be necessary in each potential indication, cost-effectiveness results in the most prevalent indications can be generalised to indications with similar expected health benefits. After exploring all potential indications, the list should be narrowed to those indications which are candidates for reimbursement by excluding indications with no assumed incremental health gain or with low public priority.

TABLE 1 | Consensus recommendations related to selected challenges of medical device HTA in late adopter countries.

Major issues	Summary of challenges and recommendations
1. Clinical value assessment (effectiveness and safety)	1.1 Lower level of evidence for MDs - Use relative effectiveness and safety assessment from joint EU work or use rigorous relative assessment from other jurisdictions - Rely on real-world evidence, when evidence from explanatory randomised clinical trials is limited - Consider coverage with evidence development, when the scientific evidence from randomised clinical trials and real-world is premature
	1.2 Limited transferability of real-world data and real-world evidence from foreign countries - Explore the feasibility of transferring real-world evidence to late technology adopter countries in a stepwise approach
	1.3 Limited transferability of surrogate endpoints - Reuse internationally validated surrogate endpoints with extensive sensitivity analyses
	1.4 Learning curve - In the introductory period of medical devices consider inferior effectiveness and safety (1) based on learning curves from other countries (2) and by using Bayesian approach
	1.5 Centre effect - Consider the relative effectiveness and safety of medical devices in large volume centres with licenced health care professionals
2. Economic value assessment (cost calculation, cost-effectiveness, budget impact)	2.1 Cost calculation of complex medical devices for cost-effectiveness and budget impact analyses - HTA for medical devices should be considered primarily for national reimbursement decisions or centralised procurement by taking into account average expected payments (e.g., fee or charges) rather than actual costs
	2.2 Limited transferability of economic evaluation of medical devices - Adapt international economic models from early technology adopter countries after transferability assessment
3. HTA process	3.1 Frequent product modifications and dynamic pricing - HTA should not be performed for a particular version of a medical device, but for the group of devices with the same (or similar) characteristics
	3.2 Diverse and numerous clinical indications of medical devices - Full scope HTA may not be necessary in each potential indication, cost-effectiveness results in the most prevalent indications can be generalised to indications with similar expected health benefits

MDs, medical devices; HTA, health technology assessment.

Full scope HTA with detailed cost-effectiveness analysis has to be conducted in the most prevalent indications. If the MD is cost-effective in the most prevalent indications, in case of limited HTA capacities, it is an acceptable compromise to generalise cost-effectiveness evidence in the most prevalent indications to other indications with similar estimated health gain. However, budget impact analysis is needed for the entire target patient population eligible for public coverage.

The recommendations described above are summarised in **Table 1**.

DISCUSSION

Those European countries which are late adopters of new technologies usually have more limited resources not only for covering high-cost innovative technologies from public resources, but also for conducting HTA to substantiate their health policy decisions with appropriate evidence base. However, delays in the uptake of new technologies create an opportunity to benefit from HTA methods and reports generated in early technology adopter countries. Therefore, considering the appropriate reuse of international HTA materials, late technology adopter countries can still implement HTA, even for MDs, which have more limited evidence base compared with pharmaceuticals.

As a general recommendation, CEE countries should be encouraged to contribute to joint work to extend the evidence base of MDs, including generation of RWE, validation of surrogate endpoints and exploratory research on learning curves. The increased participation of CEE centres in international collaborative research projects can improve the knowledge on the transferability of scientific evidence related to MDs. CED schemes should also be considered in CEE, however, further research is needed on which elements of CED schemes from early adopter countries are transferable to late adopter countries.

Recommendations given in this report may provide guidance for policy-makers on how standards for HTA of MDs should be improved in lower income countries. Although adaptation of these recommendations should remain in national competencies, if these proposals are translated into routine practise, manufacturers will have more clarity on what they need to deliver to facilitate the market access of their new MDs in countries with less market potential. Recommendations are not equally relevant for all different types of MDs.

Ultimately, we hope that applying these recommendations can lead to better care for patients with greater need for health improvement in more deprived European countries by reducing the opportunity cost of inappropriate coverage decisions.

Although in our guidance we focused mainly on lower income European countries especially in CEE, these recommendations can be valid to any lower income country outside Europe,

which are also late adopters of MDs due to their limited market potential, and also to any small country, which have limited human and financial resources to support coverage decisions with HTA evidence.

It should be noted that our recommendations are based primarily on the opinion of relatively small-sized expert group, which is the most important limitation of our study. However, focusing on commonalities and not differences across countries with highly experienced HTA professionals in the deliberative process facilitated the replicability of our conclusions. Still, our recommendations cannot be equally generalizable to all different categories of MDs. Hence future research, especially at the national level targeting different types of MDs, should test the appropriateness of our recommendations, which may necessitate additional recommendations for specific groups of MDs.

Overall, our guidance should be viewed only as a first step in a multi-stakeholder dialogue about HTA practises of MDs in lower income European countries, which can be strengthened by voluntary regional collaboration in HTA.

DATA AVAILABILITY STATEMENT

The original contributions generated in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

RD-B, SK, AZ, and ZK iteratively designed the study methodology, developed the initial guidance, reviewed and edited the materials. RD-B and SK completed on the targeted literature review. AZ and ZK undertook conceptualisation and supervised the entire project. MD, OC, RodT, CB, AT, and RosT developed those recommendations in the COMED consortium, which had to be addressed from the perspective of CEE countries interactively with the Syreon team. MD, OC, RodT, CB, and MA extensively commented each version of the manuscript. MD contributed to the conceptual aspects from the initial phase, took major role in the virtual workshops, contributed to the parts about coverage with evidence development. OC and RodT contributed to the parts about surrogate endpoints. CB contributed the parts related to early assessment of medical devices. MA contributed to the parts about learning curve. AT and RosT are the thought leaders of COMED consortium, supported the transferability exercise throughout the project with presentations and extensive review of methodological aspects and draught reports. MN, MH, OM, JY, GP, MH-V, OP, and LL took part in the validation process, reviewed the recommendations from their own stakeholder and geographical perspectives, and commented extensively each manuscript version. All authors approved the submitted paper.

FUNDING

This project received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement #779306 (COMED—Pushing the Boundaries of Cost and Outcome Analysis of Medical Technologies). The results reflect only the authors' views, and the EU is not responsible for any use that may be made of the information it contains.

ACKNOWLEDGMENTS

Contribution of Kagan Atikeler, Dragana Atanasijevic, Veronika Dóczy, Ramóna Hegyi, Rok Hren, Maria Kamusheva, Katarzyna Kolasa, Olga Löblova, Gergo Merész, Nuriya Musina, Vitaly Omelyanovskiy, Piotr Ozieranski, Tünde Péter, Benedetta Pongiglione, Cristina Rais, Bernardette Rossi, Andrea Székely, Tomas Tesar, Áron Vinczíczki, Vladimir Zah during the online validation meetings is gratefully acknowledged.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2020.612410/full#supplementary-material

REFERENCES

1. Tarricone R, Callea G, Ogorevc M, Prevolnik Rupel V. Improving the methods for the economic evaluation of medical devices. *Health Econ*. (2017) 26(Suppl. 1):70–92. doi: 10.1002/hec.3471
2. Blüher M, Saunders SJ, Mittard V, Torrejon Torres R, Davis JA, Saunders R. Critical review of european health-economic guidelines for the health technology assessment of medical devices. *Front Med*. (2019) 6:278. doi: 10.3389/fmed.2019.00278
3. MedtecHTA. *Methods for Health Technology Assessment of Medical Devices: a European Perspective (MedtecHTA)*. (2013). Available online at: www.medtechta.eu/wps/wcm/connect/site/medtechta/home (accessed July 9, 2020).
4. COMED. *Pushing the Boundaries of Cost and Outcome Analysis of Medical Technologies (COMED)*. (2018). Available online at: www.comedh2020.eu (accessed July 9, 2020).
5. Zatonski W. The east-west health gap in europe-what are the causes? *Eur J Public Health*. (2007) 17:121. doi: 10.1093/eurpub/ckm006
6. Karanikolos M, Adany R, McKee M. The epidemiological transition in Eastern and Western Europe: a historic natural experiment. *Eur J Public Health*. (2017) 27(Suppl. 4):4–8. doi: 10.1093/eurpub/ckx158
7. Kaló Z, Gheorghe A, Huic M, Csanádi M, Kristensen FB. HTA implementation roadmap in Central and Eastern European Countries. *Health Econ*. (2016) 25(Suppl. 1):179–92. doi: 10.1002/hec.3298
8. World Health Organization. *Barriers to Innovation in the Field of Medical Devices: Background Paper 6 Geneva: World Health Organization*. (2010). Available online at: https://apps.who.int/iris/handle/10665/70457 (accessed July 9, 2020).
9. Beck A, Retel VP, Bhairosing PA, van den Brekel M, van Harten WH. Barriers and facilitators of patient access to medical devices in Europe: a systematic literature review. *Health Policy*. (2019) 123:1185–98. doi: 10.1016/j.healthpol.2019.10.002

10. Ciani O, Wilcher B, Blankart CR, Hatz M, Rupel VP, Erker RS, et al. Health technology assessment of medical devices: a survey of non-European Union Agencies. *Int J Technol Assess Health Care*. (2015) 31:154–65. doi: 10.1017/S0266462315000185
11. Drummond M, Griffin A, Tarricone R. Economic evaluation for devices and drugs–same or different? *Value Health*. (2009) 12:402–4. doi: 10.1111/j.1524-4733.2008.00476_1.x
12. Fuchs S, Olberg B, Panteli D, Perleth M, Busse R. HTA of medical devices: challenges and ideas for the future from a European perspective. *Health Policy*. (2017) 121:215–29. doi: 10.1016/j.healthpol.2016.08.010
13. Schnell-Inderst P, Mayer J, Lauterberg J, Hunger T, Arvandi M, Conrads-Frank A, et al. Health technology assessment of medical devices: what is different? An overview of three European projects. *Zeitschrift fur Evidenz, Fortbildung und Qualitat im Gesundheitswesen*. (2015) 109:309–18. doi: 10.1016/j.zefq.2015.06.011
14. Tarricone R, Torbica A, Ferré F, Drummond M. Generating appropriate clinical data for value assessment of medical devices: what role does regulation play? *Expert Rev Pharmacoecon Outcomes Res*. (2014) 14:707–18. doi: 10.1586/14737167.2014.950233
15. Ciani O, Wilcher B, van Giessen A, Taylor RS. Linking the regulatory and reimbursement processes for medical devices: the need for integrated assessments. *Health Econ (United Kingdom)*. (2017) 26:13–29. doi: 10.1002/hec.3479
16. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (2017). Available online at: http://data.europa.eu/eli/reg/2017/745/oj (accessed November 9, 2020).
17. Sorenson C, Tarricone R, Siebert M, Drummond M. Applying health economics for policy decision making: do devices differ from drugs? *Europace*. (2011) 13(Suppl. 2):ii54–8. doi: 10.1093/europace/eur089
18. Gelijns AC, Russo MJ, Hong KN, Brown LD, Ascheim DD, Moskowitz AJ. Dynamics of device innovation: implications for assessing value. *Int J Technol Assess Health Care*. (2013) 29:365–73. doi: 10.1017/S0266462313000561
19. Schnell-Inderst P, Hunger T, Conrads-Frank A, Arvandi M, Siebert U. Ten recommendations for assessing the comparative effectiveness of therapeutic medical devices: a targeted review and adaptation. *J Clin Epidemiol*. (2018) 94:97–113. doi: 10.1016/j.jclinepi.2017.09.022
20. Németh B, Kaló Z. European cooperation in health technology assessment implementation: the perspective of Central and Eastern European countries. *J Comp Eff Res*. (2020) 9:599–602. doi: 10.2217/cer-2020-0062
21. Berger ML, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Pharmacoepidemiol Drug Saf*. (2017) 26:1033–9. doi: 10.1002/pds.4297
22. Tarricone R, Boscolo PR, Armeni P. What type of clinical evidence is needed to assess medical devices? *Eur Respir Rev*. (2016) 25:259–65. doi: 10.1183/16000617.0016-2016
23. Hariton E, Locascio JJ. Randomised controlled trials—the gold standard for effectiveness research: study design: randomised controlled trials. *BJOG*. (2018) 125:1716. doi: 10.1111/1471-0528.15199
24. Khosla S, White R, Medina J, Ouwens M, Emmas C, Koder T, et al. Real world evidence (RWE)—a disruptive innovation or the quiet evolution of medical evidence generation? *F1000Research*. (2018) 7:111. doi: 10.12688/f1000research.13585.1
25. Berger M, Daniel G, Frank K, Hernandez A, McClellan M, Okun S, et al. *A Framework for Regulatory Use of Real-world Evidence*. (2017). Available online at: healthpolicy.duke.edu/sites/default/files/atoms/files/rwe_white_paper_2017.09.06.pdf (accessed July 9, 2020).
26. Food and Drug Administration. *Framework for FDA's Real-World Evidence Program*. (2018). Available online at: www.fda.gov/media/120060/download (accessed July 9, 2020).
27. Mahendraratnam N, Silcox C, Mercon K, Kroetsch A, Romine M, Harrison N, et al. *Determining Real-World Data's Fitness for Use and the Role of Reliability*. (2019). Available online at: https://healthpolicy.duke.edu/sites/default/files/u31/rwd_reliability.pdf (accessed July 9, 2020).
28. Bernard A, Vaneau M, Fournel I, Galmiche H, Nony P, Dubernard JM. Methodological choices for the clinical development of medical devices. *Med Devices (Auckl)*. (2014) 7:325–34. doi: 10.2147/MDER.S63869
29. Neugebauer EAM, Rath A, Antoine SL, Eikermann M, Seidel D, Koenen C, et al. Specific barriers to the conduct of randomised clinical trials on medical devices. *Trials*. (2017) 18:360. doi: 10.1186/s13063-017-2168-0
30. Lu N, Xu Y, Yue LQ. Good statistical practice in utilizing real-world data in a comparative study for premarket evaluation of medical devices. *J Biopharm Stat*. (2019) 29:580–91. doi: 10.1080/10543406.2019.1632880
31. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence–study limitations (risk of bias). *J Clin Epidemiol*. (2011) 64:407–15. doi: 10.1016/j.jclinepi.2010.07.017
32. Schünemann HJ, Higgins JPT, Vist GE, Glasziou P, Akl EA, Skoetz N, et al. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 6*. (2019). doi: 10.1002/9781119536604.ch14
33. Sterne JAC, Hernán MA, McAleenan A, Reeves BC, Higgins JPT. Chapter 25: assessing risk of bias in a non-randomized study. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. *Cochrane*. (2019). Available online at: www.training.cochrane.org/handbook (accessed July 9, 2020).
34. Berger M, Crown W, Daniel G, Eichler H-G, Goettsch W, Graff J, et al. *Improving Transparency in Non-Interventional Research for Hypothesis Testing-WHY, WHAT, and HOW: Considerations from The Real-World Evidence Transparency Initiative*. (2019). Available online at: https://www.ispor.org/docs/default-source/strategic-initiatives/improving-transparency-in-non-interventional-research-for-hypothesis-testing_final.pdf (accessed July 9, 2020).
35. EUnetHTA. *HTA Adaptation Toolkit (Version 5)*. (2011). Available online at: https://eunethta.eu/wp-content/uploads/2011/01/EUnetHTA_adptation_toolkit_2011_version_5.pdf (accessed July 9, 2020).
36. Powers JH, 3rd, Howard K, Saretsky T, Clifford S, Hoffmann S, Llorens L, et al. Patient-reported outcome assessments as endpoints in studies in infectious diseases. *Clin Infect Dis*. (2016) 63(Suppl. 2):S52–6. doi: 10.1093/cid/ciw317
37. Grigore B, Ciani O, Dams F, Federici C, de Groot S, Möllenkamp M, et al. Surrogate endpoints in health technology assessment: an international review of methodological guidelines. *Pharmacoeconomics*. (2020) 38:1055–70. doi: 10.1007/s40273-020-00935-1
38. Tarricone R, Torbica A, Drummond M, MedtecHTA Project Group. Key recommendations from the MedtecHTA project. *Health Econ (United Kingdom)*. (2017) 26:145–52. doi: 10.1002/hec.3468
39. Rothery C, Claxton K, Palmer S, Epstein D, Tarricone R, Sculpher M. Characterising uncertainty in the assessment of medical devices and determining future research needs. *Health Econ*. (2017) 26(Suppl. 1):109–23. doi: 10.1002/hec.3467
40. Varabyova Y, Blankart CR, Schreyogg J. The role of learning in health technology assessments: an empirical assessment of endovascular aneurysm repairs in German Hospitals. *Health Econ*. (2017) 26(Suppl. 1):93–108. doi: 10.1002/hec.3466
41. Cook JA, Ramsay CR, Fayers P. Statistical evaluation of learning curve effects in surgical trials. *Clin Trials*. (2004) 1:421–7. doi: 10.1191/1740774504cn042oa
42. Hatz MH, Schreyogg J, Torbica A, Boriani G, Blankart CR. Adoption decisions for medical devices in the field of cardiology: results from a european survey. *Health Econ*. (2017) 26(Suppl. 1):124–44. doi: 10.1002/hec.3472
43. Bisschop A, van Tulder MW. Market approval processes for new types of spinal devices: challenges and recommendations for improvement. *Eur Spine J*. (2016) 25:2993–3003. doi: 10.1007/s00586-016-4606-1
44. Woods B, Revill P, Sculpher M, Claxton K. Country-level cost-effectiveness thresholds: initial estimates and the need for further research. *Value Health*. (2016) 19:929–35. doi: 10.1016/j.jval.2016.02.017
45. Skoupá J, Annemans L, Hájek P. Health economic data requirements and availability in the European Union: results of a survey among 10 European Countries. *Value Health Reg Issues*. (2014) 4:53–7. doi: 10.1016/j.vhri.2014.06.003

46. Barbieri M, Drummond M, Willke R, Chancellor J, Jolain B, Towse A. Variability of cost-effectiveness estimates for pharmaceuticals in Western Europe: lessons for inferring generalizability. *Value Health*. (2005) 8:10–23. doi: 10.1111/j.1524-4733.2005.03070.x
47. Vemer P, Rutten-van Mölken MP. Crossing borders: factors affecting differences in cost-effectiveness of smoking cessation interventions between European countries. *Value Health*. (2010) 13:230–41. doi: 10.1111/j.1524-4733.2009.00612.x
48. Sculpher MJ, Pang FS, Manca A, Drummond MF, Golder S, Urdahl H, et al. Generalisability in economic evaluation studies in healthcare: a review and case studies. *Health Technol Assess*. (2004) 8:iii–iv, 1–192. doi: 10.3310/hta8490
49. Drummond M, Augustovski F, Kalo Z, Yang BM, Pichon-Riviere A, Bae EY, et al. Challenges faced in transferring economic evaluations to middle income countries. *Int J Technol Assess Health Care*. (2015) 31:442–8. doi: 10.1017/S0266462315000604
50. Drummond M, Tarricone R, Torbica A. Incentivizing research into the effectiveness of medical devices. *Eur J Health Econ*. (2016) 17:1055–8. doi: 10.1007/s10198-016-0820-3

Conflict of Interest: RD-B, SK, AZ, and ZK were employed by Syreon Research Institute. MH is self-employed in HTA/EBM Consulting Centre and has not received any remuneration of funding from the COMED project, nor has any commercial or financial relationship that could be construed as a potential conflict of interest.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

ORIGINAL RESEARCH
published: 15 January 2021
doi: 10.3389/fpubh.2020.612976

Financial Literacy, Financial Education, and Smoking Behavior: Evidence From Japan

*Somtip Watanapongvanich, Mostafa Saidur Rahim Khan, Pongpat Putthinun, Shunsuke Ono and Yoshihiko Kadoya**

School of Economics, Hiroshima University, Hiroshima, Japan

In this study, we examine the relationship between financial literacy, financial education, and smoking behavior among the Japanese population. We hypothesize that financially literate and financially educated people, who have the ability to make more rational decisions, are less likely to smoke. Using the Preference Parameters Study of Osaka University, conducted in 2010 (N = 3,706), the probit regression results show that both financial literacy (with an emphasis on knowledge of investments) and financial education (with an emphasis on savings behavior) have a significant negative impact on smoking behavior. In addition, gender, age, education, marital status, household income and assets, risky behaviors, a myopic view of the future, risk preference, and level of happiness also significantly predict the likelihood of a person being a current smoker. This study provides empirical evidence that enhancing the rational decision-making ability of individuals through financial literacy and financial education may curtail smoking behavior.

Keywords: smoking, financial literacy, financial education, Japan, healthcare decision

OPEN ACCESS

Edited by:
Hideki Hashimoto,
The University of Tokyo, Japan

Reviewed by:
Salvador Cruz Rambaud,
University of Almeria, Spain
Rei Goto,
Keio University, Japan

****Correspondence:***
Yoshihiko Kadoya
ykadoya@hiroshima-u.ac.jp

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *01 October 2020*
Accepted: *08 December 2020*
Published: *15 January 2021*

Citation:
Watanapongvanich S, Khan MSR, Putthinun P, Ono S and Kadoya Y (2021) Financial Literacy, Financial Education, and Smoking Behavior: Evidence From Japan. Front. Public Health 8:612976. doi: 10.3389/fpubh.2020.612976

INTRODUCTION

Smoking has become an increasing public health concern in Japan over the past several decades. In 2012, the Japanese government set a target to reduce the smoking rate among adults aged 20 and over to 12% by 2022 (1). Since then, government and non-government organizations have taken several actions to curtail smoking, such as prohibiting indoor smoking, increasing the tobacco tax, and providing smoking cessation services (2). However, the prevalence of smoking in Japan has remained unchanged since 2013 (3–5). According to the Global Burden of Disease Study (GBD), Japan had the world's seventh-largest population of smokers in 2015 (6), and the 2017–2019 WHO report on the global tobacco epidemic indicates that 19% of Japanese aged 15 and over, ∼21 million people, still smoke (4, 5). These statistics raise the question of why smoking is still relatively prevalent in the country despite continued efforts to curtail it. It also seems likely that the government will not reach its smoking reduction target in the near future.

Smoking has substantially negative health [e.g., (7–13)], behavioral [e.g., (14–16)], and economic consequences [e.g., (17, 18)]. One important question is why do many people still engage in smoking despite its harmful effects. Smoking, as an addictive behavior, could be explained from three competing theoretical frameworks: rational, imperfect rational, and irrational choice (19, 20). Becker and Murphy (21) advanced the rational choice theory and postulated that smoking is a rational choice made on the basis of utility maximizing criteria. This means that people, by using available information, find that present pleasure of smoking provides more utility than the utility loss from its adverse consequences. However, the rational choice theory has been challenged

on the ground that smokers do not have all the relevant information at the time when they make smoking decisions, and they cannot objectively assess the value of information. We support the argument that smoking decisions are not rational, but rather either imperfectly rational or completely irrational.

Imperfectly rational framework postulates that smokers show bounded rationality because of their poor judgment over future health effects and time-inconsistent preferences. Compared with rational decision makers' time consistency over discounting immediate present and distant future outcomes, imperfectly rational smokers tend to be hyperbolic in discounting future consequences over current pleasure from smoking. Barlow et al. (22), MacKillop et al. (23), Ida and Goto (24), Baker et al. (21), and Mitchell (25) argued that smoking is associated with a high discount rate, meaning that smokers greatly discounted the future compared with the present outcomes. Barlow et al. (22) reviewed 69 relevant studies and argued that higher discounting was the reason for decisions to smoke and reluctance to quit even though smokers know future health consequences. O'Donoghue and Rabin (26) and Gruber and Köszegi (27) considered time-inconsistent people as biased toward the present. In other words, smokers value the pleasure they receive from smoking in the present whilst heavily discounting the negative consequences of smoking in the future (20). Besides hyperbolic discounting, several studies modeled time-inconsistent addictive behavior by the exponentiated hyperbolic discounting (28) and subadditive discounting (29, 30). In addition to time-inconsistency, smokers, particularly young smokers, are often found to have imperfect information; they lack judgment in understanding future health consequences and sometimes become addicted suddenly (31–35). The human capital model developed by Grossman (36) confirmed that smokers tend to place less emphasis on human capital as they heavily discount future health effects. Grossman (36) found that smoking destroys health capital, thereby decreasing human capital and ultimately taking a toll on productivity. Moreover, smokers often fail to recognize that the adverse effects of smoking not only affect them directly but also result in higher social costs from negative externalities and health care burdens (17, 37) compared with the social benefits of socially desirable outcomes (17). Cognitive limitation could be a possible reason why imperfectly rational people are unable to properly value the health consequences of smoking.

Finally, the conceptual framework that conjectures smoking as an irrational behavior postulates that decisions to smoke are guided by emotions, which hinder objective assessment of risks and benefits associated with the consumption. When emotions drive decisions, people lose their rationality and deviate from the objective assessment of information. Decision to smoke as an addictive good is sometimes triggered by external causes leading to impulsive behavior (20, 38). Moreover, Perikleous et al. (39) confirmed the evidence of peer influence on adolescent smoking, which confirmed that smokers do not always make unbiased decisions.

How to reduce people's tendency to smoke has been an area of great concern over the decades. One approach to prevent people from making imperfectly rational or irrational decisions is to enable them to make rational decisions. In this study, we propose that financial literacy is a rational decision-making instrument, as the people who possess this knowledge tend to behave more rationally; this is reflected in their economic and financial behavior (40–50). Moreover, financial literacy is related to improved cognitive ability (51, 52), which helps them become time-consistent in making judgment over current and future outcomes. Because of rational decision-making ability and increased cognitive ability, financially literate people are likely to make informed decision, place due importance on future outcomes, be time-consistent, and not be influenced by emotions. Thus, financially literate people, being rational agents, are less likely to be smokers because of their ability to objectively assess the benefits and health risks of smoking. In support of the ability to make rational decisions, previous studies provided evidence that financially literate people are engaged in better healthcare decision-making and health-promoting activities (53, 54). Because of the relevance of financial literacy with positive health behavior, better health outcomes, and quality of life in later years, positive reinforcement should be solicited in the society. Similar to financial literacy, financial education received in school is likely to make people aware of financial issues and help them to make informed and rational decisions (55–58).

Although previous studies show that financial literacy is associated with better health-related decision-making, few empirical studies have investigated the relationship between financial literacy and health behaviors. A recent study by Watanapongvanich et al. (59) used financial literacy and financial education to explain gambling behavior among the Japanese population. They found that financial literacy has a significantly negative relationship with gambling frequency, whilst financial education has an insignificant impact. Existing research also shows that the thought processes behind smoking and gambling have the same intertemporal relationship with the financial decision-making process (21, 25, 60, 61) and that gambling and smoking are proxies for irrational behaviors that can lead to negative consequences. Therefore, apart from gambling, it is worth investigating how financial literacy and financial education can reduce smoking behavior in Japan.

In this study, we examine the relationship between financial literacy, financial education, and smoking behavior. We hypothesize that financially literate and financially educated people, who have the ability to make more rational decisions, are less likely to be smokers. To the best of our knowledge, no existing empirical research addresses the association between financial literacy, financial education, and smoking behavior. Our study contributes to the existing literature by providing empirical evidence on the connection between financial literacy and financial education as a rational decision-making tool and how they affect irrational decisions tied to smoking behavior in Japan. The results of this study can help policymakers implement effective interventions to prevent and minimize the negative consequences of smoking.

The remainder of this paper is organized as follows. In sections Data and Methodology, the data and methodology are described, respectively. In section Results, the empirical results are summarized, in section Robustness Check, robustness of the

results are checked, and in section Discussion, the results are discussed. Section Conclusion presents the conclusion.

DATA

Data

In this study, we use data from the Preference Parameters Study (PPS) conducted by the Institute of Social and Economic Research at Osaka University. The PPS is a panel survey that collects information on socioeconomic characteristics and preferences from a representative sample of the Japanese population. In this study, we utilized data from the 2010 wave, which contained questions about smoking behavior, financial literacy, and financial education. The sample includes data from 3,706 individuals, or ∼69% of all respondents in 2010 (5,386 individuals). We excluded individuals with missing data on smoking behavior, financial literacy, financial education, and demographic variables (1,680 individuals).

Variable Definitions

The dependent variable in this study is smoking behavior. The PPS contains the question "Do you smoke?" and provides seven responses, where 1 means "do not smoke at all," 2 means "hardly smoke," 3 means "smoke sometimes," 4 means "about 10 cigarettes a day," 5 means "about a pack a day," 6 means "more than two packs a day," and 7 means "I used to smoke but have quit." We grouped these responses into a binary scale of non-smokers and current smokers by coding respondents who answered 1, 2, or 7 as 0 or non-smokers (62) and those who answered 3, 4, 5, or 6 as 1 or current smokers.

There are two main variables of interest in our study: financial literacy and financial education. To measure financial literacy, we followed the methodology proposed by Lusardi and Mitchell (63), which is simple and widely adopted in existing literature [e.g., (59, 64–71)]. It uses the following three questions.

a. Suppose you had 10,000 JPY in a savings account and the interest rate is 2% per year and you never withdraw money or interest payments. After 5 years, how much would you have in this account in total?

- More than 10,200 JPY (correct answer)
- Exactly 10,200 JPY
- Less than 10,200 JPY
- Do not know
- Refuse to answer

b. Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?

- More than today
- Exactly the same
- Less than today (correct answer)
- Do not know
- Refuse to answer

c. Please indicate whether the following statement is true or false. "Buying a company stock usually provides a safer return than a stock mutual fund."

- True
- False (correct answer)
- Do not know
- Refuse to answer

The first two questions measure the respondent's understanding of how compound interest works and the effect of inflation. Indeed, the questions help evaluate a respondent's understanding of economic concepts and basic numeracy (63). The third question evaluates respondents' understanding of the concept of risk diversification. In this study, we assigned a score of one for each correct answer and 0 for each incorrect answer. We obtained the financial literacy variable by taking the equally weighted average scores of the three questions.

For financial education, the respondents were asked, "Did you receive any compulsory financial education when you were in elementary school?" with three possible responses: yes, no, and do not know. We coded the respondents who answered yes as one and those who answered no or do not know as 0. We treated this as a binary variable.

We should note an important difference between financial literacy and financial education in Japan. A recent work by Watanapongvanich et al. (59) suggests that Lusardi and Mitchell's (63) financial literacy questions measure respondents' current level of financial knowledge from an investment perspective. However, the Japanese school curriculum includes financial education to teach students about savings using a children's bank campaign (59, 72). Therefore, financial literacy that focuses on investment concepts is somewhat different from financial education that focuses on savings behavior (59). Sekita (71), who found that receiving financial education through a children's bank campaign has no effect on the level of financial literacy among representative Japanese adults, also supports this claim. Consequently, we include both financial literacy and financial education as explanatory variables in this study.

Furthermore, we include gender, age, university degree, marital status, household members, employment status, household income, and household assets as demographic variables in our specifications. We also control for risky behaviors (exercise, drinking alcohol, and gambling behavior), myopic view of the future, risk preference, level of happiness, and anxiety about health. **Table 1** provides the definitions of all the variables.

Descriptive Statistics

The descriptive statistics in **Table 2** show that 24.26% of respondents were current smokers. On average, respondents' financial literacy scores were 0.59 and 15.33% of the sample received financial education at school. For the demographic variables, about 49.24% of the sample were men and the average age was 49.79 years. Approximately 27.06% of the sample hold a university degree, 82.25% of the sample were currently married, and 3.45% were divorced. The respondents had four household members on average, and about 84.43% of the sample had children. Only 2.40% of the sample were currently unemployed. Respondents had an annual household income of ∼6.49 million JPY on average and 13.10 million JPY in household assets in

TABLE 1 | Variable definitions.

Variables	Definitions
Smoking behavior	Binary variable: 1 = current smoker (sometimes–more than two packs a day) and 0 = non-smokers (do not smoke at all, quit, or hardly smoke)
Financial literacy	Continuous variable: number of correct answers from three financial literacy questions
Financial education	Binary variable: 1 = received compulsory financial education at school and 0 = otherwise
Male	Binary variable: 1 = male and 0 = female
Age	Respondent's age
Age squared	Age squared
University degree	Binary variable: 1 = obtained university degree and 0 = otherwise
Marriage	Binary variable: 1 = married and 0 = otherwise
Divorce	Binary variable: 1 = divorced or separated and 0 = otherwise
Household members	Continuous variable: number of people currently living in household
Children	Binary variable: 1 = have child/children and 0 = otherwise
Unemployed	Binary variable: 1 = respondent is unemployed and 0 = otherwise
Household income	Continuous variable: annual earned income before taxes and with bonuses of the entire household in 2009 (unit: JPY)
Log of household income	Log (household income)
Household assets	Continuous variable: balance of financial assets (savings, stock, insurance, etc.) of the entire household (unit: JPY)
Log of household assets	Log (household assets)
Regular exercise	Binary variable: 1 = regular exercise (exercise once a week or more) and 0 = otherwise
Current drinker	Binary variable: 1 = current drinker (drink sometimes–five cans of beer daily) and 0 = otherwise
Frequent gambler	Binary variable: 1 = frequent gambler (gamble once a week or more) and 0 = otherwise
Myopic view of the future	Binary variable: 1 = agree and completely agree with the statement "Since the future is uncertain, it is a waste to think about it" and 0 = otherwise
Level of risk preference	Continuous variable: percentage score from the question "Usually, when you go outdoors, how high does the probability of rain have to be before you take an umbrella?"
Current level of happiness	Continuous variable: percentage score from the question "Overall, how happy would you say you are currently?"
Anxiety about health	Binary variable: 1 = agree and completely agree with the statement "I have anxieties about my health" and 0 = otherwise

TABLE 2 | Descriptive statistics.

Variables	Mean	Standard deviation (SD)	Min	Max
Smoking behavior	0.2426	0.4287	0	1
Financial literacy	0.5914	0.3436	0	1
Financial education	0.1533	0.3603	0	1
Male	0.4924	0.5000	0	1
Age	49.79	12.61	20	76
Age squared	2637.63	1239.67	400	5,776
University degree	0.2706	0.4444	0	1
Marriage	0.8225	0.3822	0	1
Divorce	0.0345	0.1826	0	1
Household members	3.52	1.44	1	10
Children	0.8443	0.3626	0	1
Unemployed	0.0240	0.1531	0	1
Household income	6,486,239	3,777,635	1,000,000	20,000,000
Log of household income	15.51	0.61	13.82	16.81
Household assets	13,100,000	17,600,000	2,500,000	100,000,000
Log of household assets	15.81	1.01	14.73	18.42
Regular exercise	0.3713	0.4832	0	1
Current drinker	0.5416	0.4983	0	1
Frequent gambler	0.0942	0.2921	0	1
Myopic view of the future	0.1454	0.3526	0	1
Level of risk preference	0.4908	0.1891	0	1
Current level of happiness	0.6472	0.1822	0	1
Anxiety about health	0.4199	0.4936	0	1
Observations		3,706		

2009. For risky behaviors, 37.13% of the participants exercised regularly, whilst 54.16% were current drinkers and 9.42% were frequent gamblers. Overall, 14.54% of the respondents had a myopic view of the future and risk preferences of 49.08%; in other words, they were risk neutral. Respondents rated their current level of happiness at 64.72%, and 41.99% of the sample were anxious about their health.

Tables 3–**5** present the distribution of smoking behavior classified by age group, demographic characteristics, and risky behaviors, respectively. Our sample contained 899 current smokers; that is, 24.26% of the total sample smoke between sometimes to more than two packs of cigarettes daily, whilst the remaining 2,807 respondents were non-smokers. The results in **Table 3** indicate significant differences in smoking behavior among age groups. The proportion of current smokers in the oldest age group (age 61 and older) was 17.27%, which is less than other age groups that the proportions of current smokers are more than 22%. In **Table 4**, we see significant differences in smoking behavior between genders. Approximately 37.15% of male respondents were current smokers compared to 11.75% of female respondents. However, the differences in smoking behavior by education level and employment status are insignificant. The results in **Table 5** for risky behaviors show that about 19.11% of respondents who exercise regularly are current smokers, which is less than the sample of current smokers who do not exercise regularly (27.30%). In addition, we see considerable differences in smoking behavior between current drinkers and non-drinkers and between frequent gamblers and non-gamblers.

TABLE 3 | Distribution of smoking behavior by age group.

Smoking behavior	Age					Total
	≤30	31–40	41–50	51–60	≥61	
Non-smokers	212	496	679	716	704	2,807
	77.94%	71.78%	73.09%	74.35%	82.73%	75.74%
Current smokers	60	195	250	247	147	899
	22.06%	28.22%	26.91%	25.65%	17.27%	24.26%
Total	272	691	929	963	851	3,706
	100%	100%	100%	100%	100%	100%
Mean difference	$F = 8.51^{***}$					

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

TABLE 4 | Distribution of smoking behavior by demographic characteristic.

Smoking behavior	Gender		Education		Unemployed		Total
	Female	Male	Lower than university degree	University degree and higher	No	Yes	
Non-smokers	1,660	1,147	2,035	772	2,743	64	2,807
	88.25%	62.85%	75.29%	76.97%	75.84%	71.91%	75.74%
Current smokers	221	678	668	231	874	25	899
	11.75%	37.15%	24.71%	23.03%	24.16%	28.09%	24.26%
Total	1,881	1,825	2,703	1,003	3,617	89	3,706
	100%	100%	100%	100%	100%	100%	100%
Mean difference	$t = -18.8786^{***}$		$t = 1.0614$		$t = -0.8535$		

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

TABLE 5 | Distribution of smoking behavior by risky behaviors.

Smoking behavior	Regular exercise		Current drinkers		Frequent gamblers		Total
	No	Yes	No	Yes	No	Yes	
Non-smokers	1,694	1,113	1,411	1,396	2,603	204	2,807
	72.70%	80.89%	83.05%	69.56%	77.54%	58.45%	75.74%
Current smokers	636	263	288	611	754	145	899
	27.30%	19.11%	16.95%	30.44%	22.46%	41.55%	24.26%
Total	2,330	1,376	1,699	2,007	3,357	349	3,706
	100%	100%	100%	100%	100%	100%	100%
Mean difference	$t = 5.6374^{***}$		$t = -9.6649^{***}$		$t = -7.9829^{***}$		

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Specifically, 30.44% of current drinkers and 41.55% of frequent gamblers were current smokers.

METHODOLOGY

To investigate how financial literacy and financial education are related to smoking behavior, we first separately estimate the effects of financial literacy and financial education in Equations (1, 2), respectively. We then include both financial literacy and financial education to see the combined effect of the variables in Equation (3).

$$Y_i = f(FL_i, X_i, \varepsilon_i) \quad (1)$$

$$Y_i = f(FE_i, X_i, \varepsilon_i) \quad (2)$$

$$Y_i = f(FL_i, FE_i, X_i, \varepsilon_i) \quad (3)$$

where Y_i is the smoking behavior of the *i*th respondent (current smokers or non-smokers), FL represents the score on the

financial questions measuring financial literacy, FE represents financial education received at school, X is a vector of individual characteristics, and ε is the error term. Because the dependent variable is a binary choice, we employ a probit regression to estimate all equations.

As there is a potential for multicollinearity between the explanatory variables in the models (i.e., individuals with a high level of education could have high financial knowledge, or individuals with high net worth may have more financial knowledge because of experience with asset management), we conducted correlation and multicollinearity tests in all models (available upon request). The results show that multicollinearity between the variables is not significant, suggesting that the independent effects of explanatory variables on smoking behavior are not biased. The correlation matrix shows a weak relationship between the explanatory variables (lower than 0.70). In addition, the variance inflation factor (VIF) tests of the explanatory variables are below 10, indicating that multicollinearity is not significant in all models.

The full model specifications are

$$Smoking\ behavior_i\ (1 = current\ smokers\ and\ 0 = non-smokers) = \beta_0 + \beta_1 financial\ literacy_i + \beta_2 male_i + \beta_3 age_i + \beta_4 age\ squared_i + \beta_5 university\ degree_i + \beta_6 marriage_i + \beta_7 divorce_i + \beta_8 household\ members_i + \beta_9 children_i + \beta_{10} unemployed_i + \beta_{11} log\ of\ household\ income_i + \beta_{12}\ logofhousehold\ assets_i + \beta_{13} regular\ exercise_i + \beta_{14} current\ drinkers_i + \beta_{15} frequent\ gamblers_i + \beta_{16} myopic\ view\ of\ the\ future_i + \beta_{17} level\ of\ risk\ preference_i + \beta_{18} current\ level\ of\ happiness_i + \beta_{19} anxiety\ about\ health_i + \varepsilon_i \quad (1a)$$

$$Smoking\ behavior_i\ (1 = current\ smokers\ and\ 0 = non-smokers) = \beta_0 + \beta_1 financial\ education_i + \beta_2 male_i + \beta_3 age_i + \beta_4 age\ squared_i + \beta_5 university\ degree_i + \beta_6 marriage_i + \beta_7 divorce_i + \beta_8 household\ members_i + \beta_9 children_i + \beta_{10} unemployed_i + \beta_{11} log\ of\ household\ income_i + \beta_{12}\ logofhousehold\ assets_i + \beta_{13} regular\ exercise_i + \beta_{14} current\ drinkers_i + \beta_{15} frequent\ gamblers_i + \beta_{16} myopic\ view\ of\ the\ future_i + \beta_{17} level\ of\ risk\ preference_i + \beta_{18} current\ level\ of\ happiness_i + \beta_{19} anxiety\ about\ health_i + \varepsilon_i \quad (2a)$$

$$Smoking\ behavior_i\ (1 = current\ smokers\ and\ 0 = non-smokers) = \beta_0 + \beta_1 financial\ literacy_i + \beta_2 financial\ education_i + \beta_3 male_i + \beta_4 age_i + \beta_5 age\ squared_i + \beta_6 university\ degree_i + \beta_7 marriage_i + \beta_8 divorce_i + \beta_9 household\ members_i + \beta_{10} children_i + \beta_{11} unemployed_i + \beta_{12} log\ of\ household\ income_i + \beta_{13}\ logofhousehold\ assets_i + \beta_{14} regular\ exercise_i + \beta_{15} current\ drinkers_i + \beta_{16} frequent\ gamblers_i + \beta_{17} myopic\ view\ of\ the\ future_i + \beta_{18} level\ of\ risk\ preference_i + \beta_{19} current\ level\ of\ happiness_i + \beta_{20} anxiety\ about\ health_i + \varepsilon_i \quad (3a)$$

RESULTS

We present the results of the probit regression to estimate Equations (1)-(3) in **Tables 6**–**8**, respectively. Each table presents the results of four different specifications of the explanatory variables. The first specification (Models 1.1, 2.1, and 3.1) included controls for only the demographic variables. In the second specification (Models 1.2, 2.2, and 3.2), we added risky behaviors including exercise, drinking alcohol, and gambling. The third specification (Models 1.3, 2.3, and 3.3) included respondents' myopic views of the future and risk preferences. Finally, the fourth specification (Models 1.4, 2.4, and 3.4) included respondents' self-rated level of happiness and anxiety about health.

The results in **Table 6** show that financial literacy has a negative and strongly significant impact on smoking behavior across the models. In **Table 7**, we see that financial education also has a negative and significant impact on smoking behavior in all models except Model 2.2. However, the significance levels of the financial education variables (at 5 and 10%) are less than those of the financial literacy variables (at % and 5%). Since both financial literacy and financial education have a significant impact on smoking behavior, we regress financial literacy on financial education and the other control variables to explore the relationship between these two variables. We find that financial literacy and financial education are not correlated (results not reported here, but available upon request), consistent with the finding of Sekita (71). Therefore, we should use both financial literacy and financial education as explanatory variables in the same equation to explain smoking behavior, as shown in our final model in **Table 8**.

The results in **Table 8** show that, overall, there are no differences in the significance of the estimated parameters compared to the results in **Tables 6**, **7**. The coefficients of our variables of interest, financial literacy, and financial education are negative and statistically significant in all models except financial education in Model 3.2. In other words, respondents with a high level of financial literacy and those who received financial education were less likely to be current smokers at present.

For the demographic and other control variables, most of the signs and significance levels of the coefficients are consistent across models and specifications. Male, age, and divorce have a positive impact on being a current smoker, and the coefficients are strongly significant at the 1% level. In contrast, age squared, university degree, log of household income, and log of household assets have a negative and significant impact on being a current smoker. However, marriage, household members, children, and employment status have an insignificant impact. In terms of risky behaviors, regular exercise has a negative impact, whilst being a current drinker and frequent gambler have a positive impact on current smoking status, statistically significant at the 1% level. Furthermore, a myopic view of the future and a high level of risk preference are associated with being current smokers. Conversely, respondents with self-rated high levels of happiness are less likely to be current smokers. However, respondents' anxiety regarding health showed an insignificant impact on smoking behavior.

ROBUSTNESS CHECK

To check robustness of our results, we used an alternative classification of smoking behavior. Rather than classifying respondents as non-smokers and smokers, we classified them

TABLE 6 | Probit model regression results, financial literacy as the main explanatory variable.

Variables	Dependent variable: smoking behavior			
	Model 1.1	**Model 1.2**	**Model 1.3**	**Model 1.4**
Financial literacy	−0.228***	−0.209***	−0.181**	−0.165**
	(0.0759)	(0.0767)	(0.0772)	(0.0770)
Male	1.018***	0.940***	0.928***	0.909***
	(0.0523)	(0.0558)	(0.0559)	(0.0562)
Age	0.0765***	0.0653***	0.0673***	0.0621***
	(0.0153)	(0.0152)	(0.0153)	(0.0154)
Age squared	−0.000892***	−0.000769***	−0.000787***	−0.000750***
	(0.000153)	(0.000153)	(0.000154)	(0.000155)
University degree	−0.260***	−0.250***	−0.231***	−0.214***
	(0.0589)	(0.0593)	(0.0596)	(0.0599)
Marriage	−0.0610	−0.0466	−0.0428	0.0164
	(0.105)	(0.105)	(0.105)	(0.108)
Divorce	0.441***	0.449***	0.438***	0.453***
	(0.154)	(0.153)	(0.154)	(0.156)
Household members	−0.0154	−0.0201	−0.0212	−0.0282
	(0.0199)	(0.0200)	(0.0201)	(0.0202)
Children	0.0496	0.0452	0.0532	0.0707
	(0.105)	(0.106)	(0.106)	(0.107)
Unemployed	−0.0720	−0.0216	−0.0185	−0.0863
	(0.160)	(0.158)	(0.158)	(0.161)
Log of household income	−0.109**	−0.120**	−0.118**	−0.0829*
	(0.0472)	(0.0476)	(0.0477)	(0.0485)
Log of household assets	−0.0794***	−0.0590**	−0.0551*	−0.0421
	(0.0290)	(0.0293)	(0.0294)	(0.0296)
Regular exercise		−0.281***	−0.275***	−0.246***
		(0.0527)	(0.0528)	(0.0531)
Current drinker		0.231***	0.230***	0.235***
		(0.0523)	(0.0524)	(0.0525)
Frequent gambler		0.282***	0.265***	0.256***
		(0.0769)	(0.0769)	(0.0774)
Myopic view of the future			0.172**	0.156**
			(0.0674)	(0.0676)
Level of risk preference			0.424***	0.435***
			(0.129)	(0.130)
Current level of happiness				−0.684***
				(0.142)
Anxiety about health				0.00801
				(0.0504)
Constant	0.446	0.502	0.0839	−0.121
	(0.741)	(0.746)	(0.753)	(0.755)
Observations	3,706	3,706	3,706	3,706
Log likelihood	−1803	−1773	−1764	−1753
Chi2 statistics	466.7	505.4	516	524.4
p-value	0	0	0	0

*Robust standard errors in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1.*

as non-smokers, occasional or intermittent smokers, and regular smokers. The alternative classification allowed us to check rationality in smoking behavior elaborately. Non-smokers included respondents who answered "do not smoke at all" and "I used to smoke but had quit." Occasional smokers included respondents who answered "hardly smoke" and

TABLE 7 | Probit model regression results, financial education as the main explanatory variable.

Variables	Dependent variable: smoking behavior			
	Model 2.1	**Model 2.2**	**Model 2.3**	**Model 2.4**
Financial education	−0.139**	−0.114	−0.118*	−0.127*
	(0.0693)	(0.0698)	(0.0700)	(0.0701)
Male	0.995***	0.920***	0.910***	0.893***
	(0.0518)	(0.0553)	(0.0553)	(0.0557)
Age	0.0735***	0.0623***	0.0652***	0.0602***
	(0.0152)	(0.0151)	(0.0152)	(0.0153)
Age squared	−0.000862***	−0.000740***	−0.000765***	−0.000730***
	(0.000152)	(0.000152)	(0.000153)	(0.000154)
University degree	−0.291***	−0.279***	−0.254***	−0.234***
	(0.0584)	(0.0588)	(0.0591)	(0.0594)
Marriage	−0.0645	−0.0499	−0.0451	0.0163
	(0.105)	(0.106)	(0.106)	(0.108)
Divorce	0.427***	0.437***	0.427***	0.444***
	(0.154)	(0.153)	(0.154)	(0.155)
Household members	−0.00888	−0.0142	−0.0160	−0.0235
	(0.0199)	(0.0200)	(0.0201)	(0.0202)
Children	0.0629	0.0570	0.0639	0.0811
	(0.105)	(0.106)	(0.106)	(0.107)
Unemployed	−0.0538	−0.00506	−0.00241	−0.0731
	(0.160)	(0.158)	(0.158)	(0.160)
Log of household income	−0.131***	−0.140***	−0.134***	−0.0971**
	(0.0467)	(0.0472)	(0.0473)	(0.0481)
Log of household assets	−0.0894***	−0.0679**	−0.0625**	−0.0485*
	(0.0288)	(0.0290)	(0.0291)	(0.0293)
Regular exercise		−0.284***	−0.278***	−0.247***
		(0.0527)	(0.0528)	(0.0531)
Current drinker		0.230***	0.228***	0.233***
		(0.0523)	(0.0524)	(0.0525)
Frequent gambler		0.278***	0.260***	0.249***
		(0.0769)	(0.0768)	(0.0774)
Myopic view of the future			0.179***	0.161**
			(0.0674)	(0.0677)
Level of risk preference			0.451***	0.460***
			(0.129)	(0.130)
Current level of happiness				−0.702***
				(0.142)
Anxiety about health				0.0123
				(0.0504)
Constant	0.879	0.895	0.395	0.152
	(0.723)	(0.728)	(0.737)	(0.739)
Observations	3,706	3,706	3,706	3,706
Log likelihood	−1806	−1776	−1766	−1753
Chi2 statistics	463.4	506.3	519.1	529.7
p-value	0	0	0	0

*Robust standard errors in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1.*

"smoke sometimes." Regular smokers included respondents who answered "about 10 cigarettes a day," "about a pack a day," and "more than two packs a day." Similar to our original models, we hypothesize that respondents who are financially literate or received more financial education are less likely to be regular smokers.

TABLE 8 | Probit model regression results, financial literacy, and financial education as the main explanatory variables.

Variables	Dependent variable: smoking behavior			
	Model 3.1	**Model 3.2**	**Model 3.3**	**Model 3.4**
Financial literacy	−0.227***	−0.208***	−0.180**	−0.164**
	(0.0760)	(0.0768)	(0.0772)	(0.0771)
Financial education	−0.138**	−0.113	−0.116*	−0.125*
	(0.0692)	(0.0698)	(0.0699)	(0.0700)
Male	1.018***	0.941***	0.928***	0.909***
	(0.0524)	(0.0558)	(0.0559)	(0.0562)
Age	0.0774***	0.0661***	0.0683***	0.0631***
	(0.0153)	(0.0153)	(0.0153)	(0.0154)
Age squared	−0.000896***	−0.000773***	−0.000792***	−0.000755***
	(0.000153)	(0.000153)	(0.000154)	(0.000155)
University degree	−0.259***	−0.250***	−0.231***	−0.214***
	(0.0590)	(0.0593)	(0.0597)	(0.0600)
Marriage	−0.0570	−0.0436	−0.0395	0.0201
	(0.105)	(0.105)	(0.105)	(0.108)
Divorce	0.439***	0.448***	0.437***	0.452***
	(0.154)	(0.153)	(0.154)	(0.155)
Household members	−0.0140	−0.0188	−0.0199	−0.0269
	(0.0199)	(0.0201)	(0.0201)	(0.0202)
Children	0.0503	0.0460	0.0538	0.0716
	(0.105)	(0.106)	(0.106)	(0.107)
Unemployed	−0.0668	−0.0174	−0.0134	−0.0816
	(0.159)	(0.157)	(0.158)	(0.161)
Log of household income	−0.111**	−0.121**	−0.119**	−0.0837*
	(0.0472)	(0.0476)	(0.0476)	(0.0484)
Log of household assets	−0.0791***	−0.0589**	−0.0549*	−0.0418
	(0.0291)	(0.0294)	(0.0294)	(0.0296)
Regular exercise		−0.279***	−0.273***	−0.244***
		(0.0527)	(0.0528)	(0.0532)
Current drinker		0.229***	0.228***	0.232***
		(0.0524)	(0.0525)	(0.0526)
Frequent gambler		0.277***	0.259***	0.250***
		(0.0769)	(0.0769)	(0.0774)
Myopic view of the future			0.169**	0.152**
			(0.0676)	(0.0678)
Level of risk preference			0.433***	0.444***
			(0.129)	(0.130)
Current level of happiness				−0.689***
				(0.142)
Anxiety about health				0.0105
				(0.0505)
Constant	0.438	0.492	0.0665	−0.142
	(0.741)	(0.746)	(0.753)	(0.755)
Observations	3,706	3,706	3,706	3,706
Log likelihood	−1801	−1772	−1763	−1751
Chi2 statistics	471.4	510.4	522.3	531.9
p-value	0	0	0	0

*Robust standard errors in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1.*

We used multinomial probit model (mprobit) to investigate how financial literacy and financial education are related to smoking behavior. The results of mprobit regressions are presented in **Tables 9–11**. Each table presents the results of four different specifications of the explanatory variables where the base model is non-smokers' category. The results in **Table 9** show that financial literacy has a negative and significant impact on smoking behavior in all models except Model 4.7 in case of occasional smokers. The results in **Table 10** indicate that financial education also has a negative and significant impact on smoking behavior in regular smokers' category. We included both financial literacy and financial education as explanatory variables in our final model. The results in **Table 11** show that, overall, there are no differences in the significance of the estimated parameters compared with the results in **Tables 9**, **10**. The coefficients of our variables of interest, financial literacy and financial education, are negative and statistically significant across models in case of the regular smokers. In other words, compared with non-smokers, respondents with a high level of financial literacy and those who received financial education are less likely to be regular smokers. These results are similar to the probit estimation when smoking behavior is classified into the non-smokers and smokers categories.

DISCUSSION

As mentioned earlier, there is an important difference between financial literacy and financial education in the Japanese context. Financial literacy measures respondents' current level of financial knowledge and focuses on the investment perspective, whilst financial education focuses on the savings behavior. Our results demonstrate that both financial literacy and financial education have a significant impact on smoking behavior. Hence, the focus of our discussion here is the final model, as shown in **Table 8**.

Among Models 3.1-3.4, both financial literacy and financial education have a significantly negative impact on smoking behavior, except financial education in Model 3.2. These inverse relationships indicate that respondents with a high level of financial literacy and those who received financial education were less likely to be current smokers. The findings support our hypothesis that financially literate and financially educated people, who have the ability to make more rational decisions, are less likely to make the irrational decision to smoke. Watanapongvanich et al. (59) also found an inverse relationship between financial literacy (as a proxy for a rational decision-making tool) and gambling behavior (as a proxy for irrational decisions).

For the demographic variables, we find that being male, older (until a certain age), divorced, and having an education level below university degree, low household income, and low household assets are related to the current smoking status. Our results are consistent with those of previous studies, which found that smoking is more prevalent and progressive in males [e.g., (13, 73–78)] and persons in lower socioeconomic status (SES) [e.g., (73, 79–85)]. A plausible explanation is that persons in higher SES groups tend to have more knowledge about health risks and better access to health care resources and smoking cessation services (79, 84), which results in a lower rate of smoking prevalence compared to those in the lower SES group. In addition, persons with lower education levels may suffer from lower self-esteem compared to those who complete higher education and may be more likely to take up smoking as a self-enhancement mechanism (86). Regarding marital status, Lindström (87) and Pennanen et al. (81) found that smokers living without a spouse have higher daily smoking rates and nicotine dependence. Castrén et al. (88) explain that the dissolution of a marriage may cause people to undertake harmful activities.

For risky behaviors, respondents who exercise regularly are less likely to be current smokers, whilst respondents who are current drinkers and frequent gamblers are more likely to be current smokers. Prior studies also report a negative relationship between physical activity and smoking [e.g., (74, 89–91)]. We can predict these results because people who exercise regularly tend to be health conscious and avoid behaviors that cause health deterioration. In contrast, people who are already engaged in health-risk behaviors such as drinking alcohol and gambling tend to engage in other risky behaviors, including smoking. For example, Nichter et al. (92) argue that alcohol makes drinkers more comfortable with the experience of smoking and raises their smoking limit. Therefore, drinkers are more likely to smoke than non-drinkers [e.g., (93, 94)]. In addition, engaging in gambling also encourages smoking, especially during the game [e.g., (95–97)].

Our results related to a myopic view of the future and level of risk preference, which have a positive impact on smoking behavior, also support our findings on health-risk behaviors. Respondents who have a myopic view of the future will focus more on the present; in other words, they value the pleasant feelings from smoking today more than the adverse effects that smoking will have on their health in the future [e.g., (24, 25, 98)]. Furthermore, respondents with high levels of risk preference tend to engage in risky behaviors more than those with a low level of risk preference [e.g., (24, 99, 100)]. Therefore, respondents with a myopic view of the future and a high level of risk preference are more likely to be current smokers. Lastly, respondents who reported high self-rated levels of happiness are less likely to be current smokers. This finding is consistent with that of Chang et al. (101), who found that in Japan, France, and the UK, if people feel happier, they smoke less. As smoking gives smokers pleasure, they feel more relaxed and energetic; therefore, they use smoking as a coping mechanism to combat stress (101–103).

CONCLUSION

In this study, we examined the relationship between financial literacy, financial education, and smoking behavior among the Japanese population. We hypothesized that financially literate and financially educated people who have the ability

TABLE 9 | Multinomial probit model regression results, financial literacy as the main explanatory variable.

Variables	Occasional smoker	Regular smoker	Occasional smoker	Regular smoker	Occasional smoker	Regular smoker	Occasional smoker	Regular smoker
	Model 4.1	Model 4.2	Model 4.3	Model 4.4	Model 4.5	Model 4.6	Model 4.7	Model 4.8
Financial literacy	−0.269*	−0.316***	−0.257*	−0.291***	−0.243*	−0.247**	−0.234	−0.224**
	(0.143)	(0.109)	(0.142)	(0.110)	(0.143)	(0.110)	(0.142)	(0.110)
Male	0.643***	1.468***	0.620***	1.355***	0.612***	1.336***	0.594***	1.309***
	(0.0979)	(0.0750)	(0.106)	(0.0797)	(0.106)	(0.0798)	(0.107)	(0.0802)
Age	0.0139	0.116***	0.0103	0.0999***	0.0110	0.104***	0.00670	0.0962***
	(0.0266)	(0.0222)	(0.0267)	(0.0222)	(0.0267)	(0.0222)	(0.0269)	(0.0223)
Age squared	−0.000259	−0.00133***	−0.000218	−0.00116***	−0.000230	−0.00119***	−0.000198	−0.00114***
	(0.000271)	(0.000224)	(0.000272)	(0.000224)	(0.000271)	(0.000224)	(0.000272)	(0.000225)
University degree	−0.195*	−0.396***	−0.208*	−0.382***	−0.199*	−0.351***	−0.185	−0.327***
	(0.115)	(0.0842)	(0.115)	(0.0847)	(0.115)	(0.0852)	(0.115)	(0.0857)
Marriage	0.0299	−0.118	0.0228	−0.0954	0.0257	−0.0877	0.0736	−0.00118
	(0.173)	(0.151)	(0.176)	(0.150)	(0.175)	(0.152)	(0.177)	(0.154)
Divorce	0.482*	0.598***	0.465*	0.612***	0.455	0.597***	0.472*	0.621***
	(0.280)	(0.220)	(0.280)	(0.219)	(0.280)	(0.220)	(0.280)	(0.221)
Household member	0.0408	−0.0131	0.0365	−0.0195	0.0360	−0.0219	0.0309	−0.0320
	(0.0385)	(0.0283)	(0.0384)	(0.0286)	(0.0384)	(0.0286)	(0.0385)	(0.0287)
Children	−0.275	0.0652	−0.271	0.0562	−0.262	0.0669	−0.251	0.0924
	(0.175)	(0.151)	(0.176)	(0.151)	(0.176)	(0.152)	(0.177)	(0.153)
Unemployed	−0.315	−0.0875	−0.303	−0.0120	−0.302	−0.00560	−0.352	−0.101
	(0.331)	(0.230)	(0.332)	(0.228)	(0.333)	(0.229)	(0.339)	(0.232)
Log of household income	−0.117	−0.140**	−0.116	−0.156**	−0.115	−0.152**	−0.0861	−0.102
	(0.0928)	(0.0670)	(0.0931)	(0.0675)	(0.0932)	(0.0676)	(0.0939)	(0.0691)
Log of household assets	−0.0541	−0.115***	−0.0456	−0.0864**	−0.0439	−0.0792*	−0.0346	−0.0609
	(0.0538)	(0.0413)	(0.0543)	(0.0417)	(0.0546)	(0.0418)	(0.0549)	(0.0420)
Regular exercise			−0.145	−0.386***	−0.140	−0.376***	−0.120	−0.335***
			(0.0973)	(0.0754)	(0.0975)	(0.0756)	(0.0980)	(0.0761)
Current drinker			0.162	0.329***	0.166*	0.328***	0.169*	0.334***
			(0.101)	(0.0748)	(0.100)	(0.0750)	(0.101)	(0.0752)
Frequent gambler			−0.193	0.396***	−0.197	0.368***	−0.203	0.355***
			(0.175)	(0.109)	(0.175)	(0.110)	(0.175)	(0.110)
Myopic view of the future					0.211*	0.254***	0.200	0.233**
					(0.126)	(0.0967)	(0.126)	(0.0971)
Level of risk preference					0.0796	0.712***	0.0883	0.726***
					(0.248)	(0.185)	(0.248)	(0.186)
Current level of happiness							−0.548**	−0.980***
							(0.271)	(0.203)
Anxiety about health							−0.00979	−0.00885
							(0.0955)	(0.0721)
Constant	0.692	0.196	0.622	0.283	0.485	−0.423	0.330	−0.696
	(1.373)	(1.057)	(1.375)	(1.065)	(1.387)	(1.076)	(1.384)	(1.080)
Observations	3,706	3,706	3,706	3,706	3,706	3,706	3,706	3,706
Log likelihood	−2433	−2433	−2402	−2402	−2390	−2390	−2378	−2378
Chi2 statistics	487.4	487.4	535.1	535.1	550.6	550.6	557.3	557.3
p-value	0	0	0	0	0	0	0	0

*Robust standard errors in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1.*

TABLE 10 | Multinomial probit model regression results, financial education as the main explanatory variable.

Variables	Occasional smoker	Regular smoker	Occasional smoker	Regular smoker	Occasional smoker	Regular smoker	Occasional smoker	Regular smoker
	Model 5.1	Model 5.2	Model 5.3	Model 5.4	Model 5.5	Model 5.6	Model 5.7	Model 5.8
Financial education	0.0365	−0.215**	0.0392	−0.178*	0.0464	−0.186*	0.0426	−0.198**
	(0.128)	(0.0996)	(0.128)	(0.100)	(0.128)	(0.101)	(0.128)	(0.101)
Male	0.619***	1.436***	0.597***	1.328***	0.590***	1.313***	0.573***	1.288***
	(0.0967)	(0.0744)	(0.105)	(0.0791)	(0.105)	(0.0792)	(0.106)	(0.0797)
Age	0.00902	0.112***	0.00534	0.0958***	0.00639	0.101***	0.00224	0.0936***
	(0.0265)	(0.0221)	(0.0265)	(0.0221)	(0.0265)	(0.0221)	(0.0267)	(0.0222)
Age squared	−0.000218	−0.00129***	−0.000175	−0.00112***	−0.000191	−0.00116***	−0.000160	−0.00111***
	(0.000271)	(0.000222)	(0.000271)	(0.000222)	(0.000270)	(0.000223)	(0.000271)	(0.000224)
University degree	−0.232**	−0.439***	−0.242**	−0.421***	−0.230**	−0.382***	−0.213*	−0.354***
	(0.112)	(0.0836)	(0.113)	(0.0840)	(0.112)	(0.0846)	(0.113)	(0.0850)
Marriage	0.0211	−0.121	0.0149	−0.0977	0.0187	−0.0885	0.0681	0.00139
	(0.175)	(0.152)	(0.177)	(0.151)	(0.176)	(0.152)	(0.178)	(0.155)
Divorce	0.469*	0.580***	0.454	0.598***	0.444	0.585***	0.461*	0.611***
	(0.280)	(0.219)	(0.280)	(0.219)	(0.280)	(0.220)	(0.280)	(0.221)
Household member	0.0464	−0.00413	0.0416	−0.0114	0.0404	−0.0148	0.0351	−0.0255
	(0.0383)	(0.0283)	(0.0382)	(0.0286)	(0.0383)	(0.0286)	(0.0384)	(0.0287)
Children	−0.260	0.0835	−0.257	0.0725	−0.249	0.0812	−0.237	0.106
	(0.175)	(0.151)	(0.176)	(0.151)	(0.177)	(0.152)	(0.177)	(0.153)
Unemployed	−0.312	−0.0624	−0.299	0.0108	−0.299	0.0167	−0.350	−0.0826
	(0.328)	(0.230)	(0.329)	(0.227)	(0.330)	(0.227)	(0.336)	(0.231)
Log of household income	−0.140	−0.170**	−0.139	−0.182***	−0.136	−0.174***	−0.105	−0.121*
	(0.0921)	(0.0662)	(0.0924)	(0.0669)	(0.0925)	(0.0670)	(0.0931)	(0.0684)
Log of household assets	−0.0660	−0.129***	−0.0564	−0.0991**	−0.0540	−0.0895**	−0.0440	−0.0697*
	(0.0536)	(0.0410)	(0.0540)	(0.0413)	(0.0543)	(0.0414)	(0.0546)	(0.0417)
Regular exercise			−0.156	−0.390***	−0.150	−0.379***	−0.128	−0.336***
			(0.0975)	(0.0754)	(0.0977)	(0.0755)	(0.0981)	(0.0761)
Current drinker			0.162	0.326***	0.166*	0.324***	0.169*	0.330***
			(0.101)	(0.0747)	(0.100)	(0.0749)	(0.101)	(0.0752)
Frequent gambler			−0.189	0.391***	−0.193	0.360***	−0.200	0.346***
			(0.175)	(0.109)	(0.175)	(0.109)	(0.175)	(0.110)
Myopic view of the future					0.223*	0.262***	0.211*	0.239**
					(0.125)	(0.0968)	(0.125)	(0.0973)
Level of risk preference					0.102	0.750***	0.111	0.764***
					(0.249)	(0.184)	(0.249)	(0.186)
Current level of happiness							−0.559**	−1.005***
							(0.271)	(0.203)
Anxiety about health							−0.00679	−0.00260
							(0.0953)	(0.0721)
Constant	1.212	0.799	1.120	0.830	0.932	−0.00556	0.754	−0.332
	(1.353)	(1.031)	(1.354)	(1.040)	(1.366)	(1.054)	(1.365)	(1.058)
Observations	3,706	3,706	3,706	3,706	3,706	3,706	3,706	3,706
Log likelihood	−2435	−2435	−2404	−2404	−2391	−2391	−2379	−2379
Chi2 statistics	485.3	485.3	535.9	535.9	554.2	554.2	563.6	563.6
p-value	0	0	0	0	0	0	0	0

*Robust standard errors in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1.*

TABLE 11 | Multinomial probit model regression results, financial literacy, and financial education as the main explanatory variables.

Variables	Occasional smoker	Regular smoker	Occasional smoker	Regular smoker	Occasional smoker	Regular smoker	Occasional smoker	Regular smoker
	Model 6.1	Model 6.2	Model 6.3	Model 6.4	Model 6.5	Model 6.6	Model 6.7	Model 6.8
Financial literacy	−0.268*	−0.314***	−0.256*	−0.290***	−0.242*	−0.245**	−0.233	−0.222**
	(0.143)	(0.109)	(0.142)	(0.110)	(0.143)	(0.111)	(0.142)	(0.110)
Financial education	0.0386	−0.213**	0.0408	−0.176*	0.0480	−0.184*	0.0442	−0.196*
	(0.128)	(0.0995)	(0.128)	(0.100)	(0.128)	(0.101)	(0.128)	(0.100)
Male	0.644***	1.468***	0.621***	1.357***	0.613***	1.337***	0.595***	1.310***
	(0.0979)	(0.0750)	(0.106)	(0.0796)	(0.106)	(0.0797)	(0.107)	(0.0802)
Age	0.0138	0.118***	0.0101	0.101***	0.0107	0.105***	0.00644	0.0977***
	(0.0266)	(0.0223)	(0.0267)	(0.0222)	(0.0267)	(0.0222)	(0.0269)	(0.0224)
Age squared	−0.000260	−0.00134***	−0.000217	−0.00116***	−0.000229	−0.00120***	−0.000197	−0.00114***
	(0.000271)	(0.000224)	(0.000272)	(0.000224)	(0.000272)	(0.000224)	(0.000272)	(0.000225)
University degree	−0.197*	−0.395***	−0.208*	−0.381***	−0.200*	−0.350***	−0.185	−0.326***
	(0.115)	(0.0844)	(0.115)	(0.0848)	(0.115)	(0.0853)	(0.115)	(0.0858)
Marriage	0.0309	−0.110	0.0237	−0.0892	0.0268	−0.0810	0.0746	0.00650
	(0.174)	(0.151)	(0.176)	(0.150)	(0.175)	(0.152)	(0.177)	(0.154)
Divorce	0.484*	0.597***	0.467*	0.612***	0.457	0.597***	0.474*	0.622***
	(0.280)	(0.220)	(0.280)	(0.219)	(0.280)	(0.220)	(0.280)	(0.221)
Household member	0.0401	−0.0111	0.0359	−0.0177	0.0352	−0.0200	0.0302	−0.0301
	(0.0384)	(0.0284)	(0.0383)	(0.0287)	(0.0384)	(0.0287)	(0.0385)	(0.0288)
Children	−0.276	0.0655	−0.272	0.0566	−0.263	0.0668	−0.252	0.0927
	(0.175)	(0.151)	(0.176)	(0.151)	(0.177)	(0.151)	(0.178)	(0.153)
Unemployed	−0.322	−0.0793	−0.308	−0.00519	−0.308	0.00268	−0.357	−0.0933
	(0.330)	(0.229)	(0.331)	(0.227)	(0.333)	(0.228)	(0.338)	(0.231)
Log of household income	−0.116	−0.142**	−0.116	−0.157**	−0.114	−0.153**	−0.0856	−0.103
	(0.0924)	(0.0670)	(0.0927)	(0.0675)	(0.0928)	(0.0676)	(0.0935)	(0.0691)
Log of household assets	−0.0540	−0.114***	−0.0454	−0.0863**	−0.0437	−0.0790*	−0.0343	−0.0605
	(0.0537)	(0.0414)	(0.0542)	(0.0418)	(0.0545)	(0.0419)	(0.0547)	(0.0421)
Regular exercise			−0.146	−0.383***	−0.142	−0.373***	−0.121	−0.331***
			(0.0972)	(0.0755)	(0.0975)	(0.0756)	(0.0979)	(0.0761)
Current drinker			0.163	0.325***	0.166*	0.324***	0.170*	0.330***
			(0.101)	(0.0749)	(0.101)	(0.0751)	(0.101)	(0.0753)
Frequent gambler			−0.189	0.388***	−0.192	0.359***	−0.199	0.346***
			(0.175)	(0.109)	(0.175)	(0.110)	(0.175)	(0.110)
Myopic view of the future					0.213*	0.249**	0.202	0.227**
					(0.126)	(0.0969)	(0.126)	(0.0974)
Level of risk preference					0.0764	0.727***	0.0855	0.742***
					(0.248)	(0.185)	(0.249)	(0.186)
Current level of happiness							−0.545**	−0.988***
							(0.271)	(0.203)
Anxiety about health							−0.00764	−0.00506
							(0.0955)	(0.0722)
Constant	0.683	0.181	0.617	0.264	0.481	−0.456	0.325	−0.735
	(1.370)	(1.057)	(1.373)	(1.065)	(1.383)	(1.077)	(1.381)	(1.081)
Observations	3,706	3,706	3,706	3,706	3,706	3,706	3,706	3,706
Log likelihood	−2430	−2430	−2400	−2400	−2388	−2388	−2376	−2376
Chi2 statistics	495.7	495.7	542.8	542.8	560.3	560.3	568.4	568.4
p-value	0	0	0	0	0	0	0	0

*Robust standard errors in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1.*

to make more rational decisions, are less likely to smoke. Using data from the PPS 2010, the probit regression results show that both financial literacy and financial education have a significantly negative impact on smoking behavior. In other words, a high level of financial literacy (which emphasizes knowledge of investments) and receiving financial education (which emphasizes savings behavior) significantly reduces the probability of being a current smoker. These inverse relationships indicate that rational decision-making ability reduces the tendency to engage in irrational behavior. In addition, we find that gender, age, education, marital status, household income and assets, risky behaviors, myopic view of the future, risk preference, and level of happiness also significantly predict the likelihood of a person being a current smoker.

Our findings suggest that promoting financial literacy and financial education could help mitigate smoking behavior. In the case of Japan, the government can integrate financial knowledge that focuses on the investment perspective into the current financial education program to enhance the impact of financial education and financial literacy as countermeasures to smoking. However, further research is needed to explore the impact of financial literacy and financial education as rational decision-making tools in terms of smoking and other health-related behaviors in different contexts.

However, this study has several limitations. First, we base our measurement of financial literacy on only three questions designed by Lusardi and Mitchell (63, 70). However, other studies also use this method, which makes financial literacy internationally comparable [e.g., (48, 59, 64–69, 71, 104, 105)]. Second, we define smoking behavior only in terms of participation in smoking, but not the amount of cigarette consumption or the level of tobacco dependence because of data unavailability. Despite these limitations, this study provides empirical evidence that suggests a means to enhance the rational decision-making ability of individuals through financial literacy and financial education to curtail smoking behavior.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Osaka University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YK: study design. SW, MK, and YK: analysis and interpretation of data. SW, PP, SO, and MK: writing of the report. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by JPPS KAKENHI Grant Nos. 19K13739 (YK); 19K13684 (MK), and RISTEX, JST (YK).

ACKNOWLEDGMENTS

This research used micro data from the Preference Parameters Study of Osaka University's twenty-first Century COE Program Behavioral Macrodynamics Based on Surveys and Experiments and its Global COE project Human Behavior and Socioeconomic Dynamics. The authors acknowledge the program/project's contributions of Yoshiro Tsutsui, Fumio Ohtake, and Shinsuke Ikeda.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2020.612976/full#supplementary-material

REFERENCES

1. Ministry of Health, Labour and Welfare (MHLW) (2012). A basic direction for comprehensive implementation of national health promotion. Available online at: https://www.mhlw.go.jp/file/06-Seisakujouhou-10900000-Kenkoukyoku/0000047330.pdf (accessed July 31, 2020).
2. Ministry of Health, Labour and Welfare (MHLW). (2019). Outline of the Act on the Partial Revision of the Health Promotion Act, (No. 78 of 2018). Available online at: https://www.mhlw.go.jp/english/policy/health-medical/health/dl/201904kenko.pdf (accessed July 31, 2020).
3. World Health Organization (WHO). 2015. WHO report on the global tobacco epidemic: raising taxes on tobacco, World Health Organization. Available online at: https://www.who.int/tobacco/global_report/2015/en/ (accessed August 8, 2020).
4. World Health Organization (WHO). 2017. WHO report on the global tobacco epidemic, 2017: Monitoring tobacco use and prevention policies, World Health Organization. Available online at: https://www.who.int/tobacco/global_report/2017/en/ (accessed August 8, 2020).
5. World Health Organization (WHO). 2019. WHO report on the global epidemic 2019: Offer help to quit tobacco use, World Health Organization. Available online at: https://www.who.int/publications/i/item/who-report-on-the-global-tobacco-epidemic-2019-offer-help-to-quit-tobacco-use (accessed August 8, 2020).
6. GBD 2015 Tobacco Collaborators. Smoking prevalence and attributable disease burden in 195 countries and territories, 1990-2015: a systematic analysis from the global burden of disease study 2015. *Lancet*. (2017) 389:1885–906. doi: 10.1016/S0140-6736(17)30819-X
7. Farrell M, Howes S, Bebbington P, Brugha T, Jenkins R, Lewis G, et al. Nicotine, alcohol and drug dependence, and psychiatric comorbidity – results of a national household survey. *Int Rev Psychiatry*. (2003) 15:50–6. doi: 10.1080/0954026021000045949
8. Lawrence D, Mitrou F, Zubrick SR. Smoking and mental illness: results from population surveys in Australia and the United States.

BMC Public Health. (2009) 9:1–14. doi: 10.1186/1471-2458-9-285
9. U.S. Department of Health and Human Services (USDHHS). The health consequences of smoking: a report of the Surgeon General. *Natl Libr Med*. (2004).
10. Midya V, Pal S, Sinharoy A, Das JK, Rao H, Abu-Hasan M, et al. The association between female smoking and childhood asthma prevalence–a study based on aggregative data. *Front Public Health*. (2018) 6:295. doi: 10.3389/fpubh.2018.00295
11. Ibrahim BA, Al-Humaish S, Al-Obaide MAI. Tobacco smoking, lung cancer, and therapy in Iraq: current perspective. *Front Public Health*. (2018) 6:311. doi: 10.3389/fpubh.2018.00311
12. Grine G, Royer A, Terrer E, Diallo OO, Drancourt M, Aboudharam G. Tobacco smoking affects the salivary gram-positive bacterial population. *Front Public Health*. (2019) 7:196. doi: 10.3389/fpubh.2019.00196
13. Wen H, Xie C, Wang F, Wu Y, Yu C. Trends in disease burden attributable to Tobacco in China, 1990-2017: Findings from the global burden of disease study 2017. *Front Public Health*. (2020) 8:237. doi: 10.3389/fpubh.2020.00237
14. Bush R, Wooden M. Smoking and absence from work: Australian evidence. *Soc Sci Med*. (1995) 41:437–46. doi: 10.1016/0277-9536(94)00350-3
15. Champion KE, Mather M, Spring B, Kay-Lambkin F, Teesson M, Newton NC. Clustering of multiple risk behaviors among a sample of 18-year-old Australians and associations with mental health outcomes: a latent class analysis. *Front Public Health*. (2016) 6:135. doi: 10.3389/fpubh.2018.00135
16. Hori A, Inoue Y, Kuwahara K, Kunugita N, Akter S, Nishiura C, et al. Smoking and long-term sick leave in a Japanese working population: findings of the Japan epidemiology collaboration on occupational health study. *Nicotine Tob. Res*. (2019). doi: 10.1093/ntr/ntz204. [Epub ahead of print].
17. Ekpu VU, Brown AK. The economic impact of smoking and of reducing smoking prevalence: review of evidence. *Tob Use Insights*. (2015) 8:TUI.S15628. doi: 10.4137/TUI.S15628
18. Hajat A, Kaufman JS, Rose KM, Siddiqi A, Thomas JC. Do the wealthy have a health advantage? Cardiovascular disease risk factors and wealth. *Soc Sci Med*. (2010) 71:1935–42. doi: 10.1016/j.socscimed.2010.09.027
19. Cawley J. Reefer madness, frank the tank or pretty woman: to what extent do addictive behaviors respond to incentives? In: Sloan FA, Kasper H, editors. *Incentives and Choice in Health Care*. Cambridge, MA: MIT Press (2008). p. 163–93. doi: 10.7551/mitpress/9780262195775.003.0007
20. Sloan FA, Wang Y. Economic theory and evidence on smoking behavior of adults. *Addiction*. (2008) 103:1777–85. doi: 10.1111/j.1360-0443.2008.02329.x
21. Baker F, Johnson MW, Bickel WK. Delay discounting in current and never-before cigarette smokers: similarities and differences across commodity, sign, and magnitude. *J Abnorm Psychol*. (2003) 112:382–92. doi: 10.1037/0021-843X.112.3.382
22. Barlow P, McKee M, Reeves A, Galea G, Stuckler D. Time-discounting and tobacco smoking: a systematic review and network analysis. *Int J Epidemiol*. (2017) 46:860–9. doi: 10.1093/ije/dyx060
23. MacKillop J, Amlung MT, Few LR, Ray LA, Sweet LH, Munafò MR. Delayed reward discounting and addictive behavior: a meta-analysis. *Psychopharmacology*. (2011) 216:305–21. doi: 10.1007/s00213-011-2229-0
24. Ida T, Goto R. Simultaneous measurement of time and risk preferences: stated preference discrete choice modeling analysis depending on smoking behavior. *Int Econ Rev*. (2009) 50:1169–82. doi: 10.1111/j.1468-2354.2009.00564.x
25. Mitchell SH. Measures of impulsivity in cigarette smokers and non-smokers. *Psychopharmacology*. (1999) 146:455–64. doi: 10.1007/PL00005491
26. O'Donoghue T, Rabin M. Doing it now or later. *Am Econ Rev*. (1999) 89:103–24. doi: 10.1257/aer.89.1.103
27. Gruber J, Köszegi B. Is addiction rational: theory and evidence. *Q J Econ*. (2001) 116:1261–304. doi: 10.1162/003355301753265570
28. Rambaud SC, Torrecillas MJM, Takahashi T. Observed and normative discount functions in addiction and other diseases. *Front. Pharmacol*. (2017) 8:416. doi: 10.3389/fphar.2017.00416
29. Read D, Roelofsma PHMP. Subadditive versus hyperbolic discounting: a comparison of choice and matching. *Organ Behav Hum Decis Process*. (2003) 91:140–53. doi: 10.1016/S0749-5978(03)00060-8
30. Read D. Is time-discounting hyperbolic or subadditive? *J Risk Uncertain*. (2001) 23:5–32. doi: 10.1023/A:1011198414683
31. Orphanides A, Zervos D. Rational addiction with learning and regret. *J Polit Econ*. (1995) 103:739–58. doi: 10.1086/262001
32. Wang R. The optimal consumption and the quitting of harmful addictive goods. *B E J Econ Anal Policy*. (2007) 7:1–38. doi: 10.2202/1935-1682.1684
33. Ambrose BK, Rostron BL, Johnson SE, Portnoy DB, Apelberg BJ, Kaufman AR et al. Perceptions of the relative harm of cigarettes and e-cigarettes among U.S. *youth. Am J Prev Med*. (2014) 47:S53–60. doi: 10.1016/j.amepre.2014.04.016
34. Lee YO, Hebert CJ, Nonnemaker JM, Kim AE. Youth tobacco product use in the United States. *Pediatrics*. (2015) 135:409–15. doi: 10.1542/peds.2014-3202
35. Drovandi A, Teague P-A, Glass B, Malau-Aduli B. Australian school student perceptions of effective anti-tobacco health warnings. *Front Public Health*. (2018) 6:297. doi: 10.3389/fpubh.2018.00297
36. Grossman M. The human capital model of the demand for health. *NBER working paper No. 7078. National Bureau of Economic Research*. (1999). doi: 10.3386/w7078
37. World Health Organization (WHO). (2011). Economics of tobacco toolkit: assessment of the economic costs of smoking, World Health Organization. Available online at: https://apps.who.int/iris/handle/10665/44596 (accessed August 8, 2020).
38. Loewenstein GFA. Visceral account of addiction. In: Elster J, Skog OJ, editors. *Getting Hooked: Rationality and Addiction*. New York, NY: Cambridge University Press (1999). p. 235–64. doi: 10.1017/CBO9781139173223.010
39. Perikleous EP, Paschalis S, Emmanouil, P, Theodoros CC, Evangelia N. E-Cigarette use among adolescents: an overview of the literature and future perspectives. *Front Public Health*. (2018) 6:86. doi: 10.3389/fpubh.2018.00086
40. Braunstein S, Welch C, Financial literacy: an overview of practice, research, and policy. *Fed Res Bull*. (2002) 88:445. doi: 10.17016/bulletin.2002.88-11
41. Lusardi A, Michaud P-C, Mitchell OS. Optimal financial literacy and saving for retirement. (2011) RAND Working Paper Series No. WR-905-SSA. doi: 10.2139/ssrn.1978960
42. Yoshino N, Morgan PJ, Trinh LQ. Financial literacy in Japan: determinants and impacts, ADBI Working Paper 796 (2014).
43. Hastings J, Mitchell OS. How financial literacy and impatience shape retirement wealth and investment behaviors. *J Pension Econ Financ*. (2020) 19:1–20. doi: 10.1017/S1474747218000227
44. Meier S, Sprenger C. Present-biased preferences and credit card borrowing. *Am Econ J Appl Econ*. (2010) 2:193–210. doi: 10.1257/app.2.1.193
45. Alessie R, Van Rooij M, Lusardi A. Financial literacy and retirement preparation in the Netherlands. *J Pension Econ Financ*. (2011) 10:527–45. doi: 10.1017/S1474747211000461
46. Boisclair D, Lusardi A, Michaud PC. Financial literacy and retirement planning in Canada. *J Pension Econ Financ*. (2017) 16:277–96. doi: 10.1017/S1474747215000311
47. Khan MSR, Rabbani N, Kadoya Y. Is financial literacy associated with investment in financial markets in the United States? *Sustainability*. (2020) 12:1–14. doi: 10.3390/su12187370
48. Lusardi A, Mitchell OS. The economic importance of financial literacy: theory and evidence. *J Econ Lit*. (2014) 52:5–44. doi: 10.1257/jel.52.1.5
49. Mandell L. Financial literacy of high school students. In: Xiao JJ, editor. *Handbook of Consumer Finance Research*. New York, NY: Springer (2008). p. 163–83. doi: 10.1007/978-0-387-75734-6_10
50. Willis LE. The financial education fallacy. *Am Econ Rev*. (2011) 101:429–34. doi: 10.1257/aer.101.3.429
51. Benjamin DJ, Brown SA, Shapiro JM. Who is "behavioral"? Cognitive ability and anomalous preferences. *J Eur Econ Assoc*. (2013) 11:1231–55. doi: 10.1111/jeea.12055
52. Gill D, Prowse V. Cognitive ability, character skills, and learning to play equilibrium: a level-k analysis. *J Polit Econ*. (2016) 124:1619–76. doi: 10.1086/688849

53. James BD, Boyle PA, Bennett JS, Bennett DA. The impact of health and financial literacy on decision making in community-based older adults. *Gerontology.* (2012) 58:531–9. doi: 10.1159/000339094
54. Bennett JS, Boyle PA, James BD, Bennett DA, Correlates of health and financial literacy in older adults without dementia. *BMC Geriatr*. (2012) 12:30. doi: 10.1186/1471-2318-12-30
55. Bernheim BD, Garrett DM, Maki DM. Education and saving: the long-term effects of high school financial curriculum mandates. *J Public Econ.* (2001) 80:435–465. doi: 10.1016/S0047-2727(00)00120-1
56. Brown M, Grigsby J, Van Der Klaauw W, Wen J, Zafar B. Financial education and the debt behavior of the young. *Rev Financ Stud.* (2016) 29:2490–522. doi: 10.1093/rfs/hhw006
57. Bruhn M, Leão L de S, Legovini A, Marchetti R, Zia B. The impact of high school financial education: evidence from a large-scale evaluation in Brazil. *Am Econ J Appl Econ.* (2016) 8:256–95. doi: 10.1257/app.20150149
58. Kaiser T, Menkhoff L. Does financial education impact financial literacy and financial behavior, and if so, when? *World Bank Economic Review*. (2017) 31:611–30. doi: 10.1596/1813-9450-8161
59. Watanapongvanich S, Binnagan P, Putthinun P, Khan MSR, Kadoya Y. Financial literacy and gambling behavior: evidence from Japan. *J Gambl Stud.* (2020). doi: 10.1007/s10899-020-09936-3. [Epub ahead of print].
60. Frederick S, Loewenstein G, O'Donoghue T. Time discounting and time preference: a critical review. *Time Decis Econ Psychol Perspect Intertemporal Choice.* (2003) 40:13–86. doi: 10.1257/002205102320161311
61. Hurla R, Kim M, Singer E, Soman D. *Applying Findings From Financial Literacy to Encourage Responsible Gambling*. Toronto, ON: Gambling Research Exchange Ontario (GREO). (2017).
62. Kang M Il, Ikeda S. Time discounting and smoking behavior: evidence from a panel survey. *Behav Econ Prefer Choices Happiness.* (2016) 23:197–226. doi: 10.1007/978-4-431-55402-8_9
63. Lusardi A, Mitchell OS. Planning and financial literacy: how do women fare? *Am Econ Rev*. (2008) 98:413–7. doi: 10.1257/aer.98.2.413
64. Fornero E, Monticone C. Financial literacy and pension plan participation in Italy. *J Pension Econ Financ.* (2011) 10:547–64. doi: 10.1017/S1474747211000473
65. Kadoya Y, Khan MSR. Can financial literacy reduce anxiety about life in old age? *J Risk Res.* (2018) 21:1533–50. doi: 10.1080/13669877.2017.1313760
66. Kadoya Y, Khan MSR. What determines financial literacy in Japan? *J Pension Econ Financ.* (2020) 19:353–71. doi: 10.1017/S1474747218000379
67. Kadoya Y, Khan MSR, Hamada T, Dominguez A. Financial literacy and anxiety about life in old age: evidence from the USA. *Rev Econ Househ.* (2018) 16:859–78. doi: 10.1007/s11150-017-9401-1
68. Kadoya Y, Khan MSR, Oba H, Narumoto J. Factors affecting knowledge about the adult guardianship and civil trust systems: evidence from Japan. *J Women Aging.* (2020). doi: 10.1080/08952841.2020.1727711. [Epub ahead of print].
69. Klapper L, Panos GA. Financial literacy and retirement planning: the Russian case. *J Pension Econ Financ.* (2011) 10:599–618. doi: 10.1017/S1474747211000503
70. Lusardi A, Mitchell OS. Financial literacy around the world: an overview. *J Pension Econ Financ.* (2011) 10:497–508. doi: 10.1017/S1474747211000448
71. Sekita S. Financial literacy and retirement planning in Japan. *J Pension Econ Financ.* (2011) 10:637–56. doi: 10.1017/S1474747211000527
72. Messy F-A, Monticone C. Financial Education Policies in Asia and the Pacific. *OECD Working Paper. Financ Insur Priv Pensions.* (2016) 40:66. doi: 10.1787/5jm5b32v5vvc-en
73. Aekplakorn W, Hogan MC, Tiptaradol S, Wibulpolprasert S, Punyaratabandhu P, Lim SS. Tobacco and hazardous or harmful alcohol use in Thailand: joint prevalence and associations with socioeconomic factors. *Addict Behav.* (2008) 33:503–14. doi: 10.1016/j.addbeh.2007.10.010
74. Audrain-McGovern J, Rodriguez D, Moss HB, Smoking progression and physical activity. *Cancer Epidemiol Biomarkers Prev.* (2003) 12:1121–9.
75. Bauer T, Göhlmann S, Sinning M. Gender differences in smoking behavior. *Health Econ.* (2007) 16:895–909. doi: 10.1002/hec.1259
76. Giovino GA, Mirza SA, Samet JM, Gupta PC, Jarvis MJ, Bhala N, et al. Tobacco use in 3 billion individuals from 16 countries: an analysis of nationally representative cross-sectional household surveys. *Lancet.* (2012) 380:668–79. doi: 10.1016/S0140-6736(12)61085-X
77. Lahelma E, Pietiläinen O, Ferrie J, Kivimäki M, Lahti J, Marmot M, et al. Changes over time in absolute and relative socioeconomic differences in smoking: a comparison of cohort studies from Britain, Finland, and Japan. *Nicotine Tob Res.* (2016) 18:1697–704. doi: 10.1093/ntr/ntw004
78. Witvorapong N, Vichitkunakorn P. Investigation of tobacco and alcohol co-consumption in Thailand: a joint estimation approach. *Drug Alcohol Rev*. (2020). doi: 10.1111/dar.13128. [Epub ahead of print].
79. Fukuda Y, Nakamura K, Takano T. Socioeconomic pattern of smoking in Japan: income inequality and gender and age differences. *Ann Epidemiol.* (2005) 15:365–72. doi: 10.1016/j.annepidem.2004.09.003
80. Meijer E, Gebhardt WA, Van Laar C, Kawous R, Beijk SCAM. Socio-economic status in relation to smoking: the role of (expected and desired) social support and quitter identity. *Soc Sci Med.* (2016) 162:41–9. doi: 10.1016/j.socscimed.2016.06.022
81. Pennanen M, Broms U, Korhonen T, Haukkala A, Partonen T, Tuulio-Henriksson A, et al. Smoking, nicotine dependence and nicotine intake by socio-economic status and marital status. *Addict Behav.* (2014) 39:1145–51. doi: 10.1016/j.addbeh.2014.03.005
82. Reid JL, Hammond D, Boudreau C, Fong GT, Siahpush M. Socioeconomic disparities in quit intentions, quit attempts, and smoking abstinence among smokers in four western countries: findings from the International Tobacco Control Four Country Survey. *Nicotine Tob Res.* (2010) 12:S20–33. doi: 10.1093/ntr/ntq051
83. Schaap MM, Kunst AE. Monitoring of socio-economic inequalities in smoking: learning from the experiences of recent scientific studies. *Public Health.* (2009) 123:103–9. doi: 10.1016/j.puhe.2008.10.015
84. Vedøy TF. Tracing the cigarette epidemic: an age-period-cohort study of education, gender and smoking using a pseudo-panel approach. *Soc Sci Res.* (2014) 48:35–47. doi: 10.1016/j.ssresearch.2014.05.005
85. Wetter DW, Cofta-Gunn L, Irvin JE, Fouladi RT, Wright K, Daza P, et al. What accounts for the association of education and smoking cessation? *Prev Med.* (2005) 40:452–60. doi: 10.1016/j.ypmed.2004.07.001
86. Zhu BP, Giovino GA, Mowery PD, Eriksen MP. The relationship between cigarette smoking and education revisited: implications for categorizing persons' educational status. *Am J Public Health.* (1996) 86:1582–9. doi: 10.2105/AJPH.86.11.1582
87. Lindström M. Social capital, economic conditions, marital status and daily smoking: a population-based study. *Public Health.* (2010) 124:71–7. doi: 10.1016/j.puhe.2010.01.003
88. Castrén S, Kontto J, Alho H, Salonen AH. The relationship between gambling expenditure, socio-demographics, health-related correlates, and gambling behaviour—a cross-sectional population-based survey in Finland. *Addiction.* (2018) 113:91–106. doi: 10.1111/add.13929
89. Charilaou M, Karekla M, Constantinou M, Price S. Relationship between physical activity and type of smoking behavior among adolescents and young adults in Cyprus. *Nicotine Tob Res.* (2009) 11:969–76. doi: 10.1093/ntr/ntp096
90. Chasandra M, Goudas M, Theodorakis Y. Exercise and smoking: a literature overview. *Health (Irvine. Calif).* (2015) 07:1477–91. doi: 10.4236/health.2015.711162
91. Leatherdale ST, Wong SL, Manske SR, Colditz GA. Susceptibility to smoking and its association with physical activity, BMI, and weight concerns among youth. *Nicotine Tob Res.* (2008) 10:499–505. doi: 10.1080/14622200801902201
92. Nichter M, Nichter M, Carkoglu A, Lloyd-Richardson E. Smoking and drinking among college students: "It's a package deal." *Drug Alcohol Depend.* (2010) 106:16–20. doi: 10.1016/j.drugalcdep.2009.07.025
93. Bobo JK, Husten C. Sociocultural influences on smoking and drinking. *Alcohol Res Heal.* (2000) 24:225–32.
94. Weitzman ER, Chen YY. The co-occurrence of smoking and drinking among young adults in college: National survey results from the United States. *Drug Alcohol Depend.* (2005) 80:377–86. doi: 10.1016/j.drugalcdep.2005.05.008
95. Bussu A, Detotto C. The bidirectional relationship between gambling and addictive substances. *Int Gambl Stud.* (2015) 15:285–308. doi: 10.1080/14459795.2015.1042493

96. Lorains FK, Cowlishaw S, Thomas SA. Prevalence of comorbid disorders in problem and pathological gambling: systematic review and meta-analysis of population surveys. *Addiction*. (2011) 106:490–8. doi: 10.1111/j.1360-0443.2010.03300.x
97. Petry NM, Stinson FS, Grant BF. Comorbidity of DSM-IV pathological gambling and other psychiatric disorders: results from the national epidemiologic survey on alcohol and related conditions. *J Clin Psychiatry*. (2005) 66:564–74. doi: 10.4088/JCP.v66n0504
98. Odum AL, Madden GJ, Bickel WK. Discounting of delayed health gains and losses by current, never- and ex-smokers of cigarettes. *Nicotine Tob Res.* (2002) 4:295–303. doi: 10.1080/14622200210141257
99. Anderson LR, Mellor JM, Predicting health behaviors with an experimental measure of risk preference. *J Health Econ.* (2008) 27:1260–74. doi: 10.1016/j.jhealeco.2008.05.011
100. Pfeifer C. A note on smoking behavior and health risk taking. *Nord J Heal Econ.* (2012) 1:135–51. doi: 10.5617/njhe.195
101. Chang T, Chu HP, Deale FW, Gupta R. The causal relationship between happiness and smoking: a bootstrap panel causality test. *J Happiness Stud.* (2016) 17:1327–36. doi: 10.1007/s10902-015-9645-5
102. Kassel JD, Stroud LR, Paronis CA. Smoking, stress, and negative affect: correlation, causation, and context across stages of smoking. *Psychol Bull.* (2003) 129:270–304. doi: 10.1037/0033-2909.129.2.270
103. Weinhold D, Chaloupka FJ. Smoking status and subjective well-being. *Tob Control.* (2017) 26:195–201. doi: 10.1136/tobaccocontrol-2015-052601
104. Almenberg J, Säve-Söderbergh J, Financial literacy and retirement planning in Sweden. *J Pension Econ Financ.* (2011) 10:585–98. doi: 10.1017/S1474747211000497
105. Crossan D, Feslier D, Hurnard R. Financial literacy and retirement planning in New Zealand. *J Pension Econ Financ.* (2011) 10:619–35. doi: 10.1017/S1474747211000515

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

ORIGINAL RESEARCH
published: 20 January 2021
doi: 10.3389/fpubh.2020.582140

Associating the Change in New COVID-19 Cases to GDP per Capita in 38 European Countries in the First Wave of the Pandemic

Shahina Pardhan[1]* *and Nick Drydakis*[2,3,4,5]

[1] Vision and Eye Research Institute, School of Medicine, Anglia Ruskin University, Cambridge, United Kingdom, [2] Centre for Pluralist Economics, Faculty of Business and Law, Anglia Ruskin University, Cambridge, United Kingdom, [3] Pembroke College, University of Cambridge, Cambridge, United Kingdom, [4] Institute of Labor Economics, Bonn, Germany, [5] Global Labor Organization, Essen, Germany

COVID-19 has affected all countries globally. We explore associations between the change in new COVID-19 registered cases per million population and various macroeconomic and well-being indicators in 38 European countries over a 2-month period (1st April-31st May 2020). A statistically significant ($p = 0.002$) negative association was estimated between the change in new COVID-19 cases and GDP per capita, after controlling for key health determinants including public expenditure on health, life expectancy, smoking tobacco and sanitation. The country with the highest GDP per capita in Europe (i.e., Luxemburg) was found to experience the lowest change in new COVID-19 cases within the time period whilst the opposite was found for countries with lower GDP per capita (i.e., Ukraine, Bulgaria, and Romania). The outcomes of this study indicate that, in the first wave of the pandemic in Europe, a country's GDP per capita might be associated with a lower rate of new COVID-19 cases. The study concludes by suggesting that in European regions a country's economic performance should be a critical health priority for policy makers.

Keywords: COVID-19, GDP per capita, life expectancy, sanitation, smoking

OPEN ACCESS

Edited by:
Mirjana Ratko Jovanovic, University of Kragujevac, Serbia

Reviewed by:
Brian Godman, University of Strathclyde, United Kingdom
Simon Grima, University of Malta, Malta

***Correspondence:**
Shahina Pardhan
shahina.pardhan@anglia.ac.uk

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *15 July 2020*
Accepted: *22 December 2020*
Published: *20 January 2021*

Citation:
Pardhan S and Drydakis N (2021) Associating the Change in New COVID-19 Cases to GDP per Capita in 38 European Countries in the First Wave of the Pandemic. Front. Public Health 8:582140. doi: 10.3389/fpubh.2020.582140

INTRODUCTION

The global pandemic caused by COVID-19 has affected every county in the world (1). Various factors have been shown to influence the rate of change of infection caused by SARS CoV-2-s as confirmed by the daily change in new cases and of mortality. Individual and demographic factors including older age (2, 3), male gender (4), socio-economic conditions (5), underlying health co-morbidities (6), ethnicity (7, 8), smoking (9, 10) and obesity (11) have been identified as significant influencing factors in the spread and mortality due to COVID-19. However, it has to be recognized that these individual factors are likely to be influenced by various country specific parameters including lockdown policies (12), public expenditure in health (13) and other country-specific determinants such as level of sanitation (14), healthcare support i.e., number of acute care beds, hospital beds and number of physicians; (15), and social support (16). Grima et al. (17) indicated that monitoring of demographic features, country's activity features, economic exposure and societal vulnerability could help a country strengthen its capacity to meet the economic, social and in turn healthcare demands due to pandemic hazards such as COVID-19.

In poorer countries, households experience tighter budgets and worse economic struggles (18, 19), resulting in poor physical and mental health and decreased life expectancy (18, 20). One of the most well-known illustrations of the relationship between economic conditions and population health is the Preston curve (21) which demonstrates that people in richer countries, on average, live longer than in poorer countries. Evidence indicates that a country's economic performance and fair income distribution results in an increase in life expectancy, and a decrease in adverse health outcomes and mortality rates (22). Theoretical insights indicate that countries with higher GDP rates are able to provide their people with better living standards, public health programs, education and environmental sanitation all leading to enhanced prevention, treatment of disease, better health and life expectancy (17, 22–26).

In the US, GDP and economic growth were found to be related to a decline in mortality rate between 1901 and 2000 (27). In OECD countries, for a period ranging from 1820 to 2001, GDP and GDP per capita were estimated to have a significant positive influence on life expectancy (28). In Europe, longitudinal studies, review studies and meta-analyses indicate that reductions in GDP, national health system budgets and households' income are associated with adverse health outcomes and deterioration in people's well-being (29–31). The Great Recession studies re-highlighted the positive associations of economic prosperity and adequate national health provision to better health outcomes (29–31). In Europe, since 1950, the rapid economic growth brought a degree of prosperity which enabled effective national health systems to be established, new drugs and medical technologies to be developed that brought several infections and diseases under control (32). However, better health systems, cutting-edge medical technologies and drugs required substantial funding (32). Within the World Health Organization (WHO) European regions, for a period ranging from 1900 to 2008, life expectancy was shown to be dependent on economic activity, and mortality from cardiovascular diseases were mainly attributable to changes in national income (23). These patterns have been found in other diverse settings. For instance, a recent study of young population groups in 103 low and middle-income countries showed that higher GDP was inversely associated with all-cause, communicable and non-communicable disease mortality both in males and females (33). A study examining empirical assessments for 17 European countries between 1970 and 2010, demonstrated that countries with higher national income, higher health care expenditure, higher quality of government, and higher social transfers have smaller inequalities in mortality (34).

Given the presented theoretical considerations and empirical findings, we hypothesize that poorer countries in Europe, characterized by a lower economic performance, and limited health and fiscal capacities of governments might underperform in tackling the COVID-19 pandemic. Restricted personal and family income, and inadequate public support for critical health services could affect health prevention and the quality and quantity of services experienced during the pandemic. In the US, those affected more by COVID-19 were living in poorer regions, had lower access to healthcare, experienced intergenerational poverty, and had a higher prevalence of underlying health conditions (35, 36). A recent study utilizing data from 188 countries found that COVID-19 has mainly affected vulnerable population groups with underlying health conditions (37).

Although GDP per capita is one of the most widely used covariates in health research (38), most current papers on COVID-19 have focused on individuals' socio-economic and health characteristics (35–37). However, there are few exceptions. Roy (39) presented a diagrammatic reasoning demonstrating a negative relationship between the total number of COVID-19 cases and GDP per capita. On the other hand, Gangemi et al. (40) found a moderate positive correlation between COVID-19 cases and GDP per capita as did Lippi et al. (41) who showed a positive correlation between COVID-19 mortality and GDP per capita. The studies suggested that industrial pollution, airplane connections, obesity and social events, which are higher in developed regions, might have driven the positive correlations (40, 41).

The aim of the present study is to explore correlations and associations between the change in new COVID-19 cases per million population, in the initial wave of the pandemic, obtained on two dates that were 2 months apart (1st April and 31st May 2020), and macroeconomic and well-being indicators in 38 European countries. In this study we indicate that the total number of COVID-19 cases may not be the right indicator to be utilized because case zero of COVID-19 case differs in each country as the pandemic affected regions at different times. In addition, we hypothesize that the change in new COVID-19 cases per million population might be more detrimental in regions where prior to the pandemic, people were experiencing worse health due to greater economic hardships and insecurity.

The present paper is among the first studies to offer multivariate regressions controlling for key heterogeneities and assess associations between GDP per capita and change in new COVID-19 cases in European regions in the first wave of the pandemic. The outcomes of the study will show (i) a negative correlation between the change in new COVID-19 cases per million population and GDP per capita, public expenditure on health, sanitation facilities and life expectancy at birth, and (ii) a positive correlation between the change in new COVID-19 cases and tobacco smoking. Moreover, the study's outcomes will reveal a negative association between the change in new COVID-19 cases per million population and GDP per capita after controlling for critical health heterogeneities. The study will conclude that in Europe a country's prosperity, as it is captured by GDP per capita, might be associated with a reduction in new COVID-19 cases within the study's time period. The assigned patterns will indicate that better economically performed countries might be able to respond to a health crisis and therefore a country's economic growth and development should be of importance. Indeed, the study will present that the best performing European country in terms of GDP per capita (i.e., Luxemburg) experienced the lowest change in new COVID-19 cases per million population.

The rest of the paper is structured as follows. In the next section we describe the data set and methodology. Then, we offer a correlation and regression analysis. The last section offers a discussion and conclusions.

MATERIALS AND METHODS

In this study data on the change in new COVID-19 cases per million population were obtained for 38 European countries, as some of the European data required for correlation and multivariate analyses were only available for certain countries[1]. In June 2020, COVID-19 data were extracted from Roser et al. (42) which is a publicly available data set.

Key macroeconomic and well-being indicators for the targeted European countries were considered after reviewing major socio-epidemiological surveys (14, 22, 29, 37). These variables constitute critical health determinants such as GDP per capita, smoking tobacco, sanitation facilities, alcohol consumption, acute care and total number of hospital beds (22). The variables were extracted from major publicly available data sets[2]. For each parameter, the last recorded year's entry was used for each country (if relevant information was available). As the study utilized publicly available data, ethical clearance was not required.

We indicate that while previous studies have compared the total number of COVID-19 cases per million population registered on a certain date, this approach might be inaccurate as different European countries reported their first COVID-19 case at different times and therefore some European countries would probably show higher prevalence compared to others. In the present study we calculate the change in the numbers of new COVID-19 cases between two dates, which were 2 months apart, i.e., 1st April and 31st May 2020. As an example, the change of new COVID-19 cases in the UK within the study's time period is −1,405 COVID-19 cases (42).

RESULTS

The empirical specification of this study consists of two parts. In the first part, we present a correlation analysis, also used by Gangemi et al. (40) and Lippi et al. (41), between the change in new COVID-19 cases per million population and macroeconomic and well-being indicators. Given the nature of the variables (continuous) Pearson correlation coefficients are reported. In the second part, we offer an OLS regression analysis assessing the determinants of change in new COVID-19 cases per million population. The regressions control for critical determinants of COVID-19 disease (35–37). These are GDP per capita, public expenditure on health, life expectancy at birth, smoking tobacco, and sanitation facilities. In the regression analysis, the main interest is to assess whether GDP per capita is associated with the change in new Covid-19 cases per million population. If GDP per capita remains statistically significant after controlling for key covariates, then this feature might indicate an association between the change in new COVID-19 cases per million population and countries' economic performance in the specified region and period.

In this study we highlight that the regression outcomes should be interpreted as associations and not as causal effects. It is well-documented in the literature that a two-way relationship between a country's performance and good health might exist (22) due to the fact that health may actually drive economic performance (28, 31).

Correlation and Regression Outcomes

Correlation Analysis

Table 1 shows correlation coefficients for the change in new COVID-19 cases per million population. Statistically significant negative correlations were obtained between the change in new COVID-19 cases and GDP per capita, public expenditure on health, sanitation facilities and life expectancy at birth. The outcomes indicate a country's higher prosperity (GDP per capita), public spending on health, and well-being indicators (such as sanitation infrastructures and life expectancy at birth) is associated with a lower change in new COVID-19 per million population. A statistically significant positive correlation between the change in new COVID-19 cases and tobacco smoking is also found, indicating that smoking prevalence might deteriorate health status and/or be an underlying health co-morbidity.

Figure 1 presents a negative correlation between the change in new COVID-19 cases per million population and GDP per capita. The figure indicates that Luxemburg which was the best performing European country in terms of GDP per capita also experienced the lowest change in new COVID-19 cases per million population. On the other hand, countries facing lower GDP per capita, such as Ukraine, Bulgaria, Romania and Russia, experienced a higher level of change in new COVID-19 cases per million population. The country with the highest level of change in new COVID-19 cases per million population, within the time period, was Russia.

Regression Outcomes

Table 2 presents the regression outcomes. We include only those variables that were found to be statistically significant with the change in new COVID-19 cases in **Table 1** (i.e., GDP per capita, public expenditure on health, life expectancy at birth, smoking tobacco and sanitation facilities).

Panel I controls for public expenditure on health. The estimates indicate that GDP per capita is negatively associated

[1] 1: Albania; 2: Andorra; 3: Austria 4: Belgium 5: Bosnia and Herzegovina; 6: Bulgaria; 7: Croatia; 8: Cyprus; 9: the Czech Republic; 10: Denmark; 11: Estonia; 12: Finland; 13: France; 14: Germany; 15: Greece; 16: Hungary; 17: Iceland; 18: Ireland; 19: Italy; 20: Latvia; 21: Lithuania; 22: Luxemburg; 23: Malta; 24: Montenegro; 25: the Netherlands: 26: Norway; 27: Poland; 28: Portugal; 29: Romania; 30: Russia; 31: Serbia; 32: Slovakia; 33: Slovenia; 34: Spain; 35: Sweden; 36: Switzerland; 37: Ukraine; 38: the United Kingdom.

[2] Gross Domestic Product per capita and extreme poverty: https://ourworldindata.org/coronavirus;
Public expenditure on health: http://dmt.euro.who.int/classifications/tree/A;
Prevalence of obesity: https://gateway.euro.who.int/en/indicators/h2020_9-obesity/;
Prevalence of overweight: https://gateway.euro.who.int/en/indicators/h2020_6-overweight/;
Availability of social support: http://dmt.euro.who.int/classifications/tree/A;
Smoking prevalence: http://dmt.euro.who.int/classifications/tree/A;
Alcohol consumption: http://dmt.euro.who.int/classifications/tree/B;
Life expectancy at birth: http://dmt.euro.who.int/classifications/tree/B;
Sanitation facilities: http://dmt.euro.who.int/classifications/tree/B;
Acute care and total number of hospital beds: http://dmt.euro.who.int/classifications/tree/A;
Practicing physicians: http://dmt.euro.who.int/classifications/tree/L.

TABLE 1 | Correlation coefficients between the change in new COVID-19 cases per million population and socio-epidemiological determinants over a two-month period (1st April 2020 and 31st May 2020).

	Panel I Change in new COVID-19 cases per million population	Number of countries
GDP per capita	−0.665 (0.001)***	37[a]
Extreme poverty	0.135 (0.503)	27[b]
Public expenditure on health	−0.557 (0.001)***	38
Obesity	0.077 (0.644)	38
Overweight	−0.194 (0.314)	29[c]
Social support	−0.252 (0.126)	38
Smoking tobacco	0.344 (0.035)**	38
Alcohol	−0.127 (0.453)	37[d]
Life expectancy at birth	−0.565 (0.001)***	38
Sanitation facilities	−0.409 (0.012)**	37[e]
Acute care beds	0.210 (0.212)	37[f]
No of Hospital beds	0.231 (0.162)	38
Median Age	0.032 (0.849)	37[g]
No physicians	−0.058 (0.731)	37[h]

Pearson correlation coefficients are reported. p-values are in parentheses.
[a]Andorra is excluded due to data unavailability.
[b]Andorra, Cyprus, the Czech Republic, Finland, France, Germany, the Netherlands, Poland, Serbia, Slovenia, and Switzerland are excluded due to data unavailability.
[c]Bulgaria, Croatia, Germany, Lithuania, Luxemburg, Spain, Sweden, Switzerland and the United Kingdom are excluded due to data unavailability.
[d]Germany is excluded due to data unavailability.
[e]Montenegro is excluded due to data unavailability.
[f]Bosnia and Herzegovina is excluded due to data unavailability.
[g]Andorra is excluded due to data unavailability.
[h]Slovakia is excluded due to data unavailability.
****Statistically significant at the 1%. **Statistically significant at the 5%.*

with the change in new COVID-19 cases per million population at the 1% level. In Panel II, we control also for life expectancy at birth and smoking tobacco. The new estimates continue to indicate a negative association between GDP per capita and the change in new COVID-19 cases per million population at the 1%. In Panel III, we add controls for sanitation facilities. It is found that GDP per capita continues to be negatively associated with the change in new COVID-19 cases per million population at the 1%.

DISCUSSION

The aim of the study was to examine correlations between the change in new COVID-19 cases per million population and macroeconomic indicators, well-being indicators and healthcare systems' capacity in the first wave of the COVID-19 pandemic. Most recent studies have focused on demographic, socio-economic and health characteristics and the COVID-19 pandemic (2, 6, 7, 17). In this study we were particularly interested in assessing whether countries' economic performance, as it is captured by GDP per capita, might be associated with COVID-19 spread in 38 European countries, in the early stages. By extracting data in June 2020 and calculating the change in new COVID-19 cases over a 2-month period, between 1st April and 31st May 2020, univariate correlation analysis indicated that new COVID-19 cases per million population were (i) negatively correlated with GDP per capita, public expenditure on health, sanitation facilities and life expectancy at birth and (ii) positively correlated with tobacco smoking. The adjusted regression analysis shows a significant negative association between the change in new COVID-19 cases per million population and GDP per capita. The outcomes of this study are in line with the limited research indicating that COVID-19 cases are negatively associated with public expenditure in health and health care infrastructure and capacity (13, 15), sanitation (14) and positively associated with smoking habits (9, 10).

The study contributes to the literature by assessing both correlations and associations between COVID-19 and GDP per capita. While recent studies have examined associations between COVID-19 and individuals' income (35–37), there is little information assessing a country's economic performance with COVID-19 cases (39). Our study shows a negative relationship between the change in new COVID-19 cases and GDP per capita, in the first wave of the pandemic. We indicate that GDP per capita differences among European countries might reflect existing structural and economic factors (18). People living in European countries with lower GDP per capita might have poorer access to health services and have lower income resulting in poorer health (31). These factors might be exaggerated during periods of severe health and economic crises, negatively affecting the less developed regions (31).

Prior to the pandemic, socio-epidemiological research indicated positive associations between a country's economic performance and better health (22, 27, 28, 34), as well as positive associations between individuals' income and health status (18, 20). In the present study we indicate that a wealthier country might be able to provide their people with better living standards, public health and environmental sanitation, leading to enhanced prevention and disease treatment (22–24).

The last fifty years, European countries have experienced a massive economic growth that have enabled them to invest in health and develop effective health systems and brought several infections and diseases under control (22, 24, 32). We should note that a country's economic performance not only determines health systems' capacity and effectiveness but also its labor market and vocational relations within the country. It is possible that in advanced European countries, lockdowns to save lives might be easier since a greater number of people work in sectors where information technology infrastructure enables them to work from home (43). This might be challenging in poorer European countries where a greater number of people work in sectors where manual labor is needed (43). Such jobs make up the new COVID-19 essential workforce i.e., food service workers, bus drivers.

There are a number of limitations in the present study. It is possible that a change in the number of new COVID-19 cases may not be an accurate reflection of the true infection rate, as testing of infection may be different in various countries. An unbiased estimate of the infection rate will be only possible

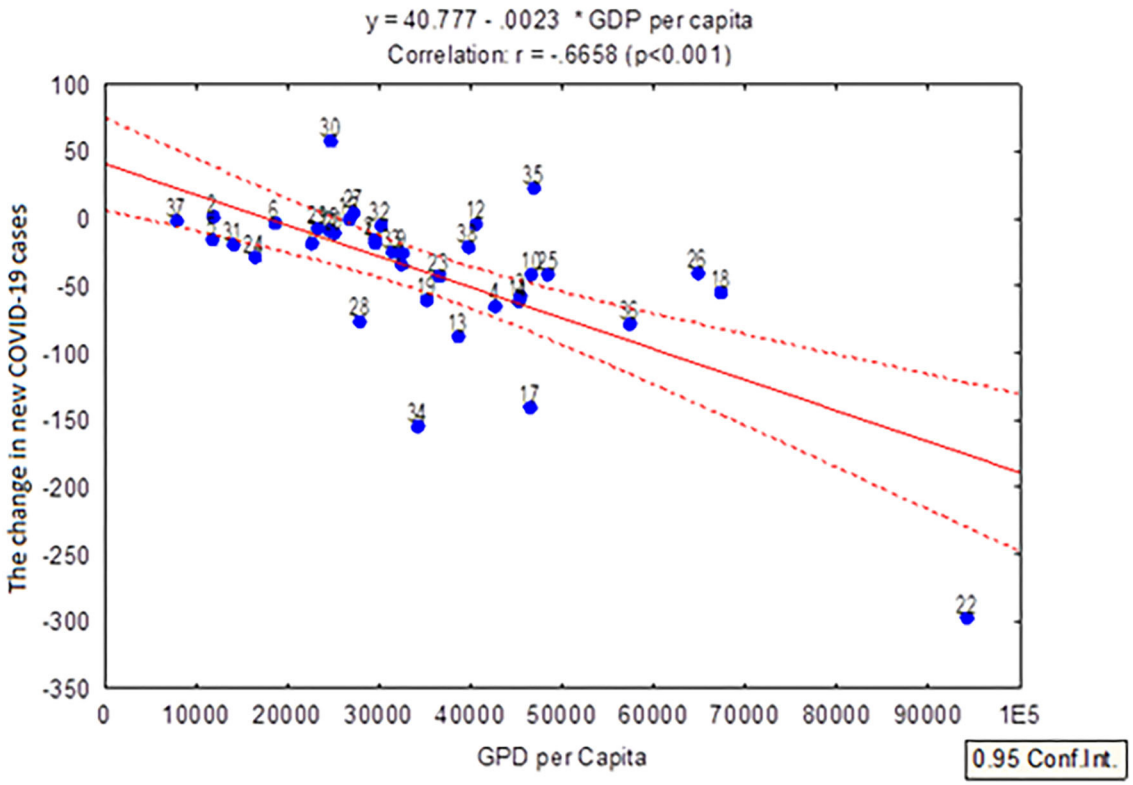

FIGURE 1 | The change in new COVID-19 cases per million population over a 2-month period (1st April and 31st May 2020) plotted against the GDP per capita in 38 European countries. The countries denoted in the figure are: 1: Albania; 2: Andorra; 3: Austria; 4: Belgium; 5: Bosnia and Herzegovina; 6: Bulgaria; 7: Croatia; 8: Cyprus; 9: the Czech Republic; 10: Denmark; 11: Estonia; 12: Finland; 13: France; 14: Germany; 15: Greece; 16: Hungary; 17: Iceland; 18: Ireland; 19: Italy; 20: Latvia; 21: Lithuania; 22: Luxemburg; 23: Malta; 24: Montenegro; 25: the Netherlands: 26: Norway; 27: Poland; 28: Portugal; 29: Romania; 30: Russia; 31: Serbia; 32: Slovakia; 33: Slovenia; 34: Spain; 35: Sweden; 36: Switzerland; 37: Ukraine; and 38: the United Kingdom. For example, point 22 is Luxembourg.

TABLE 2 | Regression outcomes: The change in new COVID-19 cases per million population.

	Panel I Change in new COVID-19 cases per million population	**Panel II Change in new COVID-19 cases per million population**	**Panel III Change in new COVID-19 cases per million population**
GDP per capita	−0.810 (0.005)***	−0.801 (0.004)***	−0.878 (0.002)***
Public expenditure on health	0.164 (0.540)	0.308 (0.296)	0.365 (0.215)
Life expectancy at birth	–	−0.346 (0.060)	−0.188 (0.383)
Smoking tobacco	–	−0.141 (0.390)	−0.120 (0.467)
Sanitation facilities	–	–	−0.218 (0.212)
Adjusted R^2	0.417	0.440	0.462
F	13.86	8.312	7.017
p	0.000	0.001	0.001

*N = 36. Andorra and Bosnia and Herzegovina are excluded due to data unavailability. p-values are in parentheses. ***Statistically significant at the 1%.*

at the end of the pandemic and/or when there are no deaths reported due to COVID-19. As the data set captured a certain time interval, the choice of the specified time-period did not take into account any lockdown effects in each country or the phase of infection in that country. In addition, the pandemic has carried on so analysis may be different for different time frames. Future studies should utilize different time series data in order to provide longitudinal evaluations.

A further limitation of the present study is that our data were from European countries. New research should consider other parts of the world for firm evaluations. In addition, in the present study, only limited number of available macroeconomic variables were analyzed. It would of interest to examine whether a country's wealth, income distribution, and saving per capita might be associated with COVID-19 spread. We should emphasize that a country's performance is not the only parameter that influences the spread of COVID-19. Other health and social determinants should be considered, as well. In the US, Millett et al. (36) found disproportionately higher rates of COVID-cases and deaths in black counties compared with other counties. This is also shown in the UK with higher prevalence of COVID-19 in Black, Asian

and Minority ethnicities (44). People living in deprived areas also experienced worse health (45, 46). Hence, it would be of importance for new research to consider interactions between economic indicators at a microeconomic and macroeconomic level and ethnicity, gender, underlying health conditions, labor characteristics and mobility during the lockdown and their associations with COVID-19 spread.

Moreover, new studies might examine possible interactions between GDP per capita and the level of industrial pollution and urban segregation, and their associations with COVID-19 cases. For instance, in Italy, mortality due to COVID-19 was found to be higher in areas with higher GDP per capita (41). The study suggested that adverse environmental factor such as higher industrial pollution, in addition to the other known risk factors such as obesity and hypertension, might have driven the outcomes (41).

It is known that many other factors influenced the rate of COVID-19 infections. In Wuhan-China, a positive impact of lockdown was found to restrain further increases of COVID-19 cases (12). In addition to the social, health and economic determinants, it is possible that a "level of preparedness" may also play a significant role in reducing COVID-19 spread. Research suggests that COVID-19 prevalence and mortality rates were lower in African and Asian countries compared to certain Western European countries and the US (47). Several factors may have contributed to this including early instigation of lockdown and border closures, younger age of the population, lack of robust reporting systems, and other unidentified genetic factors (47). It is possible that some countries might be better prepared to deal with COVID-19 spread because they have accumulated experiences from previous pandemics including SARS. For example, Ghana initiated lockdowns within weeks of the first COVID-19 cases and was ranked number one among African countries in administering tests per million people (48). Fan et al. (49) reporting raw case fatality rate of 53 countries with the highest COVID-19 death tolls, showed that 43 countries had lower raw case fatality rate estimates in the second wave of the COVID-19 pandemic, indicating that healthcare system of the countries might have been better prepared for the second phase (49, 50). The theory of better immunity in economically poor countries was considered by Roy (39) who showed that, in a sample of 46 countries, although the number of new COVID-19 cases showed a slight decline in lower-income countries, the fatality rate was independent of the financial condition of the countries in question.

Our study shows that GDP per capita might be a critical epidemiological parameter when comparing different countries within Europe. We reported a statistically significant negative association between GDP per capita and the change in new cases of COVID-19 per million population during the first wave over a 2-month period (1st April-31st May 2020) in 38 European countries. The outcomes of this study should call the attention of policy makers. If better economically performed economies in Europe display lower susceptibility to pandemics and to new infections in the early stages of the pandemic, a country's economic growth and development should be perceived as a vital priority for policy makers in future pandemics.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://ourworldindata.org/coronavirus; Public expenditure on health: http://dmt.euro.who.int/classifications/tree/A; Prevalence of obesity: https://gateway.euro.who.int/en/indicators/h2020_9-obesity/; Prevalence of overweight: https://gateway.euro.who.int/en/indicators/h2020_6-overweight/; Availability of social support: http://dmt.euro.who.int/classifications/tree/A; Smoking prevalence: http://dmt.euro.who.int/classifications/tree/A; Alcohol consumption: http://dmt.euro.who.int/classifications/tree/B; Life expectancy at birth: http://dmt.euro.who.int/classifications/tree/B; Sanitation facilities: http://dmt.euro.who.int/classifications/tree/B; Acute care and total number of hospital beds: http://dmt.euro.who.int/classifications/tree/A; Practicing physicians: http://dmt.euro.who.int/classifications/tree/L.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

Both authors contributed to the whole paper. SP carried out data analysis. SP and ND contributed equally to introduction, results, and discussion.

FUNDING

This work was supported by Anglia Ruskin University project award.

REFERENCES

1. Fong MW, Gao H, Wong JY, Xiao J, Shiu EYC, Ryu S, et al. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings. *Soc Distancing Meas*. (2020) 26:976–84. doi: 10.3201/eid2605.190995
2. Applegate WB, Ouslander JG. COVID-19 presents high risk to older persons. *J Am Geriatrics Soc*. (2020) 68:681. doi: 10.1111/jgs.16426
3. Jordan RE, Adab P, Cheng KK. COVID-19: Risk factors for severe disease and Death. *Br Med J*. (2020) 368:m1198. doi: 10.1136/bmj.m1198
4. Wenham C, Smith J, Morgan R. COVID-19: The gendered impacts of the outbreak. *Lancet*. (2020) 395:846–8. doi: 10.1016/S0140-6736(20)30526-2
5. Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Jabir AA, Iosifidis C, et al. The socio-economic implications of the coronavirus pandemic (COVID-19): a review. *Int J Surg*. (2020) 78:185–93. doi: 10.1016/j.ijsu.2020.04.018

6. Emami A, Javanmardi F, Pirbonyeh N, Akbari A. Prevalence of underlying diseases in hospitalized patients with COVID-19: a systematic review and meta-analysis. *Arch Acad Emerg Med*. (2020) 8:e35. doi: 10.22037/aaem.v8i1.600
7. Khunti K, Singh AK, Pareek M, Hanif W. Is ethnicity linked to incidence or outcomes of COVID-19? *Br Med J*. (2020) 369:m1548. doi: 10.1136/bmj.m1548
8. Pareek M, Bangash MN, Pareek N, Pan D, Sze S, Minhas JS, et al. Ethnicity and COVID-19: an urgent public health research priority. *Lancet*. (2020) 395:1421–2. doi: 10.1016/S0140-6736(20)30922-3
9. Farsalinos K, Barbouni A, Niaura R. Systematic review of the prevalence of current smoking among hospitalized COVID-19 patients in China: could nicotine be a therapeutic option? *Intern Emerg Med*. (2020) 15:845–52. doi: 10.1007/s11739-020-02355-7
10. Vardavas CI, Nikitara K. COVID-19 and smoking: a systematic review of the evidence. *Tob Induc Dis*. (2020) 18:20. doi: 10.18332/tid/119324
11. Tamara A, Tahapary DL. Obesity as a predictor for a poor prognosis of COVID-19: a systematic review. *Diabetol Metab Syndr*. (2020) 14:655–9. doi: 10.1016/j.dsx.2020.05.020
12. Lau H, Khosrawipour V, Kocbach P, Mikolajczyk A, Schubert J, Bania J, et al. The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. *J Travel Med*. (2020) 27:taaa037. doi: 10.1093/jtm/taaa037
13. Ng WL. To lockdown? When to peak? Will there be an end? A macroeconomic analysis on COVID-19 epidemic in the United States. *J Macroecon*. (2020) 65:103230. doi: 10.1016/j.jmacro.2020.103230
14. Caruso BA, Freeman MC. Shared sanitation and the spread of COVID-19: risks and next steps. *Lancet Planet Health*. (2020) 4:e173. doi: 10.1016/S2542-5196(20)30086-3
15. White DB, Lo B. A framework for rationing ventilators and critical care beds during the COVID-19 pandemic. *JAMA*. (2020) 323:1773–4. doi: 10.1001/jama.2020.5046
16. Bavel JJVK, Baicker PS, Boggio V, Capraro A, Cichocka M, Cikara MJ, et al. Using social and behavioural science to support COVID-19 pandemic response. *Nat Hum Behav*. (2020) 4:460–71. doi: 10.1038/s41562-020-0884-z
17. Grima S, Kizilkaya M, Rupeika-Apoga R, Romānova I, Dalli Gonzi R, Jakovljevic M. A country pandemic risk exposure measurement model. *Risk Manag Healthc Policy*. (2020) 13:2067–2077. doi: 10.2147/RMHP.S270553
18. OECD. *Health at a Glance: Europe 2018: State of Health in the EU Cycle*, Paris: OECD Publishing (2018).
19. Stuckler D, Basu S, Suhrcke M, Coutts A, McKee M. The public health effect of economic crises and alternative policy responses in Europe: an empirical analysis. *Lancet*. (2009) 374:315–23. doi: 10.1016/S0140-6736(09)61124-7
20. Marmot M, Allen J, Boyce T, Goldblatt P, Morrison J. *Health Equity in England: The Marmot Review 10 years on*. London: Institute of Health Equity (2020).
21. Preston SH. The changing relation between mortality and level of economic development. *Popul Stud*. (1975) 29:231–48. doi: 10.2307/2173509
22. Berkman FL, Kawachi I, Glymour MM. *Social Epidemiology*. Oxford: Oxford University Press (2014).
23. Mackenbach JP, Looman CWN. Life expectancy and national income in Europe, 1900-2008: an update of Preston's analysis. *Int J Epidemiol*. (2013) 42:1100–10. doi: 10.1093/ije/dyt122
24. Krieger N. *Epidemiology and the People's Health. Theory and Context*. Oxford: Oxford University Press (2011).
25. Kunitz SJ. *The Health of Populations. General Theories and Particular Realities*. Oxford: Oxford University Press. (2007).
26. Smith JP. Healthy bodies and thick wallets: the dual relation between health and economic status. *J Econ Perspect*. (1999) 13:145–66. doi: 10.1257/jep.13.2.145
27. Brenner HM. Commentary: Economic growth is the basis of mortality rate decline in the 20th century—Experience of the United States 1901–2000. *Int J Epidemiol*. (2005) 34:1214–21. doi: 10.1093/ije/dyi146
28. Swift R. The relationship between health and GDP in OECD countries in the very long run. *Health Econ*. (2011) 20:306–22. doi: 10.1002/hec.1590
29. Parmar D, Stavropoulou C, Ioannidis PAJ. Health outcomes during the 2008 financial crisis in Europe: systematic literature review. *British Med J*. (2016) 354:1–11. doi: 10.1136/bmj.i4588
30. Drydakis N. *The Relationship Between Recessions and Health*. Bonn: IZA World of Labor. (2016). doi: 10.15185/izawol.283
31. Drydakis N. The effect of unemployment on self-reported health and mental health in Greece from 2008 to 2013: a longitudinal study before and during the financial Crisis. *Soc Sci Med*. (2015) 128:43–51. doi: 10.1016/j.socscimed.2014.12.025
32. Mackenbach JP, Looman CWN. Changing patterns of mortality in 25 European countries and their economic and political correlates, 1955–1989. *Int J Public Health*. (2013) 58:811–23. doi: 10.1007/s00038-013-0509-7
33. Ward JL, Viner RM. The impact of income inequality and national wealth on child and adolescent mortality in low and middle-income countries. *BMC Public Health*. (2017) 17:429. doi: 10.1186/s12889-017-4310-z
34. Mackenbach JP, Bopp M, Deboosere P, Kovacs K, Leinsalu M, Martikainen P, et al. Determinants of the magnitude of socioeconomic inequalities in mortality: a study of 17 European countries. *Health Place*. (2017) 47:44–53. doi: 10.1016/j.healthplace.2017.07.005
35. Yancy CW. COVID-19 and African Americans. *JAMA*. (2020) 323:1891–2. doi: 10.1001/jama.2020.6548
36. Millett GA, Jones AT, Benkeser D, Baral S, Mercer L, Beyrer C, et al. Assessing differential impacts of COVID-19 on black communities. *Ann Epidemiol*. (2020) 47:37–44. doi: 10.1016/j.annepidem.2020.05.003
37. Clark A, Jit M, Warren-Gash C, Guthrie B, Wang HH XW, Mercer SW, et al. Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. *Lancet*. (2020) 8 e1003–17. doi: 10.1016/S2214-109X(20)30264-3
38. Spencer L, Gubbins JP, Murray CJL, Gakidou E. Developing comprehensive time series of GDP per capita for 210 countries from 1950 to 2015. *Popul Health Metr*. (2012) 10:12. doi: 10.1186/1478-7954-10-12
39. Roy S. Low-income countries are more immune to COVID-19: A misconception. *Indian J Med Sci*. (2020) 72:5–7. doi: 10.25259/IJMS_26_2020
40. Gangemi S, Billeci L, Tonacci A. Rich at risk: Socio-economic drivers of COVID-19 pandemic spread. *Clin Mol Allergy*. (2020) 18:12. doi: 10.1186/s12948-020-00127-4
41. Lippi G, Henry BM, Mattiuzzi C, Bovo C. The death rate for COVID-19 is positively associated with gross domestic products. *Acta Biomedica*. (2020) 91:224–5. doi: 10.23750/abm.v91i2.9514
42. Roser M, Ritchie H, Ortiz-Ospina E, Hasell J. *Data From: Coronavirus Pandemic (COVID-19)*. (2020). Available online at: https://ourworldindata.org/coronavirus
43. Cedefop. Future Skill Needs in Europe: Critical Labour Force Trends. Cedefop Research Paper No. 59. Luxembourg: Publications Office (2016).
44. Raisi-Estabragh Z, McCracken C, Bethell MS, Cooper J, Cooper C, Caulfield MJ, et al. Greater risk of severe COVID-19 in black, Asian and minority ethnic populations is not explained by cardiometabolic, socioeconomic or behavioural factors, or by 25(OH)-vitamin D status: Study of 1326 cases from the UK biobank. *J Public Health (Bangkok)*. (2020) 42:451–60. doi: 10.1093/pubmed/fdaa095
45. Office for National Statistics. *Health Survey for England 2018*. London: Office for National Statistics (2018).
46. Office for National Statistics. *Likelihood of Smoking Four Times Higher in England's Most Deprived Areas than Least Deprived*. London: Office for National Statistics (2019).
47. Ogunleye O, Basu D, Mueller D, Sneddon J, Seaton AR, et al. Response to the novel corona virus (COVID-19) pandemic across Africa: Successes, challenges, and implications for the future. *Front Pharmacol*. (2020) 11:205. doi: 10.3389/fphar.2020.01205

48. Afriyie DK, Asare GA, Amponsah SK, Godman B. COVID-19 pandemic in resource-poor countries: challenges, experiences and opportunities in Ghana. *J Infect Dev Ctries*. (2020) 4:838–43. doi: 10.3855/jidc.12909
49. Fan G, Yang Z, Lin Q, Zhao S, Yang L, He D. Decreased case fatality rate of COVID-19 in the second wave: a study in 53 countries. *Transboundary Emerg Dis*. (2020). doi: 10.1111/tbed.13819
50. Saito S, Asai Y, Matsunaga N, Hayakawa K, Terada M, Ohtsu H, et al. First and second COVID-19 waves in Japan: a comparison of disease severity and characteristics. *J Infect*. (2020) 23:10. doi: 10.1016/j.jinf.2020.10.033

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

ORIGINAL RESEARCH
published: 21 January 2021
doi: 10.3389/fpubh.2020.607528

Clustering and Healthcare Costs With Multiple Chronic Conditions in a US Study

Cother Hajat[1]*, *Yakir Siegal*[2] and *Amalia Adler-Waxman*[3]

[1] Public Health Institute, United Arab Emirates (UAE) University, Al Ain, United Arab Emirates, [2] Deloitte Consulting LLP, New York, NY, United States, [3] Teva Pharmaceuticals Industries Ltd., Petah Tikva, Israel

Objective: To investigate healthcare costs and contributors to costs for multiple chronic conditions (MCCs), common clusters of conditions and their impact on cost and utilization.

Methods: This was a cross-sectional analysis of US financial claims data representative of the US population, including Medicare, Medicaid, and Commercial insurance claims in 2015. Outcome measures included healthcare costs and contributors; ranking of clusters of conditions according to frequency, strength of association and unsupervised (k-means) analysis; the impact of clustering on costs and contributors to costs.

Results: Of 1,878,951 patients, 931,045(49.6%) had MCCs, 56.5% weighted to the US population. Mean age was 53.0 years (SD16.7); 393,121(42.20%) were male. Mean annual healthcare spending was $12,601, ranging from $4,385 (2 conditions) to $33,874 (11 conditions), with spending increasing by 22-fold for inpatient services, 6-fold for outpatient services, 4.5-fold for generic drugs, and 4.2-fold for branded drugs. Cluster ranking using the 3 methodologies yielded similar results: highest ranked clusters included metabolic syndrome (12.2% of US insured patients), age related diseases (7.7%), renal failure (5.6%), respiratory disorders (4.5%), cardiovascular disease(CVD) (4.3%), cancers (4.1–4.3%), mental health-related clusters (1.0–1.5%), and HIV/AIDS (0.2%). Highest spending was in HIV/AIDS clusters ($48,293), mental health-related clusters ($38,952–$40,637), renal disease ($38,551), and CVD ($37,155); with 89.9% of spending on outpatient and inpatient care combined, and 10.1% on medication.

Conclusion and Relevance: Over 57% of insured patients in the US may have MCCs. MCC Clustering is frequent and is associated with healthcare utilization. The findings favor health system redesign toward a multiple condition approach for clusters of chronic conditions, alongside other cost-containment measures for MCCs.

Keywords: cardiovascular disease, chronic disease, multimorbidity, cancer, mental health, healthcare costs, disease clustering

OPEN ACCESS

Edited by:
Mihajlo Jakovljevic, Hosei University, Japan

Reviewed by:
Gregory Lubiani, Texas A&M University Commerce, United States
Simon Grima, University of Malta, Malta
Márta Péntek, Óbuda University, Hungary

***Correspondence:**
Cother Hajat
Chajat@doctors.org.uk

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *21 September 2020*
Accepted: *02 December 2020*
Published: *21 January 2021*

Citation:
Hajat C, Siegal Y and Adler-Waxman A (2021) Clustering and Healthcare Costs With Multiple Chronic Conditions in a US Study. Front. Public Health 8:607528. doi: 10.3389/fpubh.2020.607528

HIGHLIGHTS

What is already known about the topic?

Despite one in three adults suffering from more than one chronic condition, little is known about the burden from MCCs. Some studies to date suggest markedly different disease, cost and societal burdens. Furthermore, certain conditions cluster together more frequently, however, no studies have reported on the impact that clusters have on the healthcare cost burden from MCCs in

a comprehensive manner. Recent consensus statements have called for a specific focus on multiple chronic conditions.

What does the paper add to existing knowledge?

This study is one of the most comprehensive studies investigating contributors to costs in terms of number of patients included, representativeness of the US population and inclusion of the full range of chronic conditions.

What insights does the paper provide for informing healthcare-related decision making?

Of US insured patients, over 57% may have multiple chronic conditions.

HIV/AIDS was the costliest cluster followed by clusters of mental and behavioral disorders, renal failure and CVD. Outpatient and inpatient services account for roughly 90% of health spending and medication for 10%.

Health service utilization varies by number and clusters of conditions, with potential overutilization of specialist services and underutilization of primary care and psychiatric services.

INTRODUCTION

The increasing global burden of non-communicable diseases (NCDs), accounting for three in five of global deaths (1), has long been recognized as a global priority. Less attention has been given to the issue of multiple chronic conditions (MCCs), also termed multi-morbidity, despite one in three adults suffering from more than one chronic condition (2).

Evidence from a handful of studies reporting on the burden from MCCs to date suggest that this phenomenon results in markedly different disease, cost and personal burdens. Most studies have asserted a positive association between MCC and healthcare expenditures (3), some reporting a doubling in costs with each subsequent condition (4, 5). Studies suggest more complex inpatient and outpatient care utilization and use of more prescription medications (6–9).

Certain conditions cluster together more frequently such as stroke and Alzheimer's disease, and communicable conditions such as TB and HIV/AIDS with diabetes and CVD, respectively (10, 11). However, no studies have reported on the impact that clusters have on the healthcare cost burden from MCCs in a comprehensive manner.

Recent articles have called for a specific focus on multiple chronic conditions (10–12) and the Academy of Medical Sciences in the UK identified clustering of disease as a priority research area (13).

The objective of this article is to quantify healthcare spending in patients with MCCs using a comprehensive set of chronic conditions, identify the most important clusters and identify the key contributors to costs for MCC patients. The primary research questions were:

1. What are the healthcare costs and contributors to costs for patients with multiple chronic conditions?
2. What are the most important clusters of chronic conditions?
3. How does clustering of chronic conditions impact cost and contributors to costs?

METHODS

We used a random sample of IBM MarketScan® claims-based data (14) for US adults (age 18+ years) covered by insurance for the data year 2015, including the Commercial Claims and Encounters (CCAE) Database, the Medicare Supplemental Database, and the Multi-State Medicaid Database. The IBM MarketScan® claims databases contain de-identified, patient-level health episode claims information for >200 million patients aggregated across multiple public and private health insurance providers. The claims data includes information on inpatient services, outpatient services and outpatient prescription drugs. The CCAE database covers patients with commercial insurance while the Medicare Supplemental Database and the Multi-State Medicaid Databases cover patients submitting claims through the public programmes of Medicare and Medicaid, respectively (15). The MarketScan® database utilizes standard international coding and the International Classification of Disease (16) was used to assign diagnoses. Random, representative samples of each of these three databases were sourced from IBM MarketScan® as a basis for this analysis.

Patients with two or more chronic conditions from the Agency for Healthcare Research and Quality (AHRQ) list of 69 chronic conditions which relies on ICD9 and ICD10 coding systems (17) were included. Patients with more than 11 conditions, equating to 59,432 (6.4%) of MCC patients in the sample, were omitted due to inadequate sub-group numbers.

The main purpose of this study was to understand drivers of healthcare spend for patients with key clusters of multiple chronic conditions. In the absence of agreed upon criteria and methodologies for determining key clusters, we employed a two-step process to identify and select clusters for further analysis, utilizing three different methodologies to shed light on different aspects of key clusters.

Step 1 used three different methods for ranking and characterizing clusters: Method A: Identify and rank the top 25 co-morbid condition pairs based on prevalence—i.e., highest frequency of occurrence in the sample population.

Method B: Identify and rank the top 25 co-morbid conditions pairs based on strength of association—i.e., highest probability of condition 2 being present in a patient if condition 1 is present.

Method C: Conduct an unsupervised k-means clustering analysis to identify clustering without constraining to condition pairs. K-means is an unsupervised learning method without reliance on a "dependent" variable (18). K-means clustering optimizes the within group sum of squares i.e., it assigns observations to a cluster based on how close (Euclidean distance) it is to the cluster centroid. Each patient was assigned to a disease cluster based on Euclidean distance of their disease vectors to the cluster centroids (18). The clusters are reported in the results tables characterized by the base conditions which strongly associate with the cluster (i.e., are present in >95% of patients in that cluster), as well as other conditions more moderately associated with the cluster (present in 20–95% of patients in that cluster). Step 2 was to select 10 clusters to investigate further with respect to cost driver analysis from the "top" clusters from Step 1 based on frequency, strength of association and k-means

TABLE 1 | Characteristics of patients included.

Characteristic	
Number with multiple chronic conditions	931,045 (49.6%)
Age years: mean (SD)	53.04 (16.66)
Male *n* (%)	393,121 (42.20%)
Health Coverage: *n* (%)	
Commercial (CCAE)	476,879 (51.23%)
Medicaid	270,092 (29.00%)
Medicare	184,074 (19.77%)
Number of chronic conditions:	
2 *n*(%)	201,255 (21.60%)
3 *n*(%)	167,651 (18.00%)
4 *n*(%)	134,020 (14.40%)
5 *n*(%)	103,575 (11.10%)
6 *n*(%)	78,844 (8.47%)
7 *n*(%)	60,056 (6.45%)
8 *n*(%)	45,449 (4.88%)
9 *n*(%)	34,649 (3.72%)
10 *n*(%)	26,078 (2.80%)
11 *n*(%)	20,036 (2.15%)
Annual Healthcare spending US$: mean(SD)	
Overall	12,601 (36,329)
By type of coverage:	
Commercial (CCAE)	10,571 (28,352)
Medicaid	11,729 (38,518)
Medicare	19,139 (48,596)

clustering. The process involved selecting 10 anchor conditions or condition clusters of interest based primarily on their rank in the three analyses in Step 1, and characterizing the co-morbid prevalence of other conditions in the populations defined by the anchor conditions or condition clusters. Whilst primarily relying on rank from Step, clustering around HIV/AIDS was included for further analyses as a globally-relevant cluster of a chronic *communicable* condition of relevance to global populations.

Contributors to costs that were investigated included: age, sex, number of conditions, clustering of conditions, site of care settings (inpatient, outpatient), branded and generic medications, and of type of medical services provided: specialty procedures and diagnostics, primary care, emergency visits and psychiatric services. Definitions of healthcare spending are detailed in the technical supplement.

To make the results representative of the US insured population distribution, scaling factors were used to weight samples from each coverage type to the corresponding US adult (18+) insured population from year 2015 (including 134.1 million CCAE, 43.3 million Medicare, and 29.6 million Medicaid insured population) (19).

Statistical Analysis

Analyses were conducted using the software package R, version 3.4.0 (2017-04-21). Anova testing was used to determine the statistical significance of differences in costs (for each sub-category e.g., outpatient costs, inpatient costs, medication costs) between clusters (at the 95% significance level) adjusted for age and sex. Cancers were stratified by male and female demographics of the sample and reported as male and female types, adjusted for age only. All results were reported after weighting to the US population other than the crude figures used for reporting the study sample patient characteristics (**Table 1**).

As coverage type criteria include age, comparisons between coverage types were not adjusted by demographic factors. Other patient factors such as socio-economic status and ethnicity were not available.

Patient and Public Involvement

This work forms part of a broader study in which focus group work was conducted on patients with MCCs to elicit their main struggles and concerns and the terminology used in discussing their conditions. These qualitative insights were published in a white paper (11) and the current study aims to answer the quantitative research questions.

RESULTS

Table 1 reports characteristics of the study sample. The data comprised 51.2% of CCAE, 29.0% Medicaid, and 19.8% Medicare patients. Of an initial sample of 1,878,951 subjects, at-least 2 chronic conditions were found in 931,045 (49.6%); when scaled to the US population, this would equate to 56.5%. Mean age was 53.0 years (SD16.7); 393,121 (42.20%) were male and 871,613 with between 2 and 11 conditions were further analyzed. Mean annual healthcare spending was $12,601, varying between the types of insurance at $10,571 for Medicaid, $11,730 for Medicare, and $19,139 for CCAE ($p < 0.001$).

Figure 1 shows mean healthcare spending per patient and contributors to costs, according to the number of chronic conditions. Average healthcare spending per annum for those with two conditions was $4,385, increasing 8-fold to $33,874 for those with 11 or more conditions. Whilst the absolute values increased, the percentage change was non-linear, showing a consistently gradual decrement from 33% (shift from two to three conditions) to 18% (shift from 10 to 11 conditions). Inpatient costs accounted for the greatest shift in both absolute and relative amounts. The relative increase in healthcare spending between patients with 2 and 11 conditions was 22-fold for inpatient costs, 6-fold for outpatient costs, 4.5-fold for generic drugs, and 4.2-fold for branded drugs.

Further analysis of the purely medical contributors to costs for MCC patients showed that between having 2 and 11 conditions, healthcare spending increased by 24-fold for inpatient services, 8-fold for speciality procedure and diagnostics, 6-fold for emergency visits, and 4-fold for both primary care services and psychiatric services.

Table 1 and **Supplementary Table 1** show that average spending per annum was highest for Medicare at $19,129 and considerably lower for both Medicaid at $11,783 and CCAE at $10,572 ($p < 0.001$). Younger patients on Medicaid had higher spending on male cancers, renal failure, and HIV/AIDS whilst

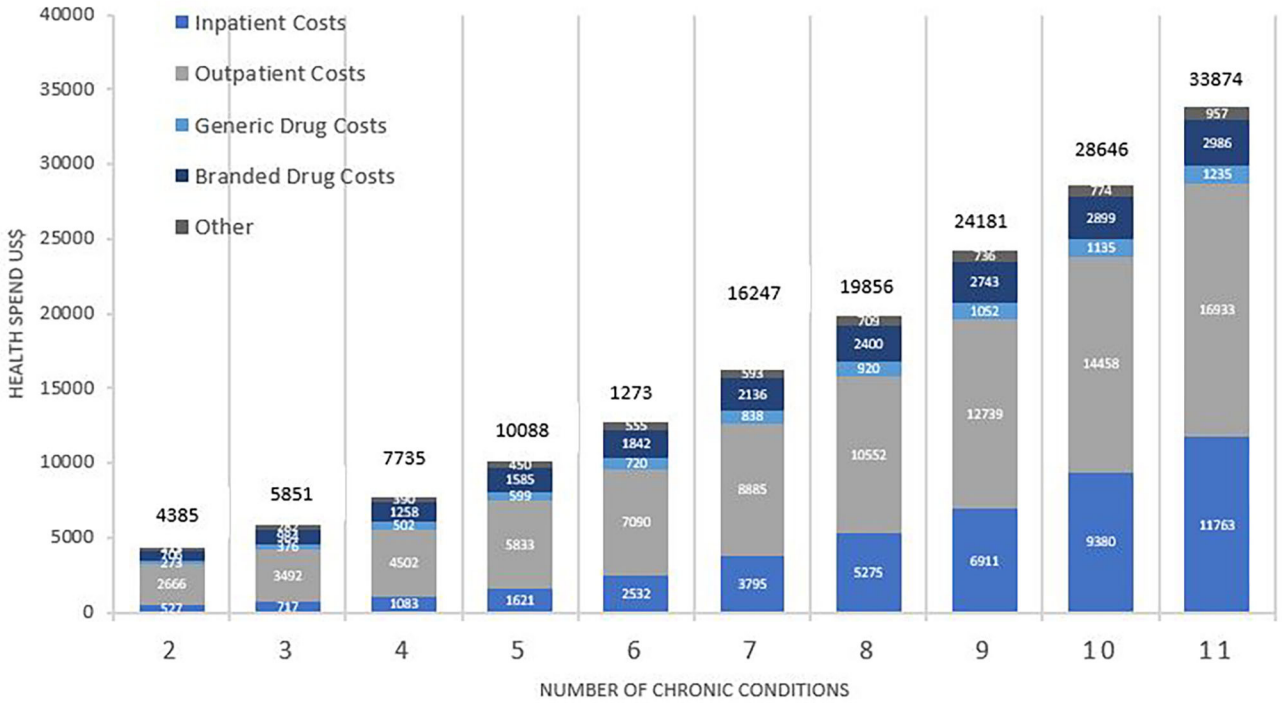

FIGURE 1 | Total average health spend and contributors to cost by number of chronic conditions. Costs are weighted to the US population and are reported in US$. Comparison between costs by number of conditions was statistically significant overall ($p < 0.001$), for inpatient ($p < 0.001$), outpatient ($p < 0.001$), branded drugs ($p < 0.001$), and generic drugs ($p < 0.001$).

older patients on Medicare had higher spending for clusters associated with mental health problems, age-related disease and respiratory problems. **Table 1** shows that the overall average spending is higher in Medicare than in CCAE, however the **Supplementary Table 1** shows that when number of conditions is held fixed, Medicare spending is lower than CCAE due to typically lower reimbursement in Medicare compared with CCAE for any given condition.

Table 2 shows the 10 clusters included in further analyses; **Supplementary Tables 2–4** show the list of top 25 clusters ranked by frequency, strength of association, and unstructured k-means clustering from which the 10 clusters were selected. Among these were metabolic syndrome, present in 12.2% of the insured population and predominantly including hypertension, high cholesterol and diabetes mellitus. Renal failure (present in 5.6%) and cancers (present in 4.1–4.3%) clustered with hypertension, high cholesterol, diabetes, and other neurological disorders, amongst other conditions. Clusters of age-related diseases (present in 7.7%) included osteoarthritis, hypertension and high cholesterol, and cardiovascular disease (CVD), present in 4.3%, included hypertension, coronary artery disease (CAD), high cholesterol and cardiomyopathy. Respiratory disorders (present in 4.5%) including chronic obstructive pulmonary disease (COPD) and allergy, clustered with hypertension, high cholesterol, peptic ulcer disease (excluding bleed), depression and other neurological disorders. Mental health disorder clusters, including depression and anxiety, were present in 1.0–1.5% and clustered with chronic pain, hypertension, high cholesterol, diabetes, other neurological disorders, and alcohol abuse. The HIV/AIDS cluster was present in 0.2% of the insured population and occurred with hypertension, high cholesterol, alcohol abuse, depression, and weight loss.

Outputs from the three different methods of ranking clusters (Step 1 in the Methods section), are reported in **Supplementary Tables 2–4**. The three methods yielded the same conditions for the majority of the top 25 clusters with clustering of high cholesterol with hypertension, diabetes with hypertension, high cholesterol with other endocrine disorders and osteoarthritis with hypertension ranking highly. Of the top 25 condition pairs by strength of association, only three individual conditions did not also appear in the top 25 by frequency, namely coagulopathy, blood loss anemia, and cystic fibrosis.

Figure 2 shows the average annual healthcare spending and contributors to costs by clusters. The highest spending was for patients with the HIV/AIDS cluster at $48,293 per patient per annum, predominantly driven by the cost of branded drugs costing on average $21866 (45.3%) of their total spending. Subsequent ranks by spending included the clusters of mental/physical debilitation at $40,637, mental health disorders at $38,952, renal disease at $38,551 and CVD at $37,155, with 45–50% of the healthcare spending on outpatient care,

TABLE 2 | Key Clusters of chronic conditions.

Cluster name	Conditions with strong associations* (>95%)	Conditions with moderate associations* (20–95%)	Sample frequency *n*(%)	US insured population frequency *n*(%)**
Metabolic syndrome	Hypertension High Cholesterol Diabetes	Eye problems; other neurological disorder; other endocrine disorder; coronary artery disease; osteoarthritis; weight-loss; obesity; chronic obstructive pulmonary disease (COPD)	110,946 (11.9%)	14,257,585 (12.2%)
Age-related disease	Osteoarthritis Hypertension High Cholesterol	Diabetes; other neurological disorder; eye problems; other endocrine disorder; peptic ulcer disease exclude bleed; coronary artery disease	65,356 (7.0%)	8,990,408 (7.7%)
Renal failure	Renal Failure	Hypertension; high cholesterol; eye problems; diabetes; other neurological disorders; weight loss; cardiac arrhythmias; coronary artery disease	50,373 (5.4%)	6,602,939 (5.6%)
Respiratory	Allergy COPD	Hypertension; high cholesterol; peptic ulcer disease exclude bleed; depression; other neurological disorders	44,486 (4.8%)	5,234,597 (4.5%)
Cardiovascular disease	Hypertension CAD Cardiomyopathy High Cholesterol	Diabetes; cardiac arrhythmias; eye problems; COPD; other neurological disorders; congestive heart failure; osteoarthritis	34,807 (3.7%)	4,982,731 (4.3%)
Cancer female	Malignant Neoplasm (females)	Hypertension; high cholesterol; other endocrine disorder; eye problems	37,062 (4.0%)	4,992,607 (4.3%)
Cancer male	Malignant Neoplasm (males)	Hypertension; high cholesterol; eye problems; diabetes	32,291 (3.5%)	4,788,380 (4.1%)
Physical/mental debilitation	Depression Chronic Pain	Hypertension; alcohol abuse; other neurological disorder; anxiety; peptic ulcer disease exclude bleed; COPD; diabetes; high cholesterol; weight loss; liver disease; enterocolitis	23,897 (2.6%)	1,802,778 (1.5%)
Mental health disorders	Depression Anxiety Hypertension	High cholesterol; other neurological disorder; alcohol abuse; peptic ulcer disease exclude bleed; COPD; diabetes; osteoarthritis	12,652 (1.4%)	1,147,493 (1.0%)
HIV/AIDS	HIV/AIDS	Hypertension; high cholesterol; alcohol abuse; depression; weight loss	2,890 (0.31%)	252,688 (0.2%)

**The second column reports the main chronic conditions that were present in over 95% of patients falling within that cluster; column three reports conditions that were present in 20–95% of patients within that cluster. **Weighted and extrapolated to the US insured population.*

40–45% on inpatient care and 10% on medication costs. In addition to these clusters, others that were associated with "other neurological disorders" accounted for higher healthcare spending in general.

Health Service Utilization was dependent on both the number and cluster of conditions. The proportion of total health spending on inpatient costs increased with each additional condition from $515(20.7%) for 2 conditions to $12,292 (47.8%) with 11 conditions. The proportion of total health spending on primary care services decreased with each additional condition from $362 (14.6%) for 2 conditions to 1,490 (5.8%) for 11 conditions. **Figure 3** shows the pattern of utilization of health services by cluster. Inpatient services accounted for over half of medical health spending for many of the clusters, followed closely by spending on specialty procedures and diagnostics. Spending on psychiatric services was low for all clusters, ranging from $58 (0.2%) in metabolic syndrome to $210 (1.4%) in the HIV/AIDS cluster and $418 (1.5%) in the mental health disorder cluster.

Healthcare spending by clusters stratified by age group followed predicted patterns of higher spending for older age groups in the clusters of metabolic syndrome, mental health disorders, physical and mental debilitation, HIV/AIDS, respiratory and age-related conditions. However other clusters such as CVD, female cancers and renal failure varied less between age groups. Mean healthcare spending for CVD was $34,485 for 18–44 year olds and $38,175 for those aged 85 and over, with similar outpatient, inpatient and generic drug costs. Spending on branded drugs accounted for a large proportion of the variance ($3,121 and $2,703 in ages 45–64 & 65–84, respectively compared with $1932 and $1,545 in ages 18–44 and >85, respectively; $p < 0.001$). Male cancers are costlier in younger age groups ($42,118 for ages 18–44 compared with $33,206 for age 85; $p < 0.001$) due to the specific types of cancer seen in younger men.

DISCUSSION

This study represents one of the most comprehensive studies to date to investigate contributors to costs in terms of number of patients included, representativeness of the US population and because it investigates the full range of, rather than selected, chronic conditions.

Key findings include:

1. In this large US insured patient sample, 50% had multiple chronic conditions.
2. Costs of healthcare spending increase non-linearly with each subsequent condition. Our study shows a relatively smaller increment in healthcare spending compared with previous smaller published studies (10), and that the relative increase tapers off with each subsequent chronic condition.

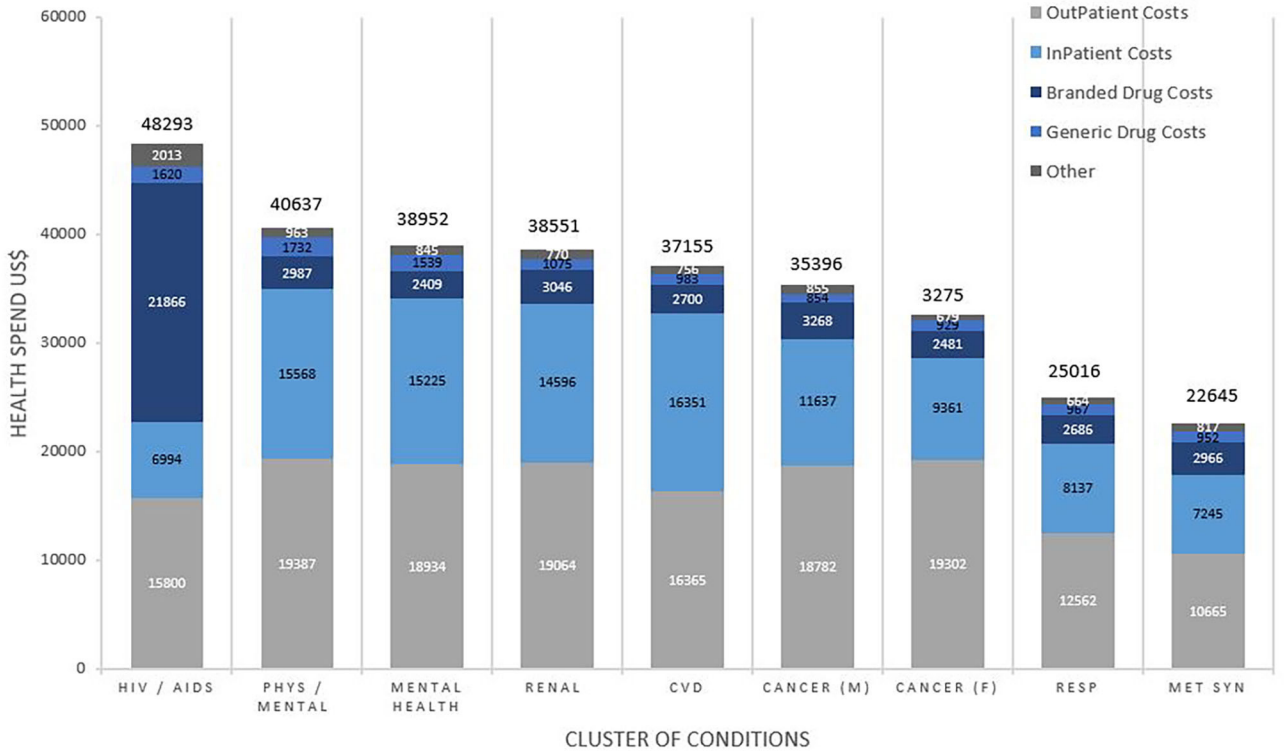

FIGURE 2 | Total average health spend and contributors to cost by cluster of conditions. Costs are weighted to the US population and reported in US$. Comparisons for cost between clusters of conditions were statistically significant overall ($p < 0.001$), for inpatient ($p < 0.001$), outpatient ($p < 0.001$), branded drugs ($p < 0.001$), and generic drugs ($p < 0.001$).

3. Overall, inpatient costs accounted for the highest increase with each subsequent condition, with a 24-fold increase between patients with 2 and 11 conditions, with considerable variation according to the "cluster" of conditions.
4. The 10 key clusters selected for more in-depth analysis included metabolic syndrome (present in 12%), chronic renal failure (8%), age related diseases (7%), respiratory disorders (4.5%), CVD (4%), cancers (4%), and mental and behavioral disorders (1–1.5%). In addition, the HIV/AIDS cluster, of particular interest to LMIC settings, was present in 0.2%.
5. Conditions occurring in over 95% of patients with these clusters included hypertension, high cholesterol, diabetes, CVD, depression, anxiety, chronic pain, osteoarthritis, allergy, COPD, male and female cancers, chronic renal failure, and HIV/AIDS.
6. HIV/AIDS was the costliest cluster due to 45% of total spending on branded drugs. This was followed by clusters of mental and behavioral disorders which were second and third most costly and, renal failure and CVD at fourth and fifth most costly, with outpatient and inpatient services accounting for roughly 90% of health spending.
7. Health service utilization varied by number of condition and clusters, with possible overutilization of specialist services and underutilization of primary care and psychiatric services in those with additional chronic conditions and certain clusters.

Frequency Rates

Our sample of 1 year of hospital episode data showed that exactly a half (56.5% scaled to the US population) of patients had at-least 2 chronic conditions. However, as this is a patient sample it does not seek to represent the US population prevalence rate which would be expected to be lower due to the presence of non-healthcare seeking adults. Nonetheless, the rates are within range of previously reported rates.

Prevalence estimates for MCC are highly heterogeneous with methodological differences such as age, the number of chronic conditions included and whether the outcomes are self-reported or verified leading to estimates that may vary up to 3-fold. Prevalence estimates for MCC range from 25.5% in the US (for 10 chronic conditions), increasing in the US to 50% for ages 45 to 65 and 81% for ages over 65 years (20), 16% in the UK (for 17 chronic conditions) to 58% (for 114 chronic conditions) (21), 45% in China to 71% in Russia in those aged over 50 (22) and just 9.4% in India (23).

Costs in Previous Literature

There is considerable variation in the magnitude of resource utilization reported between studies, health systems and data sources. Our findings align with existing evidence that MCC patients experience more complex inpatient and outpatient care scenarios leading to disproportionately high use of specialist

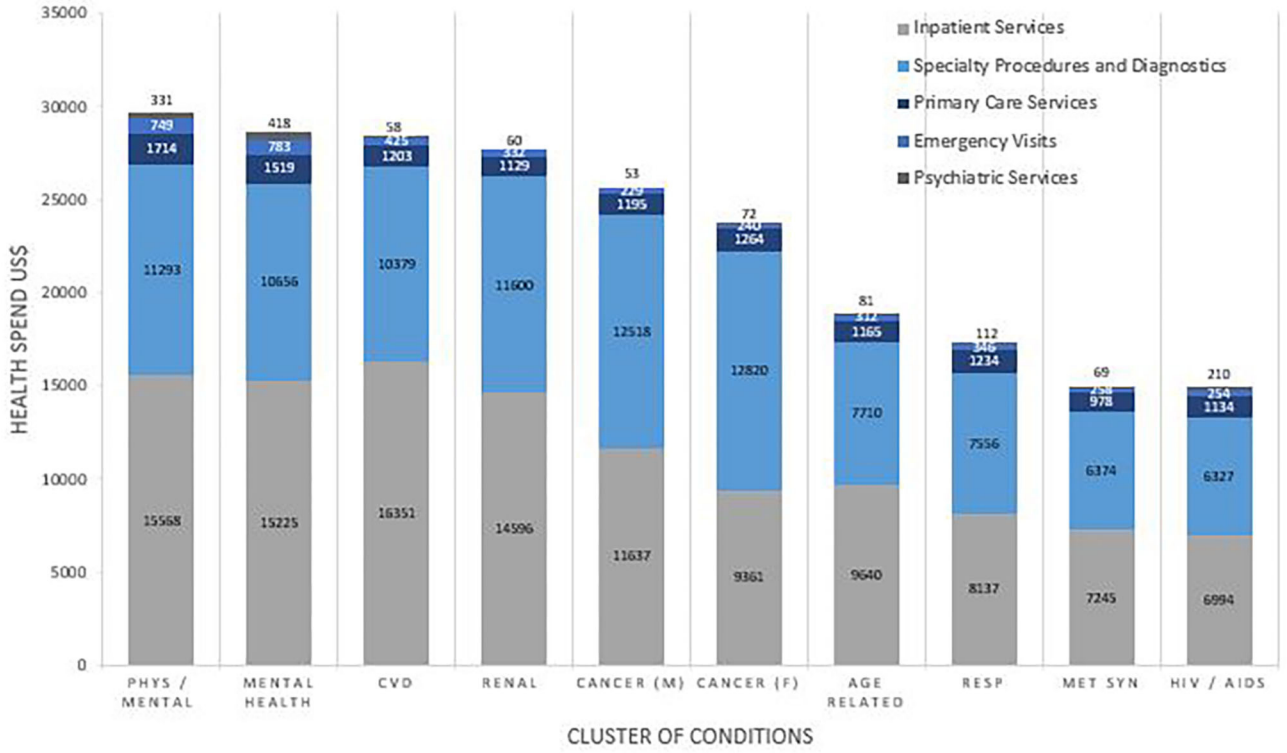

FIGURE 3 | Health Service Utilization in terms of cost, reported by Cluster. Costs are weighted to the US population and are reported in US$. Comparisons for cost between clusters of conditions were statistically significant for inpatient ($p < 0.001$), specialty procedure ($p < 0.001$), primary care ($p < 0.001$), emergency care ($p = 0.003$), and psychiatric care ($p < 0.001$).

services, visits to a multitude of physicians and confronting physicians with more problems at each visit (6, 7). MCCs patients have been reported as having more prescription medications (polypharmacy) and higher prescription drug expenditures (8, 9). However, our findings suggest that, other than for certain clusters, high utilization of medical services is the principle factor in the elevated MCC cost burden.

Patient factors previously reported to determine cost and healthcare utilization have included age, living arrangements (e.g., living alone), being female and having supplementary insurance (24–27). Our study shows that they may also be influenced by the number and clustering of chronic conditions.

Costs and Number of Conditions

Our study shows a modest increase in healthcare spending, and that the relative increase tapers off, with each subsequent chronic condition. A few previous studies reported healthcare spending doubling with each additional chronic condition (4, 5) whilst others reported smaller increments (3). As our study is much larger and includes the full spectrum of chronic conditions, it is likely to be more representative of the overall status for chronic conditions.

Clusters

This is one of the most comprehensive studies to date that identifies and quantifies clustering between the full range of chronic conditions. Conditions may cluster together by virtue of independently high prevalence rates, shared risk factors, and disease pathways or due to the causation of one condition by another, and clusters may fall into more than one of these categories (10). The commonly occurring clusters identified here were a mixture of these types: shared high prevalence rates predominantly explaining the metabolic syndrome, CVD and age-related clusters; shared risk factors for cancers and renal failure; causation of subsequent conditions for clusters with mental health disorders, chronic pain, and HIV/AIDS. Further research is required to delineate the causal pathways and also to enable the prediction of subsequent chronic conditions. Regardless of the category of clustering, the high levels of association between chronic conditions should inform healthcare redesign with a cluster-based and multiple-condition approach.

The clusters identified may vary according to the method employed such as the strength of clustering and the frequency or size of the cluster. However, in our analysis there was a high level of concordance between the three methods employed. There is no fixed methodology for defining clusters and it is an area that requires further research, including the methodology of

identifying clusters and of their categorization into concordant and discordant clusters.

Healthcare Utilization

Overall, inpatient services account for the highest increase in health spending with a 22-fold increase between having 2 and 11 conditions whilst increments in other costs such as outpatient services and medications were much lower at between 4- and 6- fold. This emphasizes the need for improved healthcare delivery to achieve greater chronic disease control and secondary prevention, as being key to cost containment in MCCs.

The value of healthcare spending per patient and patterns of utilization vary greatly by cluster regarding relative spending on medical vs. pharmaceutical costs, but less so between the medical services of primary, specialist, inpatient, psychiatric, and emergency care. The largest single category of healthcare spending was for branded drugs in patients with the HIV/AIDS cluster. Mental health clusters accounted for the second and third highest healthcare spending, higher than the CVD, cancer and renal failure clusters, and with the greatest spending, approximately half, on outpatient services.

The decrement of relative spending in primary care with each additional condition accrued suggests health system changes that enable MCC patients to be managed for longer in primary care could be hugely cost-saving. Spending on psychiatric services was low in all clusters, even in those clusters in which mental health disorders are present in over 95% of patients, suggesting underutilization of psychiatric specialist services in those who could benefit. The high overall health spending in mental health clusters may reflect poorer management of additional chronic conditions (10). Investment of resources to ensure that such patients have adequate access to healthcare for their mental health disorders is necessary to improve both the health and cost burden in such patients.

Demographic Variation

Age was not indicative of cost for all clusters; this finding for CVD and renal failure in particular are noteworthy and strengthen the case for primary prevention with a view to compression of morbidity to older ages.

Our study did not investigate variation by socio-economic status at individual level. Previous studies have shown in most countries a strong, negative relationship between SES and MCC among adults under 55 years but no consistently for adults older than 55 years (28). Inequalities in access to CVD medications have been shown both between countries and by income status within countries (29).

Implications for Health Policy

Clusters are highly amenable to large improvements in health and cost outcomes through relatively simple shifts in healthcare delivery such as the use of joint disease guidelines that tackle more than one common condition in a cluster, tailored screening and prevention. Healthcare payment mechanisms in developed countries often reward activity rather than desirable outcomes; shifting toward payment for quality or outcomes would facilitate better management of MCCs.

The variable clustering of certain chronic conditions more than others warrants urgent and careful consideration in light of the strength of such associations and the potential to have considerable impact through relatively small shifts in healthcare delivery. Our findings suggest mental health disorders may warrant particular attention through further recognition, prevention and screening practices, and disease management, due to the increased costs of clustering with mental health disorders. The relatively recent phenomenon of co-existence and clustering of chronic communicable conditions, such as HIV, with highly-prevalent NCDs, represents a serious threat for a failure of management of these conditions and increase in their prevalence, further complicated by poor healthcare access. Learnings from the successful delivery of HIV programs may be relevant to develop multiple disease frameworks, such as integrated care for NCDs and HIV in Kenya (30) and medication adherence clubs (31).

As many of the most frequent clusters identified in our study, such as metabolic syndrome, CVD, mental health issues and cancers, are highly amenable to modification, greater emphasis should be placed on the role of primary prevention and lifestyle behavior change to avoid the predicted rise in MCCs (32–34). A study in India reported MCC rates to be highest in adults with the risk factors of alcohol (12.3%), overweight (14.1%) and central obesity (17.1%) (23).

Future work on healthcare delivery toward MCC should address its many challenges of disease burden, functional health, quality of life and healthcare costs, as well as issues related to polypharmacy.

LIMITATIONS

Although the data used in this study are US based, the findings are likely to be representative for developing countries which have similar disease burdens, namely death and disability combined due to IHD, drug use disorders, back pain and other musculoskeletal disorders, COPD, DM, depression, lung cancer, and stroke (35). The contributors to costs and relative costs between clusters would be similar in developing countries as these are largely determined by the disease burden. Other findings, such as absolute healthcare spending may be less generalizable to other health systems.

Certain chronic communicable conditions, such as TB, form important clusters in LMIC (36) but without sufficient prevalence in the US for TB, were not a focus of this study with the exception of HIV/AIDS. The findings of this study are not envisaged to represent healthcare costs in LMIC where communicable conditions are prevalent.

The costs reported in this study are of total healthcare spending as it would not be feasible to distinguish costs accrued specifically from episodes directly related to chronic conditions. In addition, only costs toward healthcare appearing in financial claims were included such that other costs were not reported e.g., out of pocket expenses that have been reported to also increase for MCC patients (37).

There is no agreed taxonomy for MCCs leading to heterogeneity in the number of conditions included and whether those include symptoms and risk factors in addition to disease end-points (10–13). Furthermore, the interaction between clusters of conditions, for example concordant vs. discordant clustering, is also important to study for purposes of prediction and prevention of subsequent chronic conditions (10).

CONCLUSION

In one of the most comprehensive studies to investigate MCCs, we have reported that when applied to the US population, over half of the adult insured population have MCCs, identified the most important clusters and quantified the healthcare spending for MCCs and clusters, in a representative US patient sample. We identified that inpatient care accounts for the highest proportion of the increased spending overall but that utilization varies greatly by clusters, which is more predictive than other patient factors. Specific healthcare interventions for MCCs should take into account the local disease burden with regards to clusters. The findings emphasis the need in any long-term strategy to focus on primary prevention as the majority of the top clusters are amenable to prevention through lifestyle behavior change. In the short and medium term, health systems should focus on secondary prevention and disease control to reduce inpatient admissions. Greater reliance on specialist care may be necessary due to the greater complexity of care, however this is inefficient whilst delivered vertically for individual conditions when one in three adults have more than one chronic condition. The goal would be the delivery of care with a multi-disease framework rather than one condition at a time, in primary or secondary care. Examples of this are emerging in developing countries for HIV and CVD (30, 31, 38).

Interventions for MCCs with proven health and cost outcomes are lacking. Certain interventions have started to show early impact, including the use of fixed dose combination pills to improve medication adherence and tackle undertreatment (39), cross-condition and symptom-based management guidelines, and community models of healthcare delivery (40–42). Additional research is required to identify which interventions are impactful. Future chronic disease prevention and control approaches should be broad and patient-centric, taking into consideration healthcare payment mechanisms, the use of digital technology, tools to help with medication use and interventions to achieve positive lifestyle change, in order to avert the alarming projections of increases in MCCs rates.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: Marketscan database, available to license through Truven Health Analytics.

AUTHOR CONTRIBUTIONS

CH designed the study, oversaw the data analysis, and wrote the manuscript. YS oversaw the data analysis, designed the study, and provided input to the manuscript. AA-W reviewed and provided input to the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The study used data from the Marketscan database, available to license through Truven Health Analytics (13). This manuscript has been released as a pre-print at Medxriv (43).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2020.607528/full#supplementary-material

REFERENCES

1. Violan C, Foguet-Boreu Q, Flores-Mateo G, Salisbury C, Blom J, Freitag M, et al. Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. *PLoS ONE*. (2014) 9:e102149. doi: 10.1371/journal.pone.0102149
2. Marengoni A, Angleman S, Melis R, Mangialasche F, Karp A, Garmen A, et al. Aging with multimorbidity: a systematic review of the literature. *Ageing Res Rev*. (2011) 10:430–9. doi: 10.1016/j.arr.2011.03.003
3. McPhail SM. Multimorbidity in chronic disease: impact on health care resources and costs. *Risk Manage Healthc Policy*. (2016) 9:143–56. doi: 10.2147/RMHP.S97248
4. Schneider KM. O'Donnell BE, Dean D. Prevalence of multiple chronic conditions in the United States' Medicare population. *Health Qual Life Outcomes*. (2009) 7:82. doi: 10.1186/1477-7525-7-82
5. Wolff JL, Starfield B, Anderson G. Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. *Arch Intern Med*. (2002) 162:2269–76. doi: 10.1001/archinte.162.20.2269
6. Anderson, GF. *Chronic Care: Making the Case for Ongoing Care*. Robert Wood Johnson Foundation (2010). Available online at http://www.rwjf.org/files/research/50968chronic.care.chartbook.pdf (accessed March 3, 2018).
7. Beasley JW, Hankey TH, Erickson R, Stange KC, Mundt M, Elliott M, et al. How many problems do family physicians manage at each encounter? A WReN study. *Ann Fam Med*. (2004) 2:405–10. doi: 10.1370/afm.94
8. Mueller C, Schur C, O'Connell J. Prescription drug spending: the impact of age and chronic disease status. *Am J Public Health*. (1997) 87:1626–9.
9. Sambamoorthi U, Shea D, Crystal S. Total and out-of-pocket expenditures for prescription drugs among older persons. *Gerontologist*. (2003) 43:345–59. doi: 10.1093/geront/43.3.345
10. Hajat C, Stein E. The global burden of multiple chronic conditions: a narrative review. *Prev Med Rep*. (2018) 12:284–93. doi: 10.1016/j.pmedr.2018.10.008
11. Hajat C, Stein E, Yach D. *Multiple Chronic Conditions: The Global State*. (2017). Available online at: https://lnkd.in/etb9eAJ (accessed May 3, 2018).
12. Hajat C, Kishore SP. The case for a global focus on multiple chronic conditions. *BMJ Glob Health*. (2018) 3:e000874. doi: 10.1136/bmjgh-2018-000874

13. Academy of Medical Science. *Multimorbidity: A Priority for Global Health Research*. (2018). Available online at: https://acmedsci.ac.uk/file-download/82222577 (accessed June 27, 2018).
14. Truven Health Marketscan database. Available online at: https://marketscan.truvenhealth.com/marketscanportal/ (accessed May 3, 2018).
15. Centers for Medicare and Medicaid Services. Available online at: https://www.cms.gov/ (accessed April 1, 2019).
16. International Classification of Disease. Available online at: https://www.who.int/classifications/icd/en/ (accessed April 1, 2019).
17. The Agency for Healthcare Research and Quality Chronic Conditions list. Available online at: https://www.hcup-us.ahrq.gov/toolssoftware/chronic/chronic.jsp#codes (accessed May 3, 2018).
18. MacQueen JB. Some methods for classification and analysis of multivariate observations. In: *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability Vol 1*. Al Ain: University of California Press (1967) p. 281–97.
19. Barnett JC, Vornovitsky MS. *Health Insurance Coverage in the United States: 2015*. (2016). Available online at: https://www.census.gov/library/publications/2016/demo/p60-257.html (accessed February 1, 2018).
20. Ward BW, Schiller JS, Goodman RA. Multiple chronic conditions among US adults: a 2012 update. *Prev Chronic Dis*. (2014) 11:E62. doi: 10.5888/pcd11.130389
21. Salisbury C, Johnson L, Purdy S, Valderas JM, Montgomery AA. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract*. (2011) 61:e12–21. doi: 10.3399/bjgp11X548929
22. Garin N, Koyanagi A, Chatterji S, Tyrovolas S, Olaya B, Leonardi M, et al. Global multimorbidity patterns: a cross-sectional, population-based, multi-country study. *J Gerontol A Biol Sci Med Sci*. (2016) 71:205–14. doi: 10.1093/gerona/glv128
23. Singh K, Patel SA, Biswas S, Shivashankar R, Kondal D, Ajay VS, et al. Multimorbidity in South Asian adults: prevalence, risk factors and mortality. *J Public Health*. (2018) 41:80–9. doi: 10.1093/pubmed/fdy017
24. Rapoport J, Jacobs P, Bell NR, Klarenbach S. Refining the measurement of the economic burden of chronic diseases in Canada. *Chronic Dis Can*. (2004) 25:13–21.
25. Shelton P, Sager MA, Schraeder C. The community assessment risk screen (CARS): identifying elderly persons at risk for hospitalization or emergency department visit. *Am J Manage Care*. (2000) 6:925–33.
26. Landi F, Onder G, Cesari M, Barillaro C, Lattanzio F, Carbonin PU, et al. Comorbidity and social factors predicted hospitalization in frail elderly patients. *J Clin Epidemiol*. (2004) 57:832–6. doi: 10.1016/j.jclinepi.2004.01.013
27. Librero J, Peiró S, Ordiñana R. Chronic comorbidity and outcomes of hospital care: length of stay, mortality, and readmission at 30 and 365 days. *J Clin Epidemiol*. (1999) 52:171–9.
28. Afshar S, Roderick PJ, Kowal P, Dimitrov BD, Hill AG. Multimorbidity and the inequalities of global ageing: a cross-sectional study of 28 countries using the World Health Surveys. *BMC Public Health*. (2015) 15:776. doi: 10.1186/s12889-015-2008-7
29. Murphy A, Palafox B, O'Donnell O, Stuckler D, Perel P, AlHabib KF, et al. Inequalities in the use of secondary prevention of cardiovascular disease by socioeconomic status: evidence from the PURE observational study. *Lancet Glob Health*. (2018) 6:e292–301. doi: 10.1016/S2214-109X(18)30031-7
30. Mwangemi, F, Lamptey P. Integration of HIV and CVD services in Kenya. In: *Oral presentation, HIV and Health Systems Pre-conference*. Vienna (2010).
31. Green A. NIH project focuses on integration of HIV and NCD care. *Lancet*. (2016) 388:1869. doi: 10.1016/S0140-6736(16)31846-3
32. Islami F, Goding Sauer A, Miller KD, Siegel RL, Fedewa SA, Jacobs EJ, et al. Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. *CA Cancer J Clin*. (2018) 68:31–54. doi: 10.3322/caac.21440
33. Schuh FB, Vancampfort D, Firth J, Rosenbaum S, Ward PB, Silva ES, et al. Physical activity and incident depression: a meta-analysis of prospective cohort studies. *Am J Psychiatry*. (2018) 175:631–48. doi: 10.1176/appi.ajp.2018.17111194
34. Hajat C, Selwyn A, Harris M, Yach D. Preventive interventions for the second half of life: a systematic review. *Am J Health Promot*. (2018) 32:1122–39. doi: 10.1177/0890117117712355
35. Global Burden of Disease data. *Institute of Health Metrics and Evaluation*. Available online at: http://www.healthdata.org/united-states (accessed November 12, 2020).
36. Jeon CY, Murray MB. Diabetes mellitus increases the risk of active tuberculosis: a systematic review of 13 observational studies. *PLoS Med*. (2008) 5:e152. doi: 10.1371/journal.pmed.0050152
37. Sum G, Hone T, Atun R, Millett C, Suhrcke M, Mahal A, et al. Multimorbidity and out-of-pocket expenditure on medicines: a systematic review. *BMJ Glob Health*. (2018) 3:e000505. doi: 10.1136/bmjgh-2017-000505
38. Uebel KE, Fairall LR, van Rensburg DH, Mollentze WF, Bachmann MO, Lewin S, et al. Task shifting and integration of HIV care into primary care in South Africa: the development and content of the streamlining tasks and roles to expand treatment and care for HIV (STRETCH) intervention. *Implement Sci*. (2011) 6:86. doi: 10.1186/1748-5908-6-86
39. Rodgers A. A cure for cardiovascular disease? Combination treatment has enormous potential, especially in developing countries. *BMJ*. (2003) 326:1407–8. doi: 10.1136/bmj.326.7404.1407
40. Fairall LR, Zwarenstein M, Bateman ED, Bachmann M, Lombard C, Majara BP, et al. Effect of educational outreach to nurses on tuberculosis case detection and primary care of respiratory illness: pragmatic cluster randomised controlled trial. *BMJ*. (2005) 331750–4. doi: 10.1136/bmj.331.7519.750
41. Fairall LR, Bachmann MO, Louwagie GM, van Vuuren C, Chikobvu P, Steyn D, et al. Effectiveness of antiretroviral treatment in a South African program: a cohort study. *Arch Intern Med*. (2008) 168:86–93. doi: 10.1001/archinternmed.2007.10
42. Fairall LR, Folb N, Timmerman V, Lombard C, Steyn K, Bachmann MO, et al. Educational outreach with an integrated clinical tool for nurse-led non-communicable chronic disease management in primary care in South Africa: a pragmatic cluster randomised controlled trial. *PLoS Med*. (2016) 13:e1002178. doi: 10.1371/journal.pmed.1002178
43. Hajat C, Siegal Y, Adler-Waxman A. Clustering and healthcare costs with multiple chronic conditions in a US cross-sectional study. *medRxiv [Preprint]*. (2020) doi: 10.1101/2020.08.24.20179184

Conflict of Interest: CH and YS received funding from TEVA Pharmaceutical Industries Ltd for their time to conduct the study but did not receive funding to publish the study. AA-W is an employee of TEVA Pharmaceutical Industries Ltd. CH has received consulting fees from ECLAT, a spin-off of the University of Catania for tobacco harm reduction research, and was a paid member by Sustainability of the Advisory Panel for the Tobacco Transformation Index.

frontiers in Public Health

BRIEF RESEARCH REPORT
published: 26 February 2021
doi: 10.3389/fpubh.2021.606634

Curse or Blessing? Obesity and Income-Related Inequality in the Chinese Labor Force

Chengxiang Tang[1†], *Xiaocong Yang*[1*†], *Fei Peng*[2] *and Xianglian Hu*[3]

[1] *Department of Government Administration, School of Public Administration, Guangzhou University, Guangzhou, China,* [2] *Department of Economics, School of International Economics and Trade, Shanghai Lixin University of Accounting and Finance, Shanghai, China,* [3] *Center for Chinese Public Administration Research, School of Government, Sun Yat-sen University, Guangzhou, China*

China owns a huge labor force of around half billion workers in 2018. However, little is known about the prevalence of obesity and the association between obesity and economic status in this special population. By employing the concentration index (CI) and decomposition analysis, this paper addresses this knowledge gap by using the most recent nationally representative dataset. In specific, this study examines the prevalence of obesity and the socioeconomic gradient in the probability of obesity among Chinese workers between 16 and 65. Our results show that the prevalence of obesity is completely different by using a different measure: the overall prevalence of being general obesity (measured by body mass index, BMI $\geq$ 28) varies by gender and residency from a minimum of 5.88% to a maximum of 9.46%, whereas abdominal obesity (measured by waist circumference, WCmale $\geq$ 85 cm & WCfemale $\geq$ 80 cm) prevalence presents a socking level from 64.53% to 67.69%. Moreover, the results show a pro-rich distribution of obesity (general and abdominal) among male workers ($CI_{BMI} = 0.112$; $CI_{WC} = 0.057$) and a pro-poor distribution among female workers ($CI_{BMI} = -0.141$; $CI_{WC} = -0.166$). We also find that the direction of the contribution of socioeconomic factors to income-related inequalities in obesity differs by gender. These results have substantial implications for the measurement of socioeconomic inequality in adiposity and for improving health-related policies targeting the Chinese labor force.

Keywords: obesity, income-related, inequality, Chinese, labor force

OPEN ACCESS

Edited by:
Hideki Hashimoto,
The University of Tokyo, Japan

Reviewed by:
Mariela Deliverska,
Medical University Sofia, Bulgaria
Habib Nawaz Khan,
University of Science and Technology Bannu, Pakistan

****Correspondence:***
Xiaocong Yang
xxy980@hotmail.com

†These authors contributed equally and share first authorship

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *15 September 2020*
Accepted: *08 February 2021*
Published: *26 February 2021*

Citation:
Tang C, Yang X, Peng F and Hu X (2021) Curse or Blessing? Obesity and Income-Related Inequality in the Chinese Labor Force. Front. Public Health 9:606634. doi: 10.3389/fpubh.2021.606634

INTRODUCTION

Obesity is the most common consequence of overnutrition and has long-term negative health consequences. An increased prevalence of obesity is linked to health risks for numerous non-communicable diseases (NCDs), including cardiovascular diseases, diabetes, and some cancers (1–4). For instance, increased body mass index (BMI) is associated with an increased risk of malignancies caused by alterations in metabolism, insulin levels, and insulin-like growth factors (5). In many developed countries, obesity is a major public health concern and increases the public health burden of NCDs as well as economic costs. According to estimates, the annual medical economic costs of obesity in the United States were over US$140 billion in 2008, and if the prevalence of obesity can be held at 2010 levels, almost US$550 billion in medical costs would be saved by 2030 (4, 6, 7).

The Chinese population has a long history of undernutrition resulting in poor physical and mental health development as well as behavioral abnormalities. Over the last three decades, the Chinese economy has experienced a long period of growth accompanied by a profound transition in the population's lifestyles, wealth, and dietary habits. A wider variety of foods and drinks have become available in the Chinese market, thus increasing rates of consumption of fat-rich and energy-dense foods and sweetened carbonated beverages (8). Over the same period, rapid urbanization and motorization have contributed to decreased rates of physical activity during work and leisure activities (9, 10). Together, these factors have led to a widespread imbalance between energy intake and energy expenditure that has changed the Chinese body weight and composition of nutrition. Consequently, being overweight and obese is becoming increasingly common in both rural and urban areas. The most famous longitudinal study on the Chinese population, the China Health and Nutrition Survey (CHNS), demonstrated that the population's mean BMI has increased over recent decades (11). Specifically, the rates of being overweight and of obesity in Chinese adults were 30.6 and 12.0%, respectively, in 2010 (12).

However, a knowledge gap still exists in the related literature as limited information is available concerning overweight and obese workers in China. Our study contributes to the existing literature in two dimensions. The first contribution of our study is that we applied the most recently available dataset to examine the prevalence of obesity among Chinese workers. To the best of our knowledge, it is the first time to explore the prevalence of both general obesity and abdominal obesity in Chinese workers by using BMI and waist circumference (WC) measures. The majority of studies on obesity have focused on children and the general adult population (4, 13–18). Limited empirical evidence exists on obesity among the labor force, especially in China, which has the largest labor market in the world (19). The existing literature has documented that obesity negatively affects the labor market through two mechanisms. First, the poor health conditions caused by being overweight and by obesity increased rates of absenteeism and presenteeism and, thus, lower the productivity of the labor force (20, 21). Second, workers with higher BMI or perceived overweight-related problems may also be discriminated against by employers, consequently leading to the misallocation of human resources and decreases in workplace productivity (22, 23).

Second, this study provides a useful snapshot of the determinants of socioeconomic gradients in overweight and obesity inequalities. In addition to the prevalence of obesity, we also analyzed socioeconomic gradients by decomposing inequalities in obesity. The factors and processes that affect obesity are multiple and complex, so researchers and policymakers must often understand both the level of the problem and its determinants. A major concern in developed countries is related to the formulation of targeted and effective health policies that mitigate the disproportionate effects of adiposity on people with low incomes or socioeconomic status (SES). The rapid growth of the economy in China has influenced dietary habits, lifestyles, and other health behaviors across different socioeconomic classes.

In China, a few studies have investigated these gradients and the determinants of inequality in BMI-related variables. Using a biased sample, Tafreschi reported that the income-bodyweight gradient changes that have occurred in China from 1991 to 2009 are in line with the reversal hypothesis, which states that in poor or developing societies, obese people tend to be relatively wealthy, whereas in more developed societies, the obese are relatively poor (24). In contrast, our study focuses on a specific population of the workforce using a representative sample, which is an incremental contribution. Although previous study results may only reflect a conditional relationship rather than a causal one, they are meaningful concerning the evolution of socioeconomic health inequalities in China. This is because the shift in higher obesity rates to lower-income individuals may increase the health burden on poor people, thus worsening health inequalities. Another study, using the same dataset, demonstrated that SES is positively related to male BMI and concluded that the relationship between SES and obesity is complicated in China (25). Thus, this paper fills the gap existing in the literature regarding Chinese workers' rates of being overweight, especially in association with their SES.

The paper proceeds as follows: the next section introduces our data, materials, and methods to provide an overview of the employed dataset and a discussion of variables and methodology used in this study. Section three sets out the key results and findings and the final section presents a discussion, outlines the policy implications of the results, and offers concluding points.

METHODS

Data

This study employed the third wave (2016) of the China Labor-force Dynamics Survey (CLDS), a nationally representative dataset launched by Sun Yat-Sen University in 2012, which was the first national longitudinal survey targeted at the Chinese labor force. It offers longitudinal social survey data with rotating panel design, is conducted every 2 years, and has to date accumulated three waves of data. This dataset implements a probability-proportional-to-size sampling strategy. The population size, administrative units, and SES, which was provided by the local Bureau of Statistics, were used as the main stratification variables. The first wave of the CLDS covered 29 provinces/cities/autonomous regions (excluding Hong Kong, Macao, Taiwan, Tibet Autonomous Region, and Hainan Province) with a total number of 16,253 observations. The next two follow-up waves surveyed the same respondents, plus additional respondents, to provide 23,594 and 21,086 observations, respectively. The dataset of the third wave provides both demographic and socioeconomic information, including measures of employment status, income, household assets, and health outcomes. Of particular relevance for this research is that the CLDS not only uses objective health measurements to generate the key independent variables (e.g., BMI and WC) but also uses the employment status as well as the individual after-tax wage from all sources.

A detailed description of this survey can be found in another study (26). So far, two studies have used this data to examine income-related inequality in health variables. One focused on

health care utilization and the other investigated the self-rated health of migrant workers (27, 28). Our study used the most recent (the third) wave of the 2016 CLDS with respondents aged between 16 and 65 years who were working or seeking a job from 7,212 households within 400 communities or villages across the nation, providing a total number of 10,267 valid observations suitable for use. Thus, respondents in this study were part of the current labor force, were aged 16–65 years, and comprised 5,288 males and 4,979 females.

Dependent and Independent Variables

The first dependent variable in this study was BMI, a widely used general obesity indicator, which is calculated as weight in kilograms divided by the square of height in meters. Another obesity indicator is WC measuring individual waist circumference, which was used in many previous studies as abdominal obesity (20, 29, 30). According to the recommended criteria for Asian and Chinese people (14, 20, 29–31), a BMI of $\geq$28 kg/m^2 indicates general obesity for both sexes. Regard to the abdominal obesity, we defined respondents' WC of $\geq$ 85 cm, and WC $\geq$ 80 cm was considered as abdominal obesity for males and females, respectively. We recoded the continuous indicator BMI and WC into two respective binary health outcomes with a value of 1 indicating being obese (BMI $\geq$ 28 kg/m^2; WC of $\geq$ 85 cm for male and WC $\geq$ 80 cm for female) and a value of 0 indicating otherwise (BMI $<$ 28 kg/m$^{2;}$ WC of $<$ 85 cm for male and WC $<$ 80 cm for female).

Independent variables comprised of individual after-tax wage, age, gender, health indicators, educational level, marital status, migrant status, region, insurance status, work status, and lifestyle. The key independent variable in this study was the individual after-tax wage, which was used as a measure of individual economic status to analyze the inequality. Its natural logarithm value was employed in the empirical model to examine its effect on BMI and WC. This study also used a binary health outcome of self-rated health (0 = poor; 1 = good) to control the unobserved confounding effects on one's BMI and WC. Education level was coded as primary or below, junior secondary, senior secondary or vocational, and junior college and above. Marital status was classified as single/divorced/widowed, and currently married or cohabits. Previous studies have indicated that the living environment, working situation, and personal lifestyle may affect one's BMI and WC. Thus, we controlled for the migrant effect with four dichotomous variables comprising whether the respondents' migrant statuses were migrant (0 = N; 1 = Y), non-agricultural Hukou (0 = N; 1 = Y), or urban residence (0 = rural; 1 = urban) and if the respondents were located in the south part of China (0 = north; 1 = south). Three types of binary insurance status variables were controlled for medical, retirement, and other types of insurance. Work status was divided into five groups: seeking or willing to seek a job (current not working), employee, employer, self-employed, and agricultural work. The last set of control variables-lifestyle, comprised three binary variables of whether respondents were currently smoking, drinking and participating in physical activities in the past month, with a value of 1 indicating "Yes" and 0 indicating "otherwise."

Inequality Measurement

We constructed a concentration curve to illustrate income inequality in the probability of being obese and to calculate the concentration index (CI) to calibrate the degree and significance of income-related inequality regarding the probability of obesity. A concentration curve lying below the line of absolute equality (the 45-degree line) indicated that obesity was concentrated among richer workers, whereas a concentration curve that lies above the line indicates that obesity was concentrated among poorer workers (32). To estimate the degree and significance of inequality, we used a CI that denoted differences in obesity according to individual economic status (32). However, normalization is required because the dependent variables in this study are a binary outcome (e.g., whether the respondent was obese or not), so that the concentration index is quantified in the range -1 to $+1$. We employ the *Wagstaff-normalization-CI* not only because it has a greater emphasis on relative inequality, but also as it tends to work better for the low-frequency binary outcome (33). The Wagstaff-normalization-CI can be demonstrated by the following:

$$CI_n = Cov_w[\left(\frac{2}{N\mu}\sum\nolimits_{i=1}^{N} H_i R_i - 1\right) \div (1-\mu)] \qquad (1)$$

where CI_n denotes the *Wagstaff-normalization-CI* for the probability of being obese (BMI $\geq$ 28 kg/m^2 for both sexes, or WC of $\geq$ 85 cm for male and WC $\geq$ 80 cm for female) ranging between -1 (perfect pro-poor inequality) and 1 (perfect pro-rich inequality). The *CI* was calculated using the covariance between the probability of being obese and the fractional rank of the individual's after-tax wage. Variable H_i is a binary variable indicating whether the i^{th} individual was obese or not, μ stands for the mean rate of actually being obese for the sample, R_i is the fractional rank of the i^{th} individual according to his or her after-tax wage, for example, $i = 1$ for the respondent at the bottom of the income distribution (the poorest) and $i = N$ for those at the top (the richest). Cov_w is the covariance with sampling probability weights, which was provided by the CLDS (34, 35). The 95% confidence intervals for the CIs and associated *p*-values were obtained using the delta method (34, 36, 37). Therefore, a CI significantly smaller than 0 indicated that the greatest proportion of obesity exists among the poorest workers (or we can say poorer individuals were more likely to be obese), namely 'pro-poor' obesity inequality, whereas a CI significantly $>$0 indicated that the greatest proportion of obesity exists among the richest workers (or obesity rates occurred more within wealthier individuals), namely 'pro-rich' obesity inequality (34, 35, 38). We also plotted concentration curves and calculated CIs according to gender subsamples to investigate any related gender differences.

Decomposition of Inequality in Obesity

In the final stage of the analysis, we conducted a decomposition analysis, following previous studies, to assess the extent to which factors contribute to inequality in the probability of obesity (32, 38–40). Decomposing obesity inequality into the contributions of various explanatory factors was straightforward.

According to previously published studies, one's obesity status is not only influenced by demographic factors but also by one's social and cultural environment as well as individual economic status. These factors were classified into 10 main groups: 1) demographic factors (age, gender and self-rated health, etc.); 2) educational level (primary or below, junior secondary, senior secondary or vocational, and junior college and above); 3) marital status (single/ divorced/widowed and currently married or cohabiting); 4) immigrant status (migrant and resident register: *Hukou* status); 5) residency status (rural or urban); 6) regional status (southern or northern China); 7) insurance status (presence of medical insurance, retirement and other insurance); 8) work status (current not working but seeking a job, employee, employer, self-employed, or agricultural worker); 9) lifestyle (whether smoking, drinking, and participated in physical activities in the past month); and 10) individual after-tax wage in the past year. Because the dependent variable was a binary factor with a value between 0 and 1 (BMI and WC), we employed non-linear approximation rather than a linear method (41). Two advantages exist to using this method: (1) the approximation error of a nonlinear model tends to be smaller and (2) compared with linear decomposition, non-linear decomposition more accurately represents partial contributions. The non-linear approximation of a probit model with partial effects evaluated at means can be expressed as follows:

$$H_{BMI\ or\ WC} = \alpha^m + \sum\nolimits_j \beta_j^m X_{ij} \pm \sum\nolimits_k \gamma_k^m Z_{ik} + \delta^m y_i + \varepsilon_i \quad (2)$$

where $H_{BMIorWC}$ is the obesity indicator(s) defined above; a^m is the intercept; X_{ij} and Z_{ik} refer to the j^{th} demographic factors (age, gender, and another health indicator, etc.) and k^{th} socioeconomic factors (education, marital status, migrant status, residency, region, insurance status, work status, and lifestyle) of the i^{th} individual, respectively; y_i denotes individual economic status (measured as the logarithm of individual after-tax wage in the past year); and ε_i is the error term including approximation errors (35). Additionally, β_j^m, γ_k^m, and δ^m are the marginal effects for the aforementioned factors, dh/dx_j, dh/dz_k, and dh/dy of each demographic (x), socioeconomics (z), and individual economic factor (y), respectively, evaluated at sample means. Given Eqs. (1) and (2), the CI can be expressed as follows:

$$CI_{BMI\ or\ WC} = \left(\frac{\delta^m \overline{y}}{\mu}\right) C_y + \sum_j \left(\frac{\beta_j^m \overline{X_j}}{\mu}\right) C_j + \sum_k \left(\frac{\gamma_k^m \overline{Z_k}}{\mu}\right) C_k + \frac{GC_\varepsilon}{\mu} \quad (3)$$

where μ is the mean of $H_{BMIorWC}$. By employing Eq. (1) and adjusting it according to Wagstaff's method (33), C_y, C_j, and C_k denote the CI of y_i, x_j, and z_k, respectively. GC_ε represents the generalized CI of the error term ε. Additionally, y is the mean individual after-tax wage and $\overline{X_j}$ and $\overline{Z_k}$ represent the means of the demographic and socioeconomic factors, respectively. Moreover, the products $\left(\beta_j^m \overline{X_j}/\mu\right) C_j$, $\left(\gamma_k^m \overline{Z_k}/\mu\right) C_k$, and $\left(\delta^m \overline{y}/\mu\right) C_y$ are the contribution of demographic factor j, socioeconomics k, and individual wage y to the actual concentration index (CI), respectively. A CI was estimated for each of the factors, along with important and percentage contributions to the inequality in the probability of being obese ($CI_{BMIorWC}$). A positive (negative) contribution indicated that the given factor operated toward a pro-rich (pro-poor) distribution of being obese. We decomposed the CI for the probability of being obese according to the probit model with sample weight applied. Each CI was decomposed into the partial contributions of demographic, educational, marital, migrant, residency, insurance, work, lifestyle, and individual wage factors. Statistical software Stata version 15.1 was used for the analysis.

RESULTS

Descriptive Results

Table 1 summarizes the characteristics of the individuals surveyed by the CLDS 2016. There were 10,267 valid respondents for this study. Overall, ∼7.3% of the total labor force in China was diagnosed as general obesity, while 44.2% of the sample was diagnosed as abdominal obesity. Specifically, we find that female workers had significantly ($p < 0.001$) lower rates of general obesity (6.4%) than male workers (8.2%). Although the result shows that female workers had slightly higher rates (44.9%) of abdominal obesity than male workers (43.4%), it is statistical insignificance ($p = 0.237$). The above results imply that the measures used for obesity analysis matter and it could result in a completely different conclusion (20, 29, 30). Moreover, we find a piece of preliminary evidence that there is a huge gender wage gap as the mean of individual after-tax wage reveals that male workers have a significantly higher income (26,124.98 Yuan) than female workers (17,190.34 Yuan), and t-test results showed that this difference is statistical significance. We observe a similar pattern for most of the other variables except the rate of public social welfare status (i.e., medical and retirement insurance status), as listed in **Table 1**, indicating a significant difference among the majority factors between female and male workers in China.

Prevalence of Obesity Among Workers by Gender

In **Table 2**, the age-adjusted prevalence of being overweight and of obesity, as well as their 95% confidence intervals, are presented as totals and by gender for both rural and urban areas. Both BMI and WC measures are listed. All estimates were calculated using a weighting variable to accurately represent the working population of China; thus, the estimates were made to represent China's 454.2 million workers, comprising 218.2 million females and 236.0 million males. **Table 2** shows that the prevalence of being general overweight (BMI $\geq$ 28 kg/m^2) among the labor force was 21.98% (rural) and 24.8% (urban), whereas the prevalence of general obesity among workers was 6.265% (rural) and 8.153% (urban). Both results imply that urban workers were more likely to be obese than rural workers. In terms of the results of overall rates of abdominal overweight among Chinese workers were 19.38% (rural) and 17.19% (urban), while 64.53% (rural) and 67.69% (urban) were diagnosed as abdominal obesity. The above results

TABLE 1 | Variable statistics description, full, female, and male samples[a].

Variables[b]	Female (N = 4,979)	Male (N = 5,288)	Total (N = 10,267)	*p*-value[c]
Key variables				
General obesity (BMI≥28)	319 (6.4%)	430 (8.1%)	749 (7.3%)	0.001
Abdominal obesity4 (WCF≥80; WCM≥85)	1,334 (44.0%)	1,350 (43.6%)	2,684 (43.8%)	0.757
Individual after-tax wage				<0.001
Mean (SD)	17,190.34 (32,582.33)	26,124.98 (40,315.03)	21,792.11 (37,037.12)	
Min, Max	0.0, 800000.0	0.0, 960000.0	0.0, 960000.0	
Demographic factors				
Age				<0.001
Mean (SD)	43.27 (11.54)	44.38 (11.94)	43.84 (11.76)	
Min, Max	16.0, 65.0	16.0, 65.0	16.0, 65.0	
Gender (0 = F; 1 = M)				<0.001
Female	4,979 (100.0%)	0 (0.0%)	4,979 (48.5%)	
Male	0 (0.0%)	5,288 (100.0%)	5,288 (51.5%)	
Self-rate health (0 = Poor; 1 = Fair)	4,277 (85.9%)	4,730 (89.4%)	9,007 (87.7%)	<0.001
Socioeconomic factors				
Educational status				<0.001
Primary or below	2,035 (40.9%)	1,313 (24.8%)	3,348 (32.6%)	
Junior secondary	1,551 (31.2%)	2,161 (40.9%)	3,712 (36.2%)	
Senior secondary or vocational	627 (12.6%)	1,029 (19.5%)	1,656 (16.1%)	
Junior college and above	766 (15.4%)	785 (14.8%)	1,551 (15.1%)	
Married (0 = N; 1 = Y)	4,466 (89.7%)	4,504 (85.2%)	8,970 (87.4%)	<0.001
Migrant (0 = N; 1 = Y)	1,833 (36.8%)	753 (14.2%)	2,586 (25.2%)	<0.001
Hukou (0 = Agr;1 = Non-agr)	1,198 (24.1%)	1,427 (27.0%)	2,625 (25.6%)	0.001
Live in urban (0 = N; 1 = Y)	1,709 (34.3%)	1,947 (36.8%)	3,656 (35.6%)	0.008
South China	2,940 (59.0%)	3,173 (60.0%)	6,113 (59.5%)	0.324
Medical insurance (0 = N; 1 = Y)	4,564 (91.7%)	4,863 (92.0%)	9,427 (91.8%)	0.582
Retirement insurance (0 = N; 1 =Y)	3,303 (66.3%)	3,527 (66.7%)	6,830 (66.5%)	0.700
Other insurances (0 = N; 1 = Y)	974 (19.6%)	1,148 (21.7%)	2,122 (20.7%)	0.007
Working status				<0.001
Not working	499 (10.0%)	334 (6.3%)	833 (8.1%)	
Employee	1,857 (37.3%)	2,176 (41.1%)	4,033 (39.3%)	
Employer	72 (1.4%)	155 (2.9%)	227 (2.2%)	
Self-employ	437 (8.8%)	719 (13.6%)	1,156 (11.3%)	
Agriculture work	2,114 (42.5%)	1,904 (36.0%)	4,018 (39.1%)	
Lifestyle (N = 0; Y = 1)				
Smoking	77 (1.5%)	2,789 (52.7%)	2,866 (27.9%)	<0.001
Drinking	171 (3.4%)	1,936 (36.6%)	2,107 (20.5%)	<0.001
Physical activities	1,357 (27.3%)	1,537 (29.1%)	2,894 (28.2%)	0.041

Source: CLDS 2016.

[a]Respondents are laborforces aged 16–65 in China and results are adjusted by sampling weights; [b]for continuous variables the mean and standard deviation (in parentheses) are presented, while for categorical variables the number of respondents and percentage of the sample (in parentheses) are presented; [c]for continuous variable the ANOVA has been used while for categorical variables the chi-square test has been used to show the between-groups-difference; 4) due to data availability, there are total 6,132 respondents reported their waist circumference (WC).

not only reconfirmed that different obesity indicators produce completely different conclusions, but also pointed out that one-fifth of the labor force in China is overweight, which may be related to the rapid development of China's economic and living standard as well as the increased total caloric intake in the past decades (24).

Inequality and Decomposition Analysis

Figure 1 displays the concentration curves for the probability of being general obese (BMI ≥ 28) and abdominal obese (WC ≥ 80, F; WC ≥ 85, M) for the total, male, and female workers. We did not observe any significant pro-rich or pro-poor distribution of general obesity in the overall (total) labor force population

TABLE 2 | Prevalence of being overweight and of obesity among workers by residency.

	Pre-obesity or overweight			Obesity		
	(age-adjusted estimate) % (95% CI)			(age-adjusted estimate) % (95% CI)		
	Both sexes	**Male**	**Female**	**Both sexes**	**Male**	**Female**
BMI[a]						
Rural	21.86	23.79	20.13	6.226	6.589	5.89
	[20.36,23.36]	[21.64,25.94]	[18.23,22.03]	[5.458,6.993]	[5.480,7.698]	[4.877,6.904]
Urban	24.84	28.81	19.79	8.205	9.486	6.306
	[22.49,27.19]	[26.12,31.51]	[16.90,22.68]	[6.579,9.832]	[7.372,11.60]	[4.439,8.173]
t-test[c]	−2.092	−2.869	0.189	2.159	−2.382	−0.383
p-value	0.037	0.004	0.85	0.031	0.017	0.702
WC[b]						
Rural	19.31	19.12	19.07	64.75	65.14	64.83
	[17.73,20.89]	[17.14,21.10]	[17.05,21.10]	[62.65,66.84]	[62.47,67.82]	[61.81,67.84]
Urban	17	18.46	16.67	68.31	69.14	67.65
	[15.14,18.86]	[16.11,20.81]	[14.29,19.06]	[65.59,71.04]	[66.21,72.07]	[64.34,70.95]
t-test	1.859	0.419	1.499	−2.042	−1.995	−1.238
p-value	0.063	0.675	0.134	0.041	0.046	0.216

Sources: CLDS 2016, respondents are labor forces aged 16–65 in China and results are adjusted by sampling weights.
[a]BMI ≥ 24 kg/m^2 for overweight, BMI ≥ 28 kg/m^2 for obesity; [b]For female, WC ≥ 75 cm for overweight, WC ≥ 80 cm for obesity; for male, WC ≥ 80 cm for overweight, WC ≥ 85 cm for obesity; [c]t-test for the prevalence difference (overweight or obesity) between rural and urban subsamples.

when considering the BMI measures. However, a significant pro-rich distribution was observed in male workers (CI = 0.112, $p < 0.001$). By contrast, a pro-poor distribution was observed in female workers (CI = −0.141, $p < 0.01$) with a slightly larger CI value than existed for males ($|-0.141| > |0.112|$). The aforementioned results showed that the probability of being general obese is biased toward wealthier male workers, whereas poorer female workers are more likely to be general obese. Regarding the WC measures, the result shows that a slightly significant pro-poor distribution of abdominal obesity (CI = −0.043, $p < 0.1$) in the overall labor forces. Specifically, as same as the BMI measures, we also divided into male and female subsamples, and the result shows a significant pro-rich (CI = 0.059, $p < 0.1$) and a pro-poor (CI = −0.166, $p < 0.001$) distribution in male workers and female workers, respectively. In summary, we employed both BMI and WC to measure respondent obesity status and their particular relation with individual after-tax wage, even though the values of CIs are different, the results show that the income-related obesity distribution of male worker presents a pro-rich distribution, while for the female worker a pro-poor distribution is shown. The result suggests that male workers with higher incomes are more likely to become obese, and conversely, female workers with lower incomes are more likely to become obese.

Table 3 lists the detailed contributions of all factors to inequality in the probability of being obese for male and female workers in China. A positive (negative) partial contribution indicates that the factor increases (decreases) the total inequality in income-related obesity (CI_{BMI} *or* CI_{WC}), with positive (negative) percentages indicating increases (decreases) in percentages. We observe that individual economic status (logarithmic transferred individual after-tax wage in the past year) played a key role for pro-rich general obesity in males (77.68%) and played a similar role for pro-poor distribution in females (41.56%). However, its percentage contribution decreases significantly in the abdominal obese measures for both male workers (12.13%) and female workers (37.37%). We also find that being educated to the senior secondary or vocational and above contributed to a pro-poor probability of being obese, which was most pronounced for female workers in both generals (15.72 and 36.88%) and abdominal (9.43 and 40.86%) obesity. However, the educational factors were less of a contributing factor and played an opposite direction for male workers (from a minimum |−0.702| to a maximum |−21.06|). Moreover, some contributing factors displayed different roles for female and male workers since their CI curves operate the opposite distribution. For instance, another contributor to increased income-related inequality in general and abdominal obesity was region status for female respondents (south China, 14.99 and 7.510%), whereas this factor was decreased the income-related inequality in both obesity indexes and more influential for male workers (south China, −26.20 and −49.85%), implying that the male labor forces who were living in the south part of China play a significant role in declining the pre-rich income-related obesity distribution. In addition, one's employment status, such as being an employee, was the main contributor for both male and female respondents in general obesity (44.25 and 54.78%) but contributed less to abdominal obesity (16.80 and 7.0%). However, the results show that involvement in agriculture work decreases the income-related (general and abdominal) obesity inequality in both male and female workers, especially for female workers (−40.13 and −31.92%). Similar results with variations

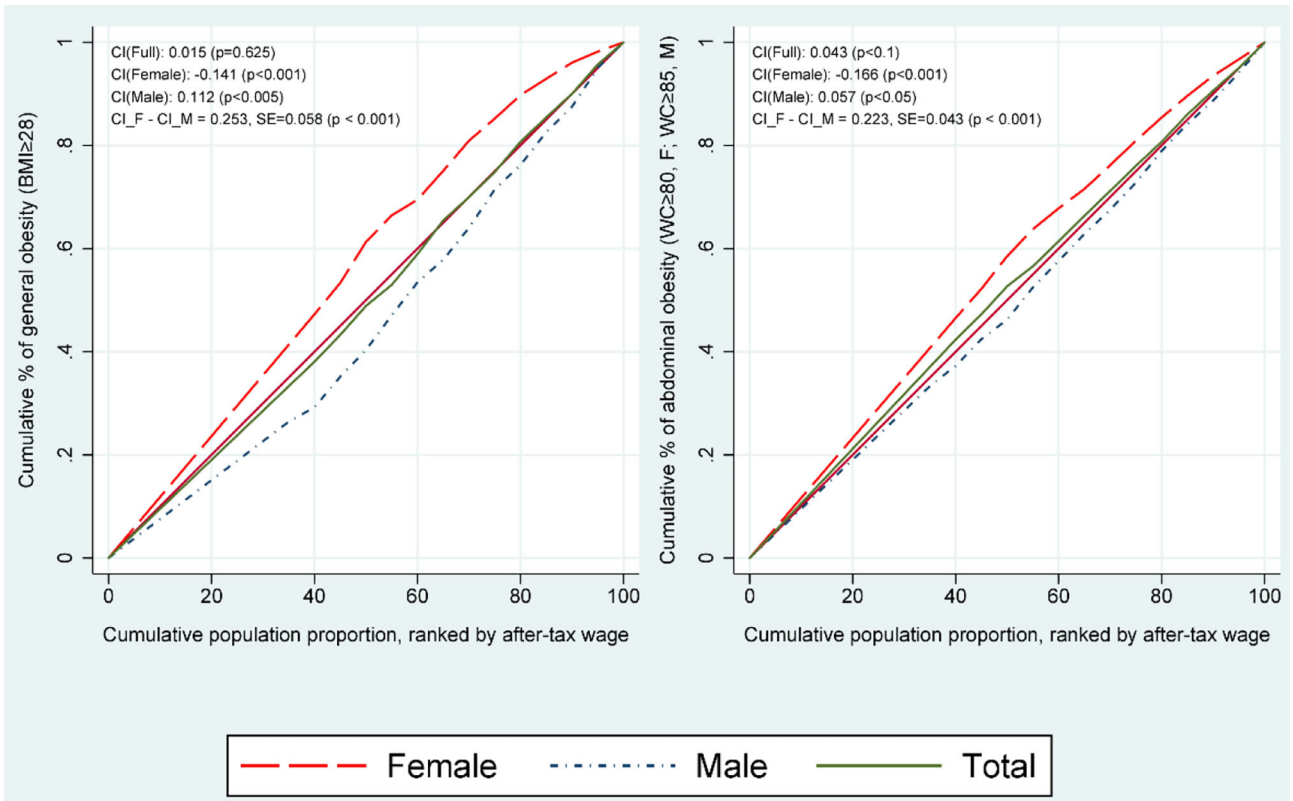

FIGURE 1 | Concentration curve for the probability of being general or abdominal obese (BMI ≥ 28; WC≥80 for females or WC≥85 for males) by total, female, and male works, China, 2016. X-axis *represents the cumulative population proportion, ranked by individual annual after-tax wage; respondents are labor forces aged 16–65 in China and results are adjusted by sampling weights*.

(positive or negative contributions) were also observed in the other group of variables, such as other migrant status indicators, insurance status indicators, and lifestyle, etc. The details are listed in **Table 3**.

Figure 2 is a graphical depiction of inequality in income-related general and abdominal obesity decomposed into 11 determinants. The results are separated by gender as per **Table 3**, and the determinants comprise demographics (age, gender, and health indicators), educational level, marital status, migrant status, residency status, region, insurance status, working status, lifestyle, individual economic status (wage), and residual terms. **Figure 2** indicates that inequality was considerably higher for females than for males, as same as results shown in **Figure 1** and **Table 3**.

Figure 2 suggests that demographic factors contributed the least to the income-related inequality in general and abdominal obesity among female and male workers. In addition, as same as listed in **Table 3**, this indicates that individual economic status and other socioeconomics factors were the main contributors to the income-related inequality in general and abdominal obesity for both female and male workers. However, some factors were associated with increased inequality in income-related obesity for male workers but with decreased inequality for females. Other socioeconomics contributors, such as insurance, immigrant status, work status, education level, work status, residency status, lifestyle, and region were also drivers of this difference. Specifically, for the female subsample, contributors such as immigrant, insurance and residency statuses accounted for reducing the pro-poor concentration of general obesity, whereas the other contributors, especial educational level, work status, wage, and unexplainable factors (residuals) operated to increase the pro-poor concentration in general obesity. For males, the distribution of pro-rich general obesity was driven by factors of residency, work status, lifestyle, and individual economic status. However, the contribution of education, immigrant, region, insurance, and marital status factors reduced the pro-rich inequality in general obesity. We can find a similar result but with some differences in abdominal obesity among male and female subsample, respectively.

DISCUSSION

This paper reveals the fact that the overall prevalence of being overweight and obesity in China is at a high-level as a developing country. The overall prevalence of being general obesity (measured by BMI) varies by gender and residency from a minimum of 5.88% to a maximum of 9.46%. In contrast,

TABLE 3 | Detailed contributions to inequality in the probability of being obese by demographic and socioeconomic factors, and female and male labor forces.

Contribution and percentage contribution	BMI				WC			
	Female		Male		Female		Male	
	Contr. to CI	% contr. to CI	Contr. to CI	% contr. to CI	Contr. to CI	% contr. to CI	Contr. to CI	% contr. to CI
CI	−0.129	100	0.112	100	−0.092	100	0.033	100
Residual	−0.056	43.193	−0.004	−3.147	−0.010	10.640	−0.008	−23.098
After-tax wage (log)	−0.054	41.555	0.087	77.676	−0.034	37.365	0.004	12.130
Demographic & socioeconomic factors								
Age	−0.003	2.107	0.018	15.649	−0.037	40.405	−0.003	−9.817
SRH	0.000	−0.381	−0.003	−2.795	−0.003	2.889	−0.004	−10.913
Education level (ref: primary or below)								
Junior secondary	0.002	−1.794	0.002	2.044	0.001	−1.097	−0.000	−0.980
Senior secondary or vocational	−0.020	15.722	−0.001	−0.702	−0.009	9.431	−0.002	−6.255
Junior college and above	−0.048	36.878	−0.024	−21.063	−0.038	40.858	−0.005	−14.255
Marital status (ref: never married/divorced/widowed)								
Currently married or cohabitating	−0.001	0.966	−0.003	−2.263	−0.000	0.034	−0.001	−2.282
Migrant status								
Migrant	0.004	−3.366	−0.005	−4.684	−0.002	1.876	−0.013	−39.840
Non-agricultural Hukou	0.028	−21.323	0.004	3.416	−0.004	4.291	0.014	41.820
Residency	0.006	−4.995	0.004	3.333	0.014	−15.187	0.009	27.701
Region	−0.019	14.994	−0.029	−26.199	−0.007	7.510	−0.016	−49.849
Insurances								
Medical insurance	0.000	−0.287	0.001	0.608	0.000	−0.225	0.001	2.002
Retirement insurance	0.002	−1.867	0.005	4.596	0.000	−0.292	0.001	2.833
Other insurances	0.037	−28.675	−0.025	−22.647	0.018	−19.213	0.004	12.074
Work status (ref: not working)								
Employee	−0.057	44.254	0.061	54.789	−0.015	16.803	0.002	7.002
Employer	−0.000	0.232	0.002	1.854	−0.000	0.074	0.000	1.201
Self-employ	0.000	−0.164	−0.002	−1.444	0.001	−1.489	0.001	2.322
Agriculture work	0.052	−40.127	−0.011	−9.814	0.029	−31.921	0.037	114.884
Lifestyle								
Smoking	−0.000	0.226	0.002	1.849	0.000	−0.330	0.000	0.231
Drinking	−0.000	0.007	0.000	0.145	−0.000	0.011	0.000	1.351
Physical activities	−0.004	2.847	0.032	28.798	0.002	−2.433	0.010	31.736

Sources: CLDS 2016.

Decomposition based on the results of probit model. Sample weights applied. CIs were calculated and adjusted by the method proposed by Wagstaff (33), as same as shown in ***Figure 1****.*

abdominal obesity (measured by WC) prevalence presents a socking level from 64.53 to 67.69%. Our study is consistent with another study that also revealed the prevalence of overweight and obesity varies greatly among different population subgroups (42). We find that there is significant income-related inequality in general and abdominal obesity thresholds for both female and male workers in China. The lower-income female workers are more likely to experience obesity than the richer female workers, whereas wealthier male workers are more likely to be obese than poorer male workers. We also noticed that the different effects for male and female workers counteract each other, thus, the overall result of the Chinese labor force does not show any pro-rich or pro-poor concentrations of income-related inequality regarding the probability of general obesity, but our result presents that there are slightly significant pro-poor concentrations of income-related inequality in abdominal obesity among all Chinese labor forces.

Additionally, the decomposition analysis revealed that educational level and individual after-tax wage in the past year were the main contributors to inequality in general and abdominal obesity for females, whereas working status, income, and lifestyle were the main contributors to inequality in general and abdominal obesity for males. Moreover, contributors to income-related obesity inequality, such as immigrant, insurance and residency statuses, were determined to reduce the pro-poor concentration of general obesity in female workers. These two factors also decreased the pro-rich concentration of general obesity in males. The aforementioned results may indicate that, besides income level, other social factors such as educational level, immigrant and insurance

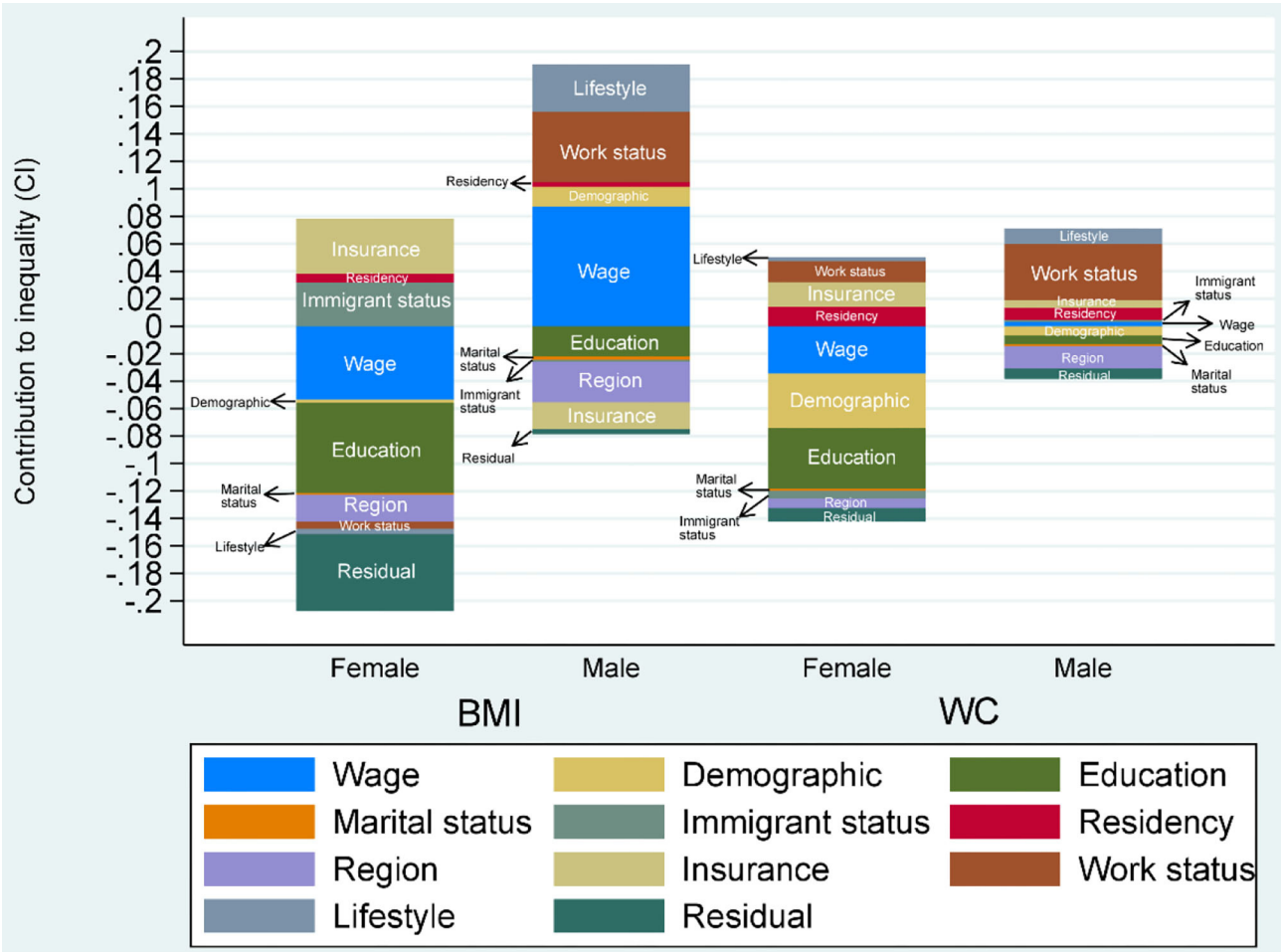

FIGURE 2 | Contributions of Demographic and socioeconomic factors to CI for the probability of obesity (BMI ≥ 28; WC ≥ 80 for females or WC ≥ 85 for males) in China by female and male workers.

statuses contributed to the general and abdominal obesity inequality significantly.

Our study represents one of the first attempts to characterize the association between obesity and socioeconomic parameters in the Chinese working population. Previous studies from other countries, such as the United States, have highlighted the medical costs and loss of productivity caused by increased rates of being overweight and obesity and have estimated that the aggregated annual cost attributable to obesity among full-time employees is US$73.1 billion (6, 43). Numerous studies have shown that obesity or being overweight among workers may cause adverse occupational consequences such as absenteeism and presenteeism, work limitations, and workplace impairment (44–46). Future studies should explore possible reasons that the labor force, which was the target population in our study, is disproportionately affected by central adiposity. Moreover, a large proportion of the working population was observed to be overweight or pre-obese, whereas the prevalence of obesity was considerably lower. This significant distinction between being overweight and being obese may indicate rapid growth in obesity in workers, rather than being an indicator of the validity of the cut-off point concerning defined body fat.

The inequalities in obesity for men were moderately less than those for women. This finding is consistent with a related theory about income and obesity (47). A related study also reported a significant pro-rich inequality for females (48), which is inconsistent with our results, indicating a pro-poor distribution among female workers. Our findings might be attributed to two factors. The first is that rich female workers might engage more in fitness-related activities and the second is that poor female workers might have an excessive intake of total carbohydrates.

Additionally, the decomposition analysis also indicated that individual income and education were the main explanatory factors of both general and abdominal obese inequality, especially for females. The possible explanation for these two significant factors might be contributed by the "one-child" policy, which to improve the educational attainment and earnings of Chinese females in the past decades, or the so-called "missing girls' effect" (49, 50). Moreover, a recent report shows that the percentage of females enrolled in masters-level postgraduate

programs increased from 50.36% in 2010 to 53.14% in 2016, while the percentage of females enrolled in undergraduate programs increased from 49.68% in 2010 to 53.44% in 2016 (51). Considering our observed samples are mainly from the latest labor force in 2016, which means that they were born in the 80s and 90s–the "one-child" policy generation. It is also explaining why the educational level at junior college and above among female workers is greater than male workers as **Table 1** shows. Therefore, compared with male workers, education plays a more important role in reducing obesity income-related inequality among female workers. Our results are also in line with those of previous studies that have indicated how income and educational attainment strongly related to levels of obesity (52). Other factors, such as working status and lifestyle, were evident contributors to inequality in obesity for male workers. It is possible as different work status and lifestyle has been proved that strongly correlated with one's weight, especial for males (53, 54). In addition, there is a related paper revealed that a majority of the social gradient concerning health inequality could be explained by the work environment and lifestyle factors (55).

This study also has a couple of limitations. First, two binary obesity indicators measured using BMI and WC may not completely present all aspects of human obesity index in the labor force (18). For example, BMI and WC do not estimate lean muscle mass. However, an advantage is that both BMI and WC are highly correlated with body fat percentage and is widely used to define obesity. Second, although the CI has been widely applied in the measurements of inequality (56), intense debate exists concerning the characteristics and value judgments of this index's spectrum (57, 58). Nevertheless, no consensus exists about the indicators that should be used to evaluate inequalities in health or health-related lifestyle (59). Third, our study presented evidence of correlation instead of causality. Finally, our data is limited to the Chinese labor force, which we considered important as it accounts for around one-fifth of the world's labor force.

CONCLUSION

To summarize, this study is the first to describe the prevalence of overweight and obesity in the Chinese labor force. Our research is relevant because we focused on the population group that is most crucial to productivity and economic development. The results also have substantial implications for the measurement of socioeconomic inequalities in adiposity. For example, income factors play opposite roles in contributing to inequalities in obesity between female and male labor forces in China. Educational level, however, plays a significant role in decreasing the obesity income-related inequality for both sexes, especially for female workers. The possible reason could be contributed by the "one-child" policy or the so-called "missing girls' effect." Apart from individual income, education, and lifestyle, we observe that work-related factors, such as work status, contribute to inequality in obesity. For health policymakers, understanding the determinants of being overweight and of obesity may help in designing interventions that promote health and fitness and, consequently, improve associated labor outcomes.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: http://www.cnsda.org/index.php?r=projects/view&id=75023529.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

CT conceived and designed the experiments. XY performed the experiments and analyzed the data. CT and XY wrote the paper. XY and FP restructured, polished, and revised the paper. XH contributed this work by providing data access. All authors read and approved the final manuscript.

FUNDING

The authors gratefully acknowledge the financial aids from the Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 71704143); the Ministry of Education in China (MOE) Project of Humanities and Social Sciences for Young Scholars (Grant No. 18YJCZH221); and the research hatch fund of Guangzhou University (Grant No. RP2020012). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

This research uses data from China Labor-force Dynamic Survey (CLDS). We thank the data collectors in the School of Sociology & Anthropology, Sun Yat-Sen University. Neither the original collectors nor the distributors of the data bear any responsibility for the analyses or interpretations presented here. All remaining errors are our own.

REFERENCES

1. Bianchini F, Kaaks R, Vainio H. Overweight, obesity, and cancer risk. *Lancet Oncol.* (2002) 3:565–74. doi: 10.1016/s1470-2045(02)00849-5
2. Kotchen TA. Obesity-related hypertension: epidemiology, pathophysiology, and clinical management. *Am J Hypertens.* (2010) 23:1170–8. doi: 10.1038/ajh.2010.172
3. Nguyen NT, Nguyen XM, Lane J, Wang P. Relationship between obesity and diabetes in a US adult population: findings from the National Health

and Nutrition Examination Survey, 1999-2006. *Obes Surg.* (2011) 21:351–5. doi: 10.1007/s11695-010-0335-4
4. Maffetone PB, Laursen PB. The Prevalence of Overfat Adults and Children in the US. *Front Public Health.* (2017) 5:90. doi: 10.3389/fpubh.2017.00290
5. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet.* (2008) 371:569–78. doi: 10.1016/S0140-6736(08)60269-X
6. Finkelstein EA, Trogdon JG, Cohen JW, Dietz W. Annual medical spending attributable to obesity: payer-and service-specific estimates. *Health Aff.* (2009) 28:w822–31. doi: 10.1377/hlthaff.28.5.w822
7. Finkelstein EA, Khavjou OA, Thompson H, Trogdon JG, Pan L, Sherry B, et al. Obesity and severe obesity forecasts through 2030. *Am J Prev Med.* (2012) 42:563–70. doi: 10.1016/j.amepre.2011.10.026
8. Wang H, Du S, Zhai F, Popkin BM. Trends in the distribution of body mass index among Chinese adults, aged 20–45 years (1989–2000). *Int J Obes.* (2007) 31:272. doi: 10.1038/sj.ijo.0803416
9. Bell AC, Ge K, Popkin BM. The road to obesity or the path to prevention: motorized transportation and obesity in China. *Obes Res.* (2002) 10:277–83. doi: 10.1038/oby.2002.38
10. Huang W-Z, Yang B-Y, Yu H-Y, Bloom MS, Markevych I, Heinrich J, et al. Association between community greenness and obesity in urban-dwelling Chinese adults. *Sci Total Environ.* (2020) 702:135040. doi: 10.1016/j.scitotenv.2019.135040
11. Popkin BM, Du S, Zhai F, Zhang B. Cohort profile: the China health and nutrition survey–monitoring and understanding socioeconomic and health change in China, 1989-2011. *Int J Epidemiol.* (2010) 39:1435–40. doi: 10.1093/ije/dyp322
12. Li XY, Jiang Y, Hu N, Li YC, Zhang M, Huang ZJ, et al. Prevalence and characteristic of overweight and obesity among adults in China, 2010. *Zhonghua Yu Fang Yi Xue Za Zhi [Chinese Journal of Preventive Medicine].* (2012) 46:683–686.
13. He Q, Li X, Wang R. Childhood obesity in China: does grandparents' coresidence matter? *Econ Hum Biol.* (2018) 29:56–63. doi: 10.1016/j.ehb.2018.02.001
14. Xi B, Liang Y, He T, Reilly KH, Hu Y, Wang Q, et al. Secular trends in the prevalence of general and abdominal obesity among Chinese adults, 1993-2009. *Obes Rev.* (2012) 13:287–96. doi: 10.1111/j.1467-789X.2011.00944.x
15. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of obesity and trends in body mass index among US children and adolescents, 1999–2010. *JAMA.* (2012) 307:483–90. doi: 10.1001/jama.2012.40
16. Yanai M, Kon A, Kumasaka K, Kawano K. Body mass index variations by age and sex, and prevalence of overweight in Japanese adults. *Int J Obes Relat Metab Disord.* (1997) 21:484–8. doi: 10.1038/0800434
17. Maffetone PB, Rivera-Dominguez I, Laursen PB. Overfat adults and children in developed countries: the public health importance of identifying excess body fat. *Front Public Health.* (2017) 5:190. doi: 10.3389/fpubh.2017.00190
18. Dreher M, Hoffmann SW, Brendel C, Heser D, Simon P. Convenience behavior and being overweight in adults: development and validation of the convenience behavior questionnaire. *Front Public Health.* (2019) 7:20. doi: 10.3389/fpubh.2019.00020
19. Bonde JP, Viikari-Juntura E. The obesity epidemic in the occupational health context. *Scand J Work Environ Health.* (2013) 39:217–9. doi: 10.5271/sjweh.3362
20. Mosca I. Body mass index, waist circumference and employment: evidence from older Irish adults. *Econ Hum Biol.* (2013) 11:522–33. doi: 10.1016/j.ehb.2013.05.002
21. Caliendo M, Lee WS. Fat chance! Obesity and the transition from unemployment to employment. *Econ Hum Biol.* (2013) 11:121–33. doi: 10.1016/j.ehb.2012.02.002
22. Larose SL, Kpelitse KA, Campbell MK, Zaric GS, Sarma S. Does obesity influence labour market outcomes among working-age adults? Evidence from Canadian longitudinal data. *Econ Hum Biol.* (2016) 20:26–41. doi: 10.1016/j.ehb.2015.09.007
23. Schmier JK, Jones ML, Halpern MT. Cost of obesity in the workplace. *Scand J Work Environ Health.* (2006) 32:5–11. doi: 10.5271/sjweh.970
24. Tafreschi D. The income body weight gradients in the developing economy of China. *Econ Hum Biol.* (2015) 16:115–34. doi: 10.1016/j.ehb.2014.02.001
25. Fang C, Liang Y. Social disparities in body mass index (BMI) trajectories among Chinese adults in 1991-2011. *Int J Equity Health.* (2017) 16:146. doi: 10.1186/s12939-017-0636-5
26. Wang J, Zhou Y, Liu S. China labor-force dynamics survey. *Chinese Sociol Dialogue.* (2017) 2:83–97. doi: 10.1177/2397200917735796
27. Lu L, Zeng J, Zeng Z. What limits the utilization of health services among china labor force? analysis of inequalities in demographic, socioeconomic and health status. *Int J Equity Health.* (2017) 16:30. doi: 10.1186/s12939-017-0523-0
28. Shao C, Meng X, Cui S, Wang J, Li C. Income-related health inequality of migrant workers in China and its decomposition: an analysis based on the 2012 China Labor-force Dynamics Survey data. *J Chin Med Assoc.* (2016) 79:531–7. doi: 10.1016/j.jcma.2016.02.009
29. Wildman RP, Gu D, Reynolds K, Duan X, He J. Appropriate body mass index and waist circumference cut-offs for categorization of overweight and central adiposity among Chinese adults. *Am J Clin Nutr.* (2004) 80:1129–36. doi: 10.1093/ajcn/80.5.1129
30. Zhou BF, Cooperative Meta-Analysis Group of the Working Group on Obesity. in: C. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults–study on optimal cut-off points of body mass index and waist circumference in Chinese adults. *Biomed Environ Sci.* (2002) 15:83–96.
31. Xu W, Zhang H, Paillard-Borg S, Zhu H, Qi X, Rizzuto D. Prevalence of overweight and obesity among chinese adults: role of adiposity indicators and age. *Obes Facts.* (2016) 9:17–28. doi: 10.1159/000443003
32. O'Donnell O, Van Doorslaer E, Wagstaff A, Lindelow M. *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and Their Implementation*. Washington, DC: The World Bank (2007).
33. Wagstaff A. The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. *Health Econ.* (2005) 14:429–32. doi: 10.1002/hec.953
34. Van Doorslaer E, Masseria C. *Income-Related Inequality in the Use of Medical Care in 21 OECD Countries*. Paris: OECD (2004).
35. Kakwani N, Wagstaff A, Van Doorslaer E. Socioeconomic inequalities in health: measurement, computation, and statistical inference. *J Econom.* (1997) 77:87–103.
36. Koolman X, van Doorslaer E. On the interpretation of a concentration index of inequality. *Health Econ.* (2004) 13:649–56. doi: 10.1002/hec.884
37. Rao CR, Statistiker M, Rao CR. *Linear statistical inference and its applications*. New York, NY: Wiley (1973).
38. Gonzalo-Almorox E, Urbanos-Garrido RM. Decomposing socioeconomic inequalities in leisure-time physical inactivity: the case of Spanish children. *Int J Equity Health.* (2016) 15:106. doi: 10.1186/s12939-016-0394-9
39. Roberts K, Cavill N, Hancock C, Rutter H. *Social and Economic Inequalities in Diet and Physical Activity*. London: University of Bath (2013).
40. Wagstaff A, Van Doorslaer E. Measuring and testing for inequity in the delivery of health care. *J Hum Resour.* (2000) 35:716–733. doi: 10.2307/146369
41. Saito E, Gilmour S, Yoneoka D, Gautam GS, Rahman MM, Shrestha PK, et al. Inequality and inequity in healthcare utilization in urban Nepal: a cross-sectional observational study. *Health Policy Plan.* (2016) 31:817–24. doi: 10.1093/heapol/czv137
42. Zhang L, Wang Z, Wang X, Chen Z, Shao L, Tian Y, et al. Prevalence of overweight and obesity in China: results from a cross-sectional study of 441 thousand adults, 2012–2015. *Obes Res Clin Pract.* (2020) 14:119–126. doi: 10.1016/j.orcp.2020.02.005
43. Finkelstein EA, DiBonaventura M, Burgess SM, Hale BC. The costs of obesity in the workplace. *J Occup Environ Med.* (2010) 52:971–6. doi: 10.1097/JOM.0b013e3181f274d2
44. Rodbard HW, Fox KM, Grandy S, Shield Study G. Impact of obesity on work productivity and role disability in individuals with and at risk for diabetes mellitus. *Am J Health Promot.* (2009) 23:353–60. doi: 10.4278/ajhp.081010-QUAN-243
45. Pollack KM, Sorock GS, Slade MD, Cantley L, Sircar K, Taiwo O, et al. Association between body mass index and acute traumatic workplace injury in hourly manufacturing employees. *Am J Epidemiol.* (2007) 166:204–11. doi: 10.1093/aje/kwm058
46. Hertz RP, Unger AN, McDonald M, Lustik MB, Biddulph-Krentar J. The impact of obesity on work limitations and cardiovascular

risk factors in the US workforce. *J Occup Environ Med.* (2004) 46:1196–203.
47. Lakdawalla D, Philipson T. The growth of obesity and technological change. *Econ Hum Biol.* (2009) 7:283–93. doi: 10.1016/j.ehb.2009.08.001
48. Ljungvall A, Gerdtham UG. More equal but heavier: a longitudinal analysis of income-related obesity inequalities in an adult Swedish cohort. *Soc Sci Med.* (2010) 70:221–31. doi: 10.1016/j.socscimed.2009.10.014
49. Chen Y, Li H, Meng L. Prenatal sex selection and missing girls in China. *J Hum Resour.* (2013) 48:36–70. doi: 10.3368/jhr.48.1.36
50. Ebenstein AY, Sharygin EJ. The Consequences of the "Missing Girls" of China. *World Bank Econ Rev.* (2009) 23:399–425. doi: 10.1093/wber/lhp012
51. Ministry of Education. Ministry of Education of the People's Republic of China, Education Statistics Annual 2010-2017. Number of Female Students of Schools by Type and Level. *Education Statistics Annual 2010-2017 Number of Female Students of Schools by Type and Level.* (2018). Available online at: http://www.moe.gov.cn/s78/A03/moe_560/jytjsj_2015/2015_qg/201610/t20161012_284506.html (accessed November 16, 2020).
52. Costa-Font J, Gil J. What lies behind socioeconomic inequalities in obesity in Spain? A decomposition approach. *Food Policy.* (2008) 33:61–73. doi: 10.1016/j.foodpol.2007.05.005
53. Brennan SL, Henry MJ, Nicholson GC, Kotowicz MA, Pasco JA. Socioeconomic status, obesity and lifestyle in men: the Geelong Osteoporosis Study. *J Men's Health.* (2010) 7:31–41. doi: 10.1016/j.jomh.2009.10.004
54. Burgess E, Hassmén P, Pumpa KL. Determinants of adherence to lifestyle intervention in adults with obesity: a systematic review. *Clin Obes.* (2017) 7:123–135. doi: 10.1111/cob.12183
55. Borg V, Kristensen TS. Social class and self-rated health: can the gradient be explained by differences in life style or work environment? *Soc Sci Med.* (2000) 51:1019–30. doi: 10.1016/s0277-9536(00)00011-3
56. Zhang Q, Wang Y. Trends in the association between obesity and socioeconomic status in U.S. adults: 1971 to 2000. *Obes Res.* (2004) 12:1622–32. doi: 10.1038/oby.2004.202
57. Erreygers G, Van Ourti T. Putting the cart before the horse. A comment on Wagstaff on inequality measurement in the presence of binary variables. *Health Econ.* (2011) 20:1161–5. doi: 10.1002/hec.1754
58. Wagstaff A. Correcting the concentration index: a comment. *J Health Econ.* (2009) 28:516–20, author reply 521-4. doi: 10.1016/j.jhealeco.2008.12.003
59. Kjellsson G, Gerdtham UG. On correcting the concentration index for binary variables. *J Health Econ.* (2013) 32:659–70. doi: 10.1016/j.jhealeco.2012.10.012

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

ORIGINAL RESEARCH
published: 15 March 2021
doi: 10.3389/fpubh.2021.629118

Political Trust Influences the Relationship Between Income and Life Satisfaction in Europe: Differential Associations With Trust at National, Community, and Individual Level

Jocelyne Clench-Aas[1]* *and Arne Holte*[2]

[1] Division of Mental and Physical Health, Norwegian Institute of Public Health, Oslo, Norway, [2] Department of Psychology, University of Oslo, Oslo, Norway

OPEN ACCESS

Edited by:
Ulrich Laaser,
Bielefeld University, Germany

Reviewed by:
Bosiljka Svetozar Djikanovic,
University of Belgrade, Serbia
Enver Envi Roshi,
University of Medicine, Tirana, Albania

***Correspondence:**
Jocelyne Clench-Aas
Jocelyne.Clench-Aas@fhi.no

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *04 December 2020*
Accepted: *08 February 2021*
Published: *15 March 2021*

Citation:
Clench-Aas J and Holte A (2021) Political Trust Influences the Relationship Between Income and Life Satisfaction in Europe: Differential Associations With Trust at National, Community, and Individual Level. Front. Public Health 9:629118. doi: 10.3389/fpubh.2021.629118

Background: A high level of well-being is associated with personal, community and national income, as well as personal, social and political trust. How these measures relate to each other within and between countries and within and across structural levels of society is largely unknown. To study this, we propose a three-layer nested socio-structural model. Each layer (individual, community, country) contains a measure of income, trust and satisfaction.

Method: With this model, we analyzed data from two waves of the European Social Survey (ESS, 2006, 2012) in 19 countries ($N = 72{,}461$; weighted $N = 73{,}307$) with multilevel techniques. Indicators were personal, community, and national income; personal, social and political trust; and personal life satisfaction, social and political satisfaction.

Results: Personal life satisfaction was associated with all income and trust variables. Greatest effect on personal life satisfaction, came from the national level, including political trust and income. However, 2/3 of the variance in personal life satisfaction came from income, that is personal, community and national. Within each socio-structural level, satisfaction was associated with income, but significantly modified by trust. When income and trust at all three levels were included, there was a significant association of the national layer on the social layer, and of the social layer on the individual layer as to the income–personal life satisfaction relationship. Consistent with the "the buffer hypothesis," all three forms of trust acted as a buffer against the effect of personal income on life satisfaction. Low-trust countries had strong income–personal life satisfaction associations and the moderating role of trust was also stronger. High- and medium-trust countries had no such associations. Likewise, direct associations between political and personal trust were much stronger in the low-trust countries.

Conclusion: The model presented in this study provides authorities with a framework for policies that will improve the general well-being of their population. Trust and income strongly influence personal life satisfaction. Money is the most important. However, trust forcefully dampens the effect of income. Politicians who want to enhance their population's personal life satisfaction, should raise the levels of trust in their electorate.

Keywords: well-being, income, trust, satisfaction, moderation, Europe

INTRODUCTION

Most people would, without doubt, prefer to live in a society with a high level of trust rather than in one with a low level (1). This concerns personal trust in terms of self-confidence and self-esteem. It concerns social trust in terms of trusting other people. It concerns political trust in terms having trust in how the country is run.

One important effect of trust is that it influences a person's well-being (1). Well-being refers partly, to an individual's long term, cognitive evaluation of one's life as a whole, life satisfaction, and partly to a more short-term, positive emotional or affective state, happiness (2–4). Life satisfaction may be divided into different facets: (1) satisfaction with one's personal life, personal life satisfaction; (2) satisfaction with one's social environment, social satisfaction; and (3) satisfaction with how the country is run, political satisfaction.

Furthermore, income seems to influence both personal, social and political trust as well as the different facets of satisfaction. This applies to personal income, community income as well as national income (5–15). For a more detailed review of studies on relationships between well-being, income and trust, see **Supplementary Material 1**.

However, two limitations to most of these studies are that they only show bivariate associations and that they rarely account for more than one or two levels of society, i.e., individual, community or national. Consequently, the relative contributions of personal, community and national income, and personal, social and political trust on the different facets of satisfaction with life is still unknown.

Examples of unanswered questions are: Which one of personal, community or national income matters the most to which facets of life satisfaction (6, 10, 16–18)? What are the roles of personal, social and political trust within these relationships? May trust play a moderating role on the relationships between income and satisfaction? May the different facets of trust (e.g., personal, social and political), have separate relative associations at different structural levels of a society (e.g., individual, community, and national)? May they have separate moderating effects on well-being, social and political satisfaction, and, respectively?

In 1991 Göran Dahlgren and Margaret Whitehead introduced a holistic multilevel model, the so called "rainbow model," to conceptualize how economic, environmental and social inequalities may determine people's risk of getting ill, their ability to prevent sickness, or their access to effective treatments (19). They placed the individual at the center of the model, with its fixed factors such as sex, age, and constitutional endowment. Surrounding them were different layers of modifiable factors that can influence health, such as individual lifestyles; social and community networks, economic, and cultural and physical environment. This framework has inspired researchers to construct a range of hypotheses about the determinants of health and to explore their relative influence on different health outcomes.

To be able to determine the relative associations between personal, community and national income and personal, social and political trust, and personal satisfaction with life, and to determine the possible moderating role of trust on the relationships between income and different facets of satisfaction, we designed a similar model.

Like Dahlgren and Whitehead's model of social determinants of health (19), we regard the society as a construction with three nested socio-structural layers. The main determinants of well-being are layered from the individual to the major structural environment and each layer can influence the well-being of the individual. We define the individual person as the basic unit (micro layer). We then regard the individual as nested into her or his local community (mezzo layer), which again is nested into the country (macro layer). Likewise, we regard the individual person's economy as nested into the economy of her or his local community, which again is nested into the economy of the country.

We then hypothesize that there is a direct association between income, trust, and satisfaction within each layer. In addition, we hypothesize that trust, i.e., personal trust, social trust, and political trust, modifies the associations between income and satisfaction (**Figure 1**). In particular, we hypothesize that personal, social and political trust acts as a buffer against the effect of personal income on personal life satisfaction ("the buffer hypothesis").

Why then is it important to investigate the effects of income and trust on satisfaction at all three levels of a society, individual, community, and country? Understanding of how these parameters relate to each other may help us getting a deeper comprehension of how societies work. Good policy development in terms of getting a population who is satisfied with life, their social environment, and how the country is run, may depend on which of these parameters have the greatest positive effects on their well-being.

But, how can including trust shed light on the multilevel relationship between well-being and income? Until we have analyzed these associations together in one and the same model,

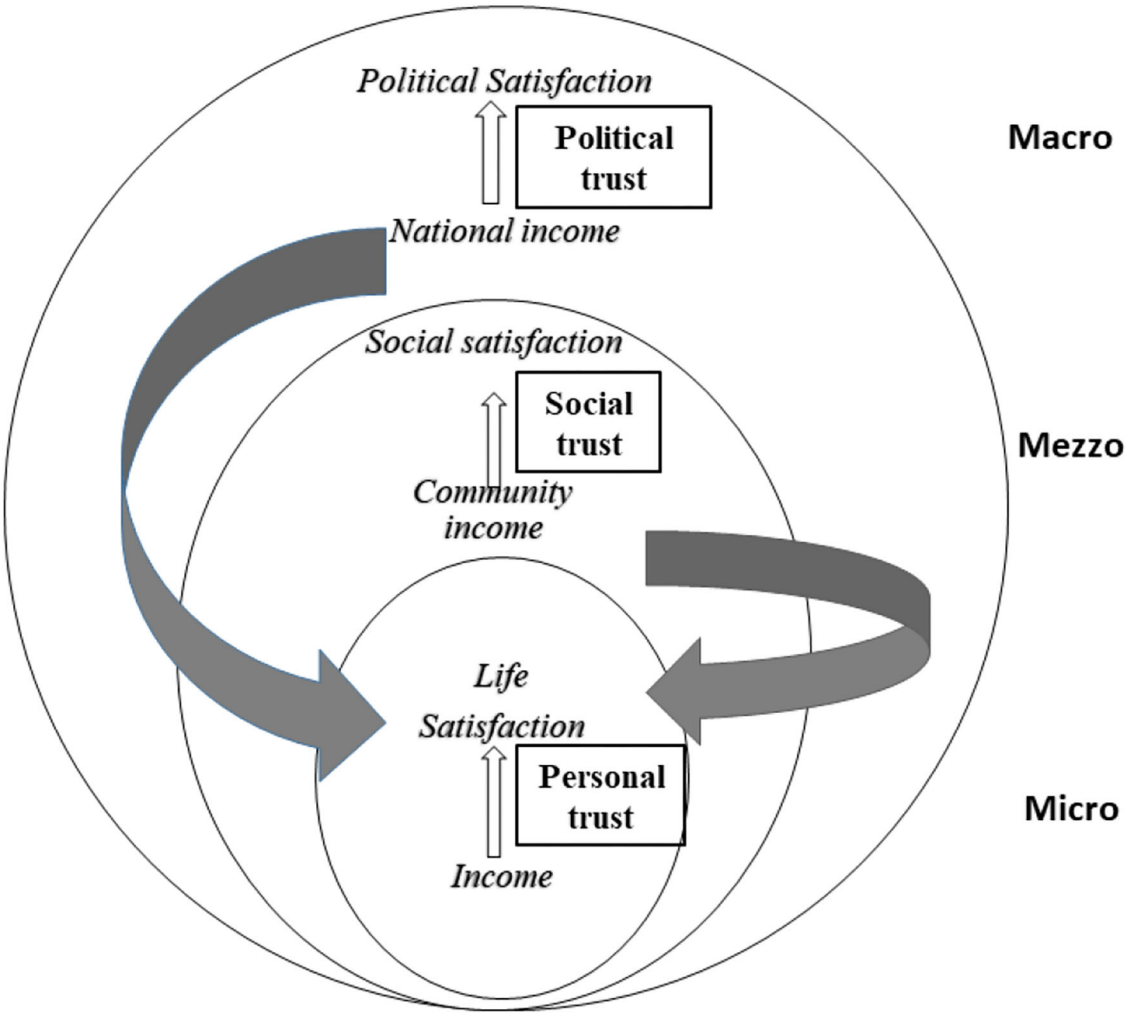

FIGURE 1 | A schematic presentation of the socio-structural model, indicating the three layers. Micro level = Individual. Mezzo level = Community. Macro level = Country. Indicators of income: Micro level = Personal income. Mezzo level = Community income. Macro level = National income. Indicators of satisfaction: Micro level: Personal life satisfaction (PLS). Mezzo level = Social satisfaction. Macro level = Political satisfaction.

taking into account effects from all three layers of society, there may be difficulties in interpreting consequences of the single relationships.

Why then could investigating the moderating role of trust in the relationship between well-being and income at the three different layers be helpful? Reaching a better comprehension of the role that different kinds of trust play in moderating effects of income on satisfaction, may be crucial in understanding basic mechanisms of society. This may again be decisive in determining which political strategies should be approached in aiming to enhance people's well-being and health. Imagine for example, that if in some countries, enhancing social trust, makes life simpler, easier, more pleasant, and friendly, this may be as effective a measure to enhance well-being and health as is increasing personal, community or national income.

Building on previous findings, we take the field one step further by integrating income, trust, and satisfaction into one holistic three-level nested model as described above, to answer the following research questions (RQ):

RQ1 What is the statistical contribution of income and trust at each socio-structural layer, i.e., micro (individual), mezzo (community), macro (country), to variation in their respective satisfaction parameter, i.e., personal life satisfaction, social satisfaction, and political satisfaction ("Within layers effects")?

RQ2 What is the statistical contribution of each of the individual parameters within each socio-structural layer to variations in personal life satisfaction ("Total holistic effect")?

RQ3 What is the relative, independent and simultaneous contribution (effect size) of each content theme at each socio-structural layer to variation in personal life satisfaction ("Effect size of holistic approach")?

RQ4 What is the relative contribution (effect size) of each socio-structural layer to variation in personal life satisfaction ("Between layers effects")?

RQ5 What is the relative contribution (effect size) of each content theme, i.e., income and trust to variation in personal life satisfaction ("The themes effects")?

RQ6 Does the relationship between income and well-being differ between countries according to their level of trust ("Variation by countries")?

RQ7 Does trust function as a moderator, in addition to its direct effect, either within the layers or in the holistic model? ("The buffer hypothesis")?

MATERIALS AND METHODS

Supplementary Material 2 provides a more detailed description of methods used in this study.

We used data from the European Social Survey (ESS). Data are highly comparable across nations, with a high response rate in all rounds.

In the present study, the data were restricted to the years complete with respect to the choice of variables. Thus, we used the cumulative dataset for rounds three and six (corresponding to 2006–2012), found on ESS web page (www.europeansocialsurvey.org). Data from the respondents in the 19 countries that participated in both rounds and included the variables of interest, were used. The final sample was $N = 72{,}461$ (Weighted-$N = 73{,}307$) and had a mean age of 48 years and 54% females (in the weighted sample 46 years and 51%. The data are freely available on the European Social Survey internet site (https://www.europeansocialsurvey.org/).

Measures

Level Defining Variables

The primary analysis is multilevel. In multilevel analysis, levels are specified prior to the analysis that defines the clusters that the analyses are performed within. For this study, we used three levels of analysis, (1) micro (individual), (2) mezzo (community), and (3) macro (country). See **Supplementary Material 2**—Methods for more details in definitions of variables.

Micro level was defined by the informant's personal number.

Mezzo level was defined by two nested variables: (a) regions within each country and (b) social class. The respondent's social class was determined using education and occupation See more details concerning these two variables in **Supplementary Material 2**–Methods. The value is a mean of the respondent and his/her partner if present. If data on occupation or education was missing for the partner, we used the respondent's education or occupation.

Macro level was defined by 19 countries: Belgium (BE), Bulgaria (BG), Cyprus (CY), Denmark (DK), Finland (FI), France (FR), Germany (DE), Ireland (IE), Netherlands (NL), Norway (NO), Poland (PL), Portugal (PT), Russia (RU), Spain (ES), Slovakia (SK), Slovenia (SI), Sweden (SE), Switzerland (CH), and United Kingdom (GB).

Income

Three income variables were used as independent variables.

At the *micro level*, we used personal income. Personal income was measured in terms of the annual household income of the individual. For further information on classification of personal income, and the necessary standardization procedures used to homogenize slightly different methods in classifying personal income, see **Supplementary Material 2**, **Table 2** in **Supplementary Material 3**.

At the mezzo level, as seen in other studies (20), we used community income. Community income was calculated for this study as the aggregate of the household income value by country, region, and social class. The aggregate value was divided by 1,000 to ease interpretation.

At the *macro level*, we used national income. National income was measured in terms of the Gross Domestic Product (GDP). The unit of measure was GDP per capita, PPP (current international $). Purchasing power parity (PPP) is a way to estimate exchange rates between currencies that account for purchasing power. GDP PPP controls for the different costs of living and price levels enabling a more accurate depiction of the different countries level of production. For the analyses in this study, we used the log of GDP (Ln GDP) per capita divided by 1,000. For further information on classification of national income, see **Supplementary Material 2**.

Satisfaction

Personal life satisfaction was used as an indicator of well-being. Personal life satisfaction was used as the dependent variable, except for research question RQ1 and RQ7, where also social satisfaction and political satisfaction were used as dependent variables.

At the *micro level,* we used personal life satisfaction. Personal life satisfaction was measured by the following item "All things considered, how satisfied are you with your life as a whole nowadays?" Responses were given on an 11-point scale ranging from 0 to 10, 0 = "extremely dissatisfied," 10 = "extremely satisfied" (21).

At the *mezzo level,* we used social satisfaction. Social satisfaction was measured by a variable constructed as the average of the responses to four questions: (1) "Do you feel close to the people in local area?", with response alternatives ranging from 1 = "Disagree strongly" to 5 = "Agree strongly"; (2) "Do you feel people treat you with respect?"; (3) "Do you feel people in local area help one another?", both of the last questions ranging in response from 0 = "Not at all" to 6 = "A great deal"; and (4) "Do you feel safe walking alone in local area after dark," with response alternatives ranging from 1 = "Very unsafe" to 4 = "Very safe," (Cronbach's alpha = 0.59). These questions cover the areas of belonging, social support and respect, as well as safety in the local area. Questions 1 and 4 were extended to conform to the range of questions 2 and 3 (22). The final variable represented the average of the four questions.

At the *macro level,* we used political satisfaction. Political satisfaction was measured by a variable constructed as the average of the responses to five questions: (1) "How satisfied are you with the present state of the economy in your country?"; (2) "How satisfied are you with the national government?"; (3) How satisfied are you with the way democracy works in your country?"; (4) "How satisfied are you with the state of education in the country nowadays?"; and (5) "How satisfied are you with

the state of health services in the country nowadays?," all with responses given on an 11-point scale ranging from 0 to 10, 0 = "Extremely dissatisfied," 10 = "Extremely satisfied" (Cronbach's alpha = 0.83). The variable political satisfaction was the sum of the answers to the five questions and ranged from 0 to 50 (22).

Trust

We used three variables of trust, both to measure direct effects of trust and to explore trust as potential moderators. These variables were developed for ESS, and have been in use since 2006 (23).

At the *micro level,* we used personal trust. Personal trust was measured by the following item: "In general I feel very positive about myself." Responses were given on a 5-point scale ranging from "Agree strongly" to "Disagree strongly" (24). The variable was recoded inversely.

At the *mezzo level,* we used social trust. Social trust was measured by the following item "Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?" Responses were given on an 11-point scale ranging from 0 to 10 (0 being "You can't be too careful." and 10 being "Most people can be trusted") (23, 25). This measure of trust has been observed to be stable and its validity confirmed (26).

At the *macro level,* we used political trust. Political trust was measured by the five following items: "How much do you personally trust the country's parliament?"; "How much do you personally trust the police?"; "How much do you personally trust the legal system?"; "How much do you personally trust the politicians?"; and "How much do you personally trust the political parties?." Responses to each were given on an 11-point scale ranging from 0 to 10 (0 being "you do not trust an institution at all" and 10 being "you have complete trust") (23, 25). The answers were added, yielding a parameter with a range of 0–50.

Confounders

The demographic variables adjusted for, in all the analyses, were year of investigation, gender, age and age^2, number of people living regularly as members of household, marital status (dummy variable), education, occupation, being permanently sick or disabled, being unemployed, and mental health. Being permanently sick or disabled and being unemployed were two alternatives in a question concerning main activity last 7 days (dummy variable). Mental health was a combination of two questions concerning feeling depressed or anxious. The two variables were recoded to either being most of the time or all of the time depressed or anxious, as opposed to less than that. The two variables were then combined so that the individual had at least one of the two conditions. Age is well-documented to have a curvilinear relationship, and therefore it is highly recommended to use the squared function (3).

Stratification of Countries

Each of the 19 countries was ranked according to its level of social and political trust separately. The resulting rankings were added together, and a new ranking performed of the combined value. The countries were then divided equally into three groups, Group 1, exhibiting the highest trust levels, included the Nordic countries of Denmark, Finland, Norway, and Sweden in addition to Switzerland and Netherlands; Group 2, exhibiting a medium-trust level, included United Kingdom, Belgium, Germany, Ireland, France, and Spain; and finally Group 3, exhibiting the lowest trust levels, included Slovenia, Cyprus, Slovakia, Russia, Portugal, Poland, and Bulgaria (See **Supplementary Material 3**, **Table 3**).

Statistical Analysis

The analyses were conducted using the Statistical Package of the Social Sciences (SPSS), version 25.0. All data were weighted in accordance with the ESS guidelines before conducting the analyses (27).

The primary method of analysis was the multilevel analysis. In SPSS this is done with the module Linear Mixed models (28). The data were weighted in these analyses using the post-stratification weight that includes a design weight. A three-level approach was used as the main method of analysis. The levels chosen were (1) the unit of measure is the individual; (2) community, which for practical purposes was defined using two variables, (a) within country region and (b) social class; and (3) country. The outcome variables were personal life satisfaction (all research questions, RQ), social satisfaction (RQ 1 and 7), and political satisfaction (RQ 1 and 7), representing the different hierarchical levels. For each of these layers, investigation year, gender, age and age^2, number of people in household, marital status, education, occupation, being permanently sick or disabled, being unemployed, and mental health, were entered as covariates. Additionally, separate economic indicators for each layer were used, personal income, community income and national income. Finally, trust variables for each layer were used (i.e., personal trust, social trust, and political trust). Unstandardized beta-coefficients with standard errors, are reported.

Moderation analyses were performed both by introducing an interaction in the multilevel analyses, and by using Andrew F. Hayes' PROCESS tool for SPSS. The latter was unfortunately unable to incorporate multilevel analyses; however, the analyses were performed on the country groups based on overall trust that were slightly more homogenous.

Analyses of relative effect was done by multiplying unstandardized coefficients with the population mean of the parameters used. The pie charts in **Figure 2** were constructed by multiplying the calculated beta with the weighted average for the entire population for each variable. Missing data were excluded listwise in the regression analysis. For number of missing values see **Table 1** in **Supplementary Material 3**. Model fit was evaluated by significant R^2 in the multiple linear regression.

Ethics

The data are available without restrictions, for not-for-profit purposes.

In accordance with the ESS ERIC Statutes (Article 23.3), the ESS ERIC subscribes to the Declaration on Professional Ethics of the International Statistical Institute. The Research Ethics Committee reviews applications for studies for which the ESS ERIC is directly responsible, that is, which it directly contracts.

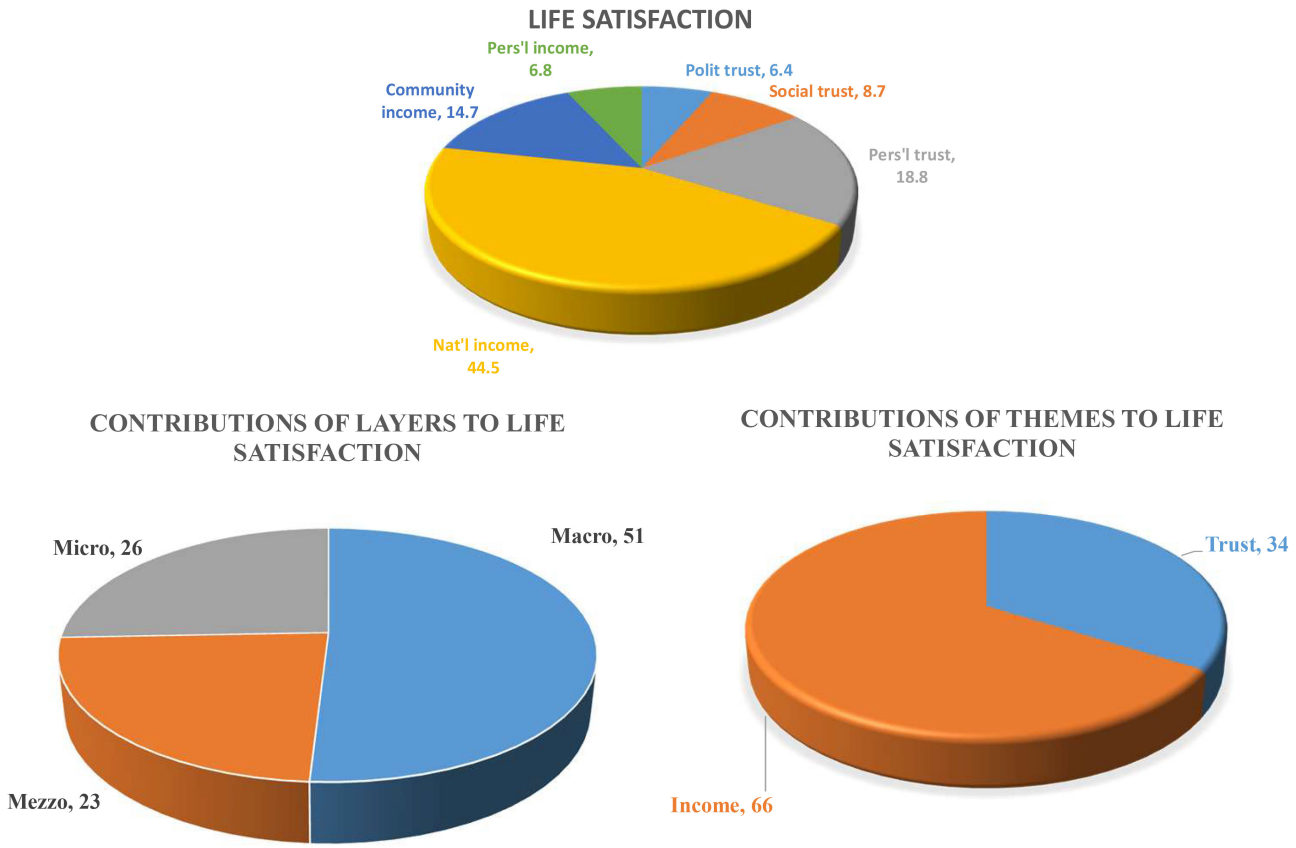

FIGURE 2 | Overview of the relative importance of personal, community, and national income; personal, social and political trust for life satisfaction, by layer and by theme. Results of multilevel analysis. Levels: individual, community, and country. Percent. Valid *N* (listwise) = 44,833. Micro level = Individual. Mezzo level = Community. Macro level = Country. Indicators of income: Micro level = Personal income. Mezzo level = Community income. Macro level = National income. Indicators of satisfaction: Micro level: Personal life satisfaction (PLS). Mezzo level = Social satisfaction. Macro level = Political satisfaction.

RESULTS

Descriptives

Satisfaction varied greatly between countries, with personal life satisfaction varying from 4.6 in Bulgaria to 8.5 in Denmark; social satisfaction varying from 4.1 in Russia to 4.9 in Norway; and finally political satisfaction varying from 2.9 in Bulgaria to 6.8 in Finland (**Table 1**).

Also income varied between countries, with personal income varying from 3,790 in Bulgaria to 72,533 in Switzerland; community income varying from 11,543 in Bulgaria to 172,606 euros in Netherlands; and finally national income varying from 14,631 in Bulgaria to 59,622 euros in Norway.

Trust varied too, but with personal trust being fairly uniform varying from 3.5 in France to 4.1 in Denmark; social trust varying from 10.8 in Bulgaria to 20.3 in Denmark; and finally, political trust varying from 11.4 in Bulgaria to 32.3 in Denmark.

Intercorrelations between the variables are presented in **Supplementary Material 3**, **Table 4**.

Layers and Themes

Answers to **RQ1** ("Within layers effects") are shown in **Table 2**. There was a significant and positive contribution of all three income variables on their respective satisfaction parameter within each layer (model 1). For personal life satisfaction, the association with personal income was strengthened upon addition of the personal trust parameter. For social satisfaction, upon the addition of social trust, the positive association between social satisfaction and community income was changed to a significant negative association. For political satisfaction, the association was substantially weakened, but still significant (model 2).

Answers to **RQ2** ("Total holistic effect") are shown in **Table 3**. Personal life satisfaction was significantly and positively associated with all three income variables and all three trust variables. Model 3 is thus an expression of the significant relationship between the macro, mezzo, and micro layers.

Answers to **RQ3** ("Effect size of holistic approach") are shown in **Figure 2**. **Figure 2** shows these results by converting the betas to relative effects using the weighted mean values for all of Europe. On a variable by variable basis, national income (44%) had the largest relative effect on variation in personal life satisfaction, followed by personal trust (19%), community income (15%), while political trust (6%), social trust (9%), and personal income (7%) had the smallest effect.

Answers to **RQ4** ("The layers effects") are also shown in **Table 3** and **Figure 2**. The country layer contributed most to the variation in personal life satisfaction (51%), followed by the

TABLE 1 | Weighted means/ranges of the primary variables of interest in the study by country.

	PLS	Social satisfaction	Political satisfaction	Personal income euros	Community income euros Range	National income euros	Personal trust	Social trust	Political trust	Regional unit NUTS
BE	7.4	4.4	6.0	33,662	77,041	39,204	3.7	15.3	24.4	2
CH	8.1	4.7	6.7	72,533	162,918	51,250	4.0	17.7	29.4	2
DE	7.2	4.5	5.0	33,582	85,784	39,111	4.1	15.7	24.2	1
DK	8.5	4.8	6.7	48,834	98,309	42,111	4.0	20.3	32.3	2
ES	7.2	4.6	4.5	24,033	171,591	31,825	4.0	15.2	19.9	2
FI	8.0	4.5	6.8	37,547	76,446	377,835	3.9	19.3	31.0	3
FR	6.5	4.4	4.7	33,455	136,793	36,303	3.5	14.9	21.5	2
GB	7.3	4.2	5.0	31,198	132,857	35,784	3.8	16.8	22.9	1
IE	7.1	4.6	4.9	40,124	113,720	46,923	4.0	17.0	22.4	3
NL	7.8	4.5	5.9	38,916	172,606	44,192	3.8	17.7	27.9	2
NO	7.9	4.9	6.5	67,575	109,959	59,622	3.7	19.7	29.2	2
PL	6.9	4.4	4.1	9,103	79,650	19,708	4.0	12.6	15.9	2
PT	5.8	4.6	3.6	14,049	97,684	26,471	4.0	12.6	16.5	2
SE	7.9	4.7	6.0	42,999	86,079	41,201	4.0	18.8	28.0	3
SI	7.0	4.6	4.4	17,909	59,609	27,190	4.0	14.0	18.0	3
CY	7.2	4.5	5.1	26,845	118,492	31,211	4.0	12.3	21.5	1
BG	4.6	4.3	2.9	3,790	11,543	14,631	4.0	10.8	11.4	3
RU	5.6	4.1	3.8	6,608	14,309	20,187	4.0	13.4	16.1	Other
SK	6.3	4.2	4.5	11,991	35,814	23,565	3.8	12.9	18.0	3

Weighted N = 73,307. Personal income is measured as yearly household income, community income per thousand is measured aggregated mean of household income for country, region and social class; national income is measured as Ln GDP (PPP) per capita per thousand. Ranges of income: Personal income, −4.45–200; Community income, −3–74; National income, 2.4–4.2; Ranges of trust: Personal trust, 1–9; Social trust, 0–30; Political trust, 0–50. Ranges of satisfaction: Life satisfaction, 0–10; Social satisfaction, 0.3–173.0; Political satisfaction, 0–10.

personal layer (26%) and finally the social layer (23%). The significant relationship between layers is further documented in model 4 (**Table 3**), by the significant negative interaction of the national income variable with the community income variable. This indicates that the effect of community income on personal life satisfaction is weaker in rich countries and stronger in poor countries. The interaction of the community and personal income variables was significant and negative. This indicates the effect of personal income on personal life satisfaction is weaker in rich communities than in poor communities. The interaction of the national with the personal income variables was not significant. This indicates no differences between either rich or poor countries in the relationship of personal income with personal life satisfaction. However, if the interaction of the national to the community income variable were removed (data not shown), the interaction between the national and personal income variables became significantly negative [−0,008 (0.001)***]. This indicates that the layer effect goes through the community layer. The pseudo R^2 values were high, both on an individual level (18% of variation explained) and especially between the levels (79% of variation explained). However, the pseudo R^2 values also indicate that the variables representing the sociodemographic parameters explain some of the variance within (10%) and especially between the levels (24%). The results of the full analysis (model 4), including confounders, is presented in **Supplementary Material 3**, **Table 5**.

Answers to **RQ5** ("The themes effects") are also shown in **Figure 2**. Thematically, income had the greatest relative association with personal life satisfaction (66%), followed by trust (34%).

Answers to **RQ6** ("Variation by countries") are shown in **Table 4** which shows results of the multilevel analysis of model 4 in **Table 3** sorted by countries grouped by their levels of trust. In Group 1 (High-trust levels) and Group 2 (Medium-trust levels) personal life satisfaction was not significantly associated with any of the three income variables. However, in Group 3 (Low-trust) all the three income variables were highly and significantly associated with personal life satisfaction. This indicates that income is an important factor in explaining variation in personal life satisfaction, but only in countries with low overall trust. The associations of the three trust parameters with personal life satisfaction were significant in all three groups, increasing in importance from high to low-trust for personal and political trust. This indicates that trust, be it personal or national, is of greater importance in countries with low overall trust. For social trust, however, this effect decreased from Groups 1 and 2 to Group 3, indicating the reverse. That is that as levels of overall trust in countries decreases, the importance of social trust also decreases. The interaction effects are an indication of the effects of the different layers. This indicates that within the group of countries with low overall trust, the relationships between personal and community income and personal life satisfaction are stronger in the wealthier countries than in

TABLE 2 | Results [Beta (SE)] of multilevel analysis of personal life satisfaction (micro level), social satisfaction (mezzo level) and political satisfaction (macro level).

Model	Level Measure of satisfaction (Mean/SE/Sig)	Micro Personal life satisfaction	Mezzo Social satisfaction	Macro Political satisfaction
1	National income			2.396 (0.050)***
	Community income		0.005 (0.000)***	
	Personal Income	0.010 (0.000)***		
	Individual level variance—within	3.560/3.206	0.777/0.720	2.110/2.051
	Three level variance—between	1.353/0.793	0.099/0.066	1.524/0.532
	Pseudo R^2 within	0.105	0.042	0.032
	Pseudo R^2 between	0.371	0.282	0.592
2	National income			0.885 (0.047)***
	Community income		−0.001 (0.001)*	
	Personal income	0.016 (0.002)***		
	Political trust			0.067 (0.006)***
	Social trust		0.042 (0.001)***	
	Personal trust	0.565 (0.017)***		
	Interaction trust × income	−0.002 (0.000)***	0.000 (0.000)***	0.010 (0.002)***
	Individual level variance–within	3.560/3.069	0.777/0.669	2.110/1.332
	Three level variance—between	1.353/0.825	0.099/0.059	1.524/0.226
	Pseudo R^2 within	0.145	0.113	0.372
	Pseudo R^2 between	0.365	0.131	0.844

All analyses include the respective income and trust variables. Weighted N = 40,219.
*Analyses controlled for gender, age, age^2, marital status, number in household, education, occupation, unemployed, being sick or disabled, mental health, and year. Personal income is measured as yearly household income, community income per thousand is measured aggregated mean of household income for country, region and social class; national income is measured as Ln GDP (PPP) per capita per thousand. Ranges of income: Personal income, −4.45–200; community income, −3–74; national income, 2.4–4.2; Ranges of trust: personal, 1–9; social, 0–30; political, 0–50. Ranges of satisfaction: Personal life satisfaction, 0–10; social satisfaction, 0.3–173.0; political satisfaction, 0–10. Levels in multilevel: individual, community (region and social class),and country. Random intercept. Restricted maximum likelihood. Residuals weighted for post-stratification weight. Significance: *p < 0.05; **p < 0.01; ***p < 0.001.*

TABLE 3 | Results [fixed effects (Beta (SE)sig) and pseudo R^2 (explained variance) for four models of multilevel analysis of personal life satisfaction against personal trust and personal income (model 1); against personal parameters and additionally social trust and community income (model 2); the preceding parameters and additionally against political trust and national income (model 3); and finally the preceding parameters with the three interactions of all three income parameters, indicating a layer effect (model 4).

Model	Null/demographic 0	Without trust	Individual 1	Community 2	Country 3	Interaction 4
Personal income		0.007 (0.000)***	0.009 (0.000)***	0.006 (0.000)***	0.006 (0.000)***	0.021 (0.006)**
Personal trust			0.507 (0.012)***	0.475 (0.011)***	0.478 (0.011)***	0.477 (0.011)***
Community income		−0.002 (0.001)^{NS}		0.010 (0.001)***	−0.003 (0.001)**	0.048 (0.010)***
Social trust				0.077 (0.002)***	0.055 (0.002)***	0.055 (0.002)***
National income		1.796 (0.055)***			1.293 (0.051)***	1.261 (0.066)***
Polit. trust					0.028 (0.001)***	0.028 (0.001)***
Int Ntlincome × Comincome						−0.012 (0.003)***
Int Ntlincome × Perinc						−0.003 (0.002)^{NS}
Int comincome × perinc						−0.000 (0.000)***
Individual level variance—within	3.560/3.322	3.203	3.070	2.969	2.916	2.911
Three level variance—between	1.353/1.028	0.399	0.825	0.496	0.306	0.284
Pseudo R^2 within	0/0.096	0.171	0.093	0.160	0.181	0.182
Pseudo R^2 between	0/0.240	0.768	0.274	0.562	0.774	0.790

*N = 44,833. Analyses controlled for gender, age, age^2, marital status, number in household, unemployed, education, occupation, being sick or disabled, mental health, and year. Personal income is measured as yearly household income, community income per thousand is measured aggregated mean of household income for country, region and social class; national income is measured as Ln GDP (PPP) per capita per thousand. Ranges of income: Personal income, −4.45–200; community income, −3–74; national income, 2.4–4.2; Ranges of trust: personal, 1–9; social, 0–30; political, 0–50. Range of personal life satisfaction, 0–10. Levels in multilevel: individual, community, and country. Random intercept. Restricted maximum likelihood. Residuals weighted for post-stratification weight. Significance: NS = Not Significant; *p < 0.05; **p < 0.01; ***p < 0.001.*

TABLE 4 | Results three models [fixed effects (Beta (SE)sig) and pseudo R^2 (explained variance)] for three models (based on country trust groups) of multilevel analysis of personal life satisfaction (at individual level) against personal, social and national income and against personal, social and political trust (model 4 in **Table 3**).

Model	Low trust 3	Medium trust 2	High trust 1
N	12,296	16,329	16,208
Personal income	0.104 (0.030)***	0.019 (0.018)^{NS}	−0.006 (0.009)^{NS}
Community income	0.133 (0.056)*	0.014 (0.039)^{NS}	0.011 (0.016)^{NS}
National income	1.824 (0.203)***	−0.739 (0.341)^{NS}	−0.267 (0.249)^{NS}
Personal trust	0.509 (0.026)***	0.482 (0.019)***	0.430 (0.015)***
Social trust	0.044 (0.004)***	0.062 (0.003)***	0.056 (0.003)***
Political trust	0.036 (0.002)***	0.028 (0.002)***	0.019 (0.001)***
Int National income × Community income	−0. 034 (0.017)*	0.003 (0.010)^{NS}	−0.002 (0.004)^{NS}
Int National income × Personal income	−0.026 (0.009)**	−0.001 (0.005)^{NS}	0.004 (0.002)^{NS}
Int Community income × Personal income	−0.000 (0.000)***	0.000 (0.000)***	−0.000 (0.000)***
Individual level variance—within	4.700/3.966	3.989/3.184	2.290/1.795
Three level variance—between	1.532/0.485	0.417/0.174	0.091/0.031
Pseudo R^2 within	0.181	0.160	0.093
Pseudo R^2 between	0.800	0.562	0.274

*Analyses controlled for gender, age, age^2, marital status, number in household, education, occupation, unemployment, being sick or disabled, mental health, and year. Personal income is measured as yearly household income, community income per thousand is measured as aggregated mean of household income for country, region, and social class; national income is measured as Ln GDP (PPP) per capita per thousand. Ranges of income: Personal income, −4.45–200; community income, −3–74; national income, 2.4–4.2; Ranges of trust: personal, 1–9; social, 0–30; political, 0–50. Range of personal life satisfaction, 0–10. Levels in multilevel: individual, community, and country. Random intercept. Restricted maximum likelihood. Residuals weighted for post-stratification weight. Significance: NS = Not Significant; *p < 0.05; **p < 0.01; ***p < 0.001.*

the poorer countries (negative interaction between national and personal income). Similarly, the layer effect of community to personal income indicates that in the richer communities, especially within those countries in the low overall trust group, higher personal income is associated with higher personal life satisfaction than it is in poor communities (negative interaction between community income and personal income).

The Buffer Hypothesis

Answers to **RQ7** ("The buffer hypothesis") are shown in **Tables 2**, **5**. **Table 2** presents the results of the analyses within layers. The interaction term between personal trust and personal income was significantly negative. This indicates that there is a moderator effect of personal trust, such that with increased personal trust the relationship between personal income and personal life satisfaction becomes less intensive, or vice versa. The interaction term between social trust and community income was significant and positive. This indicates that social trust moderates the relationship between community income and social satisfaction such that as social trust levels increase, the strength of the negative association with low community income increases. Finally, at the country layer the interaction term between political trust and national income was significant and positive. This indicates that political trust is a significant moderator of the relationship between national income and political satisfaction. Consequently, as political trust increases, the relationship between political satisfaction and national income becomes intensified. Or the opposite, if political trust is decreases, the association between national income and political satisfaction becomes weaker.

Table 5 presents the results of moderation analyses for the X–Y relationship personal income—personal life satisfaction, total and within each country group. It indicates that, for the entire population, all three forms for trust, personal, social, and national, are significant negative moderators. Although all three forms of trust are also significant negative moderators within each country group, the effect levels of the medium groups is approximately double the effect of the high trust group, whereas the effect size of the low trust group is a little <10 times the effect. Since the interaction is negative, the level of effect decreases with increasing levels of trust.

DISCUSSION AND CONCLUSION

Good public health and a high level of well-being and satisfaction with one's personal life are associated with personal, community and national income, as well as personal, social and political trust. However, how these measures relate to each other within and between countries and within and across different socio-structural layers of society, has been largely unknown. To answer these questions, we launched a three-layer nested, socio-structural model, similar to the Dahlgren and Whitehead's "rainbow model" (19).

The original "rainbow model" was launched primarily as a policy tool to handle social inequality of factors threatening, promoting and protecting health. The model stressed the importance of three socio-structural layers of society, the major structural environment, the material and social conditions, and the individual. Furthermore, it assumed that both direct and indirect effects of the layers were involved. In that respect it allowed policy makers a better perspective on the ramifications of suggested policy reforms.

Our model is constructed for research purposes. Similar to "the rainbow model," it consists of three layers (individual, community, and country). Each layer contains a corresponding measure of income (personal, community, and national), trust (personal, social and political), and satisfaction (personal, social, and national). Combined with multilevel data-analytic techniques we used this model to analyze data from the European Social Survey in 19 countries divided into three groups according to their levels of social and political trust.

A particular interest of this study was to test what we have called "the buffer hypothesis," that personal, social and political trust acts as a buffer against effects of personal income on

TABLE 5 | Results of moderation analysis (using Hayes PROCESS, which does not allow multilevel) for the moderator role of personal, social and political trust for X = personal income and Y = PLS, by country group as defined by overall trust and for the overall population.

Country group	Moderator		Effect on income parameter		
	Trust	Levels of moderator	*N*	Effect on X(SE)Sig	R^2/Interaction sig
Total	Personal	Low	55,253	0.023 (0.001)***	0.228/−0.003***
		Medium		0.020 (0.000)***	
		High		0.016 (0.001)***	
	Social	Low	54,891	0.022 (0.001)***	0.265/−0.001***
		Medium		0.016 (0.000)***	
		High		0.010 (0.000)***	
	Political	Low	52,909	0.023 (0.001)***	0.271/−0.0007***
		Medium		0.014 (0.000)***	
		High		0.007 (0.000)***	
High	Personal	Low	18,523	0.006 (0.001)***	0.186/−0.001***
		Medium		0.004 (0.000)***	
		High		0.004 (0.000)***	
	Social	Low	18,485	0.005 (0.001)***	0.186/−0.0003***
		Medium		0.004 (0.000)***	
		High		0.003 (0.000)***	
	Political	Low	17,992	0.005 (0.001)***	0.168/−0.0002***
		Medium		0.003 (0.000)***	
		High		0.003 (0.000)***	
Medium	Personal	Low	19,447	0.015 (0.001)***	0.184/−0.005***
		Medium		0.010 (0.001)***	
		High		0.010 (0.001)***	
	Social	Low	19,382	0.014 (0.001)***	0.195/−0.0009***
		Medium		0.010 (0.001)***	
		High		0.005 (0.001)***	
	Political	Low	18,843	0.013 (0.001)***	0.186/−0.0003***
		Medium		0.009 (0.001)***	
		High		0.006 (0.001)***	
Low	Personal	Low	17,283	0.045 (0.003)***	0.194/−0.006*
		Medium		0.040 (0.002)***	
		High		0.035 (0.003)***	
	Social	Low	17,024	0.050 (0.002)***	0.192/−0.001***
		Medium		0.040 (0.002)***	
		High		0.032 (0.002)***	
	Political	Low	16,074	0.055 (0.003)***	0.215/−0.0014***
		Medium		0.039 (0.002)***	
		High		0.023 (0.002)***	

*Analyses controlled for gender, age, age^2, marital status, number in household, unemployed, education, occupation, being sick or disabled, mental health, and year. Personal income is measured as yearly household income. Range of personal income −4.45 to 200. Ranges of trust: personal, 1–9; social, 0–30; political, 0–50. Range of personal life satisfaction, 0–10. Levels in multilevel: individual, community, and country. Random intercept. Restricted maximum likelihood. Residuals weighted for post-stratification weight. Significance: *p < 0.05; **p < 0.01; ***p < 0.001.*

satisfaction with one's personal life, and hence dampen effects of social inequality on the sense of well-being.

The buffer hypothesis. Our results give strong support to the buffer hypothesis. Both across European countries and within all three groups of countries, whether they are characterized by low, medium or high levels of social and political trust, the results are consistent with the buffer hypothesis. All three forms of trust dampen the effects of personal income on satisfaction with one's personal life. The more you trust yourself, your neighbors or the political and regulatory authorities in your country, the less important is personal income for how satisfied you are with your personal life. Or the other way around, the less you trust yourself, your neighbors or how the country is run, the more important is your personal income for how satisfied you are with your personal life. That is, trust, whether it is personal, social or political, compensates for the effect of low personal income on your personal satisfaction with life, and hence reduces differences in sense of well-being caused by

economic inequality, both between individuals, between local communities, and between countries.

At the community and national layer, however, we observed no such buffer effect. There we observed that high trust in the local social environment or the country's political or regulatory authorities occurred primarily in areas and countries where income was high. In these areas and countries, the relationships between community income and social satisfaction and between national income and political satisfaction were associated with social and political trust, respectively. In fact, all or nearly all of the effects of income on satisfaction either with the neighborhood or with how the country was run, were due to social and political trust, respectively. Although some literature exists indicating a moderator role of social trust on well-being (2–5), to our knowledge there are really no studies to compare these findings with.

Variation by country. Furthermore, we found that satisfaction with one's personal life is also directly associated with personal, social and political trust, irrespective of the level of social and political trust in the country. However, such trust is far more important to personal life satisfaction in countries with low levels of social and political trust as compared to countries with medium or high levels of such trust. The less social and political trust there is in a country, the more important is such trust for satisfaction with one's personal life, both directly, and indirectly by dampening the effect of income.

When all measures of income, trust, and satisfaction were controlled against each other, the results show that in countries with high- or medium levels of social and political trust, personal satisfaction with life is associated neither with national, community nor personal income. In countries with low levels of trust, however, personal satisfaction with life is strongly associated with all three measures of income.

The dampening effect of trust in countries with a medium level of trust, such as France, Germany, Spain, Ireland, and UK, is approximately the double of the effect of that in countries with a high level of trust, such as the Nordic countries. Even more, the dampening effect of overall trust in countries with a low level of trust, such as Bulgaria, Poland, Slovakia, Slovenia, Portugal, and Cyprus, is nearly 10 times as big as in the countries with high level of trust. Altogether, these findings too are consistent with "the buffer hypothesis."

A substantially higher significance of political trust to well-being in transition countries as compared to economically more developed countries in Europe, as well as a reduced association in the Nordic countries have been reported earlier (29, 30). One study reported that the relationship between personal life satisfaction and perceived quality of society was weaker in wealthier countries while the opposite was true in poorer countries (31). This could possibly be another indication of a relationship between political trust and satisfaction with one's personal life, which may indicate that the marginal effect of income is less in the high trust countries, where income is higher (29).

Total holistic effect. Another important finding of this study is that personal life satisfaction is positively linked to all three measures of income (personal, community, and national) as well as all three measures of trust (personal, social and political) when all measures of income, trust, and satisfaction are controlled against each other across countries. In our model, the strongest genuine effect on personal life satisfaction comes from national income (44%), then personal trust (19%), then community income (15%). While, may be surprising to some, political trust (6%), social trust (9%), and personal income (7%) have only minor effects.

As mentioned in the introduction, positive separate associations between several of the themes in this study, have been documented before. However, to our knowledge no study has compared the unique, relative contributions of these measures to personal life satisfaction in a multilevel model where all variables are controlled for each other.

Layer effects. When comparing effects of socio-structural layers, we found that there is a significant and substantial effect of the outer layers on the individual. Or the other way around, the individual layer was associated with the community layer which was associated with the country layer.

However, when we looked at which socio-structural layer (individual, community, and country) of the society that contributed the most to satisfaction with one's personal life, we found that more than half of the contribution comes from the country level, and that the social and personal level contribute with about a quarter each. This indicates that in the long run, what happens at the country level, may be far more important to one's personal life satisfaction than what happens at the community or individual level.

At the country level, the significance of national income, political trust and political satisfaction to personal life satisfaction, has been documented in many countries. Good government contributes to well-being by providing high quality technical delivery of goods, welfare, and democratic values. The relationship between good government and well-being involves both direct effects on individual happiness and indirect effects through educational and other reforms that help individuals realize factors important to well-being (25, 32), as well as civic nationalism (33, 34).

At the community level, it has been observed that political trust and political satisfaction may be more important than social trust in predicting personal life satisfaction (35). Several studies have also shown the significance of the community layer on well-being, emphasizing the importance of size of the geographic groups as for example the neighborhood and larger geographical units (5, 8, 13).

However, no one has ever demonstrated such layer effects within such a comprehensive multilevel context as in the current study.

Thematic effects. What then is most important, to satisfaction with one's personal life; is it income or trust? Our study shows that it's the money. In our model, two thirds of the effect come from income (personal, community, and national), while one third comes from trust (personal, social and political).

In general, our findings on the significance of personal, community, and national income to well-being are in concert with a large number of previous studies (5, 6, 8–11, 13, 14, 18,

20, 33, 36–38). Personal income, also called absolute income, seems to influence well-being in a positive direction (6, 10). Much of this discussion concerns the association of income to material consumption.

In the literature, the effect of community income may vary considerably according to the size of the "community," i.e., whether it is a state or province or a neighborhood. At the larger geographical level, e.g., provinces in Canada, a negative association with well-being has been found. However, at the smaller level, e.g., neighborhood, the majority of studies show positive associations between community income and well-being (5, 7, 8, 12, 39). At the larger level, community income is believed to reflect availability of public goods. At the smaller level, community income most likely reflects private consumption (5, 8, 9, 11, 14, 15, 39). As the size of the community level becomes smaller, also other factors such as trust and security seem to enter the picture (8).

Well-being also varies between countries. Differences between countries in national income, e.g., Gross Domestic Product (GDP), may explain some of this variation. One study reported a strong, positive relationship between national income and subjective well-being (33), but this has been contested by the Easterlin paradox (18, 40, 41). Additionally, it has been shown that national income explains more of the variation in well-being than does personal income (38).

Our study extends current knowledge by accounting for both personal, community, and national income; and personal, social and political trust in one and the same multilevel model. Even within this comprehensive model, among the income measures, it is the country's national income that counts by far the most for how satisfied one is with one's personal life. Personal income is of only minor significance.

Mechanisms. What then are the mechanisms behind these findings on trust? Being a cross-sectional study, this study cannot say anything about that. But, should we speculate based on both ours and previous findings, we would put our money on the plain thesis that trust simply makes life simpler, easier, more pleasant and friendly, brings people closer together, reduces bureaucracy, and facilitates the economy. This way, personal, social and political trust serve as a kind of lubricant for the individual, the community, and the society at large (1, 42).

In public health, social trust is strongly associated with individual happiness, altruistic attitudes, simpler collaboration between people, sense of control of one's life, and better chances in life (43–45). Economically, social trust is associated with less formalities, conflicts, legal processes, lower transaction costs in commerce and favorable conditions for investment (46–48). Politically, social and political trust seems to promote political engagement and democratic development, and to reduce criminality (49, 50). The other way around, it is also probable that economic growth generates social trust, which in turn generates further growth and the other way around (51).

If we should dare to continue speculating further and go as far as to assume that our findings reflect underlying causal mechanisms, these findings would imply that politicians, professionals and regulators who want a satisfied electorate or population, should invest in enhancing the level of trust, both personal, social and political, in the population they are serving. This is partly because there is a direct relationship between trust and satisfaction at all socio-structural layers of society, and partly because trust in its different forms plays a strong and consistent role as a buffer against the effects of personal income on our personal life satisfaction, and thus dampens the effects of economic inequality on the inhabitants' sense of well-being, and ultimately, their public mental health.

In conclusion, the results of the model presented in this study may provide the authorities with a framework for policies aimed at improving the general well-being of their population, necessary for the healthy functioning of society. Thus, we join with the Nordic Council of Ministers' "manual" on how states can act to increase social and political trust over the long term (1): "Act with openness and transparency, manage tax revenues with respect, and tackle all signs of corruption, however negligible they seem. Create a general welfare state that prevents underclasses developing in society. Support associations, not least financially. It is generally favorable if the state can have an open attitude to associations. Raise the level of education in the population. Because of the importance of retaining relative economic homogeneity in the population, it is probably particularly important to focus on those with, or at risk of, low and/or incomplete education. Counteract unemployment, particularly long-term unemployment. This particularly implies efficient integration of refugees and immigrants in the labor market"

STRENGTHS AND LIMITATIONS

One major strength of this study is that we used a combination of multilevel analysis and a three levels socio-structural model. This way, we could holistically assess how each layer-specific theme of income, trust and satisfaction is uniquely and independently associated with personal life satisfaction, controlling for all the others.

Another major strength of this study is the large sample size obtained by ESS and their use of methodological standards at all stages in the process. This makes the data ideal for comparative and cross-national analyses. The ESS team is working continuously to ensure high validity and reliability of the questionnaire and data collected. The use of strict randomized probability sampling provides a representative sample of the population, and the questionnaire used is well-tested and translated according to ESS protocols.

A third strength is that it includes comparable data across 19 European countries. This made it possible both to study the relative effect of the national layer as such and to come closer to theorizing about Europe as a whole. However, although more countries were included than in any previous similar study, the after all limited number of countries limits generalization to all Europe. Inclusion of data from other countries might have given somewhat different results, but were not available for the period examined in this paper.

This study has several other limitations too. The cross-sectional nature of the survey limits the possibility to draw causal conclusions from the findings. Consequently, we have only been able to speculate about the mechanisms that make trust such a powerful moderator of the relationship between income and satisfaction.

We used cumulative data from 2006 to 2012. This reflects particular periods in time. Joining of these time periods have been done under the assumption that the political climate and overall social discourse in these 2 years do not differ significantly.

Also because of the cross-sectional design, we were only able to examine short-term effects. In the long-term, the relationships we have revealed may be changed (52).

Furthermore, data were collected through self-report, and thus response bias might be present. However, several of the measures, such as those on trust and satisfaction, are truly subjective measures and can hardly be measured validly by other methods than by asking people.

The items used to measure satisfaction and trust at all three levels could have been more consistent across levels.

The ESS includes no standard measure on trust in oneself. However, although not 100% perfect, the concepts of self-confidence and self-esteem are logically very close. We therefore used self-esteem as a proxy for trust in oneself.

The data needed to determine measures on the community level were only available for two rounds (3 and 6), resulting in a lower samples size. At the community level, information about the neighborhood was not available. We therefore used regional level within countries to represent the local community. However, it has been shown that larger regional units represent more satisfactorily the community effect than estimates at the neighborhood level (53). To address this problem, we combined information on the regions within the country that the individuals lived in together with the social class. Although this was not ideal, it was considered an acceptable approximation.

Unfortunately, the income variable was changed in 2008 from 12 identical categories to 10 specific for each country. However, we controlled for this by imputing a personal income for each respondent using nation-specific information on the distribution by gender, age, and education.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: European Social Survey https://www.europeansocialsurvey.org/data/.

ETHICS STATEMENT

The data are available without restrictions, for not-for-profit purposes. In accordance with the ESS ERIC Statutes (Article 23.3), the ESS ERIC subscribes to the Declaration on Professional Ethics of the International Statistical Institute. The Research Ethics Committee reviews applications for studies for which the ESS ERIC is directly responsible, that is, which it directly contracts. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

JC-A and AH contributed to the conception, design of the work, contributed to further drafts, revisions of the article, and contributed to the final approval of the version to be published. JC-A was responsible for data analysis and interpretation as well as writing the first draft of the article. Both authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We would like to thank Norwegian Institute of Public Health for support and access to workspace and Senior Research Scientist, Rune Johansen at Norwegian Institute of Public Health, Division of Mental Health, for helpful advice regarding some of the statistical analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2021.629118/full#supplementary-material

REFERENCES

1. Andreasson U. *Trust–The Nordic Gold*. Copenhagen, DK: Nordic Council of Ministers (2017). p. 23.
2. Eid M, Diener E. Global judgments of subjective well-being: situational variability and long-term stability. *Soc Indic Res*. (2004) 65:245–77. doi: 10.1023/B:SOCI.0000003801.89195.bc
3. Dolan P, Peasgood T, White M. Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *J Econ Psychol*. (2008) 29:94–122. doi: 10.1016/j.joep.2007.09.001
4. Veenhoven R. The study of life satisfaction. In: Saris WE, Veenhoven R, Scherpenzeel AC, Bunting B, editors. *A Comparative Study of Satisfaction with Life in Europe*. Budapest: Eötvös University Press (1996). p. 11–48.
5. Barrington-Leigh CP, Helliwell JF. *Empathy and Emulation: Life Satisfaction and the Urban Geography of Comparison Groups*. Cambridge, MA: National Bureau of Economic Research (2008).
6. Biswas-Diener R. Material wealth and subjective well-being. In: Eid ML, Larsen, RJ, editors. *The Science Subjective Well-Being*. New York, NY: The Guilford Press (2008). p. 307–22. doi: 10.1007/978-90-481-2350-6_2
7. Blanchflower DG, Oswald AJ. Well-being over time in Britain and the USA. *J Public Econ*. (2004) 88:1359–86. doi: 10.1016/S0047-2727(02)00168-8
8. Brodeur A, Flèche S. Neighbors' income, public goods, and well-being. *Rev Income Wealth*. (2018) 65:217–38. doi: 10.1111/roiw.12367
9. Clark AE, Westergård-Nielsen N, Kristensen N. Economic satisfaction and income rank in small neighbourhoods. *J Eur Econ Assoc*. (2009) 7:519–27. doi: 10.1162/JEEA.2009.7.2-3.519

10. Diener E, Lucas RE, Oishi S. Advances and open questions in the science of subjective well-being. *Collabra Psychol*. (2018) 4:15. doi: 10.1525/collabra.115
11. Dittmann J, Goebel J. Your house, your car, your education: the socioeconomic situation of the neighborhood and its impact on life satisfaction in Germany. *Soc Indic Res*. (2010) 96:497–513. doi: 10.1007/s11205-009-9489-7
12. Graf E, Tillé Y. Imputation of income data with generalized calibration procedure and GB2 law: illustration with SILC data. In: *Proceedings 59th ISI World Statistics Congress*. Hong Kong (Session CPS203); The Hague: International Statistical Institute (2013).
13. Kingdon GG, Knight J. Community, comparisons and subjective well-being in a divided society. *J Econ Behav Organ*. (2007) 64:69–90. doi: 10.1016/j.jebo.2007.03.004
14. Knies G, Burgess S, Propper C. Keeping up with the schmidt's: an empirical test of relative deprivation theory in the neighbourhood context. *J Appl Soc Sci*. (2008) 128:75–108. doi: 10.3790/schm.128.1.75
15. Ma J, Dong G, Chen Y, Zhang W. Does satisfactory neighbourhood environment lead to a satisfying life? An investigation of the association between neighbourhood environment and life satisfaction in Beijing. *Cities*. (2018) 74:229–39. doi: 10.1016/j.cities.2017.12.008
16. Caporale GM, Georgellis Y, Tsitsianis N, Yin YP. Income and happiness across Europe: Do reference values matter? *J Econ Psychol*. (2009) 30:42–51. doi: 10.1016/j.joep.2008.06.004
17. Diener E, Tay L, Oishi S. Rising income and the subjective well-being of nations. *J Pers Soc Psychol*. (2013) 104:267. doi: 10.1037/a0030487
18. Easterlin RA. Will raising the incomes of all increase the happiness of all? *J Econ Behav Organ*. (1995) 27:35–47. doi: 10.1016/0167-2681(95)00003-B
19. Dahlgren G, Whitehead M. Policies and strategies to promote social equity in health. *Inst Future Stud*. (1991) 27:4–41.
20. Fitz BM, Lyon L, Driskell R. Why people like where they live: individual- and community-level contributors to community satisfaction. *Soc Indic Res*. (2016) 126:1209–24. doi: 10.1007/s11205-015-0922-9
21. Pavot W, Diener E, Colvin CR, Sandvik E. Further validation of the Satisfaction with Life Scale: evidence for the cross-method convergence of well-being measures. *J PersAssess*. (1991) 57:149–61. doi: 10.1207/s15327752jpa5701_17
22. Nes RB, Roysamb E, Tambs K, Harris JR, Reichborn-Kjennerud T. Subjective well-being: genetic and environmental contributions to stability and change. *Psychol Med*. (2006) 36:1033–42. doi: 10.1017/S0033291706007409
23. Huppert FA, Marks N, Clark A, Siegrist J, Stutzer A, Vittersø J, et al. Measuring well-being across Europe: description of the ESS well-being module and preliminary findings. *Soc Indic Res*. (2009) 91:301–15. doi: 10.1007/s11205-008-9346-0
24. Huppert FA, So TT. Flourishing across Europe: application of a new conceptual framework for defining well-being. *Soc Indic Res*. (2013) 110:837–61. doi: 10.1007/s11205-011-9966-7
25. Helliwell JF, Aknin LB, Shiplett H, Huang H, Wang S. Social *Capital and Prosocial Behaviour as Sources of Well-Being*. Cambridge, MA: National Bureau of Economic Research (2017).
26. Bergh A, Bjørnskov C. Trust, welfare states and income equality: sorting out the causality. *Eur J Political Econ*. (2014) 35:183–99. doi: 10.1016/j.ejpoleco.2014.06.002
27. Ganninger M. *Weighting in the ESS Cumulative Data Set*. London: European Social Survey (2007).
28. Field A. *Discovering Statistics Using IBM SPSS Statistics*. Thousand Oaks, CA: Sage (2013).
29. Glatz C, Eder A. Patterns of trust and subjective well-being across Europe: new insights from repeated cross-sectional analyses based on the European social survey 2002–2016. *Soc Indic Res*. (2019) 148:1–23. doi: 10.1007/s11205-019-02212-x
30. Rodríguez-Pose A, Von Berlepsch V. Social capital and individual happiness in Europe. *J Happiness Stud*. (2014) 15:357–86. doi: 10.1007/s10902-013-9426-y
31. Böhnke P. Does society matter? Life satisfaction in the enlarged Europe. *Soc Indic Res*. (2008) 87:189–210. doi: 10.1007/s11205-007-9169-4
32. Helliwell JF, Aknin LB. Expanding the social science of happiness. *Nat Hum Behav*. (2018) 2:1. doi: 10.1038/s41562-018-0308-5
33. Morrison M, Tay L, Diener E. Subjective well-being and national satisfaction: findings from a worldwide survey. *Psychol. Sci*. (2011) 22:166–71. doi: 10.1177/0956797610396224
34. Reeskens T, Wright M. Subjective well-being and national satisfaction: taking seriously the "proud of what?" question. *Psychol Sci*. (2011) 22:1460–2. doi: 10.1177/0956797611419673
35. Mota GL, Pereira PT. *Happiness, Economic Well-Being, Social Capital and the Quality of Institutions*. Working Papers Department of Economics 2008/40, ISEG–Lisbon School of Economics and Management, Department of Economics, Universidade de Lisboa (2008).
36. Easterlin RA. Income and happiness: towards a unified theory. *Econ J*. (2001) 111:465–84. doi: 10.1111/1468-0297.00646
37. Stevenson B, Wolfers J. *Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox*. Cambridge, MA: National Bureau of Economic Research (2008).
38. Tay L, Morrison M, Diener E. Living among the affluent: boon or bane? *Psychol Sci*. (2014) 25:1235–41. doi: 10.1177/0956797614525786
39. Kingdon G, Knight J. Unemployment in South Africa, 1995–2003: causes, problems and policies. *J Afr Econ*. (2007) 16:813–48. doi: 10.1093/jae/ejm016
40. Easterlin RA. Does economic growth improve the human lot? Some empirical evidence. *Nations Households Econ Growth*. (1974) 89:89–125. doi: 10.1016/B978-0-12-205050-3.50008-7
41. Easterlin RA, Angelescu L. *Happiness and Growth the World Over: Time Series Evidence on the Happiness-Income Paradox* (IZA Discussion Paper No. 4060). Available online at: SSRN: https://ssrncom/abstract=1369806~2009
42. Orth U, Erol RY, Luciano EC. Development of self-esteem from age 4 to 94 years: a meta-analysis of longitudinal studies. *Psychol Bull*. (2018) 144:1045. doi: 10.1037/bul0000161
43. Putnam RD, Leonardi R, Nanetti RY. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press (1994).
44. Rodríguez-Pose A. *Social Capital and Individual Happiness in Europe*. Bruges European Economic Research (BEER) Papers (2012).
45. Rothstein B. Corruption and social trust: why the fish rots from the head down. *Soc Res*. (2013) 80:1009–32. doi: 10.3390/ijerph13090924
46. De Groot HL, Linders GJ, Rietveld P, Subramanian U. The institutional determinants of bilateral trade patterns. *Kyklos*. (2004) 57:103–23. doi: 10.1111/j.0023-5962.2004.00245.x
47. Tabellini G. Culture and institutions: economic development in the regions of Europe. *J Eur Econ Assoc*. (2010) 8:677–16. doi: 10.1111/j.1542-4774.2010.tb00537.x
48. Zak PJ, Knack S. Trust and growth. *Econ J*. (2001) 111:295–321. doi: 10.1111/1468-0297.00609
49. Kennedy BP, Kawachi I, Prothrow-Stith D, Lochner K, Gupta V. Social capital, income inequality, and firearm violent crime. *Soc Sci Med*. (1998) 47:7–17. doi: 10.1016/S0277-9536(98)00097-5
50. Putnam RD. *Bowling Alone: The Collapse and Revival of American Community*. New York, NY: Simon and Schuster (2001).
51. Bjørnskov C. Economic growth. In: Svendsen GT, Svendsen GLH, editors. *Handbook of Social Capital*. London: Edward Elgar (2009). p. 3431–13.
52. Bartolini S, Sarracino F. Happy for how long? How social capital and economic growth relate to happiness over time. *Ecol Econom*. (2014) 108:242–56. doi: 10.1016/j.ecolecon.2014.10.004
53. Rickardsson J, Mellander C. *Absolute vs. Relative Income and Life Satisfaction*. Stockholm: Royal Institute of Technology, CESIS (2017).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

ORIGINAL RESEARCH
published: 06 April 2021
doi: 10.3389/fpubh.2021.654952

Market Exclusivity of the Originator Drugs in South Korea: A Retrospective Cohort Study

*Kyung-Bok Son**

School of Pharmacy, Sungkyunkwan University, Seoul, South Korea

Introduction: Generic entry is a well-known driver of competition and cost containment.

Objectives: We aim to measure the market exclusivity of originator drugs and to determine what influences the entry of generics in South Korea.

Methods: A list of originator drugs approved by the authority from 2000 to 2013 and their corresponding generics were paired. An event history model was applied for a statistical estimation for the duration until generic entry and to identify abbreviating or prolonging factors on the duration.

Results: A total of 2,061 pairs of originator and generics were identified. The market exclusivity for the originator drugs, including NDAs and non-NDAs, has not notably changed. However, competition among non-NDAs was less common than we expected. We found delayed time to entry of generics in the long run, particularly for non-NDAs in injection forms and biologics, and this finding is partially associated with market attractiveness.

Conclusion: The authority should address the delayed availability of certain types of generic drugs. The government could provide information on off-patent pharmaceuticals with no generic competition, designate their corresponding submissions as prioritized in the review process, and provide additional market exclusivity when entering the market via a long period of exclusivity.

Keywords: pharmaceutical expenditure, the first generic, market competition, market exclusivity, South Korea

OPEN ACCESS

Edited by:
Mirjana Ratko Jovanovic, University of Kragujevac, Serbia

Reviewed by:
Zafer Çaliskan, Hacettepe University, Turkey
Mariela Deliverska, Medical University Sofia, Bulgaria

***Correspondence:**
Kyung-Bok Son
sonkyungbok@gmail.com

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *19 January 2021*
Accepted: *03 March 2021*
Published: *06 April 2021*

Citation:
Son K-B (2021) Market Exclusivity of the Originator Drugs in South Korea: A Retrospective Cohort Study. Front. Public Health 9:654952. doi: 10.3389/fpubh.2021.654952

INTRODUCTION

Generic entry is a well-known driver of competition and cost containment in the pharmaceutical sector (1). Authorities approve generics that present pharmaceutical equivalence and bioequivalence on the basis of comparisons with the originator drug (2, 3). Because generic manufacturers do not have to conduct direct research, they develop generics at a lower price than that of the originator drug (4). Given the bioequivalence and lower price, economic theory suggests that generics are perfect substitutions of the originator drugs to a rational consumer in the market (5).

The entry of a generic drug will trigger competition and significantly alter the market structure (6). The entry of a generic drug will end the monopoly rent enjoyed by originator manufacturers and transform the monopoly market to an oligopoly where the originator drug and generic drugs compete (7). Thus, the duration of market exclusivity of the originator drug and the timing of the entry of generic drugs are interesting topics from the perspective of research and policy.

Previous health economics or health policy literature has emphasized the consequences of generic entrants. The literature has focused on the effect of a generic entrant from the perspectives of the price of the originator drug and generics (8–11), switching behavior by physicians (12, 13), market share of generics (14, 15), and expansion of the market within a substance category (16). Paradoxically, research has not concentrated on the market exclusivity of the originator drug and what influences the entry of generics.

However, the timing of the entry of generics and factors affecting the entry of generics are important factors in the management of pharmaceutical expenditures (17–19). This study aims to measure the market exclusivity of originator drugs and to determine what influences the entry of generics in the South Korean market. To this end, we investigated the market exclusivity of the originator drugs; determined what influences the entry of generics; and suggested policy options to rationalize pharmaceutical expenditure in South Korea.

MATERIALS AND METHODS

This study investigates the market exclusivity of the originator drugs and the timing of the first generic entry. We defined an originator drug as a pharmaceutical that was the first to be granted marketing authorization, whereas the first generic was defined as the second pharmaceutical that was granted marketing authorization after the originator and has the same active ingredients, strength, and route of administration as the originator. A list of originator drugs approved by the Ministry of Food and Drug Safety (MFDS) from 2000 to 2013 and their corresponding generics were paired with baseline information. South Korea provides 6 years of data exclusivity for new drugs (20). Thus, we excluded originator drugs approved after 2014.

Data Sources

We used two datasets provided by the Health Insurance Review and Assessment Services (HIRA) and the MFDS. First, the list of reimbursed medicines under the National Health Insurance Service (NHIS) was retrieved from the website of the HIRA. The list provides the characteristics of the pharmaceutical: generic and proprietary name of the pharmaceutical and its strength, manufacturer, and reimbursement price. Second, we extracted information on all pharmaceutical approved by the MFDS from 2000 to 2013. In particular, the Korea Pharmaceutical Information Service (KPIS) provides the similar characteristics of the pharmaceutical: generic and proprietary name of the pharmaceutical and its strength, anatomical therapeutic chemical (ATC) classification, substance type (including chemicals and biologics), manufacturer, and date of marketing approval. Using information on the generic name of the pharmaceutical and its strength, two datasets were merged.

Variables

We are mainly interested in market exclusivity of the originator drugs. Market exclusivity was measured as the year difference between the date of regulatory approval of the originator drug and that of the corresponding first generic.

We choose a set of variables to understand variations in market exclusivity of originator drugs: the characteristics of the originator drug, the manufacturer, and the market. First, we categorized originator drugs into New Drug Application (NDA) and non-NDA. An NDA refers to "a drug of new materials, a substance with a chemical structure or construction that is wholly new, or a combination drug containing new materials as effective ingredients" in South Korea (21). Second, we categorized the characteristics of the originator drug based on the pharmaceutical's ATC classification, route of administration, substance type, and year of marketing authorization. Based on the number and characteristics of identified pharmaceuticals, ATC classification was categorized into four groups: alimentary tract and metabolism/blood and blood forming organs/cardiovascular system (A/B/C), antiinfectives for systemic use/antineoplastic and immunomodulating agents (J/L), musculo-skeletal system/nervous system (M/N), and others. Pharmaceuticals belong to A/B/C group indicate medicines prescribed for chronic diseases, including hypertension and diabetes, while pharmaceuticals belong to J/L groups include cancer drugs. The year of marketing authorization was grouped into three periods to note time trends: Period I (2000–04), Period II (2005–09), and Period III (2010–13). Third, we grouped the manufacturers of the pharmaceuticals into domestic and overseas. The Ministry of Trade, Industry and Energy provides the dataset to identify the origin of the manufacturers. Finally, we identified the characteristics of the market based on reimbursement price. Price was categorized into four groups: low-price (<1,000 KRW, ∼0.86 USD), medium-price (between 1,000 and 10,000 KRW, approximately between 0.86 and 8.6 USD), high-price (between 10,000 and 100,000 KRW, approximately between 8.6 and 86 USD), and very-high-price (>100,000 KRW, ∼86 USD) medicines.

Statistical Analysis

We used two statistical analyses to understand market exclusivity of the originator. First, we used descriptive analyses to present the difference in market exclusivity between three periods, namely, period I (2000–04), period II (2005–09), and period III (2010–13). Second, we applied an event history model for a statistical estimation. The model, which is also known as a duration model, estimates the duration until an event (or generic entry) and identifies abbreviating or prolonging factors on the duration. As a univariate tool, we applied Kaplan-Meier survival estimates and conducted log-rank test to compare the generic entrance distributions of the samples. We applied the proportional hazard model as a multivariate tool. We presented two types of proportional hazard model: the simple and expanded. In the simple model, we included the characteristics of the originator drug, such as ATC classification, route of administration, substance type, and year of marketing authorization. We added characteristics of the manufacturer and market in the expanded model. Furthermore, we separated the pharmaceuticals into NDAs and non-NDAs in sub-group analyses. Data management and analysis were performed using R

TABLE 1 | Characteristics of the originator drugs.

Variables	All (n = 2,061)		Period I (n = 818)		Period II (n = 762)		Period III (n = 481)		P-value
The ATC classification									
J/L	351	17%	127	16%	133	17%	91	19%	0.0595
A/B/C	677	33%	250	31%	255	33%	172	36%	
M/N	363	18%	143	17%	140	18%	80	17%	
Others	670	33%	298	36%	234	31%	138	29%	
Route of administration									
Oral	1,035	50%	360	44%	401	53%	274	57%	<0.0001
Injection	326	34%	153	37%	120	32%	53	32%	
Others	700	16%	305	19%	241	16%	154	11%	
Type of substance									
Chemicals	1,911	7%	771	6%	693	9%	447	7%	0.0398
Biologics	150	93%	47	94%	69	91%	34	93%	
Manufacturers									
Domestic	1,181	57%	527	64%	398	52%	256	53%	<0.0001
Overseas	880	43%	291	36%	364	48%	225	47%	
New drug application									
Yes	1,733	16%	694	15%	628	8%	411	15%	0.2719
No	328	84%	124	85%	134	82%	70	85%	
Reimbursed price									
Low	814	39%	347	42%	290	38%	177	37%	<0.0001
Medium	551	27%	216	26%	227	30%	108	22%	
High	483	23%	194	24%	169	22%	120	25%	
Very high	213	10%	61	7%	76	10%	76	16%	

A, alimentary tract and metabolism; B, blood and blood forming organs; C, cardiovascular system; J, antiinfectives for systemic use; L, antineoplastic and immunomodulating agents; M, musculo-skeletal system; N, nervous system; Period I (2000–04), Period II (2005–09), Period III (2010–13).

statistical software (version 3.4.3). Statistical significance is noted by p-values < 0.05.

RESULTS

Subjects of the Study

Table 1 presents the characteristics of the subjects. During a 14-year period, a total of 2,061 pairs of originator and first generics were identified as the subjects. We categorized the subjects into three periods, namely, period I, period II, and period III, based on the approval year of the originator. Approximately, 40% (818 pairs), 37% (762 pairs), and 23% (481 pairs) of the subjects belong to periods I, II, and III, respectively. The proportion of oral forms in each period has increased, whereas that of injection and other forms has decreased. Similarly, the proportion of high- or very-high-priced pharmaceuticals has increased from 31% in period I to 41% in period III.

Market Exclusivity

Table 2 provides the market exclusivity of the originator drugs. However, our observations are right-censored, indicating that some of the originator drugs might experience generic competition over time. Thus, we separated the subject into the ongoing exclusivity group and the terminated exclusivity group and presented their market exclusivity. Terminated exclusivity indicates that generic drugs were granted marketing authorization, while ongoing exclusivity indicates that the originator drug constitutes a monopoly market without generic competition. Approximately, 45, 40, and 26% of originator drugs were grouped in the terminated group in periods I, II, and III, respectively. For pharmaceuticals belonging to the terminated group, the median of the exclusivity was 4.74, 4.32, and 2.00 years for periods I, II, and III, respectively. Similarly, the median values of the exclusivity for pharmaceuticals belonging to the ongoing group were 17.03, 12.29, and 7.34 years for periods I, II, and III, respectively.

We separated originator drugs into NDAs and non-NDAs and calculated their market exclusivity. Approximately, 59, 49, and 10% of the NDAs in periods I, II, and III were grouped in the terminated exclusivity group, respectively; the median of market exclusivity of NDAs in these periods was 8.19, 6.71, and 6.87 years, respectively. Note that South Korea provides 6 years of data exclusivity for NDAs. Thus, the median of market exclusivity for NDAs is longer than 6 years. Similarly, we calculated the market exclusivity for non-NDAs in periods I, II, and III. Approximately, 43, 39, and 29% of the non-NDAs in periods I, II, and III were grouped in the terminated exclusivity group, respectively. The median of market exclusivity of non-NDAs in these periods was 3.49, 2.51, and 1.91 years, respectively.

TABLE 2 | Market exclusivity of the originator drugs, including NDAs and non-NDAs.

	All (n = 2,061)			NDAs (n = 328)			Non-NDAs (1,733)		
	Period I	**Period II**	**Period III**	**Period I**	**Period II**	**Period III**	**Period I**	**Period II**	**Period III**
	818	762	481	124	134	70	694	628	411
Exclusivity status									
Terminate	368 45%	306 40%	127 26%	73 59%	63 49%	7 10%	295 43%	243 39%	120 29%
Ongoing	450 55%	456 60%	354 74%	51 41%	71 41%	63 90%	399 57%	385 61%	291 71%
Terminate									
Mean	5.59	4.37	2.66	9.18	6.62	6.51	4.71	3.78	2.44
Median	4.74	4.32	2.00	8.19	6.71	6.87	3.49	2.51	1.91
SD	4.77	3.58	2.44	3.45	2.89	0.73	4.64	3.51	2.32
Ongoing									
Mean	17.12	12.27	7.49	17.46	12.10	7.68	17.07	12.30	7.45
Median	17.07	12.29	7.34	17.64	12.02	7.88	17.03	12.33	7.34
SD	1.49	1.27	1.18	1.34	1.11	1.16	1.51	1.30	1.19

NDA, New Drug Application; Period I (2000–04), Period II (2005–09), Period III (2010–13).

Statistical Analysis

Kaplan-Meier Survival Analysis

Appendices 1–**3** provide a descriptive overview of the difference in durations, including all pharmaceuticals, NDAs, and non-NDAs, using Kaplan-Meier estimates. The estimates present the conditional probability that generic will enter the market after a given period. In particular, the various curves in **Appendix 1** indicate the probability that the originator drugs that will face generic competition after a specific year. The first graph in **Appendix 1** presents a curve without group comparison. The remaining graphs present curves with group comparison, including period, substance type, presence of manufacturers in South Korea, ATC classification, route of administration, reimbursed price, and designation of NDAs. Because South Korea provides 6 years of data exclusivity for new drugs, the last curve in **Appendix 1** for NDAs presents a plateau until 6 years after the marketing date of the originator drug. Additionally, the curve for NDAs went down steeply after 6 years, while the curve for non-NDAs went down smoothly during the study period. In the log-rank test, significant difference in generic entrance curves was observed in variables of substance type, ATC classification, route of administration, and reimbursed price.

Given the 6-year data exclusivity period granted to NDAs, we separated the subjects into NDAs and non-NDAs. Similar to **Appendices 1**–**3** present curves for the probability of the originator drugs that will face generic competition after a specific year. Curves with group comparisons based on substance type, including chemicals and biologics, were similar for NDAs and non-NDAs. However, other remaining curves with group comparisons were different. For instance, curves with group comparisons of route of administration were different. More specifically, the conditional probability that a generic entry will occur exhibited the order of oral, injection, and other after 15 years for NDAs. However, the same probability exhibited the order of oral, others, and injection for non-NDAs. Significant difference in generic entrance curves was observed in variables of substance type, ATC classification, route of administration, and reimbursed price in **Appendices 2**, **3**. Furthermore, the variable on period presented a significant difference in **Appendix 2**.

The Proportional Hazard Model

Table 3 provides results for the effects from the simple proportional hazard model. We fitted the simple model with four discrete factors: ATC classification, route of administration, substance type, and period based on the marketing approval year. Note that a positive coefficient indicates a short time to generic entry (timely generic competition), while a negative coefficient indicates a long time to generic entry (delayed generic competition). Thus, the time to generic entry for pharmaceuticals in injection form was delayed compared to that in oral form. Similarly, the time to generic entry for biologics was delayed compared to that for chemicals. However, the period variable was not significantly delayed or accelerated for the time to generic entry. Additionally, we separated the subjects into NDAs and non-NDAs and conducted the same analysis. Interestingly, we found that the time to generic entry for pharmaceuticals in injection form and biologics was delayed only for non-NDAs. Consistent with this result, the time to generic entry for pharmaceuticals approved in period III was delayed for NDAs.

Two variables of manufacturer and reimbursed price were added in the expanded model in **Table 4**. The expanded model produced results that were consistent with those of the simple model. Additionally, we found that the time to generic entry for medium-, high-, and very-high-price pharmaceuticals was delayed compared to that for low-price pharmaceuticals. We also found that the time to generic entry for pharmaceuticals produced by overseas manufacturers was accelerated compared to that for pharmaceuticals produced by domestic manufacturers for NDAs.

TABLE 3 | Results for the effects from the proportional hazard assumptions in the simple model.

	All (n = 2,061)			NDAs (n = 328)			Non-NDAs (1,733)		
	Coefficient	**Standard error**	***P*-value**	**Coefficient**	**Standard error**	***P*-value**	**Coefficient**	**Standard error**	***P*-value**
The ATC classification (Reference J/L)									
A/B/C	−0.2054	0.1075	0.0560	−0.3685	0.2363	0.1188	−0.2394	0.1256	0.0567
M/N	0.1776	0.1128	0.1153	0.5305	0.2350	0.0240	0.0606	0.1337	0.6501
Others	−0.3005	0.1190	0.0116	−0.1923	0.2565	0.4534	−0.3638	0.1379	0.0083
Route of administration (Reference Oral)									
Injection	−0.6440	0.0932	<0.0001	−0.2235	0.2189	0.3072	−0.7252	0.1037	<0.0001
Others	−0.1745	0.1180	0.1392	−1.2549	0.5417	0.0205	−0.1154	0.1239	0.3515
Type of substance (Reference Chemicals)									
Biologics	−0.5000	0.2027	0.0137	−0.5113	0.4278	0.2320	−0.5337	0.2316	0.0212
Period (Reference Period I)									
Period II	−0.0277	0.0794	0.7271	0.2057	0.1928	0.2861	−0.0575	0.0886	0.5160
Period III	−0.1938	0.1070	0.0702	−0.9205	0.4088	0.0244	−0.1550	0.1125	0.1682

NDA, New Drug Application; Period I (2000–04), Period II (2005–09), Period III (2010–13).

TABLE 4 | Results for the effects from the proportional hazard assumptions in the expanded model.

	All (n = 2,061)			NDAs (n = 328)			Non-NDAs (1,733)		
	Coefficient	**Standard error**	***P*-value**	**Coefficient**	**Standard error**	***P*-value**	**Coefficient**	**Standard error**	***P*-value**
The ATC classification (Reference J/L)									
A/B/C	−0.3440	0.1140	0.0025	−0.6725	0.2865	0.0189	−0.3305	0.1293	0.0105
M/N	0.04252	0.1178	0.7183	0.2646	0.2665	0.3206	−0.0157	0.1364	0.9080
Others	−0.3980	0.1229	0.0012	−0.4011	0.2748	0.1443	−0.4322	0.1406	0.0021
Route of administration (Reference Oral)									
Injection	−0.3729	0.1166	0.0013	0.2721	0.2840	0.3380	−0.4799	0.1329	0.0003
Others	−0.0113	0.1258	0.9280	−0.9886	0.5601	0.0775	−0.4785	0.1340	0.7618
Type of substance (Reference Chemicals)									
Biologics	−0.4690	0.2109	0.0261	−0.7183	0.4489	0.1095	−0.4799	0.2409	0.0464
Period (Reference Period I)									
Period II	0.0062	0.0802	0.9380	0.2439	0.1968	0.2150	−0.0202	0.0897	0.8218
Period III	−0.1305	0.1086	0.2296	−0.7988	0.4165	0.0551	−0.0958	0.1143	0.4015
Manufacturers (Reference Domestic)									
Overseas	0.0282	0.0767	0.7126	0.5818	0.2307	0.0116	−0.0396	0.0856	0.6431
Reimbursed price (Reference Low)									
Medium	−0.2435	0.0940	0.0103	−0.2739	0.2351	0.2440	−0.2438	0.1055	0.0209
High	−0.4867	0.1312	0.0002	−0.9269	0.3273	0.0046	−0.3899	0.1486	0.0086
Very high	−0.4792	0.1904	0.0118	−1.1419	0.4461	0.0104	−0.4181	0.2138	0.0505

NDA, New Drug Application; Period I (2000–04), Period II (2005–09), Period III (2010–13).

DISCUSSION

Timely entry of generic drugs is a key driver of competition and cost containment in the pharmaceutical sector. Thus, understanding the timing of the entry of generics and factors affecting their entrance are essential to rationalize pharmaceutical expenditures. To this end, we measured the market exclusivity of the originator drugs and identified what influences the entry of generics in South Korea.

Trends in Generic Entrants

According to our observations, the market exclusivity for the originator drugs has not notably changed. For instance, the period was not a significant factor in our proportional hazard model. This finding is consistent with previous literature. Son et al. (22) evaluated the effect of the patent linkage system on the patent challenge and market exclusivity of NDAs in South Korea. The authors calculated the effective market exclusivity for NDAs approved from 2007 to 2011 and concluded that the market exclusivity had not significantly changed after the introduction

of the patent linkage system. Additionally, we updated similar results for non-NDAs in the current study.

However, it is noteworthy to compare NDAs and non-NDAs from the perspectives of market competition. In the Kaplan-Meier estimates, the curve for NDAs decreased steeply after 6 years from the marketing approval of the originator drug. However, the curve for non-NDAs smoothly decreased from the marketing approval of the originator drug. Interestingly, the conditional probability that the generic entry will occur after 15 years of the approval of originator drugs was higher for NDAs than for non-NDAs. Similarly, we found that the proportion of the terminated group was higher for NDAs than for non-NDAs in periods I and II. These observations indicate that competition among non-NDAs in the long run was less common than we expected.

Factors Affecting Generic Entry

In the economic literature, it is well-documented that generic entry is driven by a variety of factors: manufacturer variables (23–25); pharmaceutical approval process in an authority (26–28); and markets attractiveness (29–34). First, manufacturer variables indicate the availability and cost of the raw materials, manufacturing processes and their corresponding cost as well as manufacturing and marketing experience with similar pharmaceutical products (23–25). Second, the pharmaceutical approval process in an authority includes quality of submissions that a manufacturer prepares and the review process that an authority provides (26–28). Finally, market attractiveness includes the size of the patient population being treated or projected profits in the market (31–33). More specifically, it was reported that the number of generic manufacturers is reduced for older orphan drugs (30), while the number of generic manufacturers entering a market is greater for pharmaceuticals with higher sales (29, 34). Furthermore, pricing and reimbursement policies of originator and/or generic drugs might influence in the timing of generic entrant (35).

Some of these factors are consistent in our study. We found that generic entries for biologics (reference chemicals) and pharmaceuticals in injection forms (reference oral forms) were delayed, indicating that manufacturer variables, including cost of the raw materials and manufacturing process, are critical factors in the entry of generics. However, interesting results were observed when we separated the subjects into NDAs and non-NDAs. Generic entries for biologics and pharmaceuticals in injection forms were delayed only for non-NDAs. The difference between NDAs and non-NDAs could be explained by their market attractiveness. During the 6-year period of data exclusivity given to NDAs, the market for the majority of NDAs continuously grew (36), indicating that projected profits for NDAs after 6 years of data exclusivity are higher than those of non-NDAs. In a similar vein, the market for NDAs manufactured by overseas manufacturers is larger than the market for NDAs manufactured by domestic manufacturers. Thus, the variables of overseas manufacturers (reference domestic manufacturers) in our proportional hazard model significantly accelerated generic entries in the market.

Policy Implications Regarding Delayed Availability of Generic Drugs

In this study, we found delayed time to entry of generic drugs in the long run, particularly for non-NDAs in injection forms and biologics, and this finding partially associated with market attractiveness. In this context, the role of the MFDS to address the delayed availability of generic drugs could be revisited (31, 37).

As previously discussed, pharmaceutical approval process in the authority matters. Thus, an expedited (or prioritized) review process for the first generic applications or products with less than three competitors might be established in South Korea to attract (additional) manufacturers enter a market (38). In the same vein, the MFDS could provide information on off-patent pharmaceuticals with "no (or inadequate) generic competition" to manufacturers and designate their corresponding submissions as prioritized in the review process (39). However, the effects of these measures in the entry of generics are not clear and require additional empirical evidence and contexts for the market (31). For instance, the main reason of fewer generics in the United States is the limited demand for additional generic drugs and their lower potential profits. It is noteworthy that generic drug prices in South Korea are higher than that of other high-income countries, indicating that lower profits for generic drugs might not be applicable in South Korea. Finally, the additional first generic exclusivity for submissions for pharmaceuticals with "no (or inadequate) generic competition" could be devised. South Korea introduced a 9-month first generic exclusivity, which is one of compartments in the patent linkage system, on March 15, 2015 (40). In the patent linkage system, the first generic applicant who has challenged a patent and obtained a favorable decision could be granted a 9-month period of market exclusivity (22). Similarly, the first generic entering the market with a long period of exclusivity of originator drugs could be granted market exclusivity, and additional market exclusivity might encourage generic entrants in the market.

Strengths and Limitations

We utilized two datasets provided by the MFDS and the HIRA and included all reimbursed pharmaceuticals in South Korea. Thus, the study findings might be generalizable to all types of pharmaceuticals. However, this study has limitations that are mainly attributed to data availability. We could not access information on the prescriptions and clinical effectiveness of the originator drug, indicating limited information on the characteristics of the originator drug and market. In particular, information on the number of prescriptions could be merged with the currently available dataset to fully understand the effect of generic entries from the perspectives of volume and/or value in the South Korean market. In a similar vein, we did not control for the effect of regulations in market exclusivity of originator drugs. Pricing and reimbursement policy could incentivize or dis-incentivize generic entrance.

CONCLUSIONS

In this study, we investigated the market exclusivity of originator drugs, and identified factors affecting the timely

availability of generic drugs in the South Korean market. Market exclusivity for originator drugs has not notably changed. However, it is noteworthy to compare market exclusivity of NDAs and non-NDAs from the perspective of competition. In particular, competition among non-NDAs in injection forms and biologics in the long run was less common than we expected. We suggested that the MFDS should address the delayed availability of certain generic drugs. The South Korean government could provide information on off-patent pharmaceuticals with no generic competition to manufacturers, designate their corresponding submissions as prioritized in the review process, and provide additional market exclusivity when entering the market via a long period of exclusivity.

DATA AVAILABILITY STATEMENT

The study analyzed data from publicly accessible datasets:https://www.hira.or.kr/bbsDummy.do?pgmid=HIRAA030014050000https://biz.kpis.or.kr/kpis_biz/index.jsp.

AUTHOR CONTRIBUTIONS

K-BS developed the concept the manuscript, undertook the analysis, and wrote the manuscript.

FUNDING

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019S1A5A8032445).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2021.654952/full#supplementary-material

Appendix 1 | Kaplan-Meier estimates for all medicines.

Appendix 2 | Kaplan-Meier estimates for NDAs.

Appendix 3 | Kaplan-Meier estimates for non-NDAs.

REFERENCES

1. Wouters OJ, Kanavos PG, McKee M. Comparing generic drug markets in Europe and the United States: prices, volumes, and spending. *Milbank Q*. (2017) 95:554–601. doi: 10.1111/1468-0009.12279
2. Davit BM, Nwakama PE, Buehler GJ, Conner DP, Haidar SH, Patel DT, et al. Comparing generic and innovator drugs: a review of 12 years of bioequivalence data from the United States Food and Drug Administration. *Ann Pharmacother*. (2009) 43:1583–97. doi: 10.1345/aph.1M141
3. Gassman AL, Nguyen CP, Joffe HV. FDA regulation of prescription drugs. *N Engl J Med*. (2017) 376:674–82. doi: 10.1056/NEJMra1602972
4. Dunne S, Shannon B, Dunne C, Cullen W. A review of the differences and similarities between generic drugs and their originator counterparts, including economic benefits associated with usage of generic medicines, using Ireland as a case study. *BMC Pharmacol Toxicol*. (2013) 14:1. doi: 10.1186/2050-6511-14-1
5. Fischer KE, Stargardt T. The diffusion of generics after patent expiry in Germany. *Eur J Health Econ*. (2016) 17:1027–40. doi: 10.1007/s10198-015-0744-3
6. Kanavos P, Costa-Font J, Seeley E. Competition in off-patent drug markets: issues, regulation and evidence. *Econ Policy*. (2008) 23:500–44. doi: 10.1111/j.1468-0327.2008.00207.x
7. Cole AL, Sanoff HK, Dusetzina SB. Possible insufficiency of generic price competition to contain prices for orally administered anticancer therapies. *JAMA Intern Med*. (2017) 177:1679–80. doi: 10.1001/jamainternmed.2017.2788
8. Alpern JD, Zhang L, Stauffer WM, Kesselheim AS. Trends in pricing and generic competition within the oral antibiotic drug market in the United States. *Clin Infect Dis*. (2017) 65:1848–52. doi: 10.1093/cid/cix634
9. Dave CV, Hartzema A, Kesselheim AS. Prices of generic drugs associated with numbers of manufacturers. *N Engl J Med*. (2017) 377:2597–8. doi: 10.1056/NEJMc1711899
10. Li DG, Joyce C, Mostaghimi A. Association between market competition and prices of generic topical dermatology drugs. *JAMA Dermatol*. (2018) 154:1441–6. doi: 10.1001/jamadermatol.2018.3798
11. Yousefi N, Mehralian G, Rasekh HR, Tayeba H. Pharmaceutical innovation and market share: evidence from a generic market. *Int J Pharm Healthc Mark*. (2016) 10:376–89. doi: 10.1108/IJPHM-06-2015-0028
12. Decollogny A, Eggli Y, Halfon P, Lufkin TM. Determinants of generic drug substitution in Switzerland. *BMC Health Serv Res*. (2011) 11:17. doi: 10.1186/1472-6963-11-17
13. Hellström J, Rudholm N. Uncertainty in the generic versus brand name prescription decision. *Empir Econ*. (2010) 38:503–21. doi: 10.1007/s00181-009-0277-7
14. Aronsson T, Bergman MA, Rudholm N. The impact of generic drug competition on brand name market shares–evidence from micro data. *Rev Indust Org*. (2001) 19:423–33. doi: 10.1023/A:1012504310953
15. Balasopoulos T, Charonis A, Athanasakis K, Kyriopoulos J, Pavi E. Why do generic drugs fail to achieve an adequate market share in Greece? Empirical findings and policy suggestions. *Health Policy*. (2017) 121:265–72. doi: 10.1016/j.healthpol.2016.12.011
16. Castanheira M, Ornaghi C, Siotis G. The unexpected consequences of generic entry. *J Health Econ*. (2019) 68:102243. doi: 10.1016/j.jhealeco.2019.102243
17. Belloni A, Morgan D, Paris V. *Pharmaceutical Expenditure and Policies*. Paris: OECD Publishing (2016).
18. Dylst P, Vulto A, Simoens S. Societal value of generic medicines beyond cost-saving through reduced prices. *Expert Rev Pharmacoecon Outcomes Res*. (2015) 15:701–11. doi: 10.1586/14737167.2015.1017565
19. Vogler S. The impact of pharmaceutical pricing and reimbursement policies on generics uptake: implementation of policy options on generics in 29 European countries–an overview. *Gener Biosimilars Initiat J*. (2012) 1:93–100. doi: 10.5639/gabij.2012.0102.020
20. Son K-B. Establishing healthy pharmaceutical regulations on statutory exclusivity: lessons from the experience in the European Union, Canada, South Korea, Australia, and the United States. *J Gener Med*. (2018) 14:167–74. doi: 10.1177/1741134318799385
21. *Pharmaceutical Affairs Act. Article 2 (Definitions)*.
22. Son K-B, Bae S, Lee T-J. Does the patent linkage system prolong effective market exclusivity? Recent evidence from the Korea-US free trade agreement in Korea. *Int J Health Serv*. (2019) 49:306–21. doi: 10.1177/0020731418822237
23. Gray A, Manasse HR Jr. Shortages of medicines: a complex global challenge. *Bull World Health Organ*. (2012) 90:158. doi: 10.2471/BLT.11.101303
24. Kweder S, Dill S. Drug shortages: the cycle of quantity and quality. *Clin Pharmacol Ther*. (2013) 93:245–51. doi: 10.1038/clpt.2012.235
25. Ventola CL. The drug shortage crisis in the United States: causes, impact, and management strategies. *Pharm Ther*. (2011) 36:740.
26. Centre for Innovation in Regulatory Science. *New Drug Approvals in Six Major Authorities 2008-2017*. London: Centre for Innovation in Regulatory Science (2018).

27. Son KB. Do free trade agreements matter to drug lag? recent evidence from Korea after the Korea–US free trade agreement. *Int J Health Ser*. (2020) 50:147–55. doi: 10.1177/0020731419898335
28. Wileman H, Mishra A. Drug lag and key regulatory barriers in the emerging markets. *Perspect Clin Res*. (2010) 1:51.
29. Grabowski HG, Kyle M, Mortimer R, Long G, Kirson N. Evolving brand-name and generic drug competition may warrant a revision of the Hatch-Waxman Act. *Health Affairs*. (2011) 30:2157–66. doi: 10.1377/hlthaff.2010.0270
30. Gupta R, Kesselheim AS, Downing N, Greene J, Ross JS. Generic drug approvals since the 1984 Hatch-Waxman Act. *JAMA Intern Med*. (2016) 176:1391–3. doi: 10.1001/jamainternmed.2016.3411
31. Gupta R, Shah ND, Ross JS. Generic drugs in the United States: policies to address pricing and competition. *Clin Pharmacol Ther*. (2019) 105:329–37. doi: 10.1002/cpt.1314
32. Moreno-Torres I, Puig-Junoy J, Borrell J-R. Generic entry into the regulated Spanish pharmaceutical market. *Rev Indus Org*. (2009) 34:373–88. doi: 10.1007/s11151-009-9214-3
33. Pauwels K, Huys I, Casteels M, Simoens S. Drug shortages in European countries: a trade-off between market attractiveness and cost containment? *BMC Health Serv Res*. (2014) 14:438. doi: 10.1186/1472-6963-14-438
34. Scott Morton FM. Entry decisions in the generic pharmaceutical industry. In: *NBER Work Paper (w6190)* (1997).
35. Zhang W, Sun H, Guh DP, Lynd LD, Hollis A, Grootendorst P, et al. The impact of tiered-pricing framework on generic entry in Canada. *Int J Health Policy Manage*. (2020). doi: 10.34172/ijhpm.2020.215. [Epub ahead of print].
36. Kwon HY. *Competitive Diffusion, Budgetary Impact and Pricing of New Drugs* (Doctorate). Seoul National University, Seoul. (2012).
37. Vokinger KN, Kesselheim AS, Avorn J, Sarpatwari A. Strategies that delay market entry of generic drugs. *JAMA Intern Med*. (2017) 177:1665–9. doi: 10.1001/jamainternmed.2017.4650
38. Toufanian M, Peters JR, Uhl K. Prioritization of generic drug review. *JAMA Intern Med*. (2017) 177:140–1. doi: 10.1001/jamainternmed.2016.7811
39. Ministry of Food and Drug Safety. *List of Off-Patent Drugs Without Competition*. Osong: Ministry of Food and Drug Safety (2020). Available online at: https://www.mfds.go.kr/brd/m_74/view.do?seq=43868&srchFr=&srchTo=&srchWord=&srchTp=&itm_seq_1=0&itm_seq_2=0&multi_itm_seq=0&company_cd=&company_nm=&page=1
40. Son K-B, Lopert R, Gleeson D, Lee T-J. Moderating the impact of patent linkage on access to medicines: lessons from variations in South Korea, Australia, Canada, and the United States. *Global Health*. (2018) 14:101. doi: 10.1186/s12992-018-0423-0

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

ORIGINAL RESEARCH
published: 20 April 2021
doi: 10.3389/fpubh.2021.645499

Donor Commitments and Disbursements for Sexual and Reproductive Health Aid in Kenya, Tanzania, Uganda and Zambia

Denis Kibira[1,2], Cornelia Asiimwe[3], Moses Muwonge[3], Hendrika A. van den Ham[1], Tim Reed[4], Hubert G. Leufkens[1] and Aukje K. Mantel-Teeuwisse[1]*

[1] Division of Pharmacoepidemiology and Clinical Pharmacology, Faculty of Science, Utrecht Centre for Pharmaceutical Policy and Regulation, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, Netherlands, [2] Coalition for Health Promotion and Social Development (HEPS-Uganda), Kampala, Uganda, [3] Samasha Medical Foundation, Kampala, Uganda, [4] Health Action International, Amsterdam, Netherlands

Background: Sexual and Reproductive Health and Rights (SRHR) investments are critical to people's well-being. However, despite the demonstrated returns on investments, underfunding of SRHR still persists. The objective of this study was to characterize donor commitments and disbursements to SRH aid in four sub-Saharan countries of Kenya, Tanzania, Uganda and Zambia and to compare trends in donor aids with SRH outcome and impact indicators for each of these countries.

Methods: The study is a secondary analysis of data from the Organization for Economic Co-operation and Development's Assistance creditor reporting system and SRH indicator data from the Global Health Observatory and country demographic health surveys for a 16-year period (2002–2017). We downloaded and compared commitments to disbursements of all donors for population policies, programs and reproductive health for the four African countries. SRH indicators were stratified into health facility level process/outcome indicators (modern contraceptive prevalence rate, unmet need for family planning, antenatal care coverage and skilled birth attendance) and health impact level indicators (maternal mortality ratio, newborn mortality rate, infant mortality rate and under five mortality rate).

Results: Donor commitments for SRH aid grew on average by 20% while disbursements grew by 21% annually between 2002 and 2017. The overall disbursement rate was 93%. Development Assistance Cooperation (DAC) countries donated the largest proportion (79%) of aid. Kenya took 33% of total aid, followed by Tanzania 26%, Uganda 23% and then Zambia (18%). There was improvement in all SRH outcome and impact indicators, but not enough to meet targets.

Conclusion: Donor aid to SRH grew over time and in the same period indicators improved, but improvement remained slow. Unpredictability and insufficiency of aid may be disruptive to recipient country planning. Donors and low- and middle-income countries should increase funding in order to meet global SRHR targets.

Keywords: sexual and reproductive health and rights, development assistance for health, official development aid, donor aid, low-and-middle-income countries

OPEN ACCESS

Edited by:
Olatunde Aremu, Birmingham City University, United Kingdom

Reviewed by:
Enver Envi Roshi, University of Medicine, Tirana, Albania
Mariela Deliverska, Medical University Sofia, Bulgaria

***Correspondence:**
Hendrika A. van den Ham
H.A.vandenHam@uu.nl

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *26 January 2021*
Accepted: *05 March 2021*
Published: *20 April 2021*

Citation:
Kibira D, Asiimwe C, Muwonge M, van den Ham HA, Reed T, Leufkens HG and Mantel-Teeuwisse AK (2021) Donor Commitments and Disbursements for Sexual and Reproductive Health Aid in Kenya, Tanzania, Uganda and Zambia. Front. Public Health 9:645499. doi: 10.3389/fpubh.2021.645499

BACKGROUND

Universal access to sexual and reproductive health and rights (SRHR) is necessary for the achievement of people's social, economic and environmental dimensions of sustainable development (1). The attainment of SRHR has not been realized as highlighted by an estimated annual death of more than 350,000 women and 5.6 million children worldwide from preventable complications related to pregnancy and childbirth (2).

Developing countries are affected disproportionately with 99% of the deaths from complications related to pregnancy and childbirth which could be mostly prevented by proper healthcare and services (3). Developing countries have the highest maternal, newborn and under-five mortality rates in the world (4). About 80 per cent of under-five deaths occur in two regions, that is sub-Saharan Africa and Southern Asia (2). **Table 1** shows SRH indicators for some of the most affected sub-Saharan countries namely Kenya, Tanzania, Uganda and Zambia.

SRHR is one of the inequities that the Millennium Development Goals (MDGs) sought to address (9). Building on the MDGs, the Sustainable Development Goals (SDGs), agreed by 193 world leaders in 2015, are a 17-point plan to end poverty, combat climate change and fight injustice and inequality. SDG 3 aims to ensure healthy lives and promote well-being for all at all ages (10).

SDG 3 sets targets by 2030 which include: reduce the global maternal mortality ratio to <70 per 100,000 live births; end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce neonatal mortality to as low as 12 per 1,000 live births and under-five mortality to as low as 25 per 1,000 live births; ensure universal access to sexual and reproductive health-care services, including family planning, information and education, and the integration of reproductive health into national strategies and programs.

To meet the above targets and improve health status, adequate health financing is essential (11). However, low- and middle-income countries (LMICs), in which resources are limited, also have inadequate health expenditure by governments (12). For example, in financial year 2009/10, the Kenyan government allocated about US$12.20 per person (equivalent to 5.4 % of the domestic budget) to health, and in Uganda the domestic budget was about US$11.20 per person equivalent to 7.4 % of the budget (13). This is against a backdrop of US$ 34 per person recommended by the WHO Commission on Macroeconomics and Health for governments to spend per year to provide a set of essential interventions (14). The limited spending on health by LMIC governments has meant that outside support is required (15). The magnitude of external funding on health as a percentage of total health expenditure has been significant, varying from 11 to 60% in over 28 sub-Saharan countries (16).

Abbreviations: SRHR, sexual and reproductive health and rights; DAC, development assistance cooperation; LMICs, low- and middle-income countries; SDGs, sustainable development goals; DAH, development assistance for health; ODA, official development assistance; OECD CRS, organization for economic co-operation and development's assistance creditor reporting system.

The United Nations (UN) Secretary-General's *Global Strategy for Women's, Children's and Adolescents' Health, 2016–2030* aims to catalyze the SDGs by mobilizing stakeholders including governments, donors/development partners, civil society, academia, healthcare providers and communities to scale up and prioritize high-impact interventions for strengthening health systems, integrating efforts across diseases and sectors as well as promoting human rights, gender equality and poverty reduction (9). In low-income countries, where much development assistance for health (DAH) is targeted, it made up 34.6% of total health spending in 2016 (17). DAH was estimated to total $37.6 billion in 2016, up 0.1% from 2015. However, after a decade of rapid growth from 2000 to 2010 (11.4% increase annually), DAH grew at only 1.8% annually between 2010 and 2016. SRHR is one of the priority areas financed by DAH from wealthier nations and international agencies (18).

In order to improve accountability for DAH, there has been increased efforts in resource tracking (19, 20). Studies have tracked trends and magnitude of donor funding to different areas of SRH that is reproductive, maternal, newborn, and child health (21), and sought to verify whether donor resources are better targeted to countries with the highest need (21). However, there is need to further explore what determines donor aid to recipient countries, priorities funded by donors within recipient countries, donor aid predictability (including whether donors disburse what they commit), how the donor aid is used by recipient countries, its effectiveness, and how donor aid influences funding of priorities by recipient countries (22).

This study sought to characterize donor predictability by examining their commitments and disbursements for SRH aid in four of the most affected countries in sub-Saharan Africa. The study therefore described the types of donors, the value and trends of their commitment and disbursement for SRH aid and matched the aid to changes in SRH indicators across the four countries in order to add to the body of knowledge on DAH accountability.

METHODS

Data Sources and Definitions

The study is a secondary analysis of data on donor aid commitments and disbursements for SRH from the Organization for Economic Co-operation and Development's Assistance creditor reporting system (OECD CRS) for a 16-year period (2002–2017).

The OECD CRS is a database to which donors of official development assistance (ODA), other official flows and private grants report their commitment and disbursement activities as described at http://www.oecd.org/dac/stats/methodology.html. The CRS is a publicly accessible web-based database on aid activities, developed and maintained by the Development Assistance Committee (DAC) of the OECD (18). OECD DAC commitments and disbursements are tracked at both the aggregate level and at the level of particular aid programmes (22).

ODA refers to grants or loans from members of the OECD DAC (a group of 30 nations including most of the West European and North American countries, the European Union, Australia,

TABLE 1 | Comparison of SRH indicators for Kenya, Tanzania, Uganda and Zambia (4–8).

Context Variable	Kenya	Tanzania	Uganda	Zambia	Sub-Saharan Africa	Global
2019 Population (in millions)	52.574	58.005	44.270	17.861	1,066,283	7,713, 468
Total Fertility Rate	3.9	5.2	5.4	4.7	4.9	2.5
Modern Contraceptive Prevalence Rate	53%	32%	38%	48%	26%	55%
Unmet need for Family Planning	18%	23%	28%	20%	24%	12%
Teenage Pregnancy Rate	15%	27%	25%	29%	–	44%
Maternal Mortality Ratio (per 100,000 live births)	362	556	336	252	546	216
Infant Mortality Rate (per 1,000 live births)	39	43	43	42	56	30.5
Under five mortality Rate (per 1,000 live births)	52	67	53	61	79	41

New Zeeland, Japan, and Korea), non-DAC bilateral donors (mostly Eastern European and Middle Eastern countries for example Croatia, Bulgaria, Turkey, Israel, United Arab Emirates, Kuwait), multilateral institutions (for example International Monetary Fund, regional development banks), global health initiatives (for example Global Fund to Fight Tuberculosis, AIDS and Malaria, Global Alliance for Vaccines and Immunization) and private philanthropists (for example Bill and Melinda Gates Foundation, Metlife Foundation, United Postcode Lotteries) with promotion of economic development and welfare as the main objective (22). In addition to financial flows, technical co-operation is included in aid (22).

Commitments refer to a firm obligation, expressed in writing and backed by the necessary funds, undertaken by an official donor to provide specified assistance to a recipient country or a multilateral organization (22). Recipients are defined by the CRS as all "developing countries" eligible to receive ODA. These include all "least developed countries" as defined by the United Nations and all LMICs defined by the World Bank, except any members of the G8, or members or agreed future members of the European Union (23).

Disbursements refer to the release of funds to or the purchase of goods or services for a recipient; by extension, the amount thus spent. Disbursements record the actual international transfer of financial resources, or of goods or services valued at the cost to the donor. In the case of activities carried out in donor countries, such as training, administration or public awareness programmes, disbursement is taken to have occurred when the funds have been transferred to the service provider or the recipient.

WHO and the United Nations Interagency Working Group set 17 population-based indicators to provide an overview of the global and national SRH situation (24). We divided these indicators into health facility level process/outcome indicators and health impact level indicators. Of the process/outcome indicators, we selected indicators that are routinely collected using country demographic health surveys conducted between 2002 and 2018. These include modern contraceptive prevalence rate (mCPR), unmet need for family planning (FP), antenatal care coverage (ANC) and percent of births attended by skilled health personnel. For impact we selected the mortality indicators, maternal mortality rate (MMR) and neonatal mortality rate (NMR), and added infant mortality rate (IMR) and under five mortality rate (U5MR).

Data Collection

We downloaded ODA data on commitments and disbursements for all donors for population policies, programs and reproductive health for four sub-Sahara African countries; Kenya, Tanzania, Uganda and Zambia from the OECD CRS for a 16-year period (2002 to 2017) on 22nd September, 2019.

OECD-CRS database has eight parameters: donors, sectors, ODA flow, channels, amount type, flow type, type of aid, and unit of aid in US million dollars. We selected data for all 110 donors reporting onto the system to the four recipient countries. Under sectors we selected code 130 with data on population policies/ programs and reproductive health and took into consideration all its subgroups which included population policy and administrative management, family planning, sexually transmitted diseases control and personnel development. We used total ODA and we considered all the different channels of fund flows including the public sector, non-government organizations (NGOs) and civil society, public-private partnerships, multilateral organizations, teaching institutions, research institutions or think tanks. On amount type, we chose constant prices in US dollars (USD) which is the amount that is adjusted for the effects of inflation. Under flow types, we considered both commitments and disbursements. We selected all types of aid including budget support, core contribution and pooled programmes, project-type interventions and technical assistance. The selected data was then exported into Microsoft Excel spreadsheet.

We collected data on the SRH indicators from the Global Health Observatory (GHO) and DHS surveys accessed from DHS StatCompiler on 22nd September, 2019. The Global Health Observatory derives this data from the United Nations Inter Agency Group (UN IAG) for Child Mortality Estimates: Levels and Trends in Child Mortality, Report 2017 (Available from: http://www.childmortality.org). Data on the MMR was derived from the World Bank Database available at http://data.worldbank.org/indicators/sh.sta.mmrt. We selected the four countries (Kenya, Tanzania, Uganda and Zambia) and filtered available data which was for the period (2002-2017) that was then exported into Microsoft Excel.

Data Analysis

We studied trends for donor commitments and disbursements of SRH aid for the period 2002–2017 to the four countries. We examined variations in: the commitments and disbursements over time by total value; the commitments and disbursements over time by different types of donors (we considered DAC countries, multilateral organizations, UN agencies and the World Bank which contributed 83.4% of funding to the four countries); and examined the commitments and disbursements over time to each of the four countries and by type of donors to each of the countries.

In a descriptive manner, we compared the time series data on donor aid disbursements to SRH indicators in each of the four countries.

RESULTS

Total Donor Commitments for SRH Aid to the Countries

Total donor commitments for SRH to the four countries (Kenya, Tanzania, Uganda and Zambia) grew annually by 20% on average between 2002 and 2017 from USD 319.14 million to 1,635.05 million. There was an increase in commitments between 2002 and 2008 but thereafter there were fluctuations. The total amount of commitments equalled USD 21,678 million over the 16-year period. Kenya received the largest donor commitments totalling USD 7,571.24 million (35%) over the sixteen-year period, followed by Tanzania at 24% amounting to USD 5,296.66 million, Uganda at 22% amounting to USD 4,837.67 million and then Zambia being the lowest at 18% amounting to USD 3,972.04 million. Despite the general growth in commitments, there were year on year fluctuations over the period with a general decline in 2010. **Figure 1** shows trends in donor commitments to the four countries.

DAC countries committed the largest proportion (82%) equivalent to USD 18,444.25 million over the sixteen-year period (2002-2017) followed by multilateral institutions, UN agencies and then the World Bank as shown in **Annex 1**.

DAC countries committed the highest amount (36% of their commitments) to a tune of USD 5,989.29 million to Kenya. Tanzania took the largest commitment of funds (USD 960.69 million, 30%) from multilateral donors. Uganda received the largest commitment of funds (USD 106.10 million, 30%) from UN agencies whereas the World Bank also committed most of its funds (USD 139.95 million, 46%) to Kenya. In contrast, Tanzania, Uganda and Zambia did not receive any commitments for SRH funds from the World Bank between 2005 and 2014. See table in **Annex 1** for details.

Total Donor Aid Disbursements to Countries

The total disbursements to the four countries over the 16-year period were USD 19,852.92 million. The overall disbursement rate over the sixteen-year period was 93%. Disbursements grew over time rising from USD 181.27 million in 2002 to 1,999.51 million in 2013, but thereafter reduced to 1,455.43 million in 2015

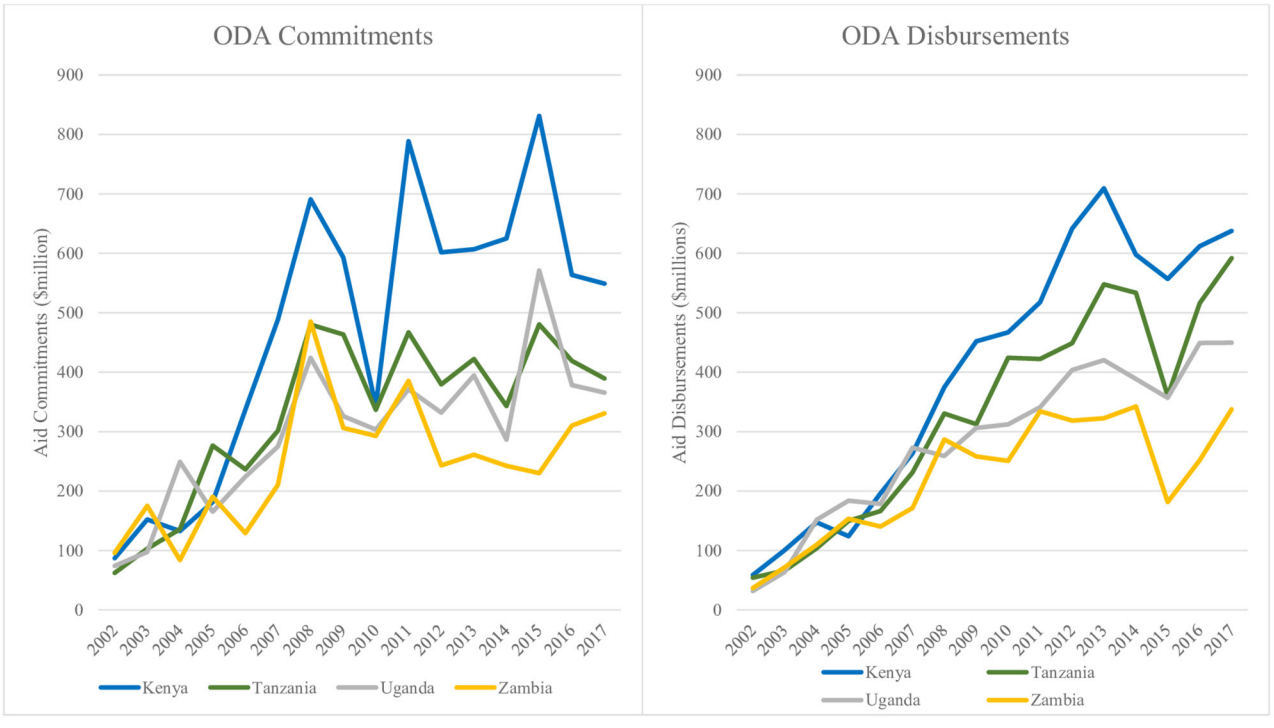

FIGURE 1 | Trends in donor commitments and disbursements for SRH aid to countries 2002–2017 (in USD millions).

Annex 1 | SRH aid donor commitments and disbursements by donor type (2002–2017), amounts are expressed in million US dollars.

Commitments																	
	2002	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**TOTAL**
DAC																	
Kenya	78.553	99.658	125.05	175.536	252.222	426.107	640.742	586.829	298.035	675.744	595.509	600.713	402.017	622.324	514.439	546.697	6640.175
Tanzania	54.354	54.638	94.037	155.153	160.298	217.03	319.881	339.207	328.858	383.507	302.409	368.666	310.246	425.458	414.982	385.601	4314.325
Uganda	68.137	45.426	128.998	159.22	217.973	267.499	338.028	315.62	294.022	350.071	276.549	386.085	278.672	382.17	374.85	362.284	4245.604
Zambia	63.155	120.596	81.331	125.684	122.639	184.706	260.846	246.388	256.685	227.943	202.837	254.779	236.784	227.983	302.879	328.914	3244.149
Total	**264.199**	**320.318**	**429.416**	**615.593**	**753.132**	**1095.342**	**1559.497**	**1488.044**	**1177.6**	**1637.265**	**1377.304**	**1610.243**	**1227.719**	**1657.935**	**1607.15**	**1623.496**	
Multilaterals																	
Kenya	8.374	52.359	7.539	5.894	84.669	62.300	50.019	6.263886	45.101221	112.72033	6.186066	5.990998	223.14406	208.90446	49.119027	2.476	931.063
Tanzania	7.858	48.412	41.935	121.148	76.179	84.234	160.242	124.30244	7.868091	83.56549	76.996327	53.665792	32.820869	54.93998	4.191553	3.810	982.173
Uganda	5.898	52.208	120.021	6.155	6.078	7.408	86.305	10.239727	9.384738	21.034499	55.47317	8.428447	7.872573	188.95425	3.232976	3.368	592.066
Zambia	32.810	54.664	2.495	64.837	6.389	25.629	224.243	59.610526	36.159785	157.35712	40.471814	6.370723	5.474929	2.0751	7.403347	1.898	727.891
Total	**54.941**	**207.645**	**171.993**	**198.035**	**173.317**	**179.572**	**520.810**	**200.416**	**98.513**	**374.677**	**179.127**	**74.455**	**269.312**	**454.873**	**63.946**	**11.554**	
UN																	
Kenya	8.374	4.35	4.1	5.895	5.038	4.4	6.458	6.264	6.173	6.119	6.186	5.991	6.486	2.49	3.301	2.476	84.101
Tanzania	6.074	6.56	4.006	5.97	6.389	7.015	6.877	7.332	7.868	7.464	6.643	7.152	7.511	4.305	4.192	3.811	99.169
Uganda	5.899	8.175	5.079	6.155	6.078	7.408	9.285	10.24	9.385	8.005	9.247	8.309	7.873	3.653	3.233	3.369	111.393
Zambia	3.992	2.758	2.495	4.793	6.39	6.391	7.46	6.445	6.394	5.454	6.607	6.371	5.475	2.075	2.822	1.898	77.82
Total	**24.339**	**21.843**	**15.680**	**22.813**	**23.895**	**25.214**	**30.080**	**30.281**	**29.820**	**27.042**	**28.683**	**27.823**	**27.345**	**12.523**	**13.548**	**11.554**	
World Bank																	
Kenya	..	..	..	..	..	57.900	..	..	38.927	..	..	..	..	..	45.817	..	142.646
Tanzania	1.784	33.951	11.171	..	..	..	..	..	..	..	..	..	..	50.634	..	..	97.542
Uganda	..	..	33.515	..	..	..	..	..	..	..	..	..	..	..	..	..	33.515
Zambia	28.818	..	..	..	..	..	..	..	..	..	..	..	..	..	4.581	..	33.400
Total	**30.603**	**33.952**	**44.687**	**0.000**	**0.000**	**57.900**	**0.000**	**0.000**	**38.928**	**0.000**	**0.000**	**0.000**	**0.000**	**50.635**	**50.400**	**0.000**	

(Continued)

Annex 1 | Continued

Disbursements

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	TOTAL
DAC																	
Kenya	41.864	73.485	103.113	108.543	188.013	231.201	334.214	428.222	405.305	478.938	562.696	625.659	533.152	445.266	523.018	543.830	5626.526
Tanzania	44.696	51.298	90.207	86.502	111.964	160.574	213.178	243.162	325.650	340.390	342.201	371.138	396.040	323.0263	346.625	415.252	3861.909
Uganda	20.978	49.379	109.961	130.106	170.654	226.643	246.705	291.195	278.517	317.901	343.652	378.290	346.411	264.3558	336.094	400.685	3911.535
Zambia	30.824	61.773	67.213	104.188	109.668	131.055	196.729	206.357	207.678	248.153	243.417	246.251	273.766	176.9099	227.847	326.241	2858.076
Total	**138.364**	**235.935**	**370.495**	**429.341**	**580.301**	**749.474**	**990.827**	**1168.937**	**1217.152**	**1385.384**	**1491.968**	**1621.340**	**1549.370**	**1209.558**	**1433.585**	**1686.01**	
Multilaterals																	
Kenya	16.768	27.301	44.336	15.280	8.251	32.406	40.386	23.681	61.785	38.921	79.041	83.738	64.511	111.624	89.1062	93.855	830.998
Tanzania	9.438	14.947	13.552	63.094	54.215	70.908	117.075	69.308	98.581	81.662	106.502	176.519	137.384	37.167	169.560	176.624	1,396.544
Uganda	10.796	14.570	41.874	53.788	7.637	46.177	12.242	14.847	33.591	23.167	60.116	41.810	42.307	92.46	112.938	48.926	657.262
Zambia	5.904	9.988	42.4543	49.238	30.526	40.504	90.173	51.818	43.251	86.356	75.061	76.099	68.535	4.619	23.948	11.437	709.917
Total	**42.907**	**66.808**	**142.218**	**181.402**	**100.630**	**189.998**	**259.878**	**159.655**	**237.211**	**230.108**	**320.722**	**378.168**	**312.739**	**245.881**	**395.554**	**330.843**	
UN																	
Kenya	8.374	4.350	4.100	5.894	5.038	4.399	6.457	6.691	6.578	6.599	7.143	6.285	6.486	7.470	7.781	5.145	98.798
Tanzania	6.074204	6.559978	4.006439	5.969597	6.388663	7.014642	6.876169	7.550227	7.926127	7.464311	6.642515	7.152426	7.510546	7.328398	6.930842	4.784	106.180
Uganda	5.898	8.175	5.079	6.155	6.078	7.408	9.285	10.239	9.384	8.004	9.246	8.309	7.872	8.276	6.896	6.126	122.437
Zambia	3.99	2.758	2.495	4.730	6.389	6.389	7.459	7.623	6.394	5.453	6.607	6.370	5.474	4.287	4.827	3.087	84.343
Total	**24.339**	**21.844**	**15.681**	**22.750**	**23.895**	**25.212**	**30.079**	**32.105**	**30.283**	**27.522**	**29.640**	**28.118**	**27.345**	**27.364**	**26.436**	**19.143**	
World Bank																	
Kenya	8.271	12.059	12.676	7.257	0.992	3.215	13.729	0.359	26.427	24.885	15.327	10.664	1.862	4.233	0	4.088	146.051
Tanzania	3.066	4.552	0.915	6.160	5.775	16.061	9.113	6.756	0.896	0	0	0	0	0	12.843	11.168	77.311
Uganda	3.455	4.163	5.796	37.420	1.332	-0.049	0	0	0	0	0	0	0	0	0	0	52.120
Zambia	1.771	3.741	10.341	9.726	5.761	15.227	1.721	0.0173	0	0	0	0	0	0	0	0.647	48.957
Total	**16.566**	**24.517**	**29.730**	**60.566**	**13.862**	**34.455**	**24.565**	**7.134**	**27.324**	**24.885**	**15.327**	**10.665**	**1.863**	**4.233**	**12.844**	**15.904**	

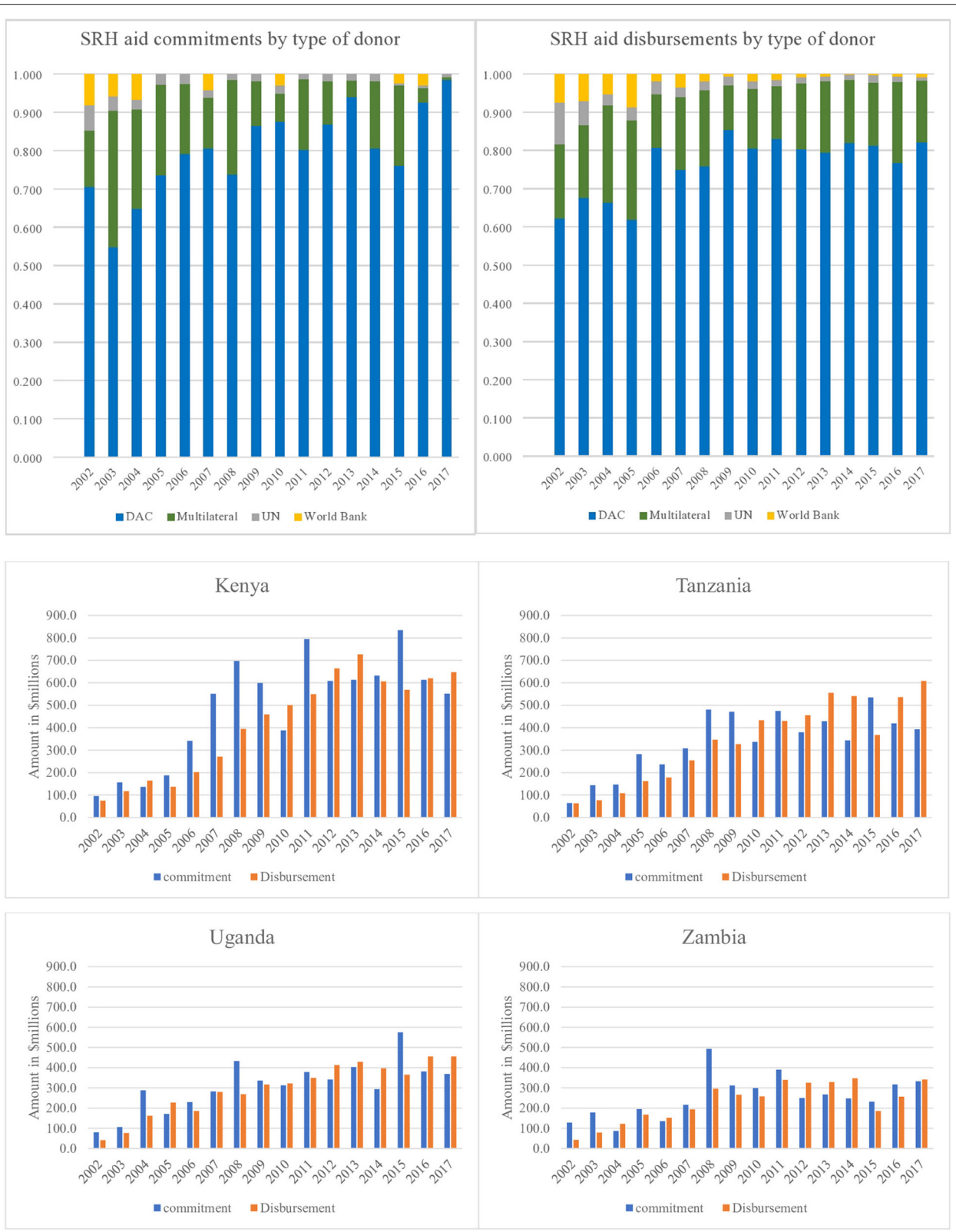

FIGURE 2 | Country specific comparisons in donor aid commitments and disbursements, 2002–2017 (USD millions).

and rising to 2,016.85 million in 2017 at an average annual growth rate of 21%. In contrast with the commitments, there was a steady increase in disbursements until 2013 and 2014 for Zambia and a drop in 2015 from where disbursements then started to rise slowly. Kenya received the largest donor disbursements totalling USD 6,457.52 million (33%) over the sixteen-year period, followed by Tanzania at 26% amounting to USD 5,258.61 million, Uganda at 23% amounting to USD 4,568.79 million and then Zambia being the lowest at 18% amounting to USD 3,567.99 million. Despite the general growth in disbursements, there was a general decline between 2013 and 2015 before picking up in 2016. Trends in donor aid disbursements to the four countries are shown in **Figure 1**.

The highest donor disbursement over the sixteen-year period (2002–2017) was from DAC countries comprising 79% of the total and rising from USD 138.36 million in 2002 to USD 1,686.01 million in 2017. Multilateral funders followed the DAC countries contributing 17% of disbursements. United Nation agencies and the World Bank contributed 2% each. Trends in donor aid disbursements for SRH by donor type in the four countries are shown in **Annex 1**.

Kenya was the biggest recipient from DAC countries getting 35% of funds worth USD 5,626.52 million over the period 2002–2017. Tanzania took the largest proportion (39%) of funds (USD 1,396.54 million) from multilateral donors; Uganda received the largest proportion 30% of funds worth USD 122.43 million from UN agencies whereas the World Bank also provided most (45%) of its funds (USD 146.05 million) to Kenya. This is detailed in table in **Annex 1**.

Country Specific Donor aid Commitments and Disbursements

Figure 2 highlights the trends i1n the donor aid commitments and disbursements to each of the four countries. The trends show that the commitments and disbursements grew mostly in line overtime but peaks in commitments were not reflected in the disbursements. While Kenya received most aid, it also had most fluctuations between amounts committed and disbursed. For Kenya 86% of commitments were disbursed compared to 100% of commitments for Tanzania, 95% for Uganda and 91% for Zambia over the total study period.

Comparison of Trends in Donor Disbursements to SRH Indicators

Figure 3 shows SRH health facility level process/outcome indicators which showed improvement over the 16-year period across the four countries. ANC improved and remained very high, mCPR increased with most pronounced increase observed in Kenya, and unmet need for FP reduced mostly in Kenya. Tanzania was much slower in improvement in the indicators. Skilled birth attendance increase was most pronounced in Uganda and Zambia. Improvement in SRH impact indicators (**Figure 4**) were most pronounced for Kenya. U5MR and IMR dropped markedly across the four countries but reduction in NMR was slow. MMR dropped across the four countries with Kenya having the most pronounced improvement. The rise in donor aid disbursements between 2005 and 2017 aligned with improved SRH outcome and impact indicators but not enough to meet SDG targets. Impact indicators reduced majorly between 2002 and 2005 and slowed thereafter especially for under five mortality.

DISCUSSION

Between 2002 and 2017, donor commitments for SRH aid to the four sub-Saharan countries of Kenya, Tanzania, Uganda and Zambia grew annually by 20% on average while disbursements grew at an average annual growth rate of 21%. DAC countries committed and disbursed the largest proportion (82 and 79%, respectively) over the sixteen-year period. Kenya received the largest proportion of aid (33%) and was most favored by DAC donors. Whereas, overall 93% of committed aid to SRH was disbursed over the 16 year period, there were year on year fluctuations in both commitments and disbursement. The study showed improvements in both SRH process/outcome indicators and impact indicators.

The trend of growth in donor aid observed in this study is in line with studies done at a global scale (18, 21, 25–29). There was an increase in both commitments and disbursements between 2002 and 2008. During 2009 to 2013 disbursements continued to grow although commitments declined. Toward the end of the Millennium Development Goals era between 2013 and 2015, a decline was observed in both commitments and disbursements for donor aid to the four countries. However, there was an increase in 2016 at the start of the SDGs era. These changes may point to some unpredictability of aid.

The peaks and dips in aid that are observed are not markedly erratic which may reinforce John Hudson's assertion that aid to health is one of the least volatile (30). However, Kenya had the largest fluctuation between funds committed and those disbursed. When donors do not disburse what they commit, it affects the recipient governments' ability to plan and therefore impacts on results as was noted by Arregoces et al. (21). Recipient countries should therefore cautiously rely on aid and track volatility in aid provided.

While there was improvement in both SRH outcome and impact indicators alongside growth in donor aid over the 16-year period, the SRH impact indicators are not reducing fast enough to meet SDG targets (2, 4). Kenya, which received most funds, also had the most promising SRH indicators. Zambia on the other hand received the least SRH aid over the period and with its population that is less than half of any of the other three countries, is struggling with SRH indicators of a similar magnitude (4–8). LMICs will require more concerted efforts to avert future maternal and child mortality.

The disbursements by the different donors showed countries of preference. DAC countries and the World Bank provided most aid to Kenya; Tanzania received most funds from multilateral donors whereas Uganda was most preferred by UN agencies. Donors have preferred countries to provide aid based on strategic interests. The magnitude of aid may differ based on various reasons, for example; delays in project

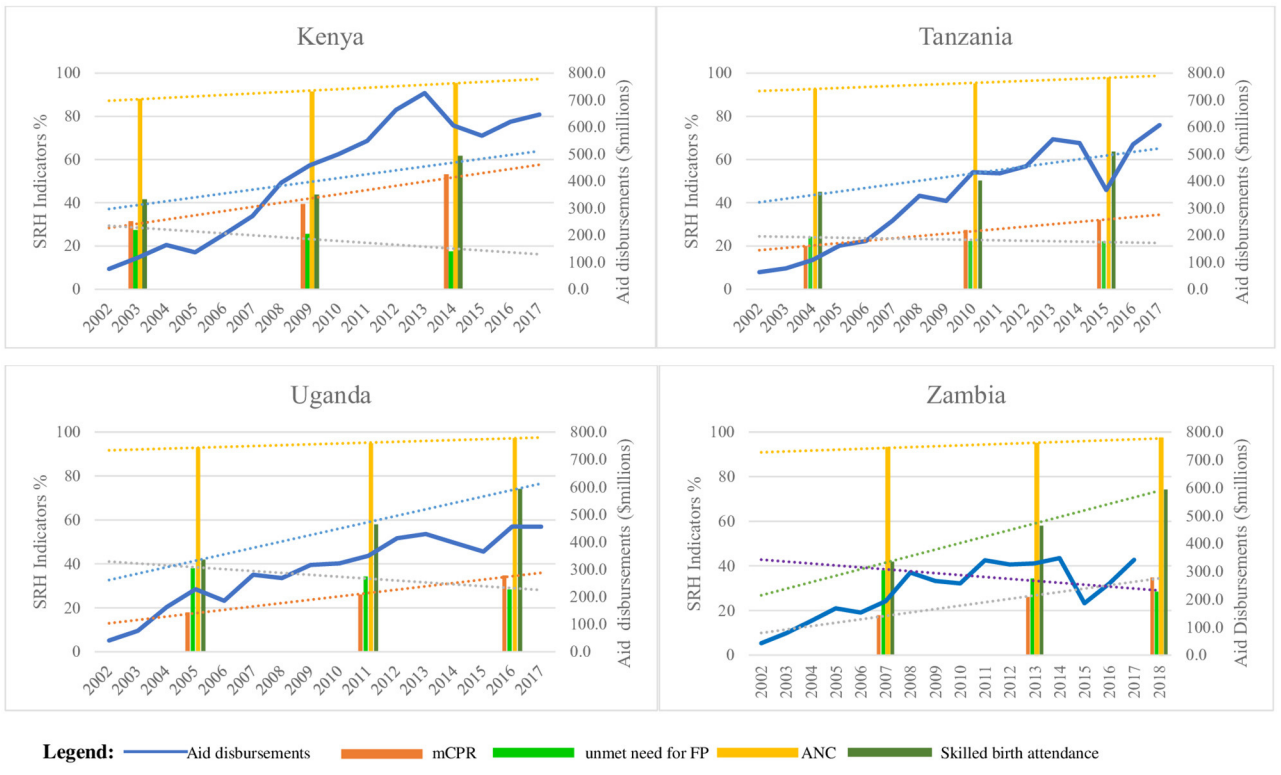

FIGURE 3 | Comparison of disbursements with selected SRH health facility level process/outcome indicators.

implementation, emergencies that call for immediate support interventions, sometimes donor countries have realized more or less than expected growth and therefore have more or less aid available, other times there are changes in donor political environment.

The preference for some countries by donors may also not be targeted to recipient national priorities or countries with most need as noted by Grollman et al. and other studies (26, 31). However, what is clear is the need for more funding to meet SDG3 targets (10) and therefore more deliberate targeting of funding to country needs and priority interventions is required (32–34). Countries have an obligation to the United Nations to spend a target of 0.7% of their gross national income (GNI) on international aid (35). Bilateral aid is a reflection of strategic interests of donors and is driven by variables that include: an obligation to protect human rights, dignity and solidarity; trade and economic relations with recipient countries; political interests including creating stability in poor countries to reduce migration; level of transparency and accountability within recipient governments (36, 37). However, politics is the ultimate determinant (38–40). As countries develop, donor countries prefer to transition from aid to trade. LMICs therefore ought to progressively move away from reliance on donor support and increase country ownership of health needs by consistently improving domestic investments in SRH (41, 42), as reflected in the 2017 Tokyo declaration on Universal Health Coverage (43).

To meet the aspirations of the United Nations (UN) Secretary-General's *Global Strategy for Women's, Children's and Adolescents' Health, 2016–2030* (9) and the Tokyo Declaration to Universal Health Coverage (43), more efforts will be required to mobilize governments, donors and other stakeholders to ensure sustained funding to SRH (42). This is especially important in the light of slower improvement in some of the SRH indicators in the latest years. Aid has been shown in other studies to have positive long-term effect on health and on development (44–46). Also, Dieleman et al. note that in the near term, increased domestic spending on health alone is unlikely to cover the gaps to meet the ambitious health goals laid out in the SDGs (47). Therefore, increased funding to DAH is required and should be in accordance with principles of the Paris Declaration on Aid Effectiveness.

SRHR investments are critical to people's well-being, the prosperity and resilience of families, communities and nations (17). These investments are cost-effective and cost-saving, freeing resources for investment in other development priorities with high pay-offs for equality and equity. Regardless of the demonstrated returns on investments, underfunding of SRHR still persists. This is a contributing factor to why the core goal of achieving universal access to sexual and reproductive services adopted by 179 governments at the International Conference on Population and Development (ICPD 1994) remains unfulfilled (17).

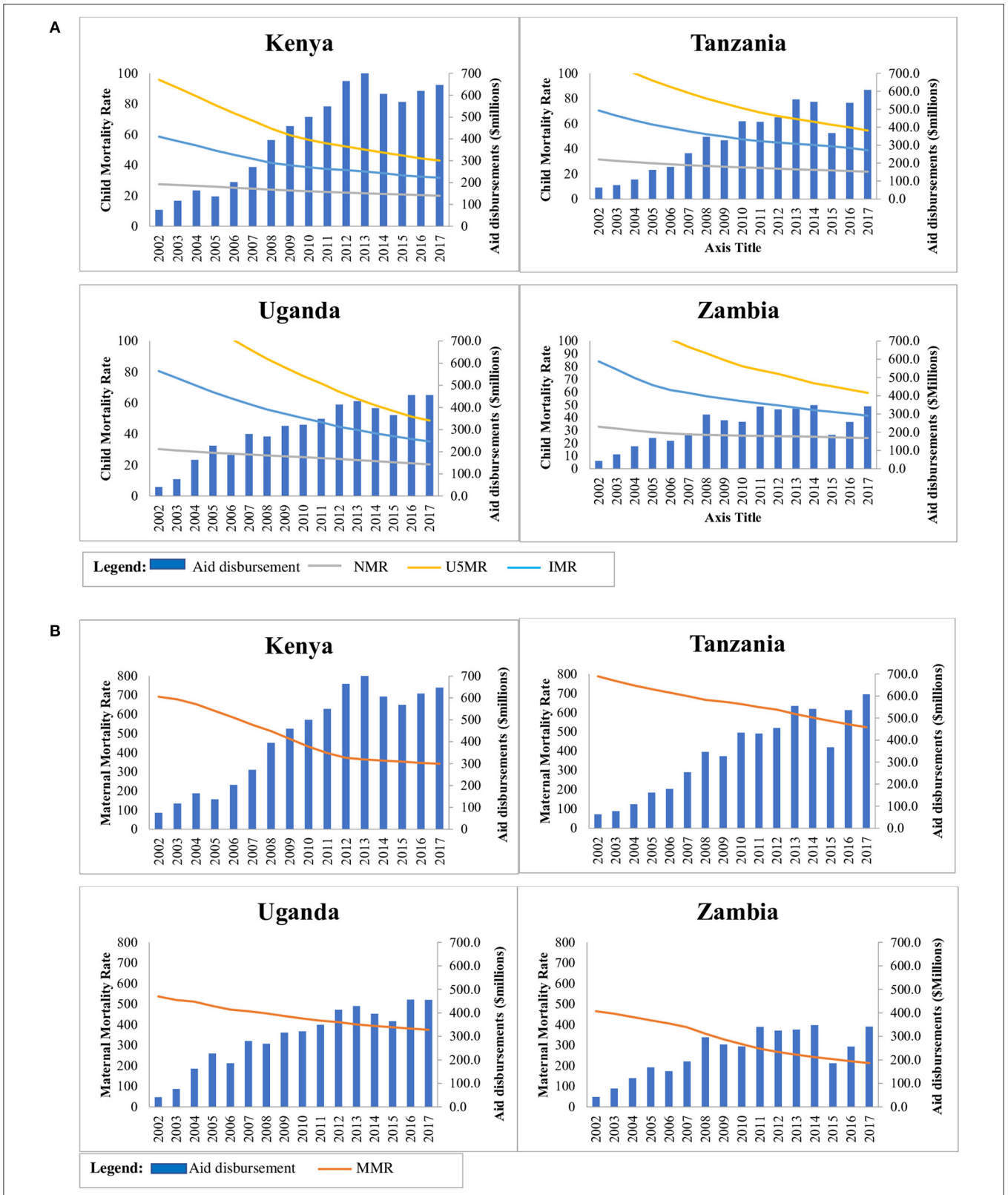

FIGURE 4 | (A) Comparison of aid disbursements with neonatal, under five and child mortality rates. **(B)** Comparison of aid disbursements with maternal mortality rates.

The Organization for Economic Co-operation and Development's Development Assistance Committee (OECD DAC) is one of the most comprehensive tracking platforms for resource flows. The OECD CRS information has been recorded, in one form or another, since 1967. It is relatively complete in terms of bilateral aid commitments since 1995. Even taking into account changes in definitions, the time series information is the most stringent and validated database on aid flows that currently exists (48). In addition, the database provides for accuracy of data as it ensures that accurate and comparable measurements of donor outflows can be derived (48).

As shown by other similar studies, donor reporting to the CRS has improved over time (49). The CRS is limited by the accuracy, completeness, consistency, and timeliness of donor reports to improve data in the system which also affect study is limited. It is important that more efforts are made to ensure improvement of data under the CRS. It is also important to note that the study only describes donor funding for SRH and how it compares with some of the most critical SRH indicators. Statistical assessment of the relation between disbursements and SRH indicators was not carried out since there are many factors which influence these indicators. The study does not presuppose that aid can be independent of both government (domestic) funding and out-of-pocket payments (49). As is noted in other studies, this study also does not explain variation or timeliness in donor aid to different countries and therefore further research is needed (50). However, the strength of this study is that it zooms in on some of the specific countries with most need, begins to assess predictability of aid by assessing commitments and disbursements to add to the body of knowledge on accountability of donor aid for SRH. Reporting on the time when both commitments and disbursements are made in CRS will help strengthen arguments around predictability of aid which is important for recipient country planning (41).

CONCLUSION

Donor commitments for SRH aid grew on average by 20% annually while disbursements grew by 21% annually between 2002 and 2017. There was improvement in SRH indicators alongside growth in donor aid but improvement is slow to meet SGD targets. There were year on year fluctuations in both commitments and disbursements. Unpredictability and insufficiency of donor aid may be disruptive to country planning and may lead to missing of global targets on SRH. Donors and LMICs should increase domestic investments in order to meet global SRHR targets.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

DK conceptualized the project and undertook data analysis and wrote the first draft of the manuscript. CA, MM, HH, TR, HL, and AM-T revised the manuscript and critically reviewed its contents. CA contributed to data collection and data analysis. AM-T and HH critically reviewed the manuscript, provided comments and guidance on all drafts of manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research used data that was publicly available from the Organization for Economic Co-operation and Development's Assistance creditor reporting system. Research time and travel was funded by Health Action International through the Health Systems Advocacy Partnership.

ACKNOWLEDGMENTS

The authors recognize the Organization for Economic Co-operation and Development's Assistance creditor reporting system for the data and Daphne Ssebugwawo who edited the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2021.645499/full#supplementary-material

REFERENCES

1. *United Nations Population Fund*. New York, NY; The State of World Population (2017).
2. United Nations Inter-Agency Group on Child Mortality Estimation. *Levels & Trends in Child Mortality*. New York, NY; United Nations Inter-Agency Group on Child Mortality Estimation (2017).
3. World Health Organisation. *Maternal Mortality Fact sheet* Geneva: WHO. (2014).
4. World Health Organisation. *Trends in maternal mortality: 1990 to 2015: Estimates by WHO, UNICEF, UNFPA*. Geneva; World Bank Group and the United Nations Population Division (2015).
5. Kenya National Bureau of Statistics. *Kenya Demographic and Health Survey* Rockville, MD: KNBS and ICF International. (2014).
6. Ministry of Health Community Development Gender Elderly Children Tanzania, et al. *Tanzania Demographic Health Survey Findings-MIS Key Findings*. Tanzania: Dar es Salaam; Rockville, MD: MoHCDGEC, MoH, NBS, OCGS, and ICF. (2016).
7. Uganda Bureau of Statistics ICF. *Uganda Demographic and Health Survey 2016: Key Indicators Report*. Kampala; Rockville, MD: Uganda Bureau of Statistics, ICF. (2017).
8. Central Statistical Office, Ministry of Health, ICF. *Zambia Demographic and Health Survey*. Rockville, MD: Central Statistical Office, Ministry of Health, and ICF International (2014).
9. Every Woman Every Child Partnership for MNCH. *Progress in Partnership: Progress Report on the Every Woman Every Child Global Strategy for Women's, Children's and Adolescents' Health*. Geneva: WHO (2017).
10. United Nations. *The Sustainable Development Goals Report*. New York, NY; United Nations Statistical Division (2016).
11. Anyanwu CJ, Erhijakpor AEO. Health expenditures and health outcomes in africa. *Afr Dev Rev*. (2007) 21:400–33. doi: 10.1111/j.1467-8268.2009.00215.x

12. World Health Organisation. *World Health Statistics*. Geneva: WHO (2010).
13. Barugahare J, Reidar KL. Obligations of low income countries in ensuring equity in global health financing. *BMC Med Ethics*. (2015) 16:59. doi: 10.1186/s12910-015-0055-3
14. Commission on Macroeconomics and Health. *Macroeconomics and Health: Investing in Health for Economic Development*. Geneva: WHO (2001).
15. *World Health Organisation and Health Financing: A strategy for the African Region*. Brazzaville: Regional Office for Africa. (2006).
16. *World Health Organisation and Statistical Information System*. Geneva: WHO (2008).
17. High-level task force for international conference on population and development. Policy considerations for financing sexual and reproductive health and rights in the post-2015 Era. *Rep Health Matters*. (2015) 23:159–63. doi: 10.1016/j.rhm.2015.07.004
18. Patel P, Roberts B, Guy S, Lee-Jones L, Conteh L. Tracking official development assistance for reproductive health in conflict- affected countries. *PLoS Med*. (2009) 6:e1000090. doi: 10.1371/journal.pmed.1000090
19. Institute for Health Metrics and Evaluation. *Financing Global Health 2013: Transition in an Age of Austerity*. Seattle, WA: IHME (2014).
20. The Partnership for Maternal Newborn & Child Health. *The PMNCH 2014 Accountability Report: Tracking Financial Commitments to the Global Strategy for Women's and Children's Health*. Geneva: WHO (2014).
21. Arregoces L, Daly F, Pitt C, Hsu J, Martinez-Alvarez M, Greco G, et al. Countdown to 2015: changes in official development assistance to reproductive, maternal, newborn, and child health, and assessment of progress between 2003 and 2012. *Lancet*. (2015) 3:E410–421. doi: 10.1016/S2214-109X(15)00057-1
22. Abe SK, Murakami Y, Morgan D, Shibuya K. Rethinking the global tracking system on development assistance to reproductive, maternal, newborn, and child health. *Lancet*. (2015) 3:PE350–E351. doi: 10.1016/S2214-109X(15)00043-1
23. Organisation for Economic Co-operation and Development. *OECD's Methodology for Calculating Sectoral Imputed Multilateral aid*. (2017). Available online at: http://www.oecd.org/dac/stats/oecdsmethodologyforcalculatingsectoralimputedmultilateralaid.htm (accessed November 04, 2019).
24. *WHO's short list of Reproductive Health Indicators for Global Monitoring*. Available online at: https://www.measureevaluation.org/prh/rh_indicators/family-planning/global/whos-short-list-of-reproductive-health-indicators-for-global-monitoring (accessed September 22, 2019).
25. International Aid Transparency Initiative. *Annual Report* Bristol, UK: Development Initiatives Poverty Research Limited. (2017).
26. Grollman C, Arregoces L, Martínez-Álvarez M, Pitt C, Mills A, Borghi J. 11 years of tracking aid to reproductive, maternal, newborn, and child health: estimates and analysis for 2003–13 from the Countdown to 2015. *Lancet*. (2017) 5:e104–14. doi: 10.1016/S2214-109X(16)30304-7
27. Hsu J, Pitt C, Greco G, Berman P, Mills A. Countdown to 2015: changes in offcial development assistance to maternal, newborn, and child health in 2009–10, and assessment of progress since 2003. *Lancet*. (2012) 380:1157–68. doi: 10.1016/S0140-6736(12)61415-9
28. Dieleman JL, Graves C, Johnson E, Templin T, Birger M, Hamavid H, et al. Sources and focus of health development assistance, 1990–2014. *JAMA*. (2015) 313:2359–68. doi: 10.1001/jama.2015.5825
29. *Institute for Health Metrics and Evaluation. Financing Global Health 2015: Development Assistance Steady on the Path to New Global Goals*. Seattle: IHME (2016).
30. Hudson J. Consequences of aid volatility for macroeconomic management and aid effectiveness. *World Dev*. (2015) 69:62–74. doi: 10.1016/j.worlddev.2013.12.010
31. Michaud C, Murray CJ. External assistance to the health sector in developing countries: a detailed analysis, 1972-90, in bulletin of the. *World Health Organ*. (1994) 639–51.
32. Keats EC, Ngugi A, Macharia W, Akseer N, Khaemba EN, Bhatti Z, et al. Progress and priorities for reproductive, maternal, newborn, and child health in kenya: a countdown to 2015 country case study. *Lancet*. (2017) 5:PE782–795. doi: 10.1016/S2214-109X(17)30246-2
33. Black RE, Levin C, Walker N, Chou D, Liu L, Temmerman M, et al. Reproductive, maternal, newborn, and child health: key messages from disease control priorities. *Lancet*. (2016) 388:2811–24. doi: 10.1016/S0140-6736(16)00738-8
34. Bhutta ZA, Das JK, Bahl R, Lawn JE, Salam RA, Paul VK, et al. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? *Lancet*. (2014) 384:347–70. doi: 10.1016/S0140-6736(14)60792-3
35. Clemens MA, Moss TJ. The ghost of 0.7 per cent: origins and relevance of the international aid target. *Int J Dev Issues*. (2007) 6:3–25. doi: 10.1108/14468950710830527
36. Burnside C, Dollar D. Aid, policies, and growth. *Am Eco Rev*. (2000) 90:847–68. doi: 10.1257/aer.90.4.847
37. Berthélemy JC, Tichit A. Bilateral donors' aid allocation decisions—a three-dimensional panel analysis. *Int Rev Eco Finance*. (2004) 13:253–74. doi: 10.1016/j.iref.2003.11.004
38. Langan M. Neo-colonialissm and donor interventions: western aid mechanisms, neo-colonialism and the poverty of 'development' in africa. *Contemporary African Polit Eco*. (2017) 61–88. doi: 10.1007/978-3-319-58571-0_3
39. Delputte S, Orbie J. The EU and donor coordination on the ground: perspectives from tanzania and zambia. *Eur J Dev Res*. (2014) 26:676–91. doi: 10.1057/ejdr.2014.11
40. Alesina A, Dollar D. Who gives foreign aid to whom and why? *J Eco Growth*. (2000) 5:33–63. doi: 10.1023/A:1009874203400
41. Sridhar D, Batniji R. Misfinancing global health: a case for transparency in disbursements and decision making. *Lancet*. (2008) 372:1185–91. doi: 10.1016/S0140-6736(08)61485-3
42. Moucheraud C, et al. Countdown to 2015 country case studies: what have we learned about processes and progress towards MDGs 4 and 5? *BMC Public Health*. (2016) 16(suppl. 2):794. doi: 10.1186/s12889-016-3401-6
43. World Bank. *Universal Health Coverage Forum*. Tokyo: The World bank, WHO, UNICEF, JICA, UHC2030 International Health Partnership. (2017).
44. Jones S, Tarp F. Does foreign aid harm political institutions. *J Dev Eco*. (2015) 118:266–81. doi: 10.1016/j.jdeveco.2015.09.004
45. Arndt C, Jones S, Tarp F. Assessing foreign aid's long-run contribution to growth and development. *World Dev*. (2015) 1:6–18. doi: 10.1016/j.worlddev.2013.12.016
46. Mishra P, Newhouse D. Does health aid matter? *Health Eco*. (2009) 28:855–72. doi: 10.1016/j.jhealeco.2009.05.004
47. Dieleman JL, Micah AE, Murray CJL. Global health spending and development assistance for health. *JAMA*. (2019) 321:2073–4. doi: 10.1001/jama.2019.3687
48. Petras R. *Comparative Study of Data Reported to the OECD Creditor Reporting System (CRS) and to the Aid Management Platform (AMP)*. Washington, DC: Development Gateway (2009).
49. Grépin KA, Pinkstaff CB, Shroff ZC, Ghaffar A. Donor funding health policy and systems research in low- and middle-income countries: how much, from where and to whom. *Health Res Policy Syst*. (2017) 15:68. doi: 10.1186/s12961-017-0224-6
50. French D. Did the millennium development goals change trends in child mortality? *Health Econ*. (2016) 25:1312–25. doi: 10.1002/hec.3218

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

BRIEF RESEARCH REPORT
published: 26 April 2021
doi: 10.3389/fpubh.2021.614915

The Impacts of Flattening Fiscal Reform on Health Expenditure in China

Jun Hao[1,2†], *Chengxiang Tang*[3,4†], *Junli Zhu*[1*] and *Jiayi Jiang*[3,4]

[1] School of Public Health, Capital Medical University, Beijing, China, [2] Medical Research and Biometrics Center, National Center for Cardiovascular Diseases, Fuwai Hospital, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China, [3] School of Public Administration, Guangzhou University, Guangzhou, China, [4] China Center for Health Economic Research, Peking University, Beijing, China

Introduction: A number of provinces have implemented a fiscal reform of flattening government since the first decade of this century in China. This study aims to quantitatively analyze the influences of this government fiscal reform on county-level health expenditure. We also bring forward policy suggestions for improving county-level fiscal system and healthcare delivery.

Methods: We collected a novel longitudinal county-level data from 2003 to 2010, including counties' socioeconomic data, fiscal revenue, and health expenditure. Jilin Province, Hebei Province, and Anhui Province were selected as representative samples for this policy evaluation. The study employed a time-varying difference-in-difference model specification to investigate the impacts of flattening fiscal reform on health expenditure.

Results: The analyses find that the fiscal system reforms of the three provinces have a significantly positive impact on the health expenditure of county-level governments. However, we find no policy effects on the proportion of health expenditure to fiscal expenditure of county-level governments. The estimation results are robust after controlling several background variables.

Conclusion: The results yield important policy insights that public finance and its reform significantly impacts health expenditures in China. The government may still need to strengthen the transfer payment system to guarantee the social welfare provision in healthcare.

Keywords: fiscal reform, health expenditure, public finance, difference-in-difference, Province-Managing-County

OPEN ACCESS

Edited by:
Mihajlo Jakovljevic, Hosei University, Japan

Reviewed by:
Brian Godman, University of Strathclyde, United Kingdom
Deepak Kumar Behera, Manipal Academy of Higher Education, India

***Correspondence:**
Junli Zhu
smallying@126.com

*These authors have contributed equally to this work

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *07 October 2020*
Accepted: *22 February 2021*
Published: *26 April 2021*

Citation:
Hao J, Tang C, Zhu J and Jiang J (2021) The Impacts of Flattening Fiscal Reform on Health Expenditure in China. Front. Public Health 9:614915. doi: 10.3389/fpubh.2021.614915

INTRODUCTION

Universal Health Coverage (UHC) has become a health strategy goal of many countries in the world. UHC requires that everyone has access to the services they need within the health system, where these services are of adequate quality to be effective, and get universal financial protection in the costs of using these services. It is the governments' responsibilities to provide domestic resources to the maximum extent possible in order to fulfill their commitments to the health and other human rights of their citizens (1, 2). China's total health expenditure consists of three parts: government budget health expenditure,

social insurance expenditure, and individual health expenditure. Government budget health expenditure refers to the financial allocation of the central and local governments for healthcare, in which the health expenditure defined in this study specifically refers to county-level governments' health expenditure. In 1997, the "Decision on Health Reform and Development" issued by the Central Committee of the Communist Party of China and the State Council clearly stated that the governments take important responsibility for the development of health services, and that the government health expenditure should be increased with the development of the economy and the growth rate should not be lower than that of fiscal expenditure. "Plan of Health China 2030" released in 2016 also emphasized adjusting and optimizing the structure of fiscal expenditure and increasing government expenditure in health sectors. China's total health expenditure report for 2019 showed that the total health expenditure accounted for 6.6% of the gross domestic product (GDP), of which the government health expenditure accounted for 26.7% of the total health expenditure. Although the governments have increased health expenditure since severe acute respiratory syndrome (SARS) in 2003, the governments' role in health expenditure is extremely weak when compared with other countries, especially Organization for Economic Cooperation and Development (OECD) countries (3). The central and local governments in China have been inactive in providing healthcare for citizens, and China's health sectors have been mainly financed by the private sectors (3).

The financial capacity of county governments is crucial to ensure adequate health services for their residents. Since 1994, China has implemented a hierarchical tax-sharing budget management system (referred to as the "tax-sharing system") clarifying the division of fiscal revenue and responsibilities between the central and local governments. China has gradually become one of the countries with a high degree of fiscal decentralization in the world since then (4). The central and provincial governments are responsible for the broader policy and strategic design and investment in the larger health infrastructure, whereas the county-level governments have practical responsibilities for implementing health programs or services (5). Data showed that more than 98% of government health expenditures in China mainly rely on local government finances (6). Especially, the county-level government plays an important role in the construction of the county-level healthcare system and undertakes important tasks, such as ensuring the equalization of basic public health services.

In China, fiscal revenue was mainly concentrated in higher-level governments, namely, the central, provincial, and prefectural governments, whereas county-level and township-level financial resources were very weak. There was a mismatch of financial power and administrative power between cities and counties. China established a hierarchical structure of governance in 1982, which consists of five layers of government—from the highest to the lowest, they are as follows: the central level, the province or municipality level, the prefecture or city level (hereafter city-level), the county level, and the township level (7). However, it is evident that the five-level administrative system has caused many problems with the rapid development of China's economy and society since the 1980s and the early 2000s, one of which is that city-level governments allocated more financial resources to cities, forming a situation of a "city-scraping county" (8). Besides, a reform of the tax-sharing fiscal system mentioned above was carried out in China in 1994, which initially clarified the fiscal and taxation relations between the central and provincial governments, whereas the relations below the province-level government were still not illuminated. As a result, fiscal revenue was mainly concentrated in the central, provincial, and prefectural governments, whereas financial resources at the county level and township level were very weak. This resulted in the mismatch of financial power and administrative power between cities and counties (9). Data showed that the proportion of county and township fiscal expenditures in China's fiscal expenditures has increased by year since 2000, and that it has become the most responsible government for fiscal expenditures since 2003. At the same time, the county and township fiscal revenues only accounted for 20% of the national fiscal revenue (10). The county and township financial system is so poor that many county governments suffered from insufficient allocation of basic public services.

Given that situation, all sectors of society began to actively explore new management systems. A fiscal decentralization reform was implemented in China's local governments, which eliminated the city-level government as the intermediate layer between the province and the county (flattening fiscal reform). In 1992, Zhejiang Province creatively put forward a flattening fiscal reform that reduced the five-level hierarchical structure to a four-level system—central level, province or municipality level, county level, and township level. After that, Zhejiang Province carried out five rounds of power expansion reform, and its county economy developed rapidly. Compared with the economically developed coastal provinces, such as Jiangsu, Guangdong, Shandong, and Fujian, Zhejiang's county economy developed to a higher and more balanced level (11). As a new exploration of fiscal decentralization mode, the flattening fiscal system has achieved remarkable results in Zhejiang Province, prompting the Ministry of Finance to carry out the flattening fiscal reform nationwide since 2002, known as the pilot work of Province-Managing-County (PMC) reform (12, 13).

The PMC was a reform to flatten governmental hierarchical structure, which made huge innovations in the expansion of economic rights of the grassroots government and the deepening of fiscal decentralization. This reform enabled county governments directly governed by the provincial government *via* the public finance transfer payment, fund dispatching, financial settlements, and work deployment, no longer subject to the prefectural finance. The purpose of the reform is to improve local economic development, augment county-level finance capacity, and equalize the supply of public services across counties (14). The reform was first piloted in counties in central and northern China, such as Anhui Province, Hebei Province, Liaoning Province, and Jilin Province. In 2006, the reform was extended to Jiangsu Province, Shaanxi Province, Sichuan Province, Gansu Province, and Qinghai Province in eastern and western China. The Ministry of Finance proposed that the PMC reform should be promoted across all provinces,

except minority autonomous regions by 2012 in the Opinions on Promoting Province-Managing-County Reform [No. 78 (2009) of The Ministry of Finance). By the end of 2012, 1,099 counties from 24 provinces, representing ∼56% of all counties across the country, had implemented the PMC reform (15).

Until now, there has been little research on the impact of the PMC reform on government health expenditure, especially from the perspective of fiscal reform, so the impact of the PMC reform on county-level government health expenditure has been unclear yet. The analysis of the factors that influence China's local government health expenditure and health service provision from a new perspective, namely, fiscal system reform, will provide a powerful support for better promoting UHC policies, such as equalization of basic public health services. Therefore, it will be necessary and interesting to study government health expenditure in these counties from the perspective of the fiscal system. This research will focus on the impact of the PMC reform on county-level government health expenditure, employing a panel data set of county-level governments in Hebei Province, Anhui Province, and Jilin Province in China during 2003–2010.

LITERATURE REVIEW

The previous studies have focused on the factors of influencing government health expenditure from multiple perspectives, such as macro-economic variables. Cantarero found that the aging population was the most important determinant of the regional health expenditure in Spain (16). Rahman showed that per capita income and literacy rate had an important impact on per capita health expenditure in India from 1971 to 1991 (17). Behera and Dash studied the long-run effects of GDP and tax revenue on public health expenditure in 16 major states of India during 1980–2014 (18). He analyzed China's 2000–2011 provincial panel data and found that per capita GDP had a significant impact on the government health expenditure, whereas population size, population structure, and urbanization had no effects (19). Lu & Wang found that economic growth, aging degree, and other factors positively affected per capita public health expenditure by an empirical research on provincial panel data of China from 2002 to 2006 (20).

Early scholars explained fiscal decentralization from the low-level government to understand the preferences of residents in the jurisdiction (21). Recently, some scholars believed that fiscal decentralization could encourage local officials to improve the welfare of residents in the jurisdiction (22, 23). Most studies have shown that fiscal decentralization benefited public health and improved public outcomes, such as widening childhood immunization coverage and reducing infant mortality rates (24–27). Since the reform of the tax-sharing system in 1994, China has gradually become one of the countries with a high degree of fiscal decentralization in the world (4). Several studies on the impact of fiscal decentralization in China have emerged, but these studies have not reached a consistent conclusion as international evidences and there are only a few articles on the impact on health. Peng & Tang proved that fiscal decentralization promoted the supply of health services (28), whereas Sun and Jin gave the evidence that fiscal decentralization did not reduce infant mortality by using provincial panel data (29).

Because the PMC reform in China involved a sudden change in vertical governmental structure and was implemented only in certain localities at a certain point in time, it is a quasi-natural experiment, providing an opportunity to evaluate the impact of fiscal decentralization on public expenditures of the local governments in China. There have been a growing number of papers in recent years on the PMC reform, especially in Chinese. A study suggested that the PMC reform may have a significant impact on the scale and composition of government expenditures, and that it increased the productive expenditure of a county government (30, 31), but the evidence for the impact on welfare expenditure has been still inconclusive. Some studies showed that the PMC reform was found to increase local welfare expenditure, although it did not change the motivation of the local government to allocate more funds to productive public services rather than to civil public services (14, 30). Another research insisted that the PMC reform decreased local welfare expenditure as more resources were available to increase productive expenditure (31, 32). Although the local government in China has taken important responsibilities for health services, there is no study that evaluates the effects of the PMC reform on the health expenditure of local governments. In one related study, the authors only chose health expenditure as one kind of county governments' welfare expenditure and found that the PMC reform has a negative effect on the proportion of health expenditure to fiscal expenditure (32).

METHODS AND DATA

Study Data

This study uses data and documents related to the PMC fiscal reform from provincial government websites and the *China Statistical Yearbooks Database* (CSYD). Considering the availability of data and the gap between the economic development level of the western provinces and the central and eastern provinces in China, the western provinces have not been included.

Taking into account the representativeness of the sample and the availability of the data, the study took counties (county-level cities) in Jilin Province (northeast region), Hebei Province (east region), and Anhui Province (central region) as the research objects (see **Figure 1**). In addition, the counties (county-level cities) that have undergone county-to-district or county-to-city administrative territorial entity adjustment during the sample period have been excluded. Therefore, the final sample set of this study was 237 counties (including 190 counties and 47 county-level cities) in Jilin Province from 2001 to 2010, Hebei Province from 2003 to 2010, and Anhui Province from 2003 to 2010, with a total of 1,976 observations. Because of the different reform time points of the three provinces, it was necessary to construct a multi-period difference-in-difference (DID) model to study the impact of the PMC fiscal reform in the three provinces on the county-level health expenditure, which requires the balanced panel data. Therefore, when analyzing the policy effect of the PMC fiscal reform in these three provinces, we treated the

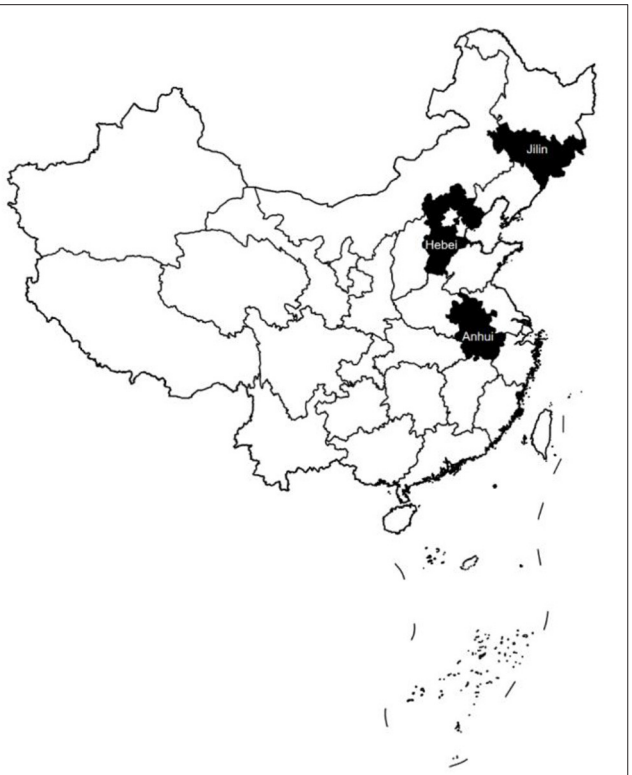

FIGURE 1 | Three provinces that implemented fiscal reform after 2005.

sample period as 2003–2010, 237 counties (county-level cities) altogether, and 1,896 observations. In January 2005, 22 counties in Hebei Province (16%, $n = 136$) were involved in the PMC fiscal reform, along with 32 counties in Jilin Province (80%, $n = 40$) in June 2005 and 57 counties in Anhui Province (93%, $n = 61$) in January 2004.

By consulting the relevant policy documents of the PMC fiscal reform in each province, we were able to see the progress of the reform and manually input the data, such as the reform time point. Social and economic data like county population, GDP, and financial data, such as fiscal revenue and expenditure, were collected by consulting the statistical yearbook and fiscal yearbook. Some missing data were obtained from city-level statistical yearbooks. Data processing and statistical analysis were performed using Stata 15.1.

Dependent Variable

In this study, the health expenditure (healthexp) and its proportion to fiscal expenditure (healthperc) were taken as dependent variables. The proportion of health expenditure to fiscal expenditure (healthperc) can be used to measure the government's investment in healthcare.

Independent Variable

In this study, the independent variable was the dummy variable of the PMC fiscal reform. One represents the value of a county (county-level cities) in the year of the fiscal reform and after the reform, and 0 represents the value of the year before the reform. Based on previous literature, this study controlled for variables that may have an impact on health expenditure. First, the level of economic development will affect the government's revenue-raising capacity, which has an impact on health expenditure, so we controlled for variables that reflect the level of the economic development—per capita gross domestic product (avgGDP) and second industry output (secindustry). Second, the population scale will affect government operation in every aspect, so the population of each county was included (pop). In order to reduce the influence of heteroscedasticity, we made log transformations for the above variables into lnavgGDP, lnpop, and lnsecindustry. Self-financing capacity (capacity) refers to the percentage of the fiscal revenue in fiscal expenditure. The self-financing capacity of the local government will affect the structure of fiscal expenditure directly, so it was included in the model. Since both counties and county-level cities have carried out PMC reforms, county-level cities have already been coordinated by provinces. Therefore, we added a dummy variable (county2) into the model in order to control the impact of different administrative units' division −1 for county-level city and 0 for county.

Identifying Strategy

The DID model is generally employed to evaluate the effect of policy implementation (33, 34), and the actual effect of policy variable is obtained through fixed-effect analysis of panel data. There have been many studies using the DID model to evaluate the effect of policy implementation, such as many empirical analyses of the policy effect of the PMC reform. Besides, the treatment structure in our study is characterized by varying policy start dates; thus, we use the following time-varying DID model specification to estimate the effect of the PMC reform on county-level health expenditure:

$$\boldsymbol{Healthexpend_{it} = \beta_1 Treatment_{it} + year_t + C_i + X_{it} + \varepsilon_{it}}. \quad (1)$$

For any observation county i in year t, $Treatment_{it}$ is a dummy variable equal to 1 for the year that a county started to introduce the PMC fiscal reform and 0 for the years after. The coefficient of interest β_1 captures the average change in the health expenditure during the process of fiscal reform. In this two-way fixed-effects model, a set of year dummies $year_t$ and county dummies C_i capture the differences fixed over years and across counties. Other control variables that change across counties and over time, X_{it}, include the log of population, the log of per capita GDP, and the log of second industry output, which may affect the health expenditure. We have reported robust standard errors clustered at the county level. For this identification strategy, we tested for assumption of common trend between the treatment group and the control group (see **Figure 2**).

RESULTS

Table 1 shows the descriptive statistics of selected variables. Relative to the treatment group, the control group had lower health expenditure; however, the proportion of health expenditure to fiscal expenditure showed little difference between

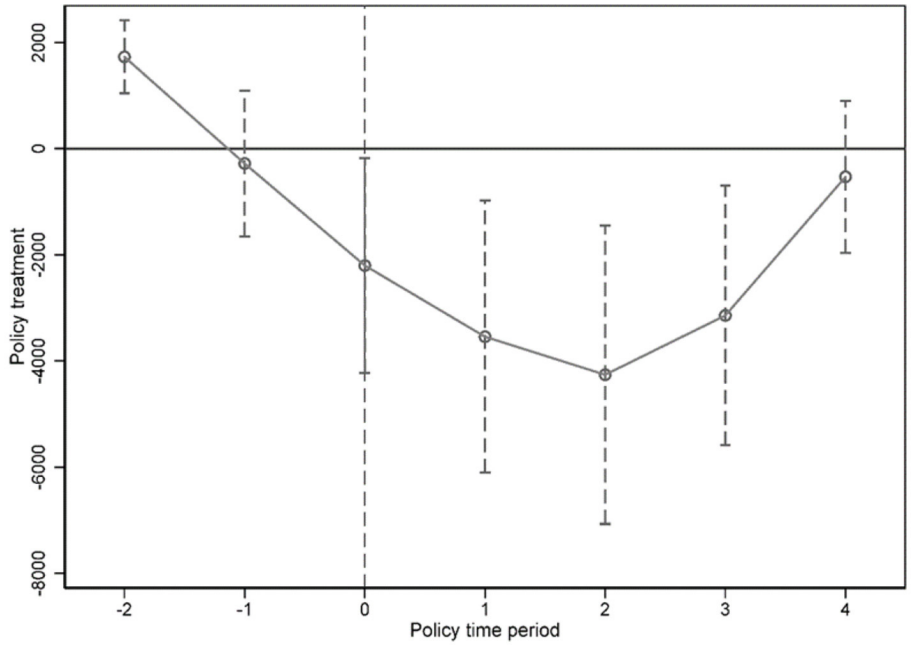

FIGURE 2 | Common trend test before and after fiscal reform.

the two groups in the outcome variables. In the control variables, the mean values of self-financing capacity, per capita GDP, population, and second industry output of the treatment group were higher than those of the control group.

Table 2 presents the regression results for health expenditure and its proportion to fiscal expenditure. All columns consist of the county-level control including self-financing capacity, per capita GDP, population, and second industry output. Besides, a set of year fixed effects and county fixed effects were included. Standard errors shown in parentheses were clustered at the county level. It was found that the PMC reform of the three provinces had a significantly positive effect on the health expenditure of county-level governments ($\beta_1 = 2{,}004.9$, $p < 0.001$), whereas we found no effects on the proportion of health expenditure to fiscal expenditure of county-level governments ($\beta_1 = 0.00$, $p > 0.05$).

To guarantee reliability, we did a robustness inspection to test the reliability of the conclusion. As is shown in **Table 3**, only the treatment variable was added in column (1). In column (2), self-financing capacity, per capita GDP, population, and second industry output were added. On the basis of column (2), we added a control variable for a dummy variable indicating whether the county is a county-level city. Columns (4), (5), and (6) were hierarchical regression. Column (4) is the regression result of counties with population higher than 52.16; column (5) is the regression result of counties with self-financing capacity lower than 0.30, and the final column is the result of counties with per capita GDP higher than 13,426.7. **Table 3** shows that the PMC fiscal reform had a significant impact on county-level health expenditure but no effects on the proportion of health expenditure to fiscal expenditure. **Table 3** shows that for health expenditure, the estimated coefficients are statistically significant in the sample period.

DISCUSSION

The PMC reform of Jilin Province, Hebei Province, and Anhui Province had a significantly positive effect on the health expenditure of the county-level government, which is the important part of civilian public service expenditure. This result is consistent with the research results based on the county-level data of Henan Province from 2000 to 2013 that the productive and civilian public service expenditure of the counties directly managed by the province has been significantly increased (35). Another study on the impact of the PMC reform on per capita basic education expenditure reached the same conclusion (14). The PMC reform can significantly improve the fiscal expenditure and economic growth of counties and alleviate the financial difficulties of county-level governments, which will lay a solid economic and financial foundation for health expenditure. At the same time, the reform reduced the levels of local finance, which alleviated the phenomenon of a "city-scraping county" and increased the fiscal autonomy of local governments. The positive effects of the reform will encourage county-level governments to increase their health expenditure. Despite the positive effects above, the government health expenditure is still insufficient, so the government transfer payment system should be improved to ensure accessibility to basic health services in underdeveloped areas.

This study shows that the policy effect of the PMC reform on the proportion of health expenditure to fiscal expenditure

TABLE 1 | Descriptive statistics of outcome and control variables in the DID model.

Variables	Treated					Control				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Healthexp	886	6,486.01	6,742.29	342	36,267	754	3,329.66	4,043.23	168	25,428
Healthperc	886	0.06	0.03	0.01	0.17	754	0.06	0.03	0.01	0.16
Capacity	888	0.31	0.16	0.06	2.05	1,008	0.29	0.14	0.05	1.31
AvgGDP	888	13,019.81	10,580.55	1,712	97,690.98	1,008	13,784.89	9,432.89	2,619.58	84,205.31
County2	888	0.28	0.45	0	1	1,008	0.13	0.33	0	1
Pop	888	66.26	39.34	8.32	220	1,008	39.74	16.90	11	121
Secindustry	888	331,480	431,992.8	10,300	4,744,940	1,008	271,758.1	248,796.5	8,716	2,097,890

Healthexp is the health expenditure of each county; Healthperc is the portion of health expenditure to fiscal expenditure; capacity is the self-financing capacity, namely, the proportion of county fiscal revenue to fiscal expenditure; AvgGDP is the per capita GDP; County2 is a dummy variable, "1" for county-level city and "0" for county; Pop is the population of each county; Secindustry is the second industry output.

TABLE 2 | Main results from the DID model.

Variables	Healthexp	Healthperc
Treatment	2,004.9**	0.00
	(6.99)	(0.59)
lnavgGDP	2,072.5	−0.02**
	(1.94)	(−5.56)
Lnpop	8,915.4*	0.07**
	(2.65)	(3.63)
Capacity	−3,067.9*	−0.01
	(−3.18)	(−0.47)
Lnsecindustry	703.0	0.00
	(1.24)	(1.37)
_cons	−58,112.4**	−0.03
	(−3.78)	(−0.31)
N	1,640	1,640

t statistics are in parentheses.
$^{*}p < 0.01$, $^{**}p < 0.001$.

of county-level governments is not significant. However, a study on 1,105 counties in the central and eastern provinces of China from 2002 to 2007 indicated that the PMC reform had a significant negative effect on the proportion of health expenditure to fiscal expenditure (32). A study on the impact of the PMC reform on the proportion of public education expenditure to the total expenditure in Henan Province also found a negative effect (35). The difference may result from the different samples used in these studies. Besides, according to the traditional decentralization theory, compared with the central government, local governments have the advantages of information and flexibility, and the provision of public services by local governments is more compatible with local residents' preferences and is more efficient (21, 36). People will respond by moving to a jurisdiction where the public provision level fits their preferences, but China has not only a household registration system that restricts the free movement of people but also a local tax system that is insensitive to the movement of people. Moreover, local governments do not have real tax legislative power in China. However, the potential benefits of decentralization also depend on the existence of decentralization of political decision-making authority, in particular effective channels for individuals to express their preferences and incentives for decision makers to respond to those preferences (37). For these reasons, many previous studies were skeptical of successful decentralization in developing countries. Obviously, these conditions do not exist in China at present. On contrary to other experiences in developing countries, political power has not been devolved yet in China: local government officials are not accountable to the local electorate but to higher-level government officials. In a word, China's fiscal federalism deviates considerably from the textbook case and thereby may yield quite different results.

Fiscal decentralization intensifies the competition among local governments and distorts their public expenditure structure, resulting in an increase in productive public service expenditure and a decrease in civilian public service expenditure (38). Fiscal decentralization can be divided into symmetric fiscal decentralization system and asymmetric fiscal decentralization system according to whether financial power and administrative power are symmetric. According to the proportion of fiscal revenue and fiscal expenditure between the central and local governments in China, fiscal decentralization in China is obviously asymmetric decentralization. The serious mismatch between financial power and administrative power also makes local governments more willing to spend their limited fiscal revenue on productive expenditure that is beneficial to them. It was also proven by the study results that the impact of asymmetric decentralization on health output in Papua Province was negative (39). Thus, only when the reforms of the administrative system and the fiscal system are carried out at the same time and perfect supervision mechanism is established, can the ideal reform effect be achieved. It is important to further improve the division of financial power and administrative power of the central and local governments at all levels. Besides, provincial governments need to strengthen the supervision and management of county-level governments and limit the irrational behavior of local governments with regard to fiscal expenditure.

The assessment of local officials focuses on indicators, such as GDP, fiscal revenue, and infrastructure (40, 41). As a result, local officials are more interested in repeated investment in

TABLE 3 | Robustness results from different specifications.

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Healthexp	Healthexp	Healthexp	Healthexp	Healthexp	Healthexp
Treatment	2,205.52*** (292.11)	2,004.90*** (286.72)	1,731.33*** (335.75)	1,442.01*** (392.02)	1,740.29*** (459.53)	2,062.57** (870.04)
lnavgGDP		2,072.50* (1,070.65)	119.10 (1,186.83)	1,014.31 (1,813.69)	−2,956.27* (1,515.01)	
Lnpop		8,915.43*** (3,360.99)	6,531.09* (3,815.37)		7,590.65 (5,879.76)	10,327.85 (9,281.13)
Lnsecindustry		702.98 (568.05)	2,450.51*** (738.22)	429.29 (828.57)	3,100.46*** (901.13)	−280.57 (396.90)
Capacity		−3,067.85*** (966.16)	−4,366.55** (1,726.93)	−1,696.22 (1,202.73)		−591.29 (2,760.82)
Constant	1,147.30*** (126.16)	−58,112.40*** (15,362.77)	−51,274.07*** (17,545.79)	−11,695.55 (11,904.66)	−36,462.23 (24,510.04)	−33,773.82 (34,821.08)
Observations	1,640	1,640	1,281	663	909	559
R-squared	0.78	0.79	0.78	0.87	0.75	0.79
Number of county	237	237	190	94	177	154
r2_w	0.78	0.79	0.78	0.87	0.75	0.79

Robust standard errors are in parentheses.

****p < 0.01, **p < 0.05, *p < 0.1.*

Columns (1) and (2) are the regression results after adding control variables. Column (1) only has a policy variable indicating the PMC reform dummy. Column (2) includes four variables of lnavgGDP, lnpop, lnsecindustry, and capacity, in which lnavgGDP is the per capita GDP (avgGDP) transformed to logarithm; and lnpop and lnsecindustry are the population and second industry output transformed to logarithm, respectively. Columns (3)–(6) are the results of hierarchical regression, where column (3) is the regression result of counties not including county-level cities; column (4) is the regression result of counties with pop >52.16; column (5) is the regression result of counties with capacity <0.30; and column (6) is the regression result of counties with avgGDP >13,426.70. All results are rounded up to two decimal places.

infrastructure and other "achievement projects" during their term of office, instead of education, health, culture, and other public services, which ultimately lead to difficulties in the transformation of a local finance system from "constructive finance" to "public finance" (42). Therefore, whether the PMC reform has no effect or a significant negative effect, it reflects the following facts. To a certain extent, the reform gives local governments greater fiscal autonomy, improves the enthusiasm of county-level governments to develop the county's economy (43), and promotes improvement in county-level governments' health expenditure (44). However, the proportion of health expenditure to fiscal expenditure has not increased (32), and local governments lack the motivation to improve the supply of health services (45). Local governments still prefer a fiscal expenditure structure that emphasizes production over people's livelihood (31). In conclusion, the reform does not change the essence of local governments' political championship. Therefore, it is necessary to adjust the incentive mechanism of county-level governments and improve the promotion system for local government officials. It is suggested to include the evaluation indexes of public services, such as education and health, into the government performance evaluation system appropriately in order to gradually reverse the expenditure tendency of "emphasizing production, neglecting people's livelihood."

Compared with previous studies on the influencing factors of government health expenditure from the perspective of macro-economy, the major contribution of this study is to study from the perspective of the fiscal system. What makes China's experience somewhat unique worldwide is the depth of fiscal decentralization on expenditure, in contrast to the recentralization of revenue since the tax-sharing system reform in 1994 (5). Comparing with previous studies on the impact of the PMC reform, we specifically studied the impact of the PMC reform on the health sector and focused on two indicators—health expenditure and its proportion to fiscal expenditure. We studied the impact of the PMC reform on not only the government expenditure but also the structure of government expenditure, which is more comprehensive. The three provinces had implemented the PMC reform earlier, so the duration is relatively long. Therefore, the research conclusion is more accurate considering the lag effect of reform.

However, this study has the following limitations. First, due to the limitation of data availability, this study only selected the three provinces of Jilin Province, Hebei Province, and Anhui Province as the research sample, which makes it difficult to represent the situation of the 24 provinces that are carrying out the PMC reform. Second, the fiscal expenditure data of counties in some years included in the study were not disclosed, and the missing proportion was 13.50%. In this study, cases containing missing values were excluded when the model was constructed, which weakens the reliability of the results. Third, the level of the aging population, changes in the spectrum of diseases of residents, and the implementation of the new rural cooperative medical care system in 2003 may have affected the level of health expenditure of county-level governments. However, due to the limited availability of data, this study did not incorporate them into the model as control variables. Therefore, future studies can increase the sample size, control other possible influencing

factors, and further explore the effect of the PMC reform on county-level health expenditure.

CONCLUSION

In this study, we evaluated the impact of the PMC decentralizing fiscal reform on health expenditure using longitudinal county-level data from 2003 to 2010 in Hebei Province, Anhui Province, and Jilin Province in China. Due to the different time points of reform, we use a time-varying DID method to estimate the effect of the PMC reform on health expenditure. We also controlled self-financing capacity, population, per capita GDP, and second industry output to account for any confounding effects on our estimates. The findings show that the fiscal system reform of the three provinces had a significantly positive effect on the health expenditure of county-level governments, whereas we found no effects on the proportion of health expenditure to fiscal expenditure of county-level governments. In order to change this situation, the government policy-making sectors need to perfect the supporting system while carrying out the PMC fiscal reform. There are four suggestions as follows: firstly, further specify the division of financial power and administrative power of the central and local governments at all levels; secondly, improve the government transfer payment system in health; thirdly, establish a scientific and rational government performance evaluation system and official promotion system; finally, strengthen the supervision and management of county-level government behaviors by provincial governments. These findings would serve as effective policy instruments aiming at achieving UHC by generating more additional resources for health sectors and minimizing the county-level disparity in the growth of health expenditure in China.

AUTHOR CONTRIBUTIONS

JZ contributed to the conception and design of the study. JH organized the database. CT performed the statistical analysis. JH, CT, JZ, and JJ wrote sections of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Natural Science Foundation of China (grant nos. 71974133, 71573182, and 71704143).

ACKNOWLEDGMENTS

The authors thank all research assistants for their works.

REFERENCES

1. Balakrishnan R. *Maximum Available Resources and Human Rights. Centre for Womens Global Leadership*. New Brunswick, NJ: Rutgers, The State University of New Jersey (2012).
2. Mcintyre D, Meheus F, Rottingen JA. What level of domestic government health expenditure should we aspire to for universal health coverage? *Health Econ Policy Law*. (2017) 12:125–37. doi: 10.1017/S1744133116000414
3. Yu Y, Zhang L, Li F, Zheng X. Strategic interaction and the determinants of public health expenditures in China: a spatial panel perspective. *Ann Reg Sci*. (2013) 50:203–21. doi: 10.1007/s00168-011-0488-7
4. Ding J. Experience, Realistic choice and future prospect of fiscal decentralization system in China. *Taxation Res*. (2010) 4:3–7 (in Chinese). doi: 10.19376/j.cnki.cn11-1011/f.2010.04.001
5. Hiroko U, Johannes PJ. Fiscal decentralization, Chinese style: good for health outcomes? *World Dev*. (2009) 37:1926–34. doi: 10.1016/j.worlddev.2009.06.007
6. Zhang S, Dai Y. Heterogeneity, fiscal decentralization and urban economic growth: a study based on panel quantile regression model. *J Finan Res*. (2012) 1:103–15 (in Chinese).
7. Wang W, Zheng X, Zhao Z. Fiscal reform and public education spending: a quasi-natural experiment of fiscal decentralization in China. *Publius*. (2012) 42:334–56. doi: 10.1093/publius/pjr039
8. Pang M. *Province-Managing-County: The Trend of the Reform of Regional Administrative System in China?* Chinese Public Administration. (2007) 6:21–5 (in Chinese). doi: 10.3782/j.issn.1006-0863.2007.06.007
9. Pang M, Li Y, Chen F. *Can Province-Managing-County solve the financial difficulties in the county and township?* Chinese Public Administration. (2009) 7:39–44 (in Chinese).
10. Lv K, He Q. Decentralization or recentralization: the impact of reforms of directly-governed counties on the inter-governmental fiscal relationship. *Mod Finan Econ J of Tianjin Univ Finan Econ*. (2016) 36:61–73 (in Chinese). doi: 10.19559/j.cnki.12-1387.2016.10.006
11. Chen G, Li Y. The innovation of local governments and the development of prosperous counties: a study based on Zhejiang's phenomenon. *J Zhejiang Univ*. (2009) 39:25–33 (in Chinese). doi: 10.3785/j.issn.1008-942X.2009.07.211
12. Fu B. The new test of Province-Managing-County reform in Zhejiang Province. *Decis Making*. (2006) 6:43–44 (in Chinese). doi: 10.3969/j.issn.1005-5940.2006.06.014
13. *Major Events of Fiscal Reform (2001-2007): Contributing to the Building of a Harmonious Socialist Society*. Available online at: http://www.mof.gov.cn/zhuantihuigu/czgg0000_1/gglc/200811/t20081107_88577.html (accessed April 12, 2021).
14. Huang B, Gao M, Xu C, Zhu Y. The impact of province-managing-county fiscal reform on primary education in China. *China Economic Review*. (2017) 45:45–61. doi: 10.1016/j.chieco.2017.06.001
15. Liu D. *Research on Province-Directly-Managing-County Fiscal System Reform* (in Chinese). Beijing: Capital University of Economics and Business (2013).
16. Cantarero D. Decentralization and health care expenditure: the Spanish case. *Appl Econ Lett*. (2005) 12:963–6. doi: 10.1080/13504850500377926
17. Rahman T. Determinants of public health expenditure: some evidence from Indian states. *Appl Econ Lett*. (2008) 15:853–7. doi: 10.1080/13504850600770970
18. Behera DK, Dash U. Impact of GDP and tax revenue on health care financing: An empirical investigation from Indian states. *Theoret Appl Econ*. (2017) XXIV:249–62.
19. He C. Study on the impact of public health expenditure of local government (in Chinese). *Finan Econ*. (2011) 4:94–100.
20. Lu W, Wang W. Study on the equalization of basic public services in China—an analysis based on the perspective of public demand and government capacity. *Public Finan Res*. (2008) 2:10–8 (in Chinese). doi: 10.19477/j.cnki.11-1077/f.2008.05.003
21. Oates WE. Oates: an essay on fiscal federalism an essay on fiscal federalism. *J Econ Literat*. (1999) 37:1120–49. doi: 10.1257/jel.37.3.1120
22. Bardhan P. Decentralization of governance and development. *J Econ Perspect*. (2002) 16:185–205. doi: 10.1257/089533002320951037
23. Weingast BR. Second generation fiscal federalism: the implications of fiscal incentives. *J Urban Econ*. (2009) 65:279–93. doi: 10.1016/j.jue.2008.12.005

24. Khaleghian P. Decentralization and public services: the case of immunization. *Social Sci Med*. (2004) 59:163–83. doi: 10.1016/j.socscimed.2003.10.013
25. Cantarero D, Pascual M. Analysing the impact of fiscal decentralization on health outcomes: empirical evidence from Spain. *Appl Econ Lett.* (2008) 15:109–11. doi: 10.1080/13504850600770913
26. Jiménez-Rubio D. Is fiscal decentralization good for your health? In: *Evidence From a Panel of OECD Countries. Health, Econometrics and Data Group (HEDG) Working Papers*. New York, NY: University of York (2010).
27. Samadi AH, Keshtkaran A, Kavosi Z, Vahedi S. The effect of fiscal decentralization on under-five mortality in Iran: a panel data analysis. *Int J Health Policy Manag.* (2013) 1:301–6. doi: 10.15171/ijhpm.2013.60
28. Peng C, Tang E. Competition among local governments: healthcare expenditure under fiscal decentralization. *J Finan Econ*. (2018) 44:94–108 (in Chinese). doi: 10.16538/j.cnki.jfe.2018.06.007
29. Sun R, Jin Y. Does fiscal decentralization improve healthcare outcomes? Empirical Evidence from China. *Public Finan Manage.* (2011) 11:234–61.
30. Li P, Lu Y, Wang J. Does flattening government improve economic performance? Evidence from China. *J Dev Econ.* (2016) 123:18–37. doi: 10.1016/j.jdeveco.2016.07.002
31. Liu J, Wu J, Wu J. The influence of the province-managing-county reform on public goods provision of counties: an empirical analysis of panel data from 136 counties and cities in Hebei Province. *Comp Econ Soc Syst*. (2012) 1:35–45 (in Chinese).
32. Chen S, Lu S. Does decentralization increase the public services expenditures:a quasi-experiment of "County Administrated by Province" in China. *China Econ Qu*. (2014) 13:1261–82 (in Chinese). doi: 10.13821/j.cnki.ceq.2014.04.001
33. Hausman J, Kuersteiner GM. Difference in difference meets generalized least squares: higher order properties of hypotheses tests. *J Econ*. (2008) 144:371–91. doi: 10.1016/j.jeconom.2008.04.003
34. Orley A, David C. Using the longitudinal structure of earnings to estimate the effect of training programs. *Rev Econ Stat*. (1985) 67:648–60. doi: 10.2307/1924810
35. Wang W, Zheng X, Zhao ZJ. Province-managing-county reform and public education spending: a quasi-natural experiment of fiscal decentralization in China. *Soc Sci Electron Pub*. (2011) 42:334–56. doi: 10.2139/ssrn.1935849
36. Stigler GJ. Perfect competition, historically contemplated. *J Polit Econ*. (1957) 65:1–17. doi: 10.1086/257878
37. Arends H. The dangers of fiscal decentralization and public service delivery: a review of arguments. *Politische Vierteljahresschrift*. (2020) 61:599–622. doi: 10.1007/s11615-020-00233-7
38. Keen M, Marchand M. Fiscal competition and the pattern of public spending. *J Public Econ.* (1997) 66:33–53.
39. Cahyaningsih A, Fitrady A. The impact of asymmetric fiscal decentralization on education and health outcomes: evidence from Papua Province, Indonesia. *Econ Sociol*. (2019) 12:48–63. doi: 10.14254/2071-789X.2019/12-2/3
40. Li H, Zhou LA. Political turnover and economic performance: the incentive role of personnel control in China. *J Public Econ.* (2003) 89:1743–62. doi: 10.1016/j.jpubeco.2004.06.009
41. Zhou LA. The incentive and cooperation of government officials in the political tournaments: an interpretation of the prolonged local protectionism and duplicative investments in China. *Econ Res J*. (2004) 6:33–40 (in Chinese).
42. Zhou LA. Governing China's local officials: an analysis of promotion tournament model. *Econ Res J*. (2007) 7:36–50 (in Chinese).
43. Zheng X, Wang H, Zhao Y. Can Province-Directly-Managing-County boost economic growth?– difference-in-Difference method. *Manag World*. (2011) 8:34–44+65 (in Chinese). doi: 10.19744/j.cnki.11-1235/f.2011.08.005
44. Li R, Wang W. The influence of decentralization reform of local governments on people-oriented public services——based on analysis of decentralization reform of counties directly managed by Province in Henan Province. *Soc Stud*. (2018) 2:65–76 (in Chinese).
45. Yin H, Zhu H. A study of productive expenditure bias in county-level finance in China. *Soc Sci China*. (2012) 33:127–47. doi: 10.1080/02529203.2012.650414

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

ORIGINAL RESEARCH
published: 04 June 2021
doi: 10.3389/fpubh.2021.682975

Dietary Salt Reduction, Prevalence of Hypertension and Avoidable Burden of Stroke in Vietnam: Modelling the Health and Economic Impacts

Leopold Ndemnge Aminde[1], Hai N. Phung[1], Dung Phung[1], Linda J. Cobiac[2] and J. Lennert Veerman[1]*

[1] School of Medicine, Griffith University, Gold Coast, QLD, Australia, [2] Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

Dietary salt reduction has been recommended as a cost-effective population-wide strategy to prevent cardiovascular disease. The health and economic impact of salt consumption on the future burden of stroke in Vietnam is not known.

Objective: To estimate the avoidable incidence of and deaths from stroke, as well as the healthy life years and healthcare costs that could be gained from reducing salt consumption in Vietnam.

Methods: This was a macrosimulation health and economic impact assessment study. Data on blood pressure, salt consumption and stroke epidemiology were obtained from the Vietnam 2015 STEPS survey and the Global Burden of Disease study. A proportional multi-cohort multistate lifetable Markov model was used to estimate the impact of achieving the Vietnam national salt targets of 8 g/day by 2025 and 7 g/day by 2030, and to the 5 g/day WHO recommendation by 2030. Probabilistic sensitivity analysis was conducted to quantify the uncertainty in our projections.

Results: If the 8 g/day, 7 g/day, and 5 g/day targets were achieved, the prevalence of hypertension could reduce by 1.2% (95% uncertainty interval [UI]: 0.5 to 2.3), 2.0% (95% UI: 0.8 to 3.6), and 3.5% (95% UI: 1.5 to 6.3), respectively. This would translate, respectively, to over 80,000, 180,000, and 257,000 incident strokes and over 18,000, 55,000, and 73,000 stroke deaths averted. By 2025, over 56,554 stroke-related health-adjusted life years (HALYs) could be gained while saving over US$ 42.6 million in stroke healthcare costs. By 2030, about 206,030 HALYs (for 7 g/day target) and 262,170 HALYs (for 5 g/day target) could be gained while saving over US$ 88.1 million and US$ 122.3 million in stroke healthcare costs respectively.

Conclusion: Achieving the national salt reduction targets could result in substantial population health and economic benefits. Estimated gains were larger if the WHO salt targets were attained and if changes can be sustained over the longer term. Future work should consider the equity impacts of specific salt reduction programs.

Keywords: sodium, blood pressure, stroke, mortality, multi-state model, healthcare costs

OPEN ACCESS

Edited by:
Hideki Hashimoto, The University of Tokyo, Japan

Reviewed by:
Magdalena Syrkiewicz-Switala, Medical University of Silesia, Poland
Amâncio António De Sousa Carvalho, University of Trás-os-Montes and Alto Douro, Portugal

****Correspondence:***
Leopold Ndemnge Aminde
l.aminde@griffith.edu.au

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *23 March 2021*
Accepted: *10 May 2021*
Published: *04 June 2021*

Citation:
Aminde LN, Phung HN, Phung D, Cobiac LJ and Veerman JL (2021) Dietary Salt Reduction, Prevalence of Hypertension and Avoidable Burden of Stroke in Vietnam: Modelling the Health and Economic Impacts. Front. Public Health 9:682975. doi: 10.3389/fpubh.2021.682975

INTRODUCTION

Hypertension is the single largest risk factor for cardiovascular disease (CVD), the leading cause of death of globally (1). A global synthesis of national population-based surveys found that the number of people with hypertension has more than doubled over the last four decades, with an estimated 1.1 billion people with the condition in 2015 (2). The majority of affected people are in the low-income and middle-income countries (LMIC), where increasing trends in age-standardised mean systolic blood pressures (SBP) have been found (2). In Vietnam, a recent systematic review of population-based studies of measured hypertension revealed that over one in five (21.1%) adults had hypertension (3). However, using a broader definition including measured SBP $\geq$ 140/90 mmHg and/or self-report physician diagnosed and/or currently on blood pressure lowering medication, the national WHO STEPS survey reported a prevalence of 30.6% in 30–69-year-old adults (4). In general, hypertension is more common in men, increases with age, with prevalence over 40% in those aged 40 years and more (4, 5). A significant proportion of people with hypertension in Vietnam are undiagnosed, with studies reporting low awareness, treatment, and control rates (5, 6).

Hypertension is described as a silent killer because it generally has no symptoms. By the time an individual is symptomatic, they most likely have developed one or more complications. These may include myocardial infarction, stroke, hypertensive heart disease, heart failure, chronic kidney disease and retinopathy (1). In Vietnam, stroke is the most frequent of these hypertension-related complications. According to a hospital-based retrospective analysis in central Vietnam, the age-standardised incidence rate of first ever stroke was 115.7 per 100,000 persons between 2009 and 2011 (7). Stroke is also the leading cause of death in the country. In 2019, over one in five (21.5%) deaths, that is, 136,000 deaths in Vietnam were due to stroke, representing an almost 2% increase from two decades earlier (8). In addition, stroke was responsible for over three million disability adjusted life years in 2019 in Vietnam, representing a 13% increase compared to a decade earlier (8). The high rates of stroke incidence and mortality in Vietnam have mostly been attributed to undiagnosed and poorly treated or uncontrolled hypertension (7).

Among the myriad of modifiable risk factors known to increase an individual's risk of developing hypertension, excessive sodium (salt) consumption is a major contributor. By raising an individual's blood pressure, it increases the risk of developing cardiovascular events (9). Mozaffarian et al. showed that sodium intake above 2 grammes/day was responsible for over 1.6 million CVD deaths worldwide in 2010, majority from stroke and ischaemic heart disease (10). Salt consumption globally is well above recommended levels. Asian countries have some of the highest levels of salt consumption. In Vietnam, a national survey using spot urine samples estimated that the average salt intakes in adults was around 9.4 grammes per day (10.5 g/day in men and 8.3 g/day in women) (4). This is nearly twice the current World Health Organisation recommended consumption of <5 g of salt per day (11).

There is strong evidence that reducing average salt intake decreases blood pressure which could prevent thousands of future cardiovascular events across populations (9, 12). A recent comprehensive review on the economic evaluation of strategies to prevent CVD in LMICs revealed that population-wide salt reduction was cost-saving. However, the authors underscored the need for more country-specific studies, given the dearth in such studies especially from Africa and Asia (13). Studies from high-income countries have shown that salt reduction at population levels saves thousands of lives and has good value for money (14, 15). However, very few of such studies have been conducted in low-income and middle-income countries (16). Given the very high levels of salt intake, as well as high rates of stroke events in Vietnam, it is likely that such strategies could yield significant health and economic benefits.

In 2018, the Prime Minister of Vietnam approved "The Healthy Vietnam Program" whose general objective is to "…improve the health, stature, longevity, and quality of life of the Vietnamese people" (17). Among the targets outlined in the first objective of this program, is to reduce salt consumption per person per day to <8 grammes by 2025 and <7 grammes by 2030. This study therefore seeks to quantify: (i) the new cases of and deaths from (absolute and relative) stroke that would be prevented; (ii) the number of health adjusted life years that would be gained; and (iii) the future healthcare costs from stroke cases that could be avoided, if these salt reduction targets were achieved.

MATERIALS AND METHODS

Study Population and Data Sources

This study included the entire cohort of Vietnamese adults aged $\geq$ 25 years in 2019 (base year), which included 61 million people (29.5 million men and 31.5 million women). Persons below 25 years were excluded given the comparatively low rates of blood pressure-related stroke events, and the absence of reliable data on the risk-outcome relation between blood pressure and stroke in this sub-population.

Data on blood pressure and salt intake was obtained from the 2015 National Survey on Risk factors of Non-communicable Diseases for Vietnam (4). This nationally representative survey used a complex multi-stage sampling process to include adults aged $\geq$18 years from across all 63 provinces/cities of Vietnam. Standard blood pressure measurement procedures were used, and 24 h urinary sodium was estimated using the INTERSALT formula from the spot urine samples collected. This survey applied the methods and tools of the World Health Organisation (WHO) STEPwise approach to NCD surveillance (STEPS) and further details can be found here (4). Data on relative risks as well as baseline epidemiological data (incidence, prevalence, mortality) for stroke in Vietnam were obtained from the Global Burden of Disease (GBD) 2019 study (18). For healthcare costs, data was from a cohort study (19) conducted at the reference Stroke Unit of the 115 People's Hospital. This is a major teaching hospital in Ho Chi Minh city and the study recruited first ever stroke patients with complete electronic records on direct medical costs. A bottom-up costing approach was used

in this study and captured the following cost components: consultant fees, bed-day fees, diagnostic imaging, laboratory tests, medication, rehabilitation, minor procedures, medical consumables, and special meals/feeding.

Modelling Framework

The multiple cohort proportional multistate lifetable (PMSLT) Markov model was used for this analysis (20). It contains three major linked parts: the risk factor (exposure) section, the main lifetable, and the disease Markov sub-models. This model framework is unique given its ability to deal with comorbidity and has been used previously in similar NCD preventive modelling (21, 22). A three-staged process is used: first, modelling the impact of sodium (salt) on blood pressure; second, the impact of changes in blood pressure on stroke incidence and mortality; and third, the impact of changes in mortality on (healthy) life years.

Risk Factor Modelling

We used evidence from a meta-analysis of long-term trials (23) on the relationship between salt intake and blood pressure to model the impact of changes (reduction) in population salt intake on mean blood pressure. In this study, a 4.4 grammes reduction in salt intake led to a 4.18 mmHg reduction in systolic blood pressure. This meta-analysis was used given that it included trials with a minimum duration of 1 month, which is the recommended duration when assessing the public health impact of salt reduction (9, 12). We accounted for the differential impact of salt reduction for people with hypertension and those without. Assuming a normal distribution, we estimated the changes in prevalence of hypertension from shifting the blood pressure distribution with and without salt reduction. The "distributions shift" method of the potential impact fraction (PIF)—was used to calculate the proportional change in the incidence of stroke due to changes in the risk factor (blood pressure) distribution. For this computation, we used relative risks from the GBD linking SBP and both forms of stroke, as shown in the **Box 1** below. The shift in the population's SBP distribution leads to the re- estimation of stroke incidence, which feeds into the disease Markov models described below.

Disease Markov and Multistate Lifetable Modelling

For this study, the two major forms of stroke—ischaemic and haemorrhagic, were modelled. The state-transition Markov model shown in **Figure 1**, demonstrates the health states considered for each of the diseases. Five health states were considered: *Healthy* (alive without stroke), *Acute death* (dead within the first 28-days of a stroke event), *Diseased* (alive with "chronic" stroke, i.e., survivors of acute stroke death), *Dead from disease* (death from "chronic" stroke, i.e., stroke death among survivors of acute death) and *dead from other causes*. Transition hazards such as incidence, 28-day case fatality, case fatality (excess mortality) in month 2 to 12, and all-cause mortality influenced the movement of proportions of the population between these five health states, with death being an absorbing state. The DISMOD II software was used to derive other epidemiological parameters like case fatality that are seldomly reported, while maintaining consistency in the data (24).

The PMSLT Markov model has two identical populations stratified by sex and age. It has an "intervention" population that receives the intervention and a reference population that continues with current risk exposures as usual. As previously described, shifts in the risk factor distribution due to the PIF result in new disease incidence. This change in incidence leads to changes in prevalence and then mortality. The resulting disease mortality then feeds into the lifetable altering overall mortality rates and explicitly calculating life years. To obtain health-adjusted life years (HALYs), the years of life are adjusted for the poor quality of life due to the disease (in this case, stroke) being modelled and comorbidity using weights from the GBD (18). The PMSLT simulates these age-group and sex cohorts of the population until everyone dies or reaches the age of 100 years. Healthcare costs due to disease prevented are explicitly modelled over the respective time-horizons. In this study, costs were linked to each incident stroke, as our healthcare cost estimates pertained to the cost of care for first-ever stroke event. The differences in health outcomes (changes in blood pressure, stroke incidence, deaths, and HALYs) and healthcare costs between the two identical populations reflects the impact of the intervention. Health and economic outcomes were discounted at 3%.

Scenario and Uncertainty Analysis

The base case analysis assessed the changes in health and economic outcomes if everyone reduced their salt to 8 g/ day by 2025 (scenario 1) and 7 g/day by 2030 (scenario 2), both discounted at 3% (25). In a third scenario, we modelled the impact of a gradual reduction of salt intake down to 5 g/day (WHO target) by 2030 (scenario 3). For sensitivity analyses, we modelled longer time horizons, i.e., gradual reduction to 8 g/day

BOX 1 | Distribution shift potential impact fraction formula.

$$PIF = \frac{\int_a^b RR(x)P(x)dx - (\int_{a1}^{b1} RR(x)P*(x)dx + \int_{a2}^{b2} RR(x)P*(x)dx)}{\int_a^b RR(x)P(x)dx}$$

Where, x = SBP exposure levels, RR(x) = relative risk, P(x) = original SBP distribution, P∗ = SBP distribution after the intervention, a = start integration limit (90 mmHg), b = end integration limit (220 mmHg), a1 = start integration limit for normotensive people (90 mmHg), b1 = end integration limit for normotensive people (139.9 mmHg), a2 = start integration limit for hypertensive people (140 mmHg), b2 = end integration limit for hypertensive people (220 mmHg). A theoretical minimum SBP of 115 mmHg, which is the lowest level considered for elevated CVD risk.

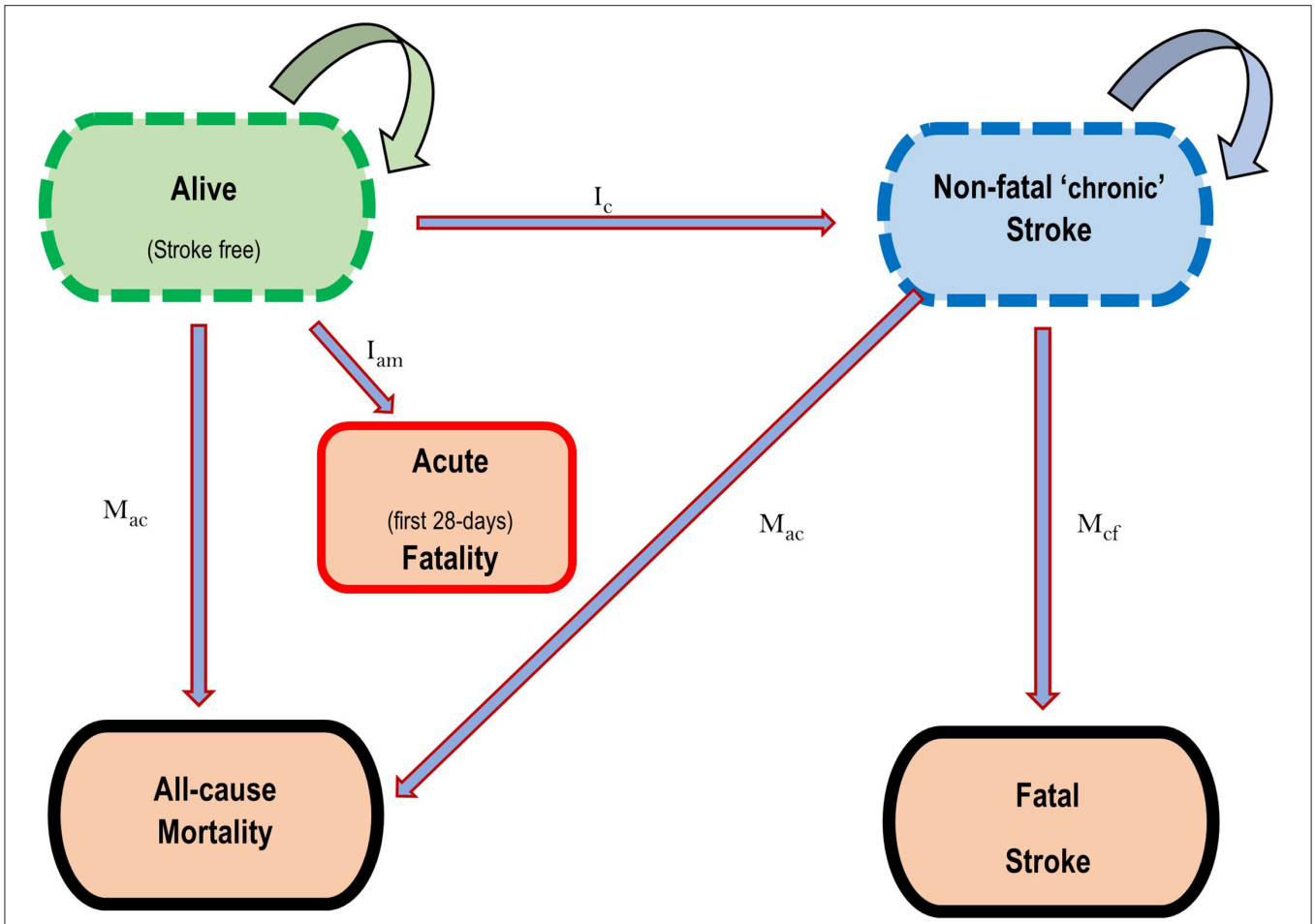

FIGURE 1 | State transition diagram depicting the five health states for the Stroke Markov model. $\mathbf{I}_c$ = Incidence of Stroke in the first year, $\mathbf{I}_{an}$ = Case fatality in the first 28 days following incident CVD, $\mathbf{M}_{cf}$ = Case fatality in the first year for survivors of first 28-days Stroke mortality, $\mathbf{M}_{ac}$ = Mortality from all causes in that year. Straight arrows represent directions of movement of proportions of the population between health states, while circular arrows represent the probability of remaining in each health state. Death is an absorbing state.

to 2025 and sustained over the remaining lifetime (scenario 4); gradual reduction to 7 g/day by 2030 and sustained over the remaining lifetime (scenario 5); gradual reduction to 5 g/day by 2030 and sustained over the remaining lifetime (scenario 6). Analyses were also rerun at varying rates of intervention decay, i.e., (i) scenario 4 above with full impact sustained for 5 years and decline to 25% 5-yearly thereafter down to zero impact beyond the year 2045 (scenario 7); (ii) scenario 5 with full impact in first 5 years, and declining by 25% every 5 years down to zero impact by 2050 (scenario 8); (iii) scenario 6 with full impact in first 5 years, then decline by 25% 5 yearly down to 0% beyond year 2050 (scenario 9). The **Figure 2** below depicts the modelled scenarios.

For parameter estimates with uncertainty, we applied the following distributions: normal (salt- BP relationship), lognormal (relative risks), and gamma (healthcare costs). Parameters were drawn randomly from these distributions in probabilistic sensitivity analyses implemented using the Ersatz software (26), with 2,000 simulations executed.

RESULTS

Estimated Effects on Blood Pressure and Hypertension

With 2019 as baseline year, our model estimates that a gradual reduction (on average, 0.383 and 0.033 g annually for 30–49-year-old men and women, respectively) of salt consumption to 8/day by 2025 could reduce mean SBP by 1.9 and 0.12 mmHg in men and women, respectively. This translates to a reduction in hypertension prevalence of 2.3% (671,000 men) and 0.2% (68,000 women) by 2025. Furthermore, gradual reductions to 7 g/day (on average, 0.300 and 0.109 g/day annually for 30–49-year-old men and women, respectively) by 2030 could lead to 927,000 (3.1%) fewer men and 284,000 (0.9%) fewer women with hypertension. If salt consumption were to go down to 5 g/day (that is, annual reductions of 0.482 and 0.291 g/day for 30–49-year-old men and women, respectively) by 2030, mean SBP could reduce by about 4.0 mmHg for men and 2.2 mmHg for women, translating to over 1.4 million (4.9%)

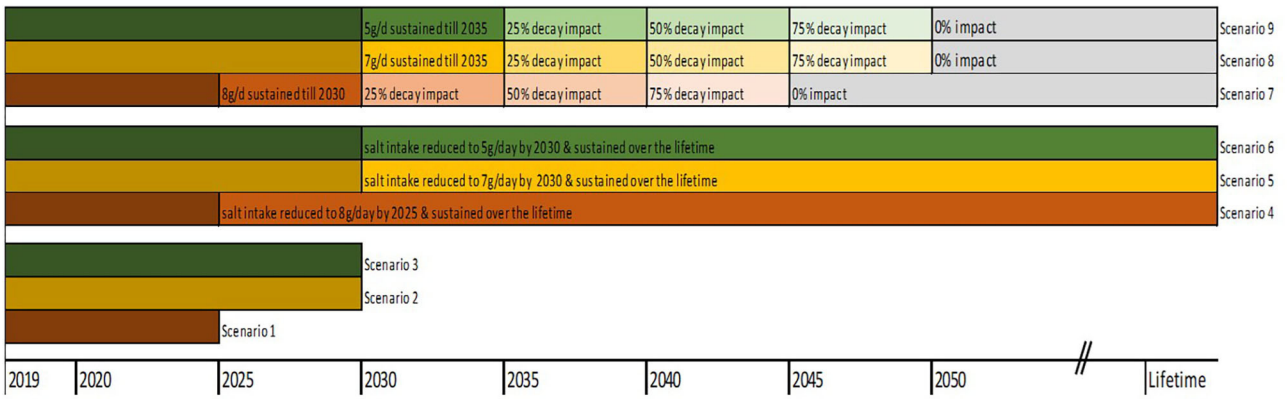

FIGURE 2 | Baseline and scenario analyses of salt reduction targets.

TABLE 1 | Estimated reductions in mean systolic blood pressure, the proportion and number of people with hypertension in Vietnam.

	Men			Women			Both		
	Mean (95%UI), mmHg	**Prevalence, % (95%UI)**	**Number In 000s (95%UI)**	**Mean (95%UI), mmHg**	**Prevalence, % (95%UI)**	**Number In 000s (95%UI)**	**Mean (95%UI), mmHg**	**Prevalence, % (95%UI)**	**Number, in 000s (95%UI)**
Scenario 1	1.9	2.3	671	0.1	0.2	68	1.0	1.2	739
	(1.4 to 2.7)	(0.8 to 4.5)	(222 to 1,332)	(0.1 to 0.2)	(0.0 to 0.6)	(2 to 205)	(0.8 to 1.4)	(0.5 to 2.3)	(282 to 1,402)
Scenario 2	2.7	3.1	927	0.7	0.9	284	1.7	2.0	1,211
	(1.9 to 3.8)	(1.0 to 6.3)	(287 to 1,867)	(0.6 to 1.2)	(0.2 to 2.1)	(55 to 651)	(1.3 to 2.4)	(0.8 to 3.6)	(501 to 2,178)
Scenario 3	4.0	4.9	1,432	2.2	2.3	726	3.2	3.5	2,158
	(3.1 to 5.9)	(1.3 to 9.8)	(395 to 2,894)	(1.7 to 3.3)	(0.5 to 5.3)	(142 to 1,674)	(2.4 to 4.3)	(1.5 to 6.3)	(886 to 3,872)

Scenario 1: Gradual reduction in salt intake to 8 g/day between 2019 and 2025; Scenario 2: Gradual reduction in salt intake to 7 g/day between 2019 and 2030; Scenario 3: Gradual reduction in salt intake to 5 g/day (WHO target) between 2019 and 2030. UI, uncertainty interval.

fewer men and 726,000 (2.0%) fewer women with hypertension (**Table 1**).

Projected Changes in Stroke Incidence

A gradual reduction in salt intake down to 8 g/day by 2025 could cumulatively avert 38,422 (95% uncertainty interval [UI]: 19,265–62,531) new cases of ischaemic stroke and 41,592 (95%UI: 21,063–66,829) new cases of haemorrhagic stroke in Vietnam. In addition, a gradual reduction in salt intake down to 7 g/day by 2030 could avert 87,955 (95% UI: 48,385–133,216) incident cases of ischaemic stroke and 92,816 (95% UI: 53,903–139,931) incident haemorrhagic strokes. If the WHO target of 5 g/day were achieved by 2030, there could be 127,522 (95% UI: 79,694–179,775) fewer new cases of ischaemic stroke and 130,150 (95% UI: 82,918–180,542) fewer new cases of haemorrhagic stroke (**Table 2**). In all three scenarios, reductions in incident strokes were greater in men compared to women (Scenario 1: 6.8% vs. 3.4% for ischaemic and 8.7% vs. 4.0% for haemorrhagic stroke; Scenario 2: 8.2% vs. 4.4% for ischaemic and 10.5% vs. 5.2% for haemorrhagic stroke; Scenario 3: 11.2% vs. 7.0% for ischaemic and 14.1% vs. 8.4% for haemorrhagic stroke (**Figure 3**).

Projected Changes in Stroke Mortality

Our model projects that a gradual reduction of population salt intake to 8 g/day by 2025, could lead to 3,663 (95% UI: 1,415–6,568) fewer deaths (1.4% reduction) from ischaemic stroke and 14,498 (95% UI: 6,701–24,304) fewer deaths (2.1% reduction) from haemorrhagic stroke in Vietnam. If the current salt consumption were gradually reduced to 7 g/day by 2030, there could be 13,030 (95% UI: 6,060–22,023) and 42,227 (95% UI: 23,259–66,480) deaths avoided from ischaemic (2.3% reduction) and haemorrhagic (4.0% reduction) stroke, respectively. A larger number of fatal events, that is, 17,249 (95% UI: 9,048–27,155) for ischaemic stroke (3.1% reduction) and 56,209 (33,117–81,734) for haemorrhagic stroke (5.3% reduction) could be avoided if average population salt consumption were reduced to 5 g/day by 2030 (**Table 3**). The estimated reductions in stroke mortality were greater for men than women (Scenario 1: 1.7% vs. 1.1% for ischaemic and 2.5% vs. 1.3% for haemorrhagic; Scenario 2: 2.8% vs. 1.7% for ischaemic and 4.8% vs. 2.4% for haemorrhagic; Scenario 3: 3.5% vs. 2.4% for ischaemic and 6.2% vs. 3.5% for haemorrhagic (**Figure 3**).

TABLE 2 | Absolute reductions in incident strokes among adults in Vietnam from 2019 to 2030 and over the lifetime.

Base analysis	Ischaemic stroke			Haemorrhagic stroke		
	Men *N* (95% UI)	Women *N* (95% UI)	Both *N* (95% UI)	Men *N* (95% UI)	Women *N* (95% UI)	Both *N* (95% UI)
Scenario 1	26,026 (10,832–45,967)	12,396 (3,147–26,644)	38,422 (19,265–62,531)	33,455 (13,535–58,264)	8,136 (2,385–16,453)	41,592 (21,063–66,829)
Scenario 2	58,250 (26,879–98,940)	29,705 (10,496–57,787)	87,955 (48,385–133,216)	73,469 (37,160–119,247)	19,348 (7,637–34,747)	92,816 (53,903–139,931)
Scenario 3	79,767 (43,483–119,706)	47,755 (22,895–79,900)	127,522 (79,694–179,775)	99,031 (57,496–146,539)	31,118 (16,561–49,017)	130,150 (82,918–180,542)
Sensitivity analysis						
Scenario 4	281,789 (122,584–474,551)	115,477 (25,682–250,683)	397,267 (206,737–623,454)	297,655 (138,099–473,912)	75,674 (17,456–157,881)	373,328 (200,838–558,219)
Scenario 5	347,259 (173,340–542,156)	179,595 (62,988–321,514)	526,854 (308,774–757,595)	360,760 (192,170–544,478)	116,356 (42,151–200,810)	477,116 (284,050–670,203)
Scenario 6	525,533 (318,991–760,985)	361,890 (195,316–555,027)	887,424 (597,364–1,185,196)	536,155 (341,998–746,822)	232,423 (124,758–350,976)	768,577 (526,175–1,019,372)
Scenario 7	108,021 (39,548–188,715)	48,833 (10,682–103,480)	156,854 (74,052–253,348)	144,293 (63,773–235,470)	33,594 (9,907–64,083)	177,887 (92,147–271,777)
Scenario 8	159,725 (71,840–260,121)	77,094 (20,222–151,785)	236,819 (133,798–360,652)	201,038 (107,628–312,018)	52,323 (18,106–94,975)	253,361 (154,421–366,041)
Scenario 9	221,723 (119,686–336,699)	129,746 (54,903–221,746)	351,469 (210,881–493,588)	279,785 (167,125–397,474)	87,231 (41,393–141,495)	367,016 (241,485–504,246)

Scenario 1 = Gradual reduction in salt intake to 8 g/day between 2019 and 2025; Scenario 2 = Gradual reduction in salt intake to 7 g/day between 2019 and 2030; Scenario 3 = Gradual reduction in salt intake to 5 g/day (WHO target) between 2019 and 2030; Scenario 4 = Scenario 1 achieved and sustained for the remaining lifetime; Scenario 5 = Scenario 2 achieved and sustained for the remaining lifetime; Scenario 6 = Scenario 3 achieved and sustained for the remaining lifetime; Scenario 7 = Scenario 1 achieved and effect gradually phases out over 20 years with zero effect beyond 2045; Scenario 8 = Scenario 2 achieved and effect gradually phases out over 20 years with zero effect beyond 2050; Scenario 9 = Scenario 3 achieved and effect gradually phases out over 20 years with zero effect beyond 2050; N, Number; UI, uncertainty interval.

Health-Adjusted Life Years and Healthcare Costs

Overall, a gradual reduction in dietary salt intake to 8 g/day by 2025 could result in 56,554 (95% UI: 25,966–93,598) health-adjusted life years (HALY) gained while reductions to 7 g/day by 2030 could result in 206,030 (95% UI: 108,209–334,276) HALYs gained. Furthermore, our model projects that gradual reductions to 5 g/day by 2030 could gain 262,170 (95% UI: 147,799–389,631) stroke-related HALYs in Vietnam (**Figure 4**).

Averted new cases of stroke are projected to result in savings in healthcare costs. Scenario 1 was projected to save US$ 18.5 (95% UI: 4.3–43.1) million cumulatively for first ever ischaemic stroke and US$ 24.1 (95% UI: 7.0–58.5) million for first ever haemorrhagic stroke treatments by 2025. Scenario 2 was projected to save US$ 38.8 (13.0–85.7) million for ischaemic stroke and US$ 49.3 (15.3–110.3) million for haemorrhagic stroke while scenario 3 could lead to savings of US$ 55.7 million and US$ 66.6 million for first ever ischaemic and haemorrhagic stroke treatments cumulatively by 2030 (**Table 4**).

Sensitivity Analyses

In addition to the 6 years (2019–2025) and 11 years (2019–2030) time horizons in the base analyses, six additional scenarios were run with lifetime horizons. As expected, for scenarios 4 to 6 in which the 2025 and 2030 salt target levels were sustained for the remaining lifetime, larger impacts were observed ranging from about 771,000 to 1.6 million cumulative new cases of both forms of stroke and about 475,000 to over a million cumulative stroke deaths averted. For scenarios 7 to 9, in which we gradually phased out the intervention effect with zero impact beyond 20- and 25-years post target achievement, health gains were almost halved compared to the sustained scenarios, but these were still about 2 to 4 times the benefits observed in the base analysis. Similar patterns were observed in for HALYs and healthcare cost offsets (**Tables 2**, **3** and **Supplementary File**).

DISCUSSION

We evaluated the potential impact of attaining three population dietary salt reduction targets stipulated in the Vietnam national and WHO recommendations. Our model projects that through modest reductions in mean systolic blood pressure, gradual reductions in average population salt consumption could avert between 80,000 to 250,000 non-fatal and 18,000 to 73,000 fatal strokes over six- to 11-year time horizons. Furthermore, there were large gains in healthy life years and savings in acute stroke direct health expenditures. These benefits were substantially larger (∼6 times baseline and ∼28 times baseline scenarios for incidence and mortality, respectively) if sustained over the lifetime. However, these lifetime impacts were reduced

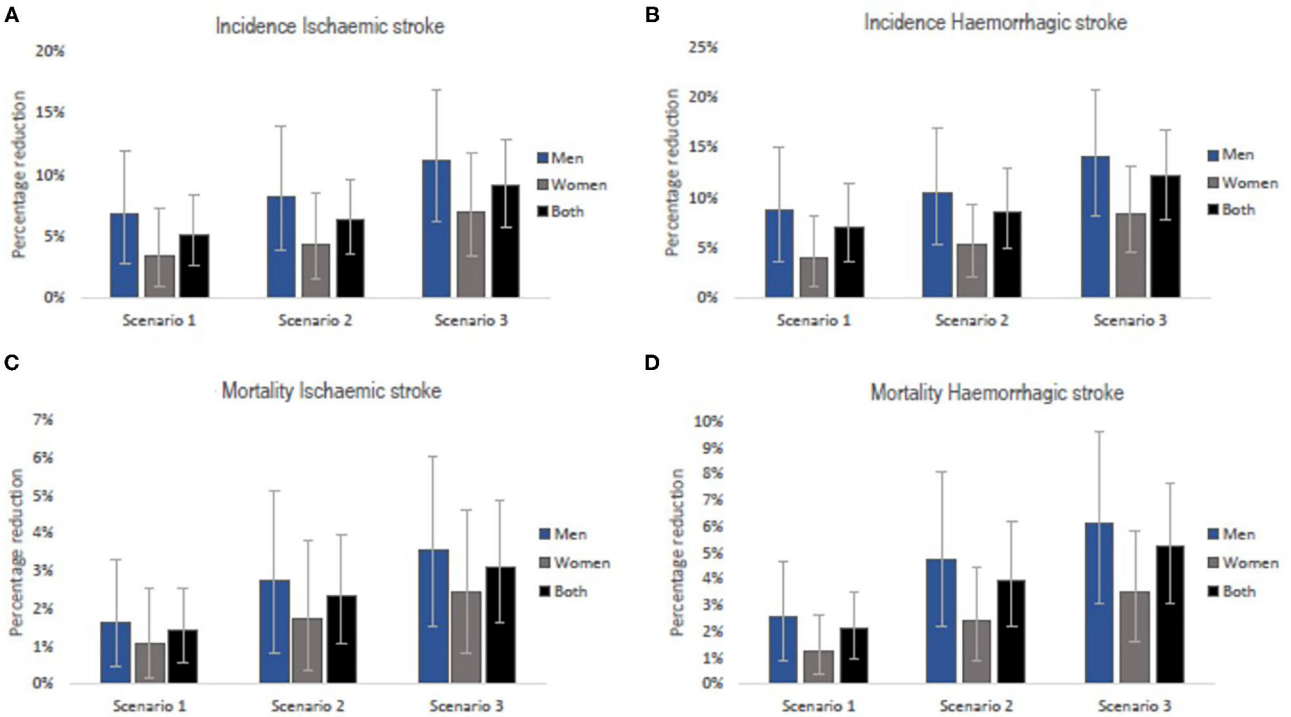

FIGURE 3 | Relative reductions in stroke incidence rates **(A,B)** and mortality rates **(C,D)** if the national (Scenarios 1 and 2) and WHO (Scenario 3) salt targets were achieved in Vietnam. The bars are the best estimates while the whiskers are the 95% uncertainty intervals, i.e., 2.5 and 97.5 percentiles.

(correspondingly ∼3 times to ∼11 times the baseline scenarios) if changes were not maintained. Compared to women, men were the bigger beneficiaries of these impacts.

Comparisons and Implications of the Findings

In a previous modelling study evaluating a suite of interventions for CVD prevention in Vietnam (27), authors found that reducing population salt intake via media campaign could avert 45,939 DALYs annually. Results from our base analysis ranged from 9,425 to 23,833 HALYs annually. On par value, this could appear lower than the Ha *et al.* findings. However, our HALYs are essentially stroke-related, as opposed to the former study which included ischaemic heart disease and stroke. It is also possible that disparities in modelling methods (28) might explain some of these differences. Basu et al. modelled the impact of a 3 g/day reduction in salt intake in India over a 30-year time horizon (29). They found that such an intervention could reduce stroke incidence rates by 9–13%, which appears on par with our work. As in our study, research from Brazil (30) and Cameroon (31) has similarly modelled the potential impact of attaining the WHO target of 5 g/day of salt intake, and equally found large reductions in CVD events and premature mortality (31, 32). Consistent with previous studies (14, 21), benefits were greater in men than in women. Evidence has shown that the impact of salt reduction increases linearly with blood pressure, with a steeper gradient in those with mean SBP > 140 mmHg and those with higher baseline salt intake (9). The Vietnam 2015 STEPS survey used in our analysis showed that men had higher average salt consumption levels and higher mean blood pressures. It is therefore unsurprising that they were more likely to accrue greater gains than the women. Given that we assume salt reduction is proportional to current consumption, hence the higher consumption in men translated to comparatively larger absolute reductions in salt intake required to meet the national targets, and therefore more gains in health impacts.

Our model projects slightly larger reductions in incidence of and mortality from haemorrhagic stroke compared to ischaemic stroke. Substantial evidence exists on the link between blood pressure and risk of stroke, demonstrating greater relative risk of haemorrhagic stroke per unit increase in blood pressure compared to ischaemic stroke (33, 34). We modelled the impact of salt reduction on stroke in a two-step process, that is, mediated via blood pressure. As a result, reduction in average population blood pressure is likely to lead to slightly bigger reductions in incidence of haemorrhagic stroke compared to ischaemic stroke. In addition, data from stroke units in Vietnam used in our study showed that acute case fatality rates in the first 28 days were higher (almost double for some age groups) for people with haemorrhagic stroke compared to those with ischaemic stroke (19). These disparities in risk of and death from either form of stroke in part might explain the slightly larger reductions in incidence and deaths from haemorrhagic stroke observed in our projections.

TABLE 3 | Absolute reductions in stroke mortality among adults in Vietnam from 2019 to 2030 and over the lifetime.

Base analysis	Ischaemic stroke			Haemorrhagic stroke		
	Men *N* (95% UI)	Women *N* (95% UI)	Both *N* (95% UI)	Men *N* (95% UI)	Women *N* (95% UI)	Both *N* (95% UI)
Scenario 1	2,521 (729–5,066)	1,143 (153–2,680)	3,663 (1,415–6,568)	11,481 (3,925–21,329)	3,017 (829–6,179)	14,498 (6,701–24,304)
Scenario 2	9,072 (2,724–16,913)	3,959 (844–8,674)	13,030 (6,060–22,023)	33,415 (15,265–57,021)	8,812 (3,301–16,253)	42,227 (23,259–66,480)
Scenario 3	11,711 (5,104–19,939)	5,539 (1,819–10,506)	17,249 (9,048–27,155)	43,243 (21,624–67,494)	12,966 (5,883–21,370)	56,209 (33,117–81,734)
Sensitivity analysis						
Scenario 4	158,035 (68,239–266,488)	63,995 (14,159–139,303)	222,030 (115,129–349,135)	205,366 (94,262–326,976)	47,683 (11,187–99,615)	253,049 (135,607–376,978)
Scenario 5	194,797 (96,741–304,265)	99,470 (34,636–178,368)	294,267 (171,682–424,978)	247,904 (131,185–372,313)	72,883 (26,701–125,848)	320,787 (189,592–450,226)
Scenario 6	295,239 (178,923–428,698)	200,407 (107,210–308,078)	495,646 (333,325–661,912)	367,211 (235,595–510,946)	144,764 (78,511–218,538)	511,975 (350,943–680,534)
Scenario 7	60,890 (22,583–106,855)	27,515 (5,994–58,867)	88,405 (41,507–143,607)	105,379 (46,826–172,341)	23,002 (6,753–43,922)	128,381 (66,205–197,377)
Scenario 8	90,018 (40,740–146,906)	43,447 (11,373–85,822)	133,464 (75,076–203,699)	145,248 (77,719–225,633)	35,431 (12,535–63,619)	180,679 (107,685–261,116)
Scenario 9	125,126 (67,057–190,961)	73,157 (30,995–125,015)	198,283 (118,213–279,685)	202,390 (121,054–287,690)	59,191 (27,451–95,818)	261,581 (173,639–358,635)

Scenario 1 = Gradual reduction in salt intake to 8 g/day between 2019 and 2025; Scenario 2 = Gradual reduction in salt intake to 7 g/day between 2019 and 2030; Scenario 3 = Gradual reduction in salt intake to 5 g/day (WHO target) between 2019 and 2030; Scenario 4 = Scenario 1 achieved and sustained for the remaining lifetime; Scenario 5 = Scenario 2 achieved and sustained for the remaining lifetime; Scenario 6 = Scenario 3 achieved and sustained for the remaining lifetime; Scenario 7 = Scenario 1 achieved and effect gradually phases out over 20 years with zero effect beyond 2045; Scenario 8 = Scenario 2 achieved and effect gradually phases out over 20 years with zero effect beyond 2050; Scenario 9 = Scenario 3 achieved and effect gradually phases out over 20 years with zero effect beyond 2050; N, Number; UI, uncertainty interval.

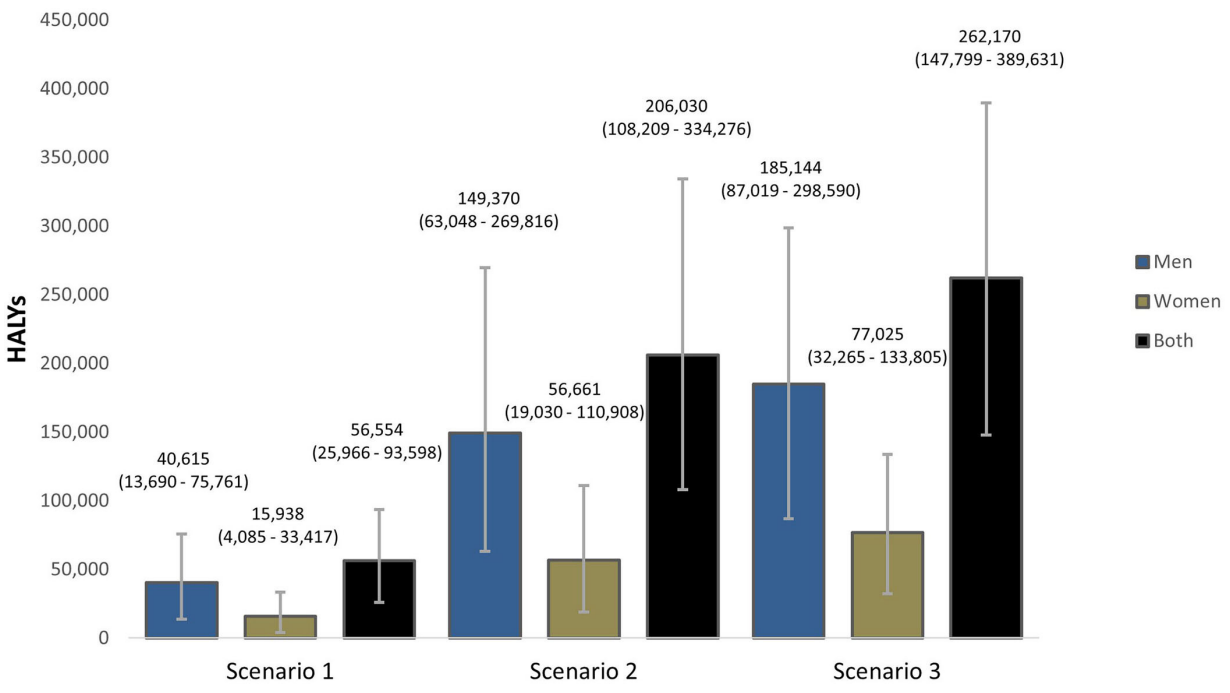

FIGURE 4 | Projected Health-adjusted life years (HALYs) that could be gained if the national (Scenarios 1 and 2) and WHO (Scenario 3) salt targets were achieved in Vietnam. The bars are the best estimates while the whiskers are the 95% uncertainty intervals, i.e., 2.5 and 97.5 percentiles.

TABLE 4 | Projected healthcare costs saved from avoided incident strokes among adults in Vietnam from 2019 to 2030 and over the lifetime.

Base analysis	Ischaemic stroke			Haemorrhagic stroke		
	Men Mean (95% UI)	Women Mean (95% UI)	Both Mean (95% UI)	Men Mean (95% UI)	Women Mean (95% UI)	Both Mean (95% UI)
In millions, US$						
Scenario 1	12.4	6.0	18.5	19.3	4.8	24.1
	(3.1 to 33.6)	(0.9 to 18.7)	(5.7 to 43.1)	(4.8 to 50.7)	(0.7 to 16.1)	(7.0 to 58.5)
Scenario 2	25.4	13.4	38.8	39.2	10.2	49.3
	(6.9 to 62.5)	(2.5 to 36.3)	(13.0 to 85.7)	(10.3 to 94.2)	(2.2 to 25.5)	(15.3 to 110.3)
Scenario 3	34.9	20.8	55.7	50.6	16.0	66.6
	(10.5 to 78.8)	(5.8 to 49.8)	(20.3 to 113.8)	(15.9 to 111.1)	(5.0 to 37.9)	(24.3 to 138.3)
Sensitivity analyses						
In millions, US$						
Scenario 4	79.3	30.6	109.9	105.8	23.8	129.6
	(22.9 to 175.9)	(5.3 to 86.4)	(38.7 to 234.6)	(26.6 to 260.6)	(4.1 to 68.8)	(37.4 to 294.8)
Scenario 5	93.6	46.4	140.0	123.6	35.9	159.5
	(29.3 to −199.7)	(9.7 to 115.2)	(51.9 to 290.6)	(34.9 to 26.8)	(7.9 to 86.0)	(50.1 to 337.2)
Scenario 6	137.9	88.8	226.7	180.9	67.4	248.4
	(53.1 to 277.7)	(31.4 to 186.7)	(95.8 to 440.9)	(60.4 to 379.2)	(23.3 to 134.4)	(97.4 to 505.4)
Scenario 7	44.7	19.4	64.1	66.5	15.3	81.8
	(10.6 to 106.4)	(3.0 to 57.8)	(21.6 to 144.6)	(15.6 to 165.0)	(2.8 to 43.6)	(22.9 to 187.1)
Scenario 8	60.6	28.2	88.7	86.9	22.3	109.2
	(18.1 to 145.2)	(5.1 to 80.9)	(32.4 to 194.2)	(25.3 to 201.9)	(4.4 to 60.3)	(35.6 to 238.3)
Scenario 9	82.3	47.5	129.8	119.1	37.5	156.6
	(26.9 to 176.2)	(11.9 to 113.0)	(50.9 to 257.9)	(36.9 to 270.8)	(11.6 to 89.0)	(55.7 to 337.5)

Scenario 1 = Gradual reduction in salt intake to 8 g/day between 2019 and 2025; Scenario 2 = Gradual reduction in salt intake to 7 g/day between 2019 and 2030; Scenario 3 = Gradual reduction in salt intake to 5 g/day (WHO target) between 2019 and 2030; Scenario 4 = Scenario 1 achieved and sustained for the remaining lifetime; Scenario 5 = Scenario 2 achieved and sustained for the remaining lifetime; Scenario 6 = Scenario 3 achieved and sustained for the remaining lifetime; Scenario 7 = Scenario 1 achieved and effect gradually phases out over 20 years with zero effect beyond 2045; Scenario 8 = Scenario 2 achieved and effect gradually phases out over 20 years with zero effect beyond 2050; Scenario 9 = Scenario 3 achieved and effect gradually phases out over 20 years with zero effect beyond 2050; UI, uncertainty interval.

There are ongoing concerns about potential negative side-effects of whole of populations reduction in salt intake down to 5–6 g/day, citing insufficient evidence (35) and potential unintended consequences to hormones (increased renin, adrenaline, aldosterone) and lipids (36). Most of these concerns are based on observational studies with less robust methods and short-term trials of large salt reductions (37). Evidence from methodologically robust studies with modest long-term reduction in salt intake has shown consistent benefits of low salt consumption on blood pressure and CVD (9, 38), thereby supporting current global recommendations of low sodium diets (11, 39, 40). With current salt consumption in Vietnam more than twice these recommendations, very high rates of stroke, and the substantial health and economic gains shown in this study, there is a strong case for population-wide reductions in salt consumption.

From a policy perspective, the accurate quantification and projections of such diet-related indicators play a critical role in developing NCD strategies and policies. It is important to acknowledge that strong national NCD policies need to match the severity of the current and future burden of NCDs. By so doing guide government action on the most critical dietary factors driving population health, as well the population groups at risk. While Vietnam has recognised the importance of addressing NCDs, there has been little rigorous analyses evaluating policies, strategies or programs monitoring the control of NCDs, particularly for salt consumption and hypertension. More so, despite strong government support for NCD prevention and control at national level, implementation of NCD programs has been lagging. Our health and economic projections therefore present an important contribution filling a critical policy-relevant knowledge gap. In addition, we hope this will set the stage for future work to understand the challenges around program implementation and guide development of innovative solutions towards improving NCD prevention and control.

Strengths and Limitations

This study has a number of strengths. First, we used local data on dietary salt consumption and blood pressure distribution obtained from the NCD surveillance WHO-STEPS study, which is a nationally representative survey. Second, our analysis accounts for the differential impact of salt reduction on blood pressure by age, sex, and hypertension status, which enhances the accuracy of our estimates. Third, by quantitatively estimating the potential health and economic impact of national public health strategies under different time horizons as well as their uncertainty, these findings have high policy-relevance.

This represents an advancement from the limited evidence from Vietnam.

Regarding limitations, first, our model is a closed cohort model, which in its current form does not allow for new births and migration. Hence, our results are likely conservative as they do not capture the benefits of younger cohorts as they age into the future. In addition, other blood pressure-related outcomes such as ischaemic heart disease, hypertensive heart disease and chronic kidney disease were not included, as these were not the focus of this work. As a result, the benefits of salt reduction are likely to be considerably larger than presented here. Thirdly, data on healthcare costs are limited to first-time hospitalisation for acute stroke and do not capture other healthcare costs that accrue post-acute hospitalisation such as rehabilitation, follow-up visits and medication. In addition, wider societal costs like carer responsibilities and productivity losses are not included, thereby underestimating the costs. Future studies should consider wider perspectives to capture these costs. Fourth, the healthcare cost data was obtained from a single tertiary centre. This may not be representative of the costs of care obtained at other peripheral hospitals. It may also overestimate costs for patients treated in primary care settings. Moreover, healthcare costs from other conditions that accrue because of people living longer from the intervention are not included. While the balance of the impact these limitations is not immediately clear, other studies incorporating these costs demonstrated net cost savings (14, 21).

In conclusion, this modelling study shows that modest reductions in population blood pressure via achieving the Vietnam national and WHO salt reduction targets could lead to substantial benefits—gains in healthy life years, reducing health expenditures, avoiding CVD events and associated misery. There is urgent need for specific salt reduction strategies in Vietnam to curb the current and impending CVD and economic burden. Future studies that account for the equity and budget impacts of such strategies would be informative for policy makers as they grapple with implementation.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

LA: conceptualisation, data curation, formal analysis, methodology, and writing—original draft. HP: conceptualisation, data curation, validation, and writing—review and editing. DP: conceptualisation, validation, and writing—review and editing. LC: methodology and writing—review and editing. JV: conceptualisation, methodology, validation, supervision, and writing—review and editing. All authors read and approved the final version of the paper.

ACKNOWLEDGMENTS

The authors are grateful to Dr. Lai Duc Truong for providing the 2015 Vietnam STEPS survey report and Healthy Vietnam program that informed the analysis contained in this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2021.682975/full#supplementary-material

REFERENCES

1. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. *Nat Rev Nephrol.* (2020) 16:223–37. doi: 10.1038/s41581-019-0244-2
2. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19.1 million participants. *Lancet*. (2017) 389:37–55. doi: 10.1016/S0140-6736(16)31919-5
3. Meiqari L, Essink D, Wright P, Scheele F. Prevalence of hypertension in Vietnam: a systematic review and meta-analysis. *Asia Pac J Public Health.* (2019) 31:101– 12. doi: 10.1177/1010539518824810
4. Ministry of Health General Department of Preventive Medicine. *National Survey on the Risk Factors of Non-Communicable Diseases (STEPS)*. Vietnam (2015).
5. Hien HA, Tam NM, Tam V, Derese A, Devroey D. Prevalence, awareness, treatment, and control of hypertension and its risk factors in (Central) Vietnam. *Int J Hypertens.* (2018) 2018:6326984. doi: 10.1155/2018/6326984
6. Hoang VM, Tran QB, Vu THL, Nguyen TKN, Kim BG, Pham QN, et al. Patterns of raised blood pressure in Vietnam: findings from the WHO STEPS survey 2015. *Int J Hypertens.* (2019) 2019:1219783. doi: 10.1155/2019/1219783
7. Yamanashi H, Ngoc MQ, Huy TV, Suzuki M, Tsujino A, Toizumi M, et al. Population- based incidence rates of first-ever stroke in central Vietnam. *PloS ONE.* (2016) 11:e0160665. doi: 10.1371/journal.pone.0160665
8. *Global Burden of Disease Study 2019 (GBD Compare)*. (2020). Available online at: https://vizhub.healthdata.org/gbd-compare/ (accessed February 1, 2021).
9. He FJ, Tan M, Ma Y, MacGregor GA. Salt reduction to prevent hypertension and cardiovascular disease: JACC state-of-the-art review. *J Am Coll Cardiol.* (2020) 75:632–47. doi: 10.1016/j.jacc.2019.11.055
10. Mozaffarian D, Fahimi S, Singh GM, Micha R, Khatibzadeh S, Engell RE, et al. Global sodium consumption and death from cardiovascular causes. *N Engl J Med.* (2014) 371:624–34. doi: 10.1056/NEJMoa1304127
11. World Health Organization. *Guideline: Sodium Intake for Adults and Children*. Geneva: World Health Organization (2012).
12. Huang L, Trieu K, Yoshimura S, Neal B, Woodward M, Campbell NRC, et al. Effect of dose and duration of reduction in dietary sodium on blood pressure levels: systematic review and meta-analysis of randomised trials. *BMJ.* (2020) 368:m315. doi: 10.1136/bmj.m315
13. Aminde LN, Takah NF, Zapata-Diomedi B, Veerman JL. Primary and secondary prevention interventions for cardiovascular disease in low-income and middle-income countries: a systematic review of economic evaluations. *Cost Eff Resour Alloc.* (2018) 16:22. doi: 10.1186/s12962-018-0108-9
14. Cobiac LJ, Vos T, Veerman JL. Cost-effectiveness of interventions to reduce dietary salt intake. *Heart.* (2010) 96:1920–5. doi: 10.1136/hrt.2010.199240
15. Bibbins-Domingo K, Chertow GM, Coxson PG, Moran A, Lightwood JM, Pletcher MJ, et al. Projected effect of dietary salt reductions on future cardiovascular disease. *N Engl J Med.* (2010) 362:590–9. doi: 10.1056/NEJMoa0907355

16. Allen LN, Pullar J, Wickramasinghe KK, Williams J, Roberts N, Mikkelsen B, et al. Evaluation of research on interventions aligned to WHO 'Best Buys' for NCDs in low-income and lower-middle-income countries: a systematic review from 1990 to 2015. *BMJ Global Health.* (2018) 3:e000535. doi: 10.1136/bmjgh-2017-000535
17. The Healthy Vietnam Program. *Vietnamese Health Programme Launched*. (2019). Available online at: https://en.nhandan.com.vn/society/health/item/7193102-vietnamese-health-programme-launched.html (accessed January 30, 2021).
18. Global Burden of Disease Collaborative Network. *Global Burden of Disease Study 2019 (GBD 2019) Results*. Seattle, WA: Institute for Health Metrics and Evaluation (2020).
19. Pham T. *Occurrence, presentation, costs and three-month outcomes of Stroke in Viet Nam*. [dissertation/Ph.D. thesis]. Hobart, TAS: University of Tasmania (2015).
20. Barendregt JJ, Van Oortmarssen GJ, Van Hout BA, Van Den Bosch JM, Bonneux L. Coping with multiple morbidity in a life table. *Math Popul Stud.* (1998) 7:29–49. doi: 10.1080/08898489809525445
21. Aminde LN, Cobiac L, Veerman JL. Cost-effectiveness analysis of population salt reduction interventions to prevent cardiovascular disease in Cameroon: mathematical modelling study. *BMJ Open.* (2020) 10:e041346. doi: 10.1136/bmjopen-2020-041346
22. Ananthapavan J, Sacks G, Brown V, Moodie M, Nguyen P, Veerman L, et al. Priority- setting for obesity prevention-The Assessing Cost-Effectiveness of obesity prevention policies in Australia (ACE-Obesity Policy) study. *Plos ONE.* (2020) 15:e0234804. doi: 10.1371/journal.pone.0234804
23. He FJ, Li J, Macgregor GA. Effect of longer term modest salt reduction on blood pressure: cochrane systematic review and meta-analysis of randomised trials. *BMJ.* (2013) 346:f1325. doi: 10.1136/bmj.f1325
24. Barendregt JJ, Van Oortmarssen GJ, Vos T, Murray CJ. A generic model for the assessment of disease epidemiology: the computational basis of DisMod II. *Popul Health Metr.* (2003) 1:4. doi: 10.1186/1478-7954-1-4
25. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press (2015). p. 219.
26. Barendregt JJ. *Ersatz*. EpiGear International Brisbane. Available online at: www.epigear.com/ (accessed January 30, 2021).
27. Ha DA, Chisholm D. Cost-effectiveness analysis of interventions to prevent cardiovascular disease in Vietnam. *Health Policy Plan.* (2011) 26:210–22. doi: 10.1093/heapol/czq045
28. Lauer JA, Röhrich K, Wirth H, Charette C, Gribble S, Murray CJ. PopMod: a longitudinal population model with two interacting disease states. *Cost Eff Resour Alloc.* (2003) 1:6. doi: 10.1186/1478-7547-1-6
29. Basu S, Stuckler D, Vellakkal S, Ebrahim S. Dietary salt reduction and cardiovascular disease rates in India: a mathematical model. *PloS ONE.* (2012) 7:e44037. doi: 10.1371/journal.pone.0044037
30. Nilson EAF, Metlzer AB, Labonté ME, Jaime PC. Modelling the effect of compliance with WHO salt recommendations on cardiovascular disease mortality and costs in Brazil. *PloS ONE.* (2020) 15:e0235514. doi: 10.1371/journal.pone.0235514
31. Aminde LN, Cobiac LJ, Veerman JL. Potential impact of a modest reduction in salt intake on blood pressure, cardiovascular disease burden and premature mortality: a modelling study. *Open Heart.* (2019) 6:e000943. doi: 10.1136/openhrt-2018-000943
32. Kontis V, Cobb LK, Mathers CD, Frieden TR, Ezzati M, Danaei G. Three public health interventions could save 94 million lives in 25 years. *Circulation.* (2019) 140:715–25. doi: 10.1161/CIRCULATIONAHA.118.038160
33. Rapsomaniki E, Timmis A, George J, Pujades-Rodriguez M, Shah AD, Denaxas S, et al. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1.25 million people. *Lancet.* (2014) 383:1899–911. doi: 10.1016/S0140-6736(14)60685-1
34. Forouzanfar MH, Liu P, Roth GA, Ng M, Biryukov S, Marczak L, et al. Global burden of hypertension and systolic blood pressure of at least 110 to 115mmHg, 1990-2015. *JAMA.* (2017) 317:165–82. doi: 10.1001/jama.2016.19043
35. O'Donnell M, Mente A, Alderman MH, Brady AJB, Diaz R, Gupta R, et al. Salt and cardiovascular disease: insufficient evidence to recommend low sodium intake. *Eur Heart J.* (2020) 41:3363–73. doi: 10.1093/eurheartj/ehaa586
36. Graudal NA, Hubeck-Graudal T, Jurgens G. Effects of low sodium diet versus high sodium diet on blood pressure, renin, aldosterone, catecholamines, cholesterol, and triglyceride. *Cochrane Database Sys Rev.* (2020) 12:Cd004022. doi: 10.1002/14651858.CD004022.pub5
37. Cook NR, He FJ, MacGregor GA, Graudal N. Sodium and health-concordance and controversy. *BMJ.* (2020) 369:m2440. doi: 10.1136/bmj.m2440
38. Webster J, Waqanivalu T, Arcand J, Trieu K, Cappuccio FP, Appel LJ, et al. Understanding the science that supports population-wide salt reduction programs. *J Clin Hypertens.* (2017) 19:569–76. doi: 10.1111/jch.12994
39. Campbell NR, Lackland DT, Chockalingam A, Lisheng L, Harrap SB, Touyz RM, et al. The International Society of Hypertension and World Hypertension League call on governments, nongovernmental organizations and the food industry to work to reduce dietary sodium. *J Hypertens.* (2014) 32:446–7. doi: 10.1097/HJH.0000000000000079
40. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol.* (2019) 74:1376–414. doi: 10.1016/j.jacc.2019.03.009

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

ORIGINAL RESEARCH
published: 01 November 2021
doi: 10.3389/fpubh.2021.743371

The Impact of the COVID-19 Pandemic on Physician Visits in Japan

*Narimasa Kumagai**

Faculty of Economics, Seinan Gakuin University, Fukuoka, Japan

Background: Emerging from the coronavirus disease 2019 (COVID-19) scenario, fears of social distancing and contagion have led to a decline in the number of physician visits in Japan, placing severe financial strain on most hospitals and clinics. In this context, this study examined the impact of the spread of COVID-19 on the utilization of outpatient services.

Methods: This study used monthly data drawn from the monthly statistics report of the social insurance medical fee payment fund in Japan and estimated fixed-effects models.

Results: The results showed that the decline in the number of physician visits because of the first state of emergency declaration in Japan was greater than that caused by COVID-19's spread during the same period. However, there was a decline in the impact of the declaration over time. After the second state of emergency declaration, the decline in the number of physician visits caused by the spread reduced by almost half. The nationwide preschool closure under the declaration of the first state of emergency also adversely impacted the number of physician visits. The reduced healthcare per capita costs of preschool children were greater among prefectures taking specific precautions. The results showed non-negligible regional differences in physician visits of preschool children during the sample period.

Conclusions: The findings imply that we should not overestimate the negative impacts of the state of emergency declaration without lockdown on physician visits. To restore the number of physician visits to its pre-pandemic level, it is crucial to facilitate a smooth transition of COVID-19 patients between hospitals and an effective compensation program for hospitals with COVID-19 patients.

Keywords: COVID-19, healthcare costs, physician visits, preschool children, state of emergency declaration

OPEN ACCESS

Edited by:
Mihajlo Jakovljevic, Hosei University, Japan

Reviewed by:
Simon Grima, University of Malta, Malta
Chie Hanaoka, Toyo University, Japan

***Correspondence:**
Narimasa Kumagai
kumanari@seinan-gu.ac.jp

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *18 July 2021*
Accepted: *29 September 2021*
Published: *01 November 2021*

Citation:
Kumagai N (2021) The Impact of the COVID-19 Pandemic on Physician Visits in Japan. Front. Public Health 9:743371. doi: 10.3389/fpubh.2021.743371

INTRODUCTION

During the early phase of the COVID-19 pandemic, studies investigated the socio-demographic and economic impacts of the pandemic's onset. Findings revealed a positive correlation between population density and COVID-19 mortality or health outcome variables (1–3). Empirical evidence has indicated that the COVID-19 pandemic might have enlarged regional disparities in the healthcare access of the population worldwide.

The COVID-19 pandemic led to a decline in outpatient visits in mid-March 2020. In the United States, the decreasing trend in outpatient visits saw the lowest number of visits in

mid-April 2020 (4). Boserup et al. (5) showed that the decline in the weekly emergency department visits on March 8, 2020 coincided with the World Health Organization's declaration on March 11, 2020 declaring COVID-19, which originated in Wuhan, China, as a pandemic (6).

In Japan, to minimize the impact of the country's population density on the spread of COVID-19, the country's first state of emergency was declared on April 7, covering seven prefectures—Tokyo, three surrounding prefectures (Saitama, Chiba, and Kanagawa), Osaka, Hyogo, and Fukuoka. On April 16, the areas covered by the state of emergency were expanded to all of Japan's 47 prefectures. The state of emergency declaration to stay indoors and close non-essential stores was legally non-binding. Furthermore, this declaration was not followed by a lockdown. The pandemic has given rise to social distancing and contagion fears, which has led to a decline in the number of physician visits. In Japan, the total claimed hospital charges also decreased by 7, 14, and 5% in April, May, and June 2020, respectively, over the same months in 2019 (7).

At the beginning of the outbreak, the extent of the decline in outpatient visits was not uniform across medical specialties. In addition, the impact has particularly been large for children. Studies in Norway have associated higher engagement in health-protective behaviors—such as increased hygiene or physical distancing—with households having more children (8). In Japan, parents with younger children also tend to engage in self-protective behaviors against the pandemic. Concerning parents' physician visits, a survey from April 30 to May 31, 2020 revealed that 45% of surveyed parents with symptoms refrained from physician visits due to the fear of contagion (9). The number of physical visits decreased by the nationwide preschool closure, and this scenario was further worsened by the leaves taken by nurses who had preschool children at home, owing to the school closure. According to a survey on nurses' responses to COVID-19 by the Japanese Nursing Association published on December 22, 2020, during the pandemic, nurses had reduced working days among 40% of hospitals. Most took absences due to the school closure. Furthermore, the shortage of nurses that ensued was due to a decrease in outpatient services provided.

Even during the second state of emergency declaration, the fear of contagion compelled several parents to avoid physician visits. On January 7, 2021, Japan's second state of emergency was declared, affecting Tokyo and three surrounding prefectures. On January 14, coverage was extended to an additional seven prefectures, including Aichi, Osaka, and Fukuoka. According to a report by the Japan Medical Association[1] published on April 28, 2021, the situation worsened the financial conditions of clinics during the months from November 2020 to January 2021. Particularly, outpatient visits to pediatricians stood at ∼62% in January 2021, relative to the same month the previous year. The report concluded that people refraining from physician visits diminished pediatricians' first visit fees.

Given this context, this study estimates the impact of the spread of COVID-19 on the utilization of outpatient services. Since responses were motivated by subjective rather than objective perceptions of risk (10), the study estimation function considers both the first and second states of emergency declaration. Thus, this study examines the differences in the effects of emergency declaration on physician visits. In this context, there is little evidence that the public management of infectious diseases is sensitive to changes in physician visits and disease avoidance. This study also analyzes the heterogeneous impact of COVID-19 across regions. It must be noted that the increased clinic closure owing to continual decline in the number of physician visits may increase regional disparities in healthcare access.

This study tests the following hypotheses: (1) The declaration of the state of emergency reduced new COVID-19 cases and fear of contagion and the tendency of patients to refrain from physician visits and (2) The heterogeneous impact of COVID-19 across regions and regional differences in the reduced healthcare per capita costs of preschool children. However, monthly data that show the real bed occupancy rate for COVID-19 patients are not available. Due to this data limitation, this study cannot consider the supply constraints of hospital admissions. The remainder of this paper is organized as follows. Section Methods presents the monthly trends in outpatient visits in Japan, empirical strategy, and data used in this study. Section Results presents the estimation results. Sections Discussion and Conclusion present the discussion and conclusion, respectively.

METHODS

Monthly Trends in Outpatient Visits and the Declaration of the State of Emergency in Japan

As of Spring 2020, the expert meeting on new coronavirus disease control (EMNCDC) of the Japanese government postulated that an 80% restriction in social activity would be necessary to contain the spread of infection. Owing to the several measures that had been taken in this direction, the Japanese public fulfilled the recommendations of the EMNCDC even before the declaration of the first state of emergency in the country on April 7 (11). In this regard, Watanabe and Yabu (12) used smartphone location data and examined the effect of the declaration of the first state of emergency and the closure of schools on the stay-at-home measure. They found that the government's declaration induced people to stay indoors. They showed that the declaration of the state of emergency led to a decline in the number of people leaving their homes by 8.5% and that a 1% increase in new infections within the prefecture reduced people's outings within that prefecture by 0.027%. In the given context, it can be stated that the government's social distancing measure acted as a barrier to healthcare access.

According to the monthly statistics report of the social insurance medical fee payment fund, in Japan, the pandemic led to a decline in the total number outpatient visits during the 12-month period in 2020, relative to the same period in the previous year. **Figure 1** shows that the COVID-19 pandemic led to a decline in outpatient visits in March 2020. The maximum

[1]The report in Japanese is available at: https://www.med.or.jp/dl-med/teireikaiken/20210428_3.pdf.

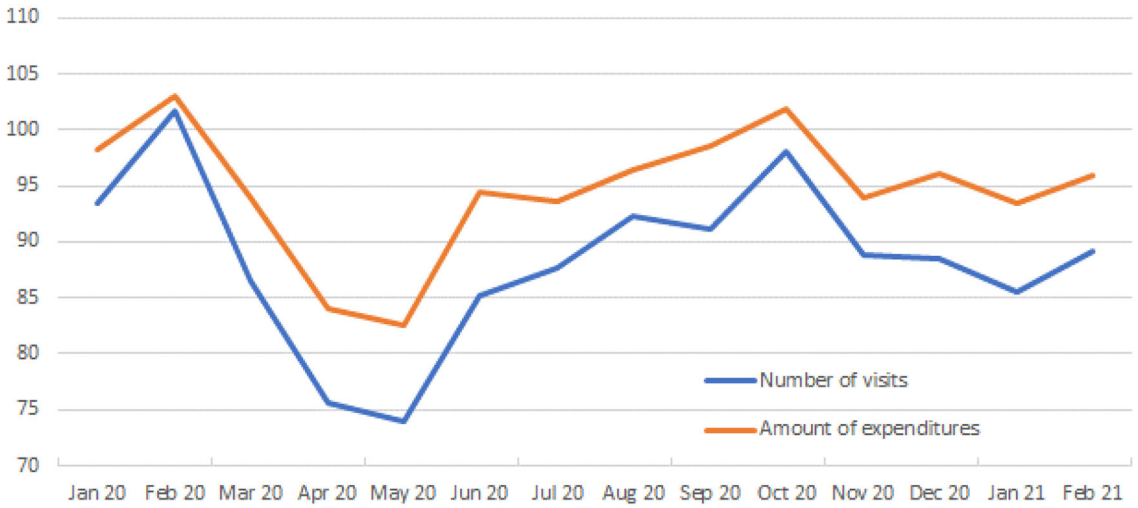

FIGURE 1 | Monthly trends in outpatient visits relative to the same months in 2020–2021. This figure reports the change rate in outpatient visits relative to the same month in the previous year (January 2020–February 2021), not controlling for flu season effects. Data are from the monthly statistics report of the social insurance medical fee payment fund (*Shiharai Kikin*; https://www.ssk.or.jp/tokeijoho/geppo/index.html).

decrease in outpatient visits peaked in May 2020, when the pandemic reduced the outpatient visits by 74% relative to the same month in the previous year. The impact on outpatient visits has been declining since June 2020; however, the decline in the number of outpatient visits exerted severe financial pressures on most hospitals and clinics. The impact on outpatient visits declined in the 5 months ensuing June. Specifically, there was no difference in the outpatient expenditure between October 2020 and the same month in the previous year. However, the spread of COVID-19 led to a decline in the number of outpatient visits over the 3-month-period from November 2020 to January 2021.

The extent of the decline in the number of physician visits during the sample period may be related to the COVID-19 prevalence in the prefectures. For example, there were variations in the monthly trends in the number of confirmed cases among the three prefectures of Hokkaido, Hyogo, and Fukuoka during the sample period, while there was no difference in the total number of confirmed COVID-19 cases. There were little variations in the monthly trends among Saitama, Chiba, and Kanagawa (see, a dummy variable for the first state of emergency, the seven prefectures shown in **Table 1**). The confirmed cases in Hokkaido peaked in November, and this spread reduced people's outings within Hokkaido (see **Figure 2**). Conversely, the number of confirmed cases in Hyogo and Fukuoka peaked in January 2021. Fukuoka had two peaks during the sample period; however, this study did not observe the spread in the summer season in Hyogo. Regional differences in the fear of contagion might have influenced the number of physician visits.

There was a sharp drop in the hospitalization rate of patients with asthma during the COVID-19 pandemic (13). Further, there was also a decline in the number of patients with other infectious diseases, such as influenza. According to a report by the National Institute of Infectious Diseases in Japan, the signs of seasonal influenza epidemics in Japan have still not been identified (14). With regard to COVID-19, area hospitals struggled to keep up with COVID-19 hospitalizations during the second state of emergency.

Empirical Strategy

The estimation function measures the direct effect of the spread of COVID-19 on physician visits and the indirect effect of the declaration of the state of emergency in reducing new COVID-19 cases by reducing the fear of contagion and practice of refraining from physician visits (see **Figure 3**). The other unobservable effect of health-protective behaviors, such as increased hygiene, might have a positive effect on physician visits. The core econometric specification of this function is as follows:

$$Y_{i,t} = C + \boldsymbol{Z}_{i,t}\beta' + \boldsymbol{D}_{i,t}\gamma' + (\boldsymbol{Z}_{i,t} \times \boldsymbol{D}_{i,t})\delta' + \boldsymbol{X}_{i,t}\theta' + \tau W_{i,t} + \mu J_{i,t} + \nu_{i,t} + \varepsilon_{i,t} \quad \text{(A)}$$

where $Y_{i,t}$ is a dependent variable, such as the number of outpatients relative to the same month in the previous year. The subscripts i and t indexes the prefecture and time periods, respectively. $\boldsymbol{Z}$ is a vector of confirmed COVID-19 cases per 100,000 people (current and lagged), and $\boldsymbol{D}$ is a vector of dummy variables for the first or second states of emergency or specific precautions. W is a dummy variable for the winter season, J is the active job openings-to-applicants ratio (seasonally adjusted), and C is a constant term. The vector $\boldsymbol{X}$ includes the reciprocal variable of the change rate in outpatient or inpatient costs and the number of inpatients or outpatients (current and lagged). To avoid the endogeneity problem in which physician visits and healthcare costs are determined simultaneously, the study uses lagged reciprocal variables. β', γ', δ', θ', τ, and μ are the coefficients to be estimated. The error term is $\nu_{i,t} + \varepsilon_{i,t}$; $\nu_{i,t}$ is

TABLE 1 | Monthly summary statistics by prefecture.

Variables	*N*	Mean	SD	Minimum	Maximum
Confirmed COVID-19 cases per 100,000 people	658	14.289	28.483	0	289.97
Outpatient visits relative to the same month in the previous year	658	0.895	0.070	0.66	1.22
Outpatient costs relative to the same month in the previous year	658	0.946	0.053	0.77	1.08
Number of inpatients relative to the same month in the previous year	658	0.970	0.075	0.76	1.25
Inpatient costs relative to the same month in the previous year	658	0.981	0.052	0.84	1.15
Active job openings-to-applicants ratio (seasonally adjusted)	658	1.264	0.247	0.72	2.17
Preschool children					
Healthcare costs per preschool children (1,000 yen)	658	14.694	2.505	8.14	20.92
Population of preschool children, excluding infants (1,000 person)	658	101.192	105.854	22	535
Doctor's consultation rate (inpatients included)	658	1.470	0.307	0.73	2.24
Healthcare costs per physician visits (inpatients included)	658	10.133	1.134	7.45	14.47
Monthly change rate in healthcare costs per preschool children (*hcpc*)	611	−0.001	0.145	−0.47	0.47
Monthly change rate in doctor's consultation rate	611	−0.001	0.178	−0.46	0.63
Monthly change rate in *hcpc* per physician visits	611	0.009	0.101	−0.29	0.52
Declaration of state of emergency					
Dummy variable for the first state of emergency in seven prefectures	658	0.023	0.149	0	1
Dummy variable for the specific precautions in six prefectures	658	0.018	0.134	0	1
Dummy variable for the second state of emergency in 11 prefectures	658	0.033	0.180	0	1

Health data were collected from the monthly statistics report of social insurance medical fee payment fund (Shiharai Kikin). The active job openings-to-applicants ratio was collected from the Employment Referrals for General Workers, the Ministry of Health, and Labor and Welfare.
Children below 5 years, as of October 2019, were used as a proxy variable of the population of preschool children, excluding infants during the sample period. 14 months × 47 prefectures = 658.

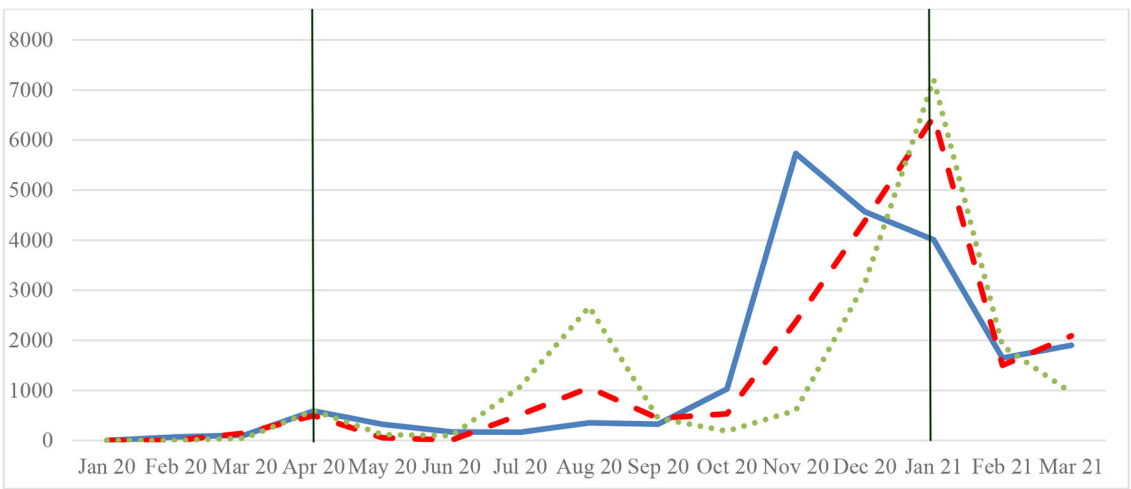

FIGURE 2 | Trends in confirmed COVID-19 cases among the three prefectures. Hokkaido (blue bold line), Hyogo, and Fukuoka (dotted line), January 2020–March 2021. The daily numbers of confirmed COVID-19 cases are taken from the NHK (Nippon Hoso Kyokai); monthly data were created by the author. The green vertical line shows the declaration of the state of emergency.

the unit-specific constant error term; all the unobserved time-invariant components, $\varepsilon_{i,t}$, are assumed to be independent of covariates in Equation (A).

The fixed-effects model considers two aspects—the positive and negative effects of the declaration of the state of emergency on physician visits. Under the declaration of the state of emergency, some people tended to stay indoors and refrain from physician visits. Conversely, people have less fear of contagion when there is a decline in the number of confirmed COVID-19 cases. To capture the interaction effect of the confirmed COVID-19 cases and the declaration of the state of emergency on physician visits, the study used the interaction term of those variables ($\boldsymbol{Z}_{i,t} \times \boldsymbol{D}_{i,t}$). When there is a decline in the fear of contagion and an increase in the number of outings, relative to staying indoors, the interaction term exerts a positive effect on physician visits (a square box in **Figure 3**). In this regard,

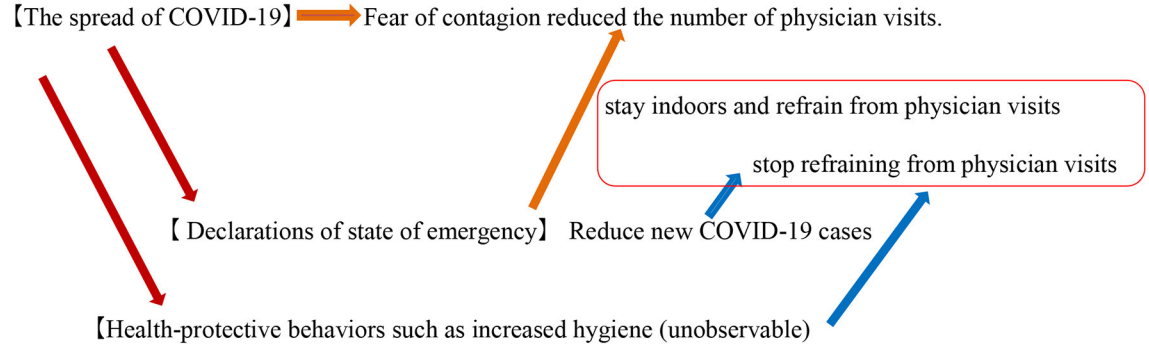

FIGURE 3 | Influence of the state of emergency declaration.

Boone and Ladreit (15) showed that an aggregate measure of the level of stringency of non-pharmaceutical interventions was associated with greater reductions in mobility in advanced member countries of the Organization for Economic Co-operation and Development (OECD) during the first wave. The aggregate measure, an index similar to that of the Oxford Stringency Index, includes school closure, workplace closure, the cancelation of public events, the restrictions on gatherings and internal movements, the closure of public transport, stay-at-home requirements, and international travel controls. Based on their results, it is considered that the nationwide preschool closure, under the declaration of the state of emergency, adversely influences the number of physician visits. This impact was heterogeneous across regions owing to the degree of reduction in mobility. In this context, this study investigates regional differences in the reduced healthcare per capita costs of preschool children.

Data

This study uses monthly data from January 2020 to February 2021 and estimates the fixed-effects models. Data were drawn from the monthly statistics report of the social insurance medical fee payment fund (*Shiharai Kikin* in Japanese) in Japan. During the sample period, relative to the same period in the previous year, there was an almost 10 and 5% decline in the number of outpatients and their healthcare costs, respectively (**Table 1**). The minimum value of the monthly change rate in healthcare costs per preschooler and doctor's consultation rate for preschool children were −0.47, and −0.46, respectively. With regard to new COVID-19 cases, the mean of the monthly confirmed COVID-19 cases per 100,000 people during the sample period was 14.29, with a standard deviation of 28.48. Therefore, COVID-19 cases per 100,000 people exhibited a fat-tailed distribution because the upper interval of one standard deviation from the mean was (SD = 42.78, 289.97).

In the given context, it must be noted that the COVID-19 pandemic in Japan led to a deterioration of the employment situation amid the social distancing and temporary business closures. The ratio of active job openings-to-applicants was 1.20 (seasonally adjusted value) in May 2020—a year-on-year decrease of 0.12% points. This was the country's second-largest monthly decline in history, following the 0.20% point fall in February 1974. The ratio of active job openings-to-applicants indicates the number of jobs from companies per active job seeker. For the first state of emergency, the seven prefectures of Saitama, Chiba, Tokyo, Kanagawa, Osaka, Hyogo, and Fukuoka took a value of 1 as a dummy variable, and 0 otherwise. For the specific precautions in the prefectures, the six prefectures of Hokkaido, Ibaragi, Ishikawa, Gifu, Aichi, and Kyoto took a value of 1 as a dummy variable, and 0 otherwise. For the second state of emergency, the 11 prefectures of Tochigi, Saitama, Chiba, Tokyo, Kanagawa, Gifu, Aichi, Kyoto, Osaka, Hyogo, and Fukuoka took a value of 1 as a dummy variable, and 0 otherwise.

RESULTS

This study used the Hausman's specification test to compare the results of the random-effects model with that of the fixed-effects model. The test supported the initial hypothesis that the prefectural-level effects are adequately modeled by a fixed-effects model at the 1% level (not shown). **Table 2** shows the estimation results for the number of outpatients during the periods April–May 2020 and January–March 2021, relative to the same months in the previous years. Equation (2) includes the active job openings-to-applicants ratio as a variable of the business cycle, which was not significant at the 5% level. The major findings from Equation (1) are summarized as follows: First, when there were 110 new confirmed cases per 100,000 people in the past month (Stage 4: explosive spread of infection), the spread of COVID-19 had reduced 10% of the number of physician visits relative to that in the previous year. There is comparability in the magnitude of the following two impacts on the number of physician visits: the first state of emergency declaration in the seven prefectures and the 82 new confirmed cases per 100,000 people in a month (Stage 3: rapid increases in infected people). Second, after the second state of emergency declaration, the decline in the number of physician visits caused by the spread reduced by almost half (−0.555 = 0.000539/−0.000971). However, the second declaration had a negative impact on the number of inpatients.

TABLE 2 | Number of outpatients.

Variables	(1)	(2)
Dependent variable: number of outpatients relative to the same month in the previous year		
Confirmed COVID-19 cases per 100,000 people	−0.000971***	−0.00102***
	(0.000115)	(0.000129)
Confirmed COVID-19 cases per 100,000 people (lagged)	−0.000542***	−0.000574***
	(0.000118)	(0.000123)
Dummy variable for the first state of emergency (April–May 2020)	−0.0803***	−0.0787***
	(0.0195)	(0.0196)
Dummy variable for prefectures taking specific precautions (April–May 2020)	−0.0614***	−0.0599***
	(0.0207)	(0.0208)
Dummy variable for the second state of emergency (January–March 2021)	0.00315	0.00360
	(0.0235)	(0.0235)
Confirmed COVID-19 cases per 100,000 people × the first state of emergency	−0.000896	−0.000955
	(0.00180)	(0.00180)
Confirmed COVID-19 cases per 100,000 people × the specific precautions	0.000196	0.000180
	(0.00241)	(0.00241)
Confirmed COVID-19 cases per 100,000 people × the second state of emergency	0.000539***	0.000569***
	(0.000189)	(0.000192)
1/change rate in outpatient costs (lagged)#	−0.289***	−0.308***
	(0.0451)	(0.0496)
Number of inpatients#	0.630***	0.621***
	(0.0431)	(0.0445)
Number of inpatients (lagged)#	−0.232***	−0.226***
	(0.0494)	(0.0498)
Dummy variable for winter season	0.0280***	0.0328***
(1 = December, January, and February, and 0 otherwise)	(0.00539)	(0.00763)
Active job openings-to-applicants ratio (Seasonally adjusted, workplace)		−0.0179
		(0.0201)
Constant term	0.828***	0.874***
	(0.0872)	(0.101)
N	611	611
R-squared	0.585	0.585

*Standard errors are in the parenthesis. ***p < 0.01.*
#Relative to the same month in the previous year. 13 months × 47 prefectures = 611.

TABLE 3 | Number of inpatients.

Variables	(3)	(4)
Dependent variable: number of inpatients relative to the same month in the previous year		
Confirmed COVID-19 cases per 100,000 people	0.000636***	0.000313**
	(0.000111)	(0.000124)
Confirmed COVID-19 cases per 100,000 people (lagged)	0.000852***	0.000667***
	(0.000107)	(0.000110)
Dummy variable for the first state of emergency (April–May 2020)	−0.0172	−0.00974
	(0.0184)	(0.0180)
Dummy variable for prefectures taking specific precautions (April–May 2020)	−0.0325*	−0.0240
	(0.0194)	(0.0190)
Dummy variable for the second state of emergency (January–March 2021)	−0.0464**	−0.0421**
	(0.0219)	(0.0214)
Confirmed COVID-19 cases per 100,000 people × the first state of emergency	−0.00293*	−0.00305*
	(0.00167)	(0.00163)
Confirmed COVID-19 cases per 100,000 people × the specific precautions	−0.000958	−0.000947
	(0.00224)	(0.00219)
Confirmed COVID-19 cases per 100,000 people × the second state of emergency	−0.000326*	−0.000137
	(0.000178)	(0.000177)
1/change rate in inpatient costs (lagged)#	0.0259	−0.0252
	(0.0433)	(0.0434)
Number of outpatients#	0.455***	0.422***
	(0.0338)	(0.0336)
Number of outpatients (lagged)#	0.250***	0.348***
	(0.0355)	(0.0393)
Dummy variable for winter season	0.0234***	0.0507***
(1 = December, January, and February, and 0 otherwise)	(0.00496)	(0.00706)
Active job openings-to-applicants ratio (Seasonally adjusted, workplace)		−0.105***
		(0.0197)
Constant term	0.287***	0.408***
	(0.0672)	(0.0695)
N	611	611
R-squared	0.672	0.688

*Standard errors are in the parenthesis. ***p < 0.01, **p < 0.05, and *p < 0.1.*
#Relative to the same month in the previous year. 3 months × 47 prefectures = 611.

The interaction term of the second state of emergency was positive and significant, whereas the other coefficients of the interaction terms were not significant at the 5% level. The positive sign implies that a decline in the fear of contagion among people reduced their tendency to stay at home during the second state of emergency[2]. The total coefficient of the second state of emergency was positive, while it was −0.08 in the first state of emergency. This indicates a decline in the impact of the declaration over time.

The results also show a significant change in the active job openings-to-applicants ratio in Equation (4). A 0.2% decrease in the active job openings-to-applicants ratio increased the number of inpatients by 2.1%, relative to the same month in the previous year. Concerning inpatients, the number of inpatients during the winter season was 5% greater than that in the other seasons, though the coefficient was 2.8% for the number of outpatients (see Equation 1). When there were 110 new confirmed cases per 100,000 people (Stage 4), the spread of COVID-19 had increased by 3% relative to that in the previous year.

The mean lengths of stay of the confirmed and suspected COVID-19 patients between April and June 2020 were 16.9 and 14.1 days, respectively (7). However, mild or non-severe COVID-19 patients, who did not have fever, did not require a longer duration of hospitalization during the second state of emergency, relative to the severe patients. The estimation result indicated that the rate of decline in the number of inpatients during the second state of emergency was almost the same as that in Stage 4 when

[2]The daily liquidity population data collected by the Agoop Corp. show that there was no decline in population within 500 m of the major terminal station during the second state of emergency, relative to those of the first state of emergency (https://corporate-web.agoop.net/pdf/covid-19/agoop_analysis_coronavirus.pdf).

TABLE 4 | Healthcare costs of preschool children.

Variables	(5)	(6)
Dependent variable: healthcare costs per preschool children		
Confirmed COVID-19 cases per 100,000 people	−0.0166***	−0.00722*
	(0.00352)	(0.00388)
Confirmed COVID-19 cases per 100,000 people (lagged)	−0.0120***	−0.00590
	(0.00348)	(0.00359)
Dummy variable for the first state of emergency (April–May 2020)	−1.501**	−1.926***
	(0.591)	(0.583)
Dummy variable for prefectures taking specific precautions (April–May 2020)	−2.050***	−2.471***
	(0.625)	(0.616)
Dummy variable for the second state of emergency (January–March 2021)	0.403	0.228
	(0.717)	(0.701)
Confirmed COVID-19 cases per 100,000 people × the first state of emergency	−0.211***	−0.209***
	(0.0547)	(0.0534)
Confirmed COVID-19 cases per 100,000 people × the specific precautions	−0.0226	−0.0264
	(0.0735)	(0.0719)
Confirmed COVID-19 cases per 100,000 people × the second state of emergency	0.00174	−0.00390
	(0.00580)	(0.00577)
1/change rate in outpatient costs (lagged)[#]	−11.48***	−9.080***
	(0.663)	(0.795)
Dummy variable for winter season	2.290***	1.486***
(1 = December, January, and February, and 0 otherwise)	(0.150)	(0.213)
Active job openings-to-applicants ratio (Seasonally adjusted, workplace)		3.127***
		(0.599)
Constant term	27.20***	20.70***
	(0.747)	(1.444)
N	611	611
R-squared	0.578	0.598

*Standard errors are in the parenthesis. ***p < 0.01, **p < 0.05, and *p < 0.1.*
[#]Relative to the same month in the previous year. 13 months × 47 prefectures = 611.

TABLE 5 | Doctor's consultation rate for preschool children.

Variables	(7)	(8)
Dependent variable: doctor's consultation rate for preschool children (inpatients included)		
Confirmed COVID-19 cases per 100,000 people	−0.00313***	−0.00150***
	(0.000420)	(0.000445)
Confirmed COVID-19 cases per 100,000 people (lagged)	−0.00227***	−0.00117***
	(0.000415)	(0.000415)
Dummy variable for the first state of emergency (April–May 2020)	−0.201***	−0.269***
	(0.0704)	(0.0672)
Dummy variable for prefectures taking specific precautions (April–May 2020)	−0.213***	−0.283***
	(0.0747)	(0.0712)
Dummy variable for the second state of emergency (January–March 2021)	0.0811	0.0538
	(0.0856)	(0.0811)
Confirmed COVID-19 cases per 100,000 people × the first state of emergency	−0.0201***	−0.0199***
	(0.00653)	(0.00618)
Confirmed COVID-19 cases per 100,000 people × the specific precautions	−0.00301	−0.00331
	(0.00879)	(0.00832)
Confirmed COVID-19 cases per 100,000 people × the second state of emergency	0.000294	−0.000651
	(0.000692)	(0.000665)
1/change rate in outpatient costs (lagged)[#]	−2.137***	−1.572***
	(0.113)	(0.127)
Dummy variable for winter season	0.300***	0.158***
(1 = December, January, and February, and 0 otherwise)	(0.0179)	(0.0243)
Active job openings-to-applicants ratio (Seasonally adjusted, workplace)		0.547***
		(0.0674)
Constant term	3.708***	2.443***
	(0.120)	(0.193)
N	611	611
R-squared	0.600	0.643

*Standard errors are in the parenthesis. ***p < 0.01.*
[#]Relative to the same month in the previous year. 13 months × 47 prefectures = 611.

there were 135 new confirmed cases per 100,000 people (**Table 3**). The estimated coefficient of the dummy variable for the second state of emergency (-0.0421) was equivalent to 0.000313 × 135 confirmed cases per 100,000 people. This implies that the average impact of the second state of emergency on the reduction in hospital admissions was the same in the specific prefecture where infection spread explosively under the first state of emergency declaration. The same pattern was not observed during the first state of emergency.

Concerning the preschool children, since the sole data of the preschoolers' total healthcare costs can be used for estimation, we obtained a series of healthcare costs per preschooler (1,000 yen) by dividing the total healthcare costs by the population of preschool children. Healthcare costs per preschool child include the costs of outpatients and inpatients. Owing to limited available information, children below 5 years (as of October 2019) were used as a proxy variable for the population of preschool children, excluding infants, during the sample period.

The active job openings-to-applicants ratio positively influenced the healthcare costs of preschool children (**Table 4**). Related to a good/bad economy, the results showed non-negligible regional differences in the physician visits of preschool children during the sample period. Among the six prefectures taking specific precautions, the reduced healthcare costs per preschooler were greater than those of the seven prefectures where the first state of emergency was declared. Since the ratio of the estimated coefficients was 1.28 (=2.471/1.926), the results showed a non-negligible regional difference in the effect of the declaration on the reduced healthcare costs per preschooler. However, there was a relatively small regional difference in the doctor's consultation rate (1.05 = 0.283/0.269, Equation 8 in **Table 5**), suggesting a regional difference in preschool children's healthcare costs per visit. However, the reason for this could not be determined because of the lack of patient diagnosis information.

Frequent handwashing after outdoor activities and before touching the mouth/nose area reduced the risk of infection by 97.9 and 69.7%, respectively (16). According to a survey by Watanabe and Yoshikawa (17), the number of children with infectious diseases (e.g., influenza or pneumonia) has reduced by

half. Handwashing habits contributed toward reducing children's morbidity in Japan; however, the interaction terms of the first state of emergency in Equations (6, 8) were negative and significant at the 1% level. This indicates that the staying-at-home effect led parents to refrain from physician visits, which can be attributed to their voluntary self-protection behavior in response to the epidemic (10).

DISCUSSION

The decline in the number of physician visits owing to the first state of emergency declaration was greater than that caused by the spread during the same period. Conversely, the second state of emergency declaration did not have a statistically significant impact on the number of physician visits. The declaration's impact declined over time, and the staying-at-home effect did not persist. Given this finding, we should not overestimate the negative impacts of the state of emergency declaration on physician visits. Conversely, in the current and past months, the number of inpatients exerted a net positive impact on the number of physician visits, indicating that the number of physician visits will return to its pre-pandemic level early in the areas reporting a smooth transition of COVID-19 patients between hospitals.

Incentive and payment systems for private healthcare providers contribute to the quality and performance of health service delivery and play key roles in pandemic preparedness (18). The effect of these systems is evident in countries where the private sector has played a significant role in the health service delivery. However, Shin et al. (7) pointed out that the compensation program focusing on patients with severe COVID-19 did not contribute effectively toward improving the Japanese healthcare system. The study used 2,739,878 inpatient and 53,479,658 outpatient cases from 195 hospitals and showed that an increase in payments for severe COVID-19 patients did not compensate for the decline in the income of hospitals with COVID-19 patients. The emergency comprehensive support grant (ECSG) per COVID-19 patient to medical institutions has not been proportional to the prefectural proportion of the cumulative COVID-19 cases. For example, according to Ito (19), Tokyo received 16.4% of the total ECSG for 26% confirmed COVID-19 cases. The support grants to medical institutions should be proportionally distributed according to COVID-19 patients' severity or density.

During the state of emergency period, the regional differences in physician visits can be attributed to the lack of compensation programs for hospitals with COVID-19 patients. Indeed, the decline in the number of physician visits has a negative effect for medical institutions from a financial perspective. However, this may also contribute to avoiding unnecessary health care use, especially in children. Since parents who have children might update their information on infections based on the number of new infections, they may have refrained from seeing a doctor for fear of infection during the second state of emergency. Thus, children who can receive medical expense subsidies that local governments generally provide might reduce unnecessary physician visits during the pandemic. This study, however, cannot explore the impact of COVID-19 on behavioral change in health care use, which should be examined by future studies.

CONCLUSION

Social distancing and the fear of contagion led to a decline in the number of physician visits in Japan. In this context, this study estimated the impact of the spread of COVID-19 on the utilization of outpatient services using monthly data drawn from the monthly statistics report of the social insurance medical fee payment fund in Japan. The results showed that the decline in the number of physician visits owing to the first state of emergency declaration was greater than that caused by the spread of COVID-19 during the same period; however, the declaration's impact declined over time. The nationwide preschool closure during the first state of emergency also adversely affected the number of physician visits. The results showed the non-negligible regional differences in the physician visits of preschool children during the sample period. Hence, we should not overestimate the negative impacts of the state of emergency declaration without lockdown on physician visits. Information on the spread of infection is more important for parents that have children than the state of emergency declaration.

As stated earlier, there has been a lack of effectiveness in the compensation program for hospitals with severe COVID-19 patients. Given this, a smooth transition of COVID-19 patients between hospitals and an effective compensation program for hospitals with COVID-19 patients would play crucial roles in restoring the number of physician visits to the pre-pandemic level. Future studies should also be conducted to determine the reason for the regional difference in the healthcare costs per visit during the COVID-19.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: the monthly statistics report of the Social Insurance Medical Fee Payment Fund (https://www.ssk.or.jp/tokeijoho/geppo/index.html).

AUTHOR CONTRIBUTIONS

NK was responsible for the conceptualization of the study, the formal study analysis, the writing of the original draft, the review of the tasks throughout the study, and he was accountable for the content of the work.

FUNDING

This study was supported by the Japanese Ministry of Education, Culture, Sports, Science, and Technology (Grant No. 20K01739). The funder had no role in the design of the study; the collection, analysis, and interpretation of the data; and the writing of the manuscript.

ACKNOWLEDGMENTS

The author is grateful for the helpful comments of Professor Yoko Ibuka and the reviewers. Any remaining errors in this work are my own responsibility.

REFERENCES

1. Sá F. Socioeconomic determinants of COVID-19 infections and mortality: evidence from England and Wales. *Covid Economics, Vetted and Real-Time Papers*. (2020) 22:47–58. Available online at: https://cepr.org/content/covid-economics-vetted-and-real-time-papers-0
2. Desmet K, Wacziarg R. Understanding spatial variation in Covid-19 across the United States. In: *NBER Working Paper 27329*. Cambridge, MA: National Bureau of Economic Research (2020).
3. Verwimp P. The spread of COVID-19 in Belgium: a municipality-level analysis. In: *ECARES Working Paper 2020-2025*. Brussels: Université Libre de Bruxelles (2020).
4. Chatterji P, Li Y. Effects of the COVID-19 pandemic on outpatient providers in the United States. *Med Care.* (2021) 59:58–61. doi: 10.1097/MLR.0000000000001448
5. Boserup B, McKenney M, Elkbuli A. The impact of the COVID-19 pandemic on emergency department visits and patient safety in the United States. *Am J Emerg Med*. (2020) 38:1732–6. doi: 10.1016/j.ajem.2020.06.007
6. World Health Organization. *General's Opening Remarks at the Media Briefing on COVID-19 - 11 March 2020*. (2020). Available online at: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-mediabriefing-on-covid-19-11-march-2020 (accessed May 1, 2020).
7. Shin JH, Takada D, Morishita T, Lin H, Bun S, Teraoka E, et al. Economic impact of the first wave of the COVID-19 pandemic on acute care hospitals in Japan. *PLoS One.* (2020) 15:e0244852. doi: 10.1371/journal.pone.0244852
8. Zickfeld JH, Schubert TW, Herting AK, Grahe J, Faasse K. Correlates of health-protective behavior during the initial days of the COVID-19 outbreak in Norway. *Front Psychol.* (2020) 11:564083. doi: 10.3389/fpsyg.2020.564083
9. National Center for Child Health and Development. *National Online Survey of Children's Well-Being During the COVID-19 Pandemic in Japan (1st-5th Wave)*. (2021). Available online at: https://www.ncchd.go.jp/en/news/2021/210408.html (accessed May 30, 2021).
10. Fenichel EP, Kuminoff NV, Chowell G. Skip the trip: air travelers' behavioral responses to pandemic influenza. *PLoS One.* (2013) 8:e58249. doi: 10.1371/journal.pone.0058249
11. Okazawa M, Suzuki S. Japanese tactics for suppressing COVID-19 spread. *Public Health.* (2020) 186:6–7. doi: 10.1016/j.puhe.2020.07.012
12. Watanabe T, Yabu T. Japan's voluntary lockdown. *PLoS One.* (2021) 16:e0252468. doi: 10.1371/journal.pone.0252468
13. Abe K, Miyawaki A, Nakamura M, Ninomiya H, Kobayashi Y. Trends in hospitalizations for asthma during the COVID-19 outbreak in Japan. *J Allergy Clin Immunol Pract.* (2021) 9:494–6.e1. doi: 10.1016/j.jaip.2020.09.060
14. Hayashi T, Konishi I. Significant decrease in seasonal influenza in the COVID-19 era: impact of global movement restrictions? *J Clin Med Res.* (2021) 13:191–4. doi: 10.14740/jocmr4450
15. Boone L, Ladreit C. Fear of COVID and non-pharmaceutical interventions: an analysis of their economic impact among 29 advanced OECD countries. *Covid Econ.* (2021) 73:1–40.
16. Lio CF, Cheong HH, Lei CI, Lo IL, Yao L, Lam C, et al. Effectiveness of personal protective health behaviour against COVID-19. *BMC Public Health.* (2021) 21:827. doi: 10.1186/s12889-021-10680-5
17. Watanabe S, Yoshikawa A. *Iryo Hokai No Shinjitsu (Truth of the Covid-19 Has Overwhelmed Health Care System)*. Tokyo: MdN Corporation (2021) (in Japanese).
18. Kwon S, Kim E. Universal Health Coverage and COVID-19 preparedness and response in Asia and the Pacific. In: *JCER Working Paper AEPR Series No. 2021-1-6* Tokyo (2021).
19. Ito Y. *Corona Ga Shimeshita Kadai (Problems the COVID-19 Indicated)*. Morning Edition of Nihon Keizai Shimbun, July 6, 2021 (2021) (in Japanese).

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a past co-authorship with the author.

frontiers in Public Health

ORIGINAL RESEARCH
published: 25 November 2021
doi: 10.3389/fpubh.2021.766385

A Research on the Crisis Spillover Effect of Food Safety Incidents on Competitive Firms: The Influence of Political Connections and Charitable Donations

Rong Xiang, Mengqi Wang, Li Lin and Dongxia Wu†*

School of Business Administration, Zhejiang Gongshang University, Hangzhou, China

Taking the perspective of corporate social responsibility and institutional theory, this research establishes an innovative relationship between variables such as charitable donation, political connection and crisis spillover effect of firms through quantitative analysis using the event study method, regression analysis and the Heckman two-stage model. Taking 8 food safety incidents from 2011 to 2016 as research samples, this paper studies the impact of food safety incidents on the market value of both firms under crisis and their competitive firms, as well as the influence of political connection and charitable donation. Based on the current situation that the product crisis or reputation crisis of a firm will, inevitably, affect the market performance and value of its competitive firms in the same industry, this paper attempts to answer questions such as "what kind of firms are capable of minimizing this negative influence?" "will the political connection of competitive firms exert a positive or negative impact?" and "can actions taken before the crisis, such as charitable donation of competitive firms, help these firms in reducing the harm?" The conclusions are as follows: first, the occurrence of food safety incidents not only has a negative impact on the market value of the crisis firm, but also has a negative spillover effect on the competitive firm; second, charitable donations made by the competitive firm before the crisis demonstrates a positive competitive effect on the competitive firm, and the intensity of such charitable donations is positively correlated with this positive competitive effect; third, the political connection of the competitive firm has no significant impact on the crisis spillover effect. These findings provide enlightenment for the operation and management of firms in the food industry.

Keywords: crisis spillover, political connection, charitable donation, abnormal returns, food safety incidents

OPEN ACCESS

Edited by:
Mihajlo Jakovljevic,
Hosei University, Japan

Reviewed by:
Luigi Aldieri,
University of Salerno, Italy
Concetto Paolo Vinci,
University of Salerno, Italy

****Correspondence:***
Li Lin
kantarlily@126.com

†Present address:
Zhejiang E-Commerce Bank Co., Ltd, Hangzhou, China

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *29 August 2021*
Accepted: *01 November 2021*
Published: *25 November 2021*

Citation:
Xiang R, Wang M, Lin L and Wu D (2021) A Research on the Crisis Spillover Effect of Food Safety Incidents on Competitive Firms: The Influence of Political Connections and Charitable Donations. Front. Public Health 9:766385. doi: 10.3389/fpubh.2021.766385

INTRODUCTION

Taking the perspective of corporate social responsibility and institutional theory, this research establishes an innovative relationship between variables such as charitable donation, political connection and crisis spillover effect of firms through quantitative analysis using the event study method, regression analysis and the Heckman two-stage model. Taking 8 food safety incidents from 2011 to 2016 as research samples, this paper studies the impact of food safety incidents on

the market value of both firms under crisis and their competitive firms, as well as the influence of political connection and charitable donation. The conclusions are as follows: first, the occurrence of food safety incidents not only has a negative impact on the market value of the crisis firm, but also has a negative spillover effect on the competitive firm; second, charitable donations made by the competitive firm before the crisis demonstrates a positive competitive effect on the competitive firm, and the intensity of such charitable donations is positively correlated with this positive competitive effect; third, the political connection of the competitive firm has no significant impact on the crisis spillover effect. These findings provide enlightenment for the operation and management of firms in the food industry.

As a matter of fact, crisis management of firms is a research subject that is constantly updated and iterated. Researches on individual firms are gradually extended to researches on the industry, and crisis spillover has gradually become a new research field in crisis management of firms. Faced with highly-developed social media networks and closely-connected supply chains, firms can hardly fight any battle alone in the context of industrial crisis. This paper mainly explores the following four major issues as regard to this circumstance. First, how would market value of the firms under crisis change with the occurrence of food safety incidents. Second, with such incidents happening, whether competitive firms will be affected by the crisis spillover effect; and if so, will contagion effect or competition effect take place? Third, how would competitive firms' political connection influence such crisis spillover effect? And finally, how would charitable donations made by competitive firms impact the crisis spillover effect?

HYPOTHESES

Food Safety Incidents and Market Value of Firms

It is found that the biggest and most harmful food safety incident in American history happened in the well-known fast-food chain Jack-in-the-box and has caused illnesses to more than 700 people and death of 4 children across America, due to excessive addition of Escherichia coli in food. This fast-food chain has not only lost tens of billions of dollars in sales, but also continued to pay the price for such incidents many years later (1). The event study method is often used to study stock market changes during the crisis window, which is measured by the abnormal rate of return (hereinafter referred to as AR). During the crisis, AR is the difference between the expected rate of return and the actual rate of return, which reflects the positive or negative direction of the spillover effect (2–4). A positive AR shows that the event is satisfactory and the future rate of return of the firm is also positive. While a negative AR means that the event is not welcomed and the firm's earnings will be negatively affected in the future. Extant studies have shown that, food safety incidents, in general, will lead to negative impact on the market value of the crisis firm and will also cause positive or negative influence to the competitive firms depending on different scenarios (5, 6). Generally speaking, crisis spillover effect can be further divided into the competition effect, which is positive, and the contagion effect, which is negative. Contagion effect refers to the damage to the market value of peer firms due to the impact of crisis events. On the contrary, competition effect refers to the profits gained by firms in the same industry due to the crisis of the competitors (7). According to the categorization and priming theories, a high degree of overlap fosters assimilation and a low degree of overlap fosters contrast (8). So, when a crisis occurs, the similarity between the firms involved in the incident and other firms will affect consumers' activation of either the assimilation effect or contrast effect in their brains. If the similarity between the two brands is small, the contrast effect will be activated, which will trigger a positive spillover effect; on the contrary, if the similary range of the two is large, the assimilation effect is more likely to be activated, which will cause negative spillover effect, that is, the contagion effect. And it is usually acknowledged that the similarity among competitive firms in the food industry is fairly high. Therefore, the following hypotheses are proposed.

H1a: Incident of food safety crisis has a negative impact on the market value of the crisis firm.

H1b: With the passage of the incident, its negative impact on the market value of the crisis firm generally shows a weakening trend.

H1c: The occurrence of food safety incident creates crisis spillover effect on the competitive firms, and the contagion effect is greater than the competition effect.

The Relationship Between Political Connection and Crisis Spillover

First, the relationship between political connection and crisis spillover is two-sided. On one hand, political connection may relieve the contagion effect of the crisis spillover. When a crisis occurs to the firm, investors may tend to believe that firms with political connection will be protected by the government from being involved in the crisis (9). Similarly, investors may believe that political ties may help competitive firms gain competitiveness in the presence of crisis spillover (10). On the other hand, political ties may aggravate the contagion effect of the crisis spillover. A large number of studies have shown that firms can benefit from political ties in that it is easier for firms with political ties to obtain scarce resources and legitimacy. Also, firms with political ties will be paid more attention by investors in the market. And once concerned by more people in the market, these firms will be criticized more, which makes it easier to form a contagion effect and affect the performance of these firms.

Second, firms will usually adopt strategies that are in consistency with their own institutional advantages. There is a mutually dependent relationship between politically connected firms and the government. In order to achieve its economic goals, the government will exert pressure on the firms, and in the meanwhile, it will also deregulate or offer certain preferential policies to such firms. In this scenario, consumers' perception of the product quality of the firm will be affected by the deviation of its political connection, which will create a bad impression of the firm. In addition, people tend to have a natural distrust of politically connected firms and believe that such firms are likely

to engage in unethical behaviors (9). The crisis spillover effect mentioned above is directly reflected in the average cumulative abnormal rate of return of the competitive firms. Therefore, the following hypotheses are proposed.

H2a: Political connection of the competitive firm has no significant impact on the average cumulative abnormal rate of return of the competitive firm.
H2b: Political connection of the competitive firm has a positive impact on the average cumulative abnormal rate of return of the competitive firm.
H2c: Political connection of the competitive firm has a negative impact on the average cumulative abnormal rate of return of the competitive firm.

The Relationship Between Charitable Donation and Crisis Spillover

Charitable donation is an important part of corporate social responsibility, while the demonstration of corporate social responsibility is very likely to weaken the negative spillover effect of the crisis (11). On one hand, actions related to corporate social responsibility taken before the crisis shows that the firm has some idle resources. Generally speaking, the more charitable donations a firm makes, the more idle resources it would have (12). After the occurrence of the crisis, for the unaffected competitive firm, its stock price will decrease if it is found to be related to such negative crisis. Therefore, before the occurrence of the crisis, the idle resource signal sent by the competitive firm demonstrated through its corporate charitable donation is not enough to weaken the spillover effect of the negative crisis incident. On the other hand, corporate social responsibility has a strong relationship with consumer attribution, which will affect consumers' purchase and evaluation of the products of the affected firm (13). No matter how consumers attribute the crisis, the action of charitable donation of the competitive firm before the crisis plays an important role in reducing the crisis spillover effect (2).

A crisis may be attributed to factors related to moral issues or capability issues. A moral-issue related crisis involves acts of firms that are contrary to the existing moral standards of consumers, such as those with dishonest conducts; while a capability-related crisis involves products offered by the firm which cannot meet the perceived expectations of consumers. If the crisis incident is considered to be a moral crisis, the halo effect formed by the previous actions of the competitive firm representing its corporate social responsibilities will be helpful in separating it from the impact of the crisis. If the crisis is attributed to capability issues, then, charitable donations can effectively alleviate the crisis spillover effect by affecting consumers' perception of the brand or firm under crisis. Some scholars have proposed that consumers' evaluation of the brand or firm under crisis will form a hypothesis-confirming context, and it is based on this background knowledge that consumers will understand the crisis; thus, this hypothesis-confirming context will help reduce the negative evaluation of the brand or the firm (14). More importantly, it is worth mentioning that these evaluations generally are made based on actions related to corporate social responsibilities (15).

Charitable donation is an important act representing corporate social responsibility, which affects consumers' attribution, and in turn, consumers' perception of the brand and the firm. Once a crisis occurs, due to information asymmetry in the market, consumers tend to use their existing rigid knowledge to evaluate relevant competitive firms, which will indirectly affect the spillover effect of the crisis (16). The crisis spillover effect described above is directly reflected by the average cumulative abnormal rate of return of the competitive firm. Therefore, the following hypotheses are put forward.

H3a: Charitable donation made by the competitive firm before the crisis will positively influence the average cumulative abnormal rate of return of the competitive firm.
H3b: The intensity of the charitable donation made by the competitive firm before the crisis has a significant influence on the average cumulative abnormal rate of return of the competitive firm.
H3c: The greater the intensity of the charitable donation made by the competitive firm before the crisis, the greater the positive influence it will have on the average cumulative abnormal rate of return of the competitive firm.

RESEARCH DESIGN

Sample Selection and Data Source

This paper takes 8 firms under crisis related to food safety incidents as the research object, while selecting their corresponding competitive firms, respectively, with the same or similar range of business operations, and studies the spillover effect of the food safety incidents based on the stock price changes of both the crisis firms and the competitive firms.

After sorting out news reports and government regulatory releases on food safety incidents in the food industry from 2010 to 2016, and excluding food safety incidents of non-listed companies and non-mainland A-share listed companies, eight food safety incidents of listed companies were eventually selected as the research object, with a total of 70 research samples.

Variables

Three kinds of variables are involved in this study, namely, the core independent variables, dependent variables and control variables. Based on the extant literature, these three types of variables, which are used in subsequent empirical study, are defined as below. **Table 1** provides the description of each specific variable.

EMPIRICAL TEST AND RESULT ANALYSIS

Quantitative Analysis Based on the Event Study Method

Defining the Events

According to the extant literature, the event study method typically applies to two main types of events. The first type of events in all samples are of the same nature, but maybe with

TABLE 1 | Description of the research variables.

Type	Name	Measurement
Core independent variables	If the firm has donated (ifDon)	Take the amount of charitable donations of the firm in the previous year. If there are donations, assign 1; if there is no donation, assign 0.
	Intensity of charitable donation (Don)	Take the natural logarithm of the donation expenditure out of the non-operating expenditure from the annual report data of the listed company in the previous year
	Political connection (Pol)	Take whether the key executives (chairman or general manager) of the company currently or previously served in the central government, local government, military, CPPCC (Chinese People's Political Consultative Conference) and National People's Congress in the year of the incident. If yes, assign 1; if no, assign 0.
Dependent variables	Average cumulative abnormal rate of return of the competitive firm (PCAR)	AR = ER-Actual rate of return CAR = ΣAR CAR = CAR/N
Control variables	Firm size (size)	The natural logarithm of total assets at the end of the previous year
	Asset liability ratio (debt)	Percentage of total corporate liabilities in total corporate assets in the annual report of the previous year
	Cashflow (CF)	Cash and cash equivalents/total operating income in the previous year's annual report
	Ownership concentration (OC)	Shareholding ratio of top ten shareholders in the previous year
	Incident identifier (incident)	Assign values of 1, 2, 3, 4, 5, 6, 7, and 8 to each specific incident

different dates of occurrence. The second type is a single event, that is, the events and times to be studied in all samples are the same. This paper studies eight food safety events in crisis firms, all of which are associated with the spillover effects of food safety events. So they are typically similar events, but with inconsistent time of occurrence.

Defining the Estimation Period and Window Period

The estimation period and time window period need to be defined to reflect the influence of events with the changes of stock prices during the studied period. The date of occurrence of the event should be included in the event period, which includes a period of time before and after the event. The stock price some time before the event can capture the responses of the capital market on the eve of the event, which may affect the stock price due to leakage of information and other reasons. The stock price for a period of time after the event can clearly reflect the attitude of the capital market toward such event. The estimation period is defined to estimate the rate of return before the event occurs by using the data of the estimation period. The expected rate of return is subtracted from the actual rate of return after the event to obtain the abnormal rate of return brought by the event. In this study, the event window period is defined as 21 days before and after the event day, i.e., the event window is [- 10,10], and the estimation period is defined as 70 days to 11 days before the event, i.e., the estimation window is [−70, −11].

Determining the Research Object

This paper takes 8 A-share listed firms encountering crisis with food safety events as the main research object, and studies the spillover effect of food safety events based on the stock price changes of crisis firms and competitive firms by associating the crisis firms with different numbers of competitive firms within the same or similar business. After excluding issues related to incomplete data and so on, 70 research samples were eventually obtained according to classification based on the similarity of main businesses among the firms.

Calculation of Expected Rate of Return

Event study method is a research method used to analyze the impact of specific events on the company's stock price, which offers a means to measure the change of the firm value. In a rational financial market, the impact of an event will be immediately reflected in the stock price, so the impact of the event can often be considered through the change of the stock's return on equity (17). Usually, an appropriate estimation model needs to be selected and the stock price information in the estimation period needs to be used to determine the expected rate of return. According to the extant literature, there are mainly three estimation models, namely, the mean-adjusted returns model, the marketing-adjusted returns model and the OLS market model. This study adopts the market model to carry out the subsequent research.

The market model is the most complex one among the three models, and it is also the most widely used research model in the event study method (18, 19). The market model assumes that there is a stable linear relationship between market return and individual stock return. The market model is based on the data in the estimation period [−70, −11] and the event window period [−10, 10], and the regression model is established using the least square method (see as follows).

$$R_{it} = \alpha_i + \beta_i R_{mt} + \xi_{it}$$
$$E(\xi_{it}) = 0 \quad Var(\xi_{it}) = \sigma^2_{\xi_{it}}$$

In the equations shown above, R_{it} is the expected return of stock i on day t; R_{mt} is the market rate of return on day t;α_i and β_i are both estimating parameters; ξ_{it} is the residual term. In this paper, different indicators were selected to represent the market rate of return according to different indicators actually used by different exchanges for the listed companies. To be specific, for listed

companies in Shanghai Stock Exchange, Shanghai composite index rate of return is used as the indicator of the market rate of return; and for listed companies in Shenzhen Stock Exchange, Shenzhen component index rate of return is used as the indicator of the market rate of return.

Calculation of Abnormal Rate of Return

In the event study method, calculating the abnormal rate of return of stocks is a key step. Abnormal rate of return refers to the difference between the actual rate of return and the expected rate of return, which is shown as follows.

$$AR_{it} = R_{it} - E(R_{it})$$

In the above equation, AR_{it} is the abnormal rate of return for stock *i* on day t of the incident period; R_{it} is the actual rate of return for stock *i* on day t of the incident period; and $E(R_{it})$ is the expected rate of return for stock *i* on day t of the incident period.

$$AAR_{it} = \overline{AR_{it}} = \frac{1}{N}\sum_{i=1}^{N} AR_{it}$$

In order to more intuitively show the impact of the incidents on stock value, while studying the abnormal rate of return of stocks during the incident period, it is also necessary to calculate the cumulative abnormal rate of return of sample firms $CAR_i(t_1,t_2)$,$T_1 < t_1 \leq t_2 \leq T_2$, according to accumulated period of time (shown as follows).

$$CAR_i\left(t_1, t_2\right) = \sum_{t=t_1}^{t_2} \overline{AR_{it}}$$

According to the known average abnormal rate of return, the time can be further aggregated to obtain the cumulative average abnormal rate of return of individual stocks in the incident period (shown as follows).

$$CAAR = \overline{CAR_i(t_1, t_2)} = \frac{1}{N}\sum_{t=t_1}^{t_2} \overline{AR_{it}}$$

In the above equation, CAAR represents the cumulative average abnormal rate of return, and N represents the number of days.

Through the above calculation, the cumulative abnormal rate of return and average cumulative abnormal return of the sample in the incident period can be obtained (see **Table 2**).

Figure 1 shows the average abnormal rate of return and cumulative average abnormal rate of return of crisis firms during the incident period. It can be obviously seen that the average abnormal rate of return fluctuates around the value zero, and the fluctuation range of negative values is greater than that of the positive values. Therefore, the occurrence of food safety incidents has a negative impact on the stock price earnings of crisis firms. The change in the rate of return on the day of the incident fell into the lowest trough, indicating that the market responded quickly after receiving the crisis information. However, from the overall sample of crisis firms, with efforts made in public relations, clarification and announcement by the crisis firms, the negative impact gradually decreases, and the cumulative average abnormal rate of return begins to rise gradually after T7.

Figure 2 shows the average abnormal rate of return and cumulative average abnormal rate of return of competitive firms during the incident period. Due to the fact that competitive firms share the same or similar range of business with crisis

TABLE 2 | AAR and CAAR of two types of sample firms during the incident period.

Day of the incident	Crisis firm		Competitive firm	
	Average abnormal rate of return (AAR)	Cumulative average abnormal rate of return (CAAR)	Average abnormal rate of return (AAR)	Cumulative average abnormal rate of return (CAAR)
T-10	-0.00715	-0.00715	-0.00079	-0.00079
T-9	-0.00145	-0.0086	0.000559	-0.00023
T-8	0.01028	0.001681	-0.0039	-0.00412
T-7	0.012063	0.013744	0.004702	0.000579
T-6	0.013732	0.027476	0.007657	0.008236
T-5	-0.008	0.019477	-0.00426	0.003978
T-4	-0.00484	0.014634	0.000593	0.004571
T-3	0.011081	0.025715	0.001348	0.005919
T-2	-0.04837	-0.02265	0.001277	0.007196
T-1	-0.08899	-0.11165	-0.02452	-0.01733
T0	-0.15027	-0.26191	-0.03778	-0.05511
T+1	-0.12493	-0.38684	-0.05655	-0.11166
T+2	-0.0558	-0.44192	-0.03268	-0.14434
T+3	-0.0195	-0.46177	-0.01214	-0.15648
T+4	-0.0095	-0.47142	0.007338	-0.14914
T+5	-0.0158	-0.48722	-0.00328	-0.15242
T+6	-0.01418	-0.50139	0.004599	-0.14782
T+7	-0.0233	-0.52492	-0.00015	-0.14797
T+8	0.01805	-0.50692	-0.00638	-0.15435
T+9	0.005681	-0.50124	-0.00301	-0.15735
T+10	-0.00976	-0.511	0.00039	-0.15693

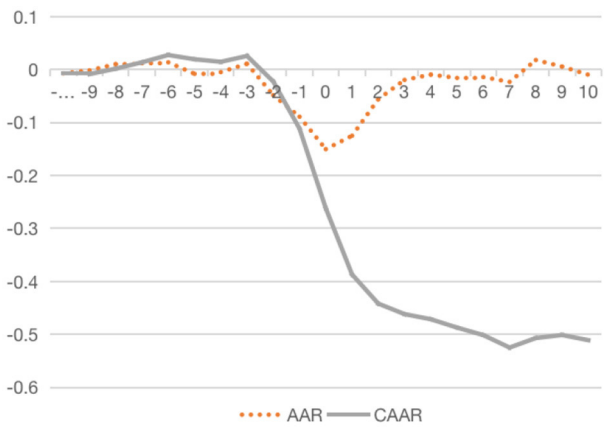

FIGURE 1 | AAR and CAAR of crisis firms during the incident period.

firms, these competitive firms also, unpreventably, underwent suspicions by the market after the crisis incident; however, it can be seen that the lowest average abnormal rate of return of the competitive firm appears on day T1, which is 1 day after the appearance of the lowest value experienced by the crisis firm. In the meanwhile, the relative recovery period for the competitive firms is also shorter, and the average cumulative abnormal rate of return goes back to positive value on day T4.

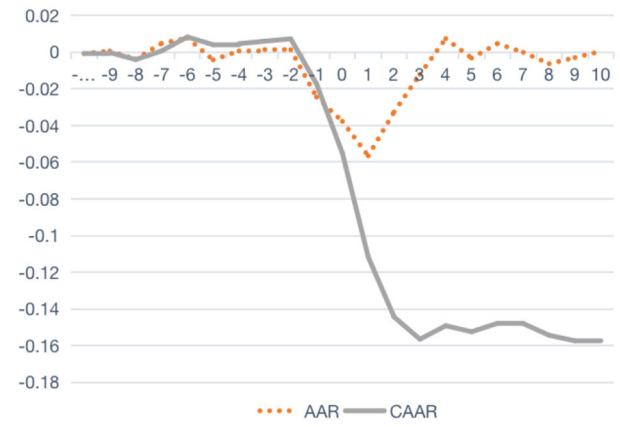

FIGURE 2 | AAR and CAAR of the competitive firm during the incident period.

T-Test

After obtaining the average abnormal rate of return and the cumulative average abnormal rate of return, a statistical test should be carried out to verify whether the incident has an impact on the stock value at a certain significance level. This paper uses *t*-test to determine whether the average cumulative abnormal rate of return is significantly different from 0. The results are shown in **Table 3**.

In the above table, mean value represents the standardized average abnormal rate of return, T value represents the standardized average abnormal rate of return, and *P*-value represents the level of significance. Through the test of the above statistics, it is found that during the incident period, food safety incidents have a significant impact on the stock prices of both crisis firms and competitive enterprises, and there is a negative contagion effect presenting on the competitive firms. Within [T-3, T3] days of the incident period, the normalized average abnormal rate of return is significantly different from 0 at the 90% confidence level; within [T-2, T2] days of the incident period, the normalized average abnormal rate of return is significantly different from 0 at the 95% confidence level, and even more significantly different from 0 at the 99% confidence level within the day before and after the incident. Therefore, hypothesis H1a is verified, i.e., food safety crisis of the firm has a negative impact on the market value of the crisis firm. As can be seen from **Table 3**, for competitive firms, the average cumulative abnormal rate of return shows an upward trend over time, so hypothesis H1b is also supported.

TABLE 3 | Summary of significance of statistics T_{AR} of two types of sample firms during the incident period.

Day of the incident	SAR of crisis firm			SAR of competitive firm		
	Mean (%)	T value	*P*-value	Mean (%)	T Value	*P*-value
T-10	−0.7147	−1.757	0.122	−0.079	−0.264	0.793
T-9	−0.1452	−0.321	0.824	0.056	0.209	0.835
T-8	1.028	0.828	0.435	0.340	−1.010	0.317
T-7	1.206	1.392	0.207	3.065	1.208	0.232
T-6	1.373	0.790	0.456	0.766	1.906	0.061
T-5	−0.799	−1.048	0.330	−0.425	−1.237	0.221
T-4	−0.484	−1.219	0.262	0.059	0.207	0.837
T-3	1.108	1.842	0.009**	0.200	0.532	0.597
T-2	−4.837	−6.017	0.001***	0.128	0.536	0.594
T-1	−8.899	−12.581	0.000***	−2.452	−8.336	0.000***
T0	−15.03	−10.051	0.000***	−3.778	−6.996	0.000***
T1	−12.49	−6.843	0.001***	−5.655	−8.932	0.000***
T2	−5.508	−5.138	0.004**	−3.268	−10.775	0.000***
T3	−1.985	−4.342	0.010*	−1.214	−3.562	0.001***
T4	−0.964	−0.512	0.627	0.734	2.007	0.049**
T5	−1.579	−1.115	0.308	−0.328	−1.125	0.265
T6	−3.352	−3.605	0.100	0.460	1.588	0.118
T7	1.801	0.799	0.455	−0.015	−0.50	0.960
T8	0.986	0.713	0.548	−0.637	−2.292	0.025**
T9	0.568	0.688	0.517	−0.301	−0.928	0.357
T10	0.028	0.039	0.970	−1.744	−3.386	0.001

****P* < 0.001, ***P* < 0.05, **P* < 0.10.

It can be noticed through the performance of the competitive firms that, within [T-1, T4] days of the crisis period, the average abnormal rate of return of the competitive firms after standardization is significantly different from 0 at the 95% confidence level, and is significantly negative. For these competitive firms, the average cumulative abnormal rate of return also shows an obvious negative relationship. Therefore, H1C is supported, i.e., the occurrence of food safety incidents of crisis firms has a spillover effect on their competitive firms, and seen from the overall picture, the contagion effect is greater than the competition effect.

The above quantitative research mainly explores our first research question, that is, after the occurrence of food safety incident, whether the crisis firm will bring about a spillover effect on the competitive firms. The results show that the crisis firm causes an obvious contagion effect on the competitive firms; however, with the development of the incident and the measures taken by the competitive firms, this contagion effect weakens in a gradual manner. Unfortunately, the utilization of the event study method can only explore issues such as the market value of competitive firms caused by the food safety incident; while factors affecting the change of market value of competitive firms involve various comprehensive facets. Therefore, based on the above research, our next step is to further study the impact of the characteristics and actions of competitive firms on their market value, i.e., whether the charitable donation of competitive firms before the crisis can regulate the crisis spillover effect; and if yes, how this spillover effect can be regulated, and whether the greater the donation amount is, the stronger this effect will be; in addition, in what direction the political connection of competitive firms will regulate the crisis spillover effect? Based on these two research questions, an empirical study will be carried out as below.

Empirical Analysis

Model Design

In order to study the impact of charitable donations and political connections of competitive firms before the crisis on their average cumulative abnormal rate of return, a regression equation model is built for empirical analysis. The detailed model is shown as follows.

$$\begin{aligned} PCAR_i &= \alpha + \beta_1 * Pol + \beta_2 * ifDon + \beta_3 * Pol * ifDon \\ &+ \beta_4 * size + \beta_5 * debt + \beta_6 * CF + \beta_7 * OC \\ &+ \beta_8 * event + \xi \end{aligned}$$

In this equation, $PCAR_i$ represents the average cumulative abnormal rate of return of competitive firms, where $i = 1,10$; Pol indicates the political connections of key executives of the competitive firm during the crisis period; if Don represents whether the competitive firm had made charitable donations in the previous year before the crisis occurred; size indicates the size of the competitive firm; Debt represents the asset liability ratio of the competitive firm; CF represents the cash flow of the competitive firm; OC represents the equity concentration of the competitive firm; incident is the incident identifier. β_i represents the regression coefficient to be estimated, where $i = 1,2,3,\ldots,8$; ε is the random residual term.

TABLE 4 | Descriptive statistics of variables of sample firms.

Variable	Mean	Std. Dev.	Min	Max
PCAR1	−0.007	0.011	−0.037	0.024
PCAR10	−0.001	0.009	−0.018	0.021
Don	7.202	6.347	0.000	17.925
IFDON	0.597	0.495	0.000	1.000
Pol	0.581	0.497	0.000	1.000
Size	21.826	1.105	19.411	24.332
Debt	0.429	0.187	0.058	1.063
cf	0.290	0.288	0.000	1.455
oc	58.273	17.091	17.820	88.410
CASE	4.952	2.161	1.000	8.000

Descriptive Analysis

This paper uses EViews 7.0 to carry out ADF test. It is found that the individual stock returns and relevant market returns of the sample firms remain series stationarity, therefore, corresponding correlation regression analysis can be carried out. Before the econometric analysis, descriptive statistics is conducted to observe the sample data.

Table 4 is based on the descriptive statistics of the variable data of 62 competitive firms. It is found that the average cumulative abnormal rate of return of competitive firms in the 1-day window period is significantly higher than that in the 10-day window period, indicating that with the continuous development of the incident, the negative crisis effect of the incident on competitive firms continues to weaken. From the data as regard to whether firms had made donations or not, the average value exceeds 0.5, indicating that most firms are willing to make charitable donations. However, from the perspective of donation intensity, different firms have shown relatively great disparity according to the data on previous donations.

Correlation Analysis

Before regression analysis of the equation, the correlation of each variable must be determined through correlation analysis. Correlation analysis aims to test whether there is a relationship between each variable, and uses correlation coefficient to quantify the strength of this relationship.

According to **Table 5** showing the correlation coefficient between sample variables, it is found that the core explanatory variables ("donate or not" and "donation intensity") are significantly and positively correlated with the explained variable (the average cumulative abnormal rate of return of competitive firms) in the two incident windows at the confidence level of 99%; also, political connection is positively correlated with the explanatory variable in the 1-day window period, and negatively correlated with the explanatory variable in the 10-day window period, but not significant enough. After adding other control variables, we further study the relationship between political connection and average cumulative abnormal rate of return

TABLE 5 | Correlation coefficients of variables of sample firms.

Variable	PCAR1	PCAR10	Don	IFDON	Pol	Size	Debt	cf	oc	CASE
PCAR1	1.000									
PCAR10	0.392***	1.000								
Don	0.480***	0.295**	1.000							
IFDON	0.474***	0.301**	0.940***	1.000						
Pol	0.128	−0.114	0.253**	0.234*	1.000					
Size	−0.020	−0.142	0.186	0.096	0.310**	1.000				
Debt	0.029	0.009	0.083	0.066	0.284**	0.188	1.000			
cf	−0.205	−0.134	−0.083	−0.091	−0.013	−0.023	−0.517***	1.000		
oc	0.078	−0.016	0.085	0.115	0.003	0.371***	−0.108	0.038	1.000	
CASE	−0.042	−0.362***	−0.207	−0.141	0.072	−0.094	0.171	−0.030	0.075	1.000

****P < 0.001, **P < 0.01, *P < 0.05, two-tailed test.*

TABLE 6 | The results of regression analysis of [−1,1] and [−10,10] window period.

	[−1,1] window period			[−10,10] window period		
PCAR1	**Model 1**	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
IFDON	0.00324**		0.00310**	0.00565***		0.00421***
	(0.00021)		(0.00025)	(0.00013)		(0.00023)
Pol		0.00132	0.00099		0.00180	0.00096
		(0.00141)	(0.00175)		(0.00092)	
Size	0.00019**	0.00023**	0.00031***	0.00029**	0.00027**	0.00031***
	(0.00009)	(0.00009)	(0.00009)	(0.00018)	(0.00098)	(0.00009)
Debt	−0.00344	−0.00524	−0.00504	−0.00101**	0.0117**	0.0118**
	(0.00823)	(0.0109)	(0.00866)	(0.00433)	(0.00486)	(0.00494)
cf	0.00173*	0.00212*	0.00184*	0.00200**	0.00185*	0.00197*
	(0.00117)	(0.00152)	(0.00164)	(0.00117)	(0.00135)	(0.00134)
oc	0.0008**	0.0010*	0.0012**	0.0012**	0.0009***	0.0011**
	(0.00003)	(0.00007)	(0.00047)	(0.00043)	(0.00001)	(0.00052)
Constant	0.0256	0.0305	0.0314	0.00271	−0.00386	−0.00360
	(0.0255)	(0.0289)	(0.0246)	(0.0185)	(0.0211)	(0.0212)
Case effect	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.574	0.491	0.581	0.633	0.610	0.646

****P < 0.001, **P < 0.01, *P < 0.05, two-tailed test, two tailed test; the values listed in the chart are regression coefficients; the values in parentheses are standard errors.*

of competitive firms. The average cumulative abnormal rate of return of competitive firms in the 1-day window period is significantly and positively correlated with that in the 10-day window period, and there is also a significant positive correlation between "donate or not" and "donation intensity."

Regression Analysis

It can be seen from **Table 5** that the correlation coefficients of all core explanatory variables, explained variables and control variables to be regressed are <0.700 except for the two variables "donate or not" and "donation intensity"; however, these two variables ("donate or not" and "donation intensity") will not enter the regression equation together, therefore, the problem of multicollinearity between each variable is not serious. In order to further diagnose the multicollinearity problem of the model, tests of tolerance and variance expansion factor are conducted. The results show that the maximum value of VIF is no more than 4.8, and the value of TOL is far >0.1, so it can be determined that the model has no multicollinearity problem, and regression analysis can be carried out. In order to explore the impact of political connections and charitable donations on the average cumulative abnormal rate of return of competitive firms in different incident window periods, a 1-day window period was selected to test the timeliness of market response and a 10-day window period is selected to strengthen the robustness.

Table 6 summarizes the regression results of the model in the 1-day and 10-day window periods. Different incident identifiers are all controlled in the model to avoid the deviation of results caused by sample data not being in the same incident. Model 1 shows the regression result of the core explanatory variable ("donate or not") and the explained variable (average cumulative abnormal rate of return of competitive firms) after controlling

each control variable. It is found that for competitive firms, the coefficients of "donate or not" before the crisis and their average cumulative abnormal rate of return are positively significant at the levels of 95% and 99%, respectively. This means that whether the competitive firms have made charitable donations will effectively weaken the crisis spillover effect during the event period. Thus, Hypothesis H3a holds. In addition, all control variables have passed the significance test, which shows that these control variables have played an effective role in the equation. Model 2 is the model showing relationship between political connection as the core explanatory variable and the average cumulative abnormal rate of return of competitive firms as the explained variable. The results show that there is no significant relationship between the political connection of competitive firms and their average cumulative abnormal rate of return, which is consistent with what is proposed in hypothesis H2a. Model 3 shows the result when the core explanatory variable (political connection) and "donate or not" are entered into the model at the same time, the result of which is generally consistent with the previous conclusions.

The Heckman Two-Stage Model

The above regression analysis has verified that charitable donation actions of competitive firms will alleviate the negative effects of crisis spillover. However, the difference made by the positive roles of different levels of donation intensity remains to be explored. In view of this, this study applies the Heckman two-stage model for regression analysis. The application of the Heckman two-stage model mainly involves consideration of the following two aspects. First, this model takes the virtual variable of "donate or not" as the explained variable. When there is donation made by the competitive firm before the crisis, the variable is set to be 1, otherwise it is set to be 0. According to the above, a probit model recorded as model 1 is established to study the impact of donation made by the competitive firm before the crisis on the crisis spillover effect. Second, model 2 or 3 are established on the premise that the competitive firm has made charitable donation, that is, the virtual variable of ifDon is 1; then the variable of real charitable donation intensity "Don" is used as the explanatory variable to examine the specific impact of the intensity of the charity donation made by the competitive firm on the spillover effect.

The regression results are shown in **Table 7**. Model 1 shows that the political connection of the competitive firm is positively correlated with whether it donates at the 95% significance level, which means that political connection of the competitive firm will enhance the probability of its charitable donation. Firms may carry out public welfare donation with its own strategic concerns in order to establish effective ties with the government. Model 2 shows that in the 1-day window period, the intensity of charitable donations is positively correlated with the average cumulative abnormal rate of return of competitive firms at a significant level of 95%, which means in the 1-day window period, the more charitable donations made by the competitive firms, the more effective it is to alleviate the negative crisis spillover effect. Model 3 shows similar result to that of model 2 in that the intensity of charitable donation is positively correlated with the average cumulative abnormal rate of return of competitive firms at a significant level of 99% in the 10 day window period, which means compared with the 1-day window period, the effect of the intensity of donation of competitive firms on alleviating the negative spillover effect of the crisis is more obvious in the 10 day window period. Therefore, H3b and H3c both hold.

TABLE 7 | Summary of the regression results of Heckman two-stage model.

	Model 1	Model 2	Model 3
VARIABLES	**IFDON**	**PCAR1**	**PCAR10**
Don		0.00375**	0.00591***
		(0.00124)	(0.00001)
Pol	0.681**	0.00164	0.00175*
	(0.072)	(0.00113)	(0.00107)
Size	0.0327**	0.00036***	0.00026**
	(0.0173)	(0.00007)	(0.00014)
Debt	−0.372	0.0198	0.0284
	(1.110)	(0.379)	(0.179)
cf	0.554**	0.00173*	0.00243**
	(0.182)	(0.00154)	(0.00087)
oc	0.00996***	0.0011**	0.0013***
	(0.0009)	(0.00016)	(0.00003)
Constant	0.318	0.116	0.0731
	(3.487)	(1.186)	(0.559)
Case effect	No	Yes	Yes

****P < 0.001, **P < 0.05, *P < 0.10.*

In terms of control variables, it can be seen in model 1 that the size, cash flow and equity concentration of the firm are significantly positively correlated with whether the firm donates, indicating that the larger the firm size, the more sufficient the cash flow and the higher the equity concentration, the more likely the firm is to make charitable donations. In models 2 and 3, firm size, cash flow and equity concentration are positively correlated with the average cumulative abnormal rate of return of competitive firms, which is consistent with the results of previous research.

CONCLUSIONS AND IMPLICATIONS

Research Conclusions

Through methods such as event study, regression analysis and Heckman two-stage model, this paper explores the crisis spillover effect of food safety incidents on firms in the food industry, and conducts an in-depth study on the impact of political connections and charitable donations of competitive firms on their average cumulative abnormal rate of return. It is found that the occurrence of food safety incidents will have a negative impact on the market value of crisis firm, but with the passage of the incident, the effect of the negative impact continues to weaken. The main reason is that after the incident, crisis firms will take effective countermeasures, such as clarification, apology, distinction, etc.; in addition to the fact that consumers' attention is limited after all, both consumers and the media will eventually pay less attention to the incident; and with involvement of other influencing factors, the negative effects will be weakened with

the passage of time. Furthermore, the occurrence of food safety incidents will also exert an impact on competitive firms inside and outside the industry. The results show that the overall crisis spillover effect of food safety incidents on competitive firms is a negative one, to be exact, a contagion effect. It is noticed from the overall sample that, in the first 3 days of the incident, due to the possible early disclosure of the incident information and the small-scale spread of information related to the incident, consumers will have a sense of distrust of competitive firms in the same industry. However, with the gradual unfolding of the incident, consumers and investors will come to understand that the incident may only be caused by only an individual firm or a small group of firms due to some illegal operation. After such clarification of facts, the negative impact of the incident will be gradually weakened.

In addition, our research results show that the political connection of competitive firms before the crisis has no significant effect on their average cumulative abnormal rate of return during the crisis period. In other words, political connections of competitive firms play no significant positive or negative role in influencing average cumulative abnormal rate of return during the crisis period. This may due to the fact that there are too few samples collected in this paper, and the implicit impact of political connection on the yield of the competitive firms is yet to be accurately captured. However, the donation actions made by the competitive firms before the crisis will increase their cumulative abnormal rate of return during the crisis, meaning that the charitable donation actions of competitive firms will cause significant positive effect. The greater the intensity of such charitable donation made by the competitive firms before the crisis, the more positive the impact on their average cumulative abnormal rate of return; and the significance of this impact increases with the extension of the incident window.

Managerial Implications

On the basis of quantitative and empirical research, this paper verifies the existence of crisis spillover effect, discusses the relationship between political connection, charitable donation and crisis spillover effect, and reaches a few practical conclusions. Although only the food industry is explored, this study still provides useful implications for firms in various industries in avoiding negative spillover effect of crisis and striving for positive spillover effect in the future. To be specific, this study provides managerial implications at both industrial and firm levels.

At the industrial level, the industry should strengthen the norms of industrial autonomy. Food industry, especially, is an industry with a high incidence of crisis. Once a food safety incident occurs, it will not only put the crisis firm into business meltdown, but also affect a large number of competitive firms in the same industry and bring a significant impact to the whole industry. In detail, trade associations should play an active role in strengthening the autonomy and standardization of the industry, constructing the review mechanism of the industry, and guiding the whole industry to form a high-quality business model, which takes quality as the foundation and services as the guarantee.

Seen from the firm level, first of all, firms should lay absolute emphasis on product quality. Quality is the foundation of a firm's survival, and ensuring product quality should be considered the most fundamental and prioritized obligation for a firm. To be specific, firms should establish an effective safety management system to ensure product quality from various aspects such as the sources of raw materials, production, processing, distribution and so on. Second, attention should be paid to information related to industrial crisis. Firms should not develop themselves behind closed doors; instead, attention should always be paid to the dynamics of competitive firms through sensing and collecting all kinds of information related to potential and existing crisis, grasping the public opinion guidance of the media and the public, and formulating and taking effective responses in a timely manner. Third, firms should actively undertake corporate social responsibilities. Our research conclusion shows that actions related to corporate social responsibilities in such form as charitable donation can effectively reduce the negative impact during the crisis. Through actively shouldering social responsibilities, firms are able to establish a decent corporate image, play a positive external role in the crisis, thus, protecting or buffering the damage of the crisis.

Limitations and Future Prospects

The limitations of this study are mainly as follows. From the perspective of sample size, only a limited number of incidents of listed companies are used as our samples, which has brought difficulty especially in the observation stage. Also, since most of the incidents studied are individual and independent ones, so the specific performance of systematic incidents in different incident windows can seldom be displayed. Therefore, according to the above-mentioned issues, subsequent research can expand its scope to include not only the domestic A-share market, so as to obtain more samples. In addition, we need to control the types of events in different classifications in order to further study the corresponding disparities.

Furthermore, from the perspective of research scope, the sample of competitive firms selected in this paper are firms with the same or similar main businesses in the industry. However, the scope of spillover effect is actually far more than this. Upstream and downstream firms in the same industry and related firms outside the industry will also be implicitly implicated. Future studies can be conducted through expanding the research scope and classifying the type of incidents in order to explore more distinctive scenarios. In addition, the impact of upstream and downstream firms within the industry and relevant firms outside the industry can also be studied. From the perspective of research methods, this paper mainly uses the event study method to test the spillover effect, and then verifies the existence of such spillover effect through constructed statistics, namely, standardized abnormal rate of return. However, due to the immaturity of China's stock market, our research results may also be affected by other factors within the relatively clean event window. For example, there may be a deviation in calculating the cumulative abnormal rate of return of individual stocks. For the sake of robustness, this issue should be further considered in the future. Therefore, other statistical methods are suggested to be utilized in the future.

Last but not the least, from the perspective of research content, this paper also studies the impact of such variables as political

connection and charitable donation on crisis spillover effect; other factors that may also cause fluctuations in stock prices should be taken into account such as the impact scope and significance of the food safety incidents, the respective industrial status of crisis firms and competitive firms and so on. However, due to the difficulty of obtaining these control variables, these variables are not added in this paper for the sake of time and energy. In the future, more firms should be considered and appropriate indicators should be selected for classification-based researches.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

RX was in charge of the formulation of the general research topic, the construction of the research framework, and the proposing of the theoretical hypothesis. MW generally contributes to the construction of the theoretical framework based on the in-depth accumulation of a large volume of literature reading and analysis and contributes a lot in the data analysis process. LL was in responsible for the overall development of this study, including the selection of research angle, research dimensions, the planning of sample collection, data analysis and proof-reading, and polishing of the whole paper. DW was in charge of data collection, analysis of this study, and responsible for all the procedures taken during data collection. All authors contributed to the article and approved the submitted version.

FUNDING

This research is financially supported by the National Natural Science Foundation of China (Grant number: 71773115), the National Natural Science Foundation of China (Grant number: 71973129), and the Major Bidding Project of National Social Science Fund of China (Grant number: 20&ZD124).

REFERENCES

1. Seo S, Jang SCS, Almanza B, Miao L, Behnke C. The negative spillover effect of food crises on restaurant firms: did Jack in the Box really recover from an *E.* coli scare? *Int. J. Hosp. Manag.* (2014) 39:107–21. doi: 10.1016/j.ijhm.2014.02.011
2. Liu Y, Shankar V, Yun W. Crisis management strategies and the long-term effects of product recalls on firm value. *J Mark.* (2017) 81:30–48. doi: 10.1509/jm.15.0535
3. Shah R, Ball GP, Netessine S. Plant operations and product recalls in the automotive industry: an empirical investigation. *Manage Sci.* (2017) 63:2439–59. doi: 10.1287/mnsc.2016.2456
4. Wei J, Wang Q, Yu Y, Zhao D. Public engagement in product recall announcements: an empirical study on the Chinese automobile industry. *J Mark Commun.* (2019) 25:343–64. doi: 10.1080/13527266.2016.1251487
5. Chang S-C, Chang H-Y. Corporate motivations of product recall strategy: exploring the role of corporate social responsibility in stakeholder engagement. *Corp Soc Responsib Environ Manag.* (2015) 22:393–407. doi: 10.1002/csr.1354
6. Murphy F, Pütz F, Mullins M, Rohlfs T, Wrana D, Biermann M. The impact of autonomous vehicle technologies on product recall risk. *Int J Prod Res.* (2019) 57:6264–77. doi: 10.1080/00207543.2019.1566651
7. Dahlén M, Fredrik L. A disaster is contagious: how a brand in crisis affects other brands. *J Advert Res.* (2006) 46:388–97. doi: 10.2501/S0021849906060417
8. Meyers-Levy J, Sternthal B. A Two-factor explanation of assimilation and contrast effects. *J Mark Res.* (1993) 30:359–68. doi: 10.1177/002224379303000307
9. Jia M, Zhang Z. How long does the influence of organizational deviance have on innocent firms? *J Bus Res.* (2016) 69:2649–63. doi: 10.1016/j.jbusres.2015.11.001
10. Yu X, Zhang P, Zheng Y. Corporate governance, political connections, and intra-industry effects: evidence from corporate scandals in China. *Financ Manag.* (2015) 44:49–80. doi: 10.1111/fima.12064
11. Noack D, Miller DR, Smith D. Let me make it up to you: understanding the mitigative ability of corporate social responsibility following product recalls. *J Bus Ethics.* (2019) 157:431–46. doi: 10.1007/s10551-017-3639-7
12. Surroca J, Tribó JA, Waddock S. Corporate responsibility and financial performance: the role of intangible resources. *Strateg Manag J.* (2010) 31:463–90. doi: 10.1002/smj.820
13. Klein J, Dawar N. Corporate social responsibility and consumers' attributions and brand evaluations in a product–harm crisis. *Int J Res Mark.* (2004) 21:203–17. doi: 10.1016/j.ijresmar.2003.12.003
14. Dawar N, Pillutla MM. Impact of product-harm crises on brand equity: the moderating role of consumer expectations. *J Mark Res.* (2000) 37:215–26. doi: 10.1509/jmkr.37.2.215.18729
15. Kim J. Sustainability in social brand communities: influences on customer equity. *J Glob Sch Mark Sci.* (2015) 25:246–58. doi: 10.1080/21639159.2015.1041783
16. Wei J, Zhao M, Wang F, Zhao D. The effects of firm actions on customers' responses to product recall crises: analyzing an automobile recall in China. *J Risk Res.* (2016) 19:425–43. doi: 10.1080/13669877.2014.983946
17. Malkiel BG. The efficient market hypothesis and its critics. *J Econ Perspect.* (2003) 17:59–82. doi: 10.1257/089533003321164958
18. Gao H, Xie J, Wang Q, Wilbur KC. Should ad spending increase or decrease before a recall announcement? the marketing-finance interface in product-harm crisis management. *J Mark.* (2015) 79:80–99. doi: 10.1509/jm.14.0273
19. Law C, Cornelsen L, Adams J, Penney T, Rutter H, White M, et al. An analysis of the stock market reaction to the announcements of the UK Soft Drinks Industry Levy. *Econ Hum Biol.* (2020) 38:100834. doi: 10.1016/j.ehb.2019.100834

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Public Health

SYSTEMATIC REVIEW
published: 25 November 2021
doi: 10.3389/fpubh.2021.727669

Efficiency of Iranian Hospitals Before and After Health Sector Evolution Plan: A Systematic Review and Meta-Analysis Study

Saeed Amini[1,2], *Behzad Karami Matin*[3], *Mojtaba Didehdar*[4], *Ali Alimohammadi*[5], *Yahya Salimi*[6], *Mohammadreza Amiresmaili*[7] *and Ali Kazemi-Karyani*[3]*

[1] *Department of Health Services Management, Arak University of Medical Sciences, School of Health, Arak, Iran,* [2] *Molecular and Medicine Research Center, Khomein University of Medical Sciences, Khomein, Iran,* [3] *Research Center for Environmental Determinants of Health, Health Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran,* [4] *Department of Medical Parasitology and Mycology, School of Medicine, Arak University of Medical Sciences, Arak, Iran,* [5] *Department of Forensic Medicine, School of Medicine, Arak University of Medical Sciences, Arak, Iran,* [6] *Social Development and Health Promotion Research Center, Health Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran,* [7] *Department of Health Management, Policy Making and Economics, Kerman University of Medical Sciences, Kerman, Iran*

OPEN ACCESS

Edited by:
Mihajlo Jakovljevic, Hosei University, Japan

Reviewed by:
Martin Dlouhy, University of Economics, Czechia
Mariela Deliverska, Medical University Sofia, Bulgaria

***Correspondence:**
Ali Kazemi-Karyani
alikazemi.k20@gmail.com

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *19 June 2021*
Accepted: *27 September 2021*
Published: *25 November 2021*

Citation:
Amini S, Karami Matin B, Didehdar M, Alimohammadi A, Salimi Y, Amiresmaili M and Kazemi-Karyani A (2021) Efficiency of Iranian Hospitals Before and After Health Sector Evolution Plan: A Systematic Review and Meta-Analysis Study. Front. Public Health 9:727669. doi: 10.3389/fpubh.2021.727669

Purpose: Aging, chronic diseases, and development of expensive and advanced technologies has increased hospitals costs which have necessitated their efficiency in utilization of resources. This systematic review and meta-analysis study has assessed the efficiency of Iranian hospitals before and after the 2011 Health Sector Evolution Plan (HSEP).

Methods: Internal and external databases were searched using specified keywords without considering time limitations. The retrieved articles were entered into EndNote considering inclusion and exclusion criteria, and the final analysis was performed after removing duplicates. Heterogeneity between the studies was assessed using Q and I^2 tests. A forest plot with 95% confidence intervals (CI) was used to calculate different types of efficiency. The data were analyzed using STATA 14.

Results: Random pooled estimation of hospitals technical, managerial, and scale efficiencies were 0.84 (95%CI = 0.78, 0.52), 0.9 (95%CI = 0.85, 0.94), and 0.88 (95%CI = 0.84, 0.91), respectively. Sub-group analysis on the basis of study year (before and after HSEP in 2011) indicated that random pool estimation of technical (0.86), managerial (0.91), and scale (0.90) efficiencies of Iranian hospitals for 2011 and before were better than technical (0.78), managerial (0.86), and scale (0.74) efficiencies after 2011.

Conclusion: Type of hospital ownership was effective on hospital efficiency. However, HSEP has not improved hospital efficiency, so it is necessary for future national plans to consider all aspects.

Keywords: efficiency, hospital, ownership, costs, Iran

INTRODUCTION

Hospitals have an undeniable role in providing healthcare services to society but their increasing costs have become an important challenge for many countries. In other words, utilization of technologies and new methods of diagnosis and treatment of diseases and also increasing numbers of elderly citizens, increasing chronic diseases, increasing demands for healthcare services and specialists, and hospital errors have increased health system costs (1, 2). Because of these issues and problems, hospitals always encounter human and financial resource constraints which have necessitated efficiency in consuming resources more than ever (3).

The efficiency concept has been created from the combination of technical and allocative efficiencies. Technical efficiency means using the lowest amount of input to produce a specified amount of output or using a specified amount of input to produce more output. Allocative efficiency means using the correct amount of input in terms of prices to produce a specified amount of output. Technical efficiency, on the other hand, was created by multiplying scale efficiency and managerial efficiency. Scale efficiency is the ability of an organization unit to perform in or near the most profitable scale to prevent loss in resources. Lastly, managerial efficiency means hard working, correct policymaking, application of the correct number of employees, and the correct combination of production factors (4).

One of the most widely used methods in assessment of different decision-making units (DMUs) such as hospitals and other organizations in terms of the components of efficiency (e.g., technical, scale, and managerial efficiency) is the data envelopment analysis (DEA) method. It is possible, through this method, to create a logical framework to distribute human and financial resources between different wards and sections of studied organizations (5). The DEA method, as a non-parametric programming technique, has been used since the mid 1980s to measure DMU efficiency (6). In other words, linear and multiple programming models are used in this method to assess the

Abbreviations: HSEP, Health Sector Evolution Plan; CI, confidence interval; DEA, data envelopment analysis; DMUs, decision making units; SSO, Social Security Organization; ISI, Institute for Scientific Information; SID, Scientific Information Database; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

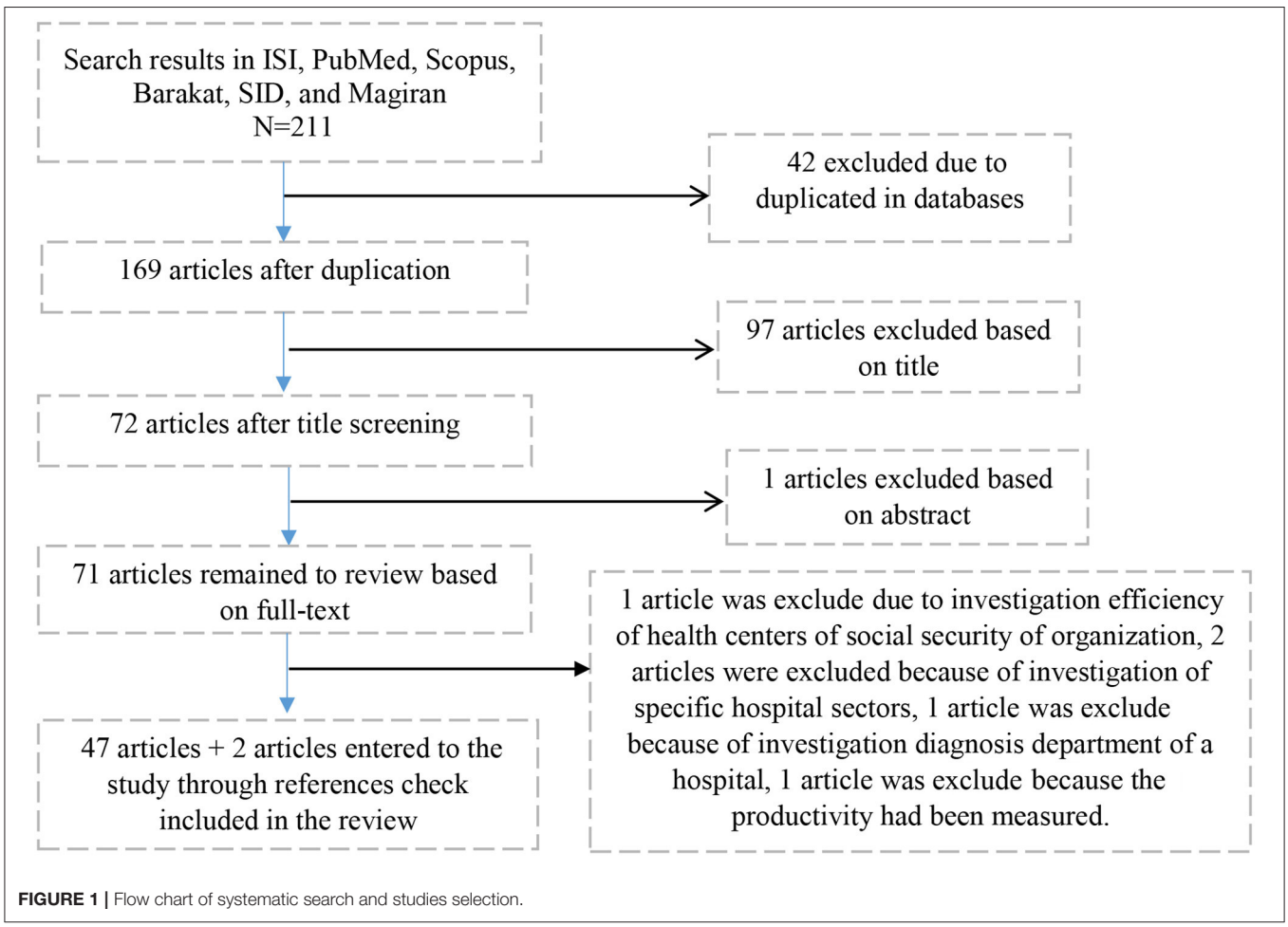

FIGURE 1 | Flow chart of systematic search and studies selection.

TABLE 1 | Characteristics of the studies included in the systematic review and meta-analysis.

Row	Authors	Years of data collection	Language	Location	Affiliation of hospitals	Number of hospital	Inputs	Outputs	Model of DEA
1	Joshan et al. (24)	2011–12	Persian	Tehran	TUMS	14	Number of beds, operation rooms, physicians, nurses, and support forces	Number of surgeries, outpatients, patients, bed occupancy rate, bed day, and admission-inpatient rate	VRS, input-oriented
2	Joshan et al. (24)	2011–12	Persian	Tehran	IUMS	8	Number of beds, operation rooms, physicians, nurses, and support forces	Number of surgeries, outpatients, patients, bed occupancy rate, bed day, and admission-inpatient rate	VRS, input-oriented
3	Joshan et al. (24)	2011–12	Persian	Tehran	SBMU	10	Number of beds, operation rooms, physicians, nurses, and support forces	Number of surgeries, outpatients, patients, bed occupancy rate, bed day, and admission-inpatient rate	VRS, input-oriented
4	Sepehrdost et al. (25)	2007–08	Persian	Iran	SSO	28	Number of medical staff, nurses, other sources, and active beds	Number of outpatients, inpatients, surgeries, and bed turnover	CRS, input-oriented
5	Sepehrdost et al. (25)	2007–08	Persian	Iran	SSO	37	Number of medical staff, nurses, other sources, and active beds	Number of outpatients, inpatients, surgeries, and bed turnover	CRS, input-oriented
6	Ghaderi et al. (26)	2005–09	Persian	Tehran & Alborz	IUMS	26	Number of beds, nurses, and others	Number of surgeries, outpatients, hospitalization day, and occupied bed day ratios	VRS, input-oriented
7	Karimi et al. (27)	2005–06	Persian	Isfahan	MUI	23	Number of physicians, nurses, and beds	Mean patient stay, bed turnover, bed occupancy, number of outpatients, and hospital income	VRS, input-oriented
8	Mohammadi Ardakani et al. (28)	2004–06	Persian	Yazd	SSO	12	Number of physicians, paramedics, and active beds	Number of inpatients and outpatients, occupied bed day	Input- and output-oriented
9	Pourreza et al. (29)	1996–98	Persian	Tehran	TUMS	12	Number of beds, nurses, physicians, and others	Number of outpatients, hospitalization-day, number of surgeries	VRS, input-oriented
10	Aboulhalaj et al. (30)	2009	Persian	Iran	MHH	122	Number of beds, physicians, paramedics, and others	Income and admission	VRS, input-oriented
11	Salehzade et al. (31)	2007	Persian	Qom	MUQ & Self-administered	8	Number of physicians, paramedics, and active beds	Number of outpatients and inpatients	VRS, input-oriented
12	Salehzade et al. (31)	2007	Persian	Qom	MUQ & Self-administered	8	Number of physicians, paramedics, and active beds	Number of inpatients and outpatients	CRS, input-oriented
13	Asadi et al. (32)	2008	Persian	Yazd	SSO	13	Costs, education hours, and staff relocation	SERVQUAL score, ratios, outpatient, inpatient, and emergency patients to physicians score	Input- and output-oriented (input-oriented)
14	Askari et al. (33)	2001–08	Persian	Yazd	SSO	13	Number of active beds, nurses, physicians, and others	Number of inpatients, bed occupancies, surgeries	VRS, input-oriented
15	Ilbeigi et al. (34)	2009	Persian	Mashhad	MUMS	17	Number of beds, physicians, nurses, paraclinical staff, and support forces	Inpatient-bed day, outpatients, and surgeries	VRS_CRS (VRS), input-oriented
16	Rahimi et al. (35)	2009	Persian	W. Azarbaijan	UMSU	23	Number of beds, physicians, and others	Occupied bed-day, outpatient admission	VRS, input-oriented
17	Najarzadeh et al. (36)	2006–10	Persian	Ahvaz	AJUMS	13	Number of physicians, nurses, and beds	Occupied bed day, number of surgeries, outpatients, inpatients, mean patient stay	VRS, input-oriented

(Continued)

TABLE 1 | Continued

Row	Authors	Years of data collection	Language	Location	Affiliation of hospitals	Number of hospital	Inputs	Outputs	Model of DEA
18	Akbari et al. (37)	2005–08	Persian	Tabriz	TBZMED	20	Number of physicians and others, beds, and hospital costs	Number of patient admissions and surgeries, bed occupancy rate	VRS, input-oriented
19	Azar et al. (38)	2009–11	Persian	Tehran	TUMS	22	Number of beds, physicians, paramedics, and others	Number of outpatients, emergencies, inpatients, and surgeries, bed occupancy rate	VRS, input-oriented
20	Safi Aryan et al. (39)	2009	Persian	Hamadan	UMSHA	16	Number of beds, physicians, nurses, and others	Number of surgeries and outpatients, bed occupancy rate, mean patient stay, inpatient bed stay	VRS, input-oriented
21	Kazemi et al. (40)	2006–08	Persian	East of Iran	Medical Universities, SSO	11	Number of beds and all employees	Occupied bed day, outpatient admission	VRS, input-oriented
22	Raeisian et al. (41)	2007–11	Persian	Ahvaz	AJUMS & SSO, Private & Charity	8	Number of beds, physicians, nurses, and others	Number of patients and surgeries, bed occupancy rate	VRS, input-oriented
23	Raeisian et al. (41)	2007–11	Persian	Ahvaz	AJUMS & SSO, Private & Charity	8	Number of beds, physicians, nurses, and others	Number of patients and surgeries, bed occupancy rate	VRS, input-oriented
24	Mohebifar et al. (42)	2006–10	Persian	Guilan	GUMS	19	Number of beds, physicians, nurses, and others	Number of outpatients, inpatients, surgeries, and inpatient days	VRS, input-oriented
25	Fazeli et al. (43)	2009–11	Persian	Ilam	MEDILAM	9	Number of beds, physicians, and others	Number of clinical, paraclinical, and outpatient services	Input-oriented
26	Fazeli et al. (43)	2009–13	Persian	Ilam	MEDILAM	9	Number of beds, physicians, and others	Number of clinical, paraclinical, and outpatient services	Input-oriented
27	Mahfoozpor et al. (44)	2013–14	Persian	Tehran	SBMU	10	Number of physicians and nurses	Number of discharged patients	VRS, input-oriented
28	Mahfoozpor et al. (44)	2013–14	Persian	Tehran	SBMU	10	Number of physicians and nurses	Surgery room function	VRS, input-oriented
29	Mahfoozpor et al. (44)	2013–14	Persian	Tehran	SBMU	10	Number of physicians and nurses	Number of discharged patients	VRS, input-oriented
30	Mahfoozpor et al. (44)	2013–14	Persian	Tehran	SBMU	10	Number of physicians and nurses	Surgery room function	VRS, input-oriented
31	Ghasemi et al. (45)	2005–11	Persian	Kermanshah	KUMS	7	Number of beds, physicians, nurses, and others	Number of outpatients, inpatients, occupied bed days, and surgeries	VRS, input-oriented
32	Firouzi et al. (46)	NA	Persian	Tehran	TUMS	40	Number of beds, physicians, paramedics, some costs	Number of contracted insurances, access to emergency, confront with hospital infections, anesthesia problems, employee consent, bed occupancy rate, employee to bed ratio	VRS, input-oriented
33	Amozadeh et al. (47)	2012, 13, 15	Persian	Mazandaran and Babul	Mazandaran & Babol UMS	21	Number of beds, physicians, nurses, and others	Number of emergencies, outpatients, and surgeries	VRS, input-oriented
34	Youzi et al. (48)	2016	Persian	Tehran	TUMS	21	Number of beds, physicians, and nurses	Percentage of active beds, bed occupancy rate, mean stay, and bed turnover	VRS, input-oriented
35	Lotfi et al. (49)	2007–2011	English	Ahvaz	Affiliated and non-affiliated with AJUMS	16	Number of beds, physicians, nurses, and others	Bed occupancy rate, number of patients and operations	Input-oriented

(Continued)

TABLE 1 | Continued

Row	Authors	Years of data collection	Language	Location	Affiliation of hospitals	Number of hospital	Inputs	Outputs	Model of DEA
36	Nabilou et al. (9)	2009–2014	English	Tehran	TUMS	17	Number of beds, nurses, physicians, and others	Number of outpatient admission, occupied bed days, surgical operations	Input-oriented, variable return to scale
37	Rezapour et al. (50)	2009–2012	English	Tehran	IUMS & TUMS	19	Human resources, capital resources	Number of inpatients and admissions and inpatient bed occupancy rate	VRS, input-oriented
38	Torabipour et al. (51)	2007–2010	English	Ahvaz	University, Charity, Private	12	Number of nurses, beds, and physicians	Number of outpatients and inpatients, mean hospital stay, number of major operations	Input-oriented
39	Kiadaliri et al. (19)	2006	English	Ahvaz	AJUMS	19	Human resources, number of beds	Number of outpatient and inpatient visits, number of surgeries and percentage of occupied beds	VRS, input-oriented
40	Nabilou et al. (52)	2013–2014	Persian	Urmia	UMSU	23	Number of nurses, physicians, beds, and others	Number of discharges, surgeries, and bed occupancy percentage	VRS, input-oriented
41	Rezaei et al. (53)	2007–2011	English	Kurdistan	MUK	12	Number of beds, nurses, physicians, and others	Number of inpatient admissions and occupied bed days	VRS, input-oriented
42	Goudarzi et al. (54)	2001–07	Persian	Lorestan	LUMS	13	Number of beds, nurses, physicians, and others	Number of outpatients, inpatients, surgeries, bed days, and bed occupancy rate	VRS, input-oriented
43	Askari et al. (55)	2001–11	English	Yazd	SSU	13	Number of beds, nurses, physicians, and non-clinical staff	Number of admissions and surgeries, bed occupancy percentage	VRS, input-oriented
44	Sabermahani et al. (56)	2011	English	Kerman	KMU	13	Full-time physicians and nurses, administrative personnel	Number of outpatient clients, surgeries, and beds per day	VRS, input-oriented
45	Jahangiri et al. (11)	2011–13	Persian	Arak	IAU-ARAK	31	Number of day-beds, working days, physicians, and other staff	Number of admissions, surgeries, child birth, and inpatient days	CRS, input-oriented
46	Najafi et al. (57)	2001–06	Persian	Ardabil	TUMS	10	Number of beds and physicians	Number of admissions and inpatient beds	VRS, input-oriented
47	Hatam et al. (58)	2006–2008	Persian	Iran	SUMS	18	Number of beds and all full-time staff, hospital budget	Bed-day, active to fixed bed, patient mean stay, bed turnover, death, and costs	CRS, input-oriented
48	Rezapour et al. (50)	1998–07	Persian	Qazvin	QUMS	4	Number of beds, physicians, nurses, and others	Number of discharges, surgeries, admissions, emergencies, bed turnover, patient days	VRS, input-oriented
49	Hadian et al. (59)	2006–11	Persian	Tehran	IUMS & TUMS	19	Number of beds, nurses, physicians, and others	Number of outpatient admissions, inpatient days, occupied bed days, surgeries	VRS, input-oriented
50	Mehrtak et al. (60)	NA	English	E. Azarbaijan	IUMS	18	Number of beds, physicians, and nurses	Number of discharges, surgeries, bed occupancy rate	VRS, input-oriented

TUMS, Tehran University of Medical Sciences; IUMS, Iran University of Medical Sciences; SBMU, Shaheed Beheshti University of Medical Sciences; SSO, Social Security Organization; MUI, Isfahan University of Medical Sciences; SSU, Yazd University of Medical Sciences; TUMS, Tehran University of Medical Sciences; MHH, Ministry of Health' hospitals; MUQ, Qom university of Medical Sciences; MUMS, Mashhad University of Medical Sciences; UMSU, Urmia University of Medical Sciences; AJUMS, Ahvaz Jundishapour University of Medical Sciences; TBZMED, Tabriz University of Medical Sciences; UMSHA, Hamedan University of Medical Sciences; GUMS, Guilan University of Medial Sciences; MEDILAM, Ilam University of Medical Sciences; KUMS, Kermanshah University of Medical Sciences; MUK, Kurdistan University of Medical Sciences; LUMS, Lorestan University of Medical Sciences; KMU, Kerman University of Medical Sciences; IAU-ARAK, Islamic Azad University Branch of Arak; SUMS, Shiraz University of medical sciences; QUMS, Qazvin University of Medical Sciences.

relative efficiency of a field, section, unit, or an organization, as a DMU, using multiple input and output indices (7).

Numerous studies have assessed the efficiency of hospital efficiency using the DEA method. These studies can be divided into four categories. In the first category, the efficiency of university, teaching, and public hospitals, as the main providers of healthcare and therapeutic services, has been assessed in studies by Kalhor et al. (8) and Nabi lou et al. (9). In the second category, the efficiency of private hospitals has been studied and their efficiency has been compared with the first-category hospitals (10, 11). The third category includes studies on hospitals affiliated with special entities such as Social Security Organization (12, 13) and Armed Forces (14). The last category measures the efficiency of hospital wards such as radiology (15), dentistry (4), intensive care unit (16), and emergency (17) departments. Because the latter category studies wards of hospitals rather than the hospitals in their entirety and also have not assessed the technical, managerial, and scale efficiency of hospitals wholly, this category was excluded from the current study.

Although many studies have assessed the efficiency of hospitals using the DEA method in Iran, there has been no systematic review and meta-analysis study in this regard to present the final situation of hospital efficiency in Iran. By determining technical, managerial, and scale efficiency of Iranian hospitals, policymakers and planners can improve hospital efficiency through improving distribution and consumption of resources.

The extensive review of the literature by the authors of the current study has resulted in four systematic review and meta-analysis studies on Iranian hospital efficiency using the DEA method. The first study assessed studies in terms of the provinces where they were performed, whether they were input- or output-oriented, and whether they were fixed or variable return to scale models (18). The researchers in another two systematic and meta-analysis studies discussed the methods used to assess hospital efficiency (19, 20). The last study only included a small number of studies on hospital efficiency and did not mention the efficiency subdimensions namely scale, managerial, and technical efficiency (21). As previous systematic review and meta-analysis studies have not assessed hospital efficiency using its subcategories, the current study assessed technical, managerial, and scale efficiency of hospitals through systematic review and meta-analysis.

Regarding PICOS framework or questions, the study included hospitals in Iran which had previously had their efficiency assessed and were entered into the study depending on the inclusion and exclusion criteria. The intervention framework was the assessment of the effect of HSEP on hospital efficiency, comparisons included comparing hospital efficiency before and after HSEP, outcomes included the amount of hospital efficiency, and finally the study design included assessment of hospital efficiency through systematic review and meta-analysis.

MATERIALS AND METHODS

Search Strategy

The international databases of the Institute for Scientific Information (ISI), PubMed, Scopus, Google Scholar, and Persian databases of Scientific Information Database (SID), Magiran, and Barakat were searched using the combination of "efficiency," "hospital," "data envelopment analysis," "DEA," and "Iran" keywords in 2018. The references of the retrieved articles were searched to increase the study credibility and precision.

Inclusion and Exclusion Criteria

All published Persian and English language articles about hospital efficiency with a score between 8 and 12 were entered into the study without considering a time limit. If several formats of a research were published (such as a book, article, report, and so on), only one of them was entered into the study. Input-oriented studies were entered into the study. Short reports, letters to editors or editorial comments, one study that was available in two languages, studies on health care facilities other than hospitals, and studies on internal parts of hospitals were removed from the study. Two researchers assessed and extracted data from the studies independently and a third researcher resolved disagreements if they appeared.

This systematic review and meta-analysis utilized Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to minimize potential sources of bias (22).

Data Collection

A researcher-made checklist in an Excel spreadsheet was created to extract the studies' data including the first author name, year of data collection, place of study, language, sample size, and the score of technical, managerial, and scale efficiency. Another checklist designed previously, whose credibility has been proved by numerous studies, was used to assesses the studies' quality (19, 21). This checklist includes 12 questions regarding the study aim, method, data collection, sample size, and study

TABLE 2 | Egger's test for small-study effects to examine the publication bias.

		Coefficient	S.E.	*P*-value	95% confidence interval	Test of H0: no small-study effects
Technical efficiency	Slope	0.56	0.00	0.000	(0.40, 0.73)	$P = 0.005$
	Bias	−1.01	0.34	0.005	(0.32, 1.70)	
Managerial efficiency	Slope	1.00	0.00008	0.000	(0.99, 1.00)	$p < 0.001$
	Bias	−1.22	0.26	0.000	(−1.76, −0.68)	
Economics of scale efficiency	Slope	1.00	0.00009	<0.001	(0.99, 1.00)	$p < 0.001$
	Bias	−1.28	0.19	<0.001	(−1.67, −0.0.90)	

TABLE 3 | The results of pool estimation for technical efficiency among Iranian hospitals.

Study	Authors	Estimation	95% confidence intervals	Weight
1	Joshan et al. (24)	0.93	(0.66, 1)	1.99
2	Joshan et al. (24)	0.88	(0.47, 1)	1.58
3	Joshan et al. (24)	0.9	(0.55, 1)	1.75
4	Sepehrdost et al. (25)	0.86	(0.67, 0.96)	2.42
5	Sepehrdost et al. (25)	0.89	(0.75, 0.97)	2.55
6	Ghaderi et al. (26)	0.88	(0.7, 0.98)	2.38
7	Karimi et al. (27)	0.91	(0.72, 0.99)	2.31
8	Alimohammadi Ardakani et al. (28)	0.75	(0.43, 0.95)	1.88
9	Pourreza et al. (29)	0.92	(0.62, 1)	1.88
10	Aboulhalaj et al. (30)	0.43	(0.34, 0.53)	2.92
11	Salehzade et al. (31)	0.75	(0.35, 0.97)	1.58
12	Salehzade et al. (31)	0.88	(0.47, 1)	1.58
13	Asadi et al. (32)	0.92	(0.64, 1)	1.94
14	Askari et al. (33)	0.92	(0.64, 1)	1.94
15	Ilbeigi et al. (34)	0.76	(0.5, 0.93)	2.12
16	Rahimi et al. (35)	0.57	(0.34, 0.77)	2.31
17	Najarzadeh et al. (36)	0.69	(0.39, 0.91)	1.94
18	Akbari et al. (37)	0.95	(0.75, 1)	2.22
19	Azar et al. (38)	0.86	(0.65, 0.97)	2.28
20	Safi Aryan et al. (39)	0.88	(0.62, 0.98)	2.08
21	Kazemi et al. (40)	0.82	(0.48, 0.98)	1.82
22	Raeisian et al. (41)	0.88	(0.47, 1)	1.58
23	Raeisian et al. (41)	1	(0.63, 1)	1.58
24	Mohebifar et al. (42)	0.89	(0.67, 0.99)	2.19
25	Fazeli et al. (43)	0.78	(0.4, 0.97)	1.67
26	Fazeli et al. (43)	0.78	(0.4, 0.97)	1.67
27	Mahfoozpor et al. (44)	0.4	(0.12, 0.74)	1.75
28	Mahfoozpor et al. (44)	0.4	(0.12, 0.74)	1.75
29	Mahfoozpor et al. (44)	0.3	(0.07, 0.65)	1.75
30	Mahfoozpor et al. (44)	0.5	(0.19, 0.81)	1.75
31	Ghasemi et al. (45)	0.86	(0.42, 1)	1.49
32	Firouzi Jahantigh et al. (46)	0.93	(0.8, 0.98)	2.59
33	Amozadeh et al. (47)	0.9	(0.7, 0.99)	2.25
34	Youzi et al. (48)	0.86	(0.64, 0.97)	2.25
35	Lotfi et al. (49)	0.88	(0.62, 0.98)	2.08
36	Nabilou et al. (9)	0.94	(0.71, 1)	2.12
37	Rezapour et al. (50)	0.84	(0.6, 0.97)	2.19
38	Torabipour et al. (51)	0.92	(0.62, 1)	1.88
39	Ahmad Kiadaliri et al. (19)	0.89	(0.67, 0.99)	2.19
40	Nabilou et al. (52)	0.87	(0.66, 0.97)	2.31
41	Rezaei et al. (53)	0.83	(0.52, 0.98)	1.88
42	Goudarzi et al. (54)	0.92	(0.64, 1)	1.94
43	Askari et al. (55)	0.92	(0.64, 1)	1.94
44	Sabermahani et al. (56)	0.85	(0.55, 0.98)	1.94
45	Jahangiri et al. (11)	0.97	(0.83, 1)	2.47
46	Najafi et al. (57)	1	(0.69, 1)	1.75
47	Hatam et al. (58)	0.89	(0.65, 0.99)	2.16
48	Rezapour et al. (50)	0.75	(0.19, 0.99)	1.1
49	Hadian et al. (59)	0.95	(0.74, 1)	2.19
50	Mehrtak et al. (60)	0.78	(0.52, 0.94)	2.16
Random pooled estimation		0.84	(0.78, 0.52)	100

Heterogeneity chi^2 = 156.97 (d.f. = 49) p = 0.000.
I^2 (variation in ES attributable to heterogeneity) = 68.78%.
Estimate of between-study variance Tau^2 = 0.12.

population. Each question has the score between 0 and 1 and the score for each study is calculated by summing the scores of questions. So that the studies with scores between 8 and 12 were entered into the final analysis. The study protocol was approved by the Ethical Committee of Kerman University of Medical Sciences.

Data Analysis

Efficiency types were considered as a proportion in this study. Therefore, the numerator was the sum of technical, managerial, and efficiency scores and the denominator was the number of study hospitals. Heterogeneity between the studies was assessed using Q and I^2 tests. A *P*-value lower than 0.05 for the Q-test and an I^2 higher than 50% were considered as the measures of studies' heterogeneity. Because the studies were heterogeneous, the random effect model was used to estimate hospital efficiency. A forest plot with 95% confidence intervals (CI) was used to calculate different types of efficiency. Egger's and Begg's tests were used to assess publication bias. In order to assess the effect of the 2011 Iran Health Sector Evolution Plan (HSEP) (23) on hospital efficiency, the studies before and after it were compared. The data were entered into Excel 2016 to be edited and then transmitted and analyzed using STATA v.14.2.

RESULTS

Each one of the scientific databases were searched on the basis of a recommended search strategy by the databases themselves using defined keywords. For example, in the PubMed database, 23 articles were retrieved after placing the search query. Search query used for PubMed was: (((data envelopment analysis) OR DEA) AND hospital) AND Iran))). Among retrieved articles, nine articles had assessed efficiency in other areas such as radiology units, intensive care units, and health centers which were excluded from the study. So, finally 14 articles from the PubMed database were entered into the EndNote software. In other databases, after adjusting the search query on the basis of the database guide and then removing unrelated retrieved articles through reading titles, abstracts, and texts, 25 articles from Scopus, 41 from Google Scholar, 8 from Web of Science, 16 from Barakat, 14 from Magiran, and 7 from SID were entered into the EndNote software. After combination of these articles in the EndNote software and removing duplicate articles, 47 final

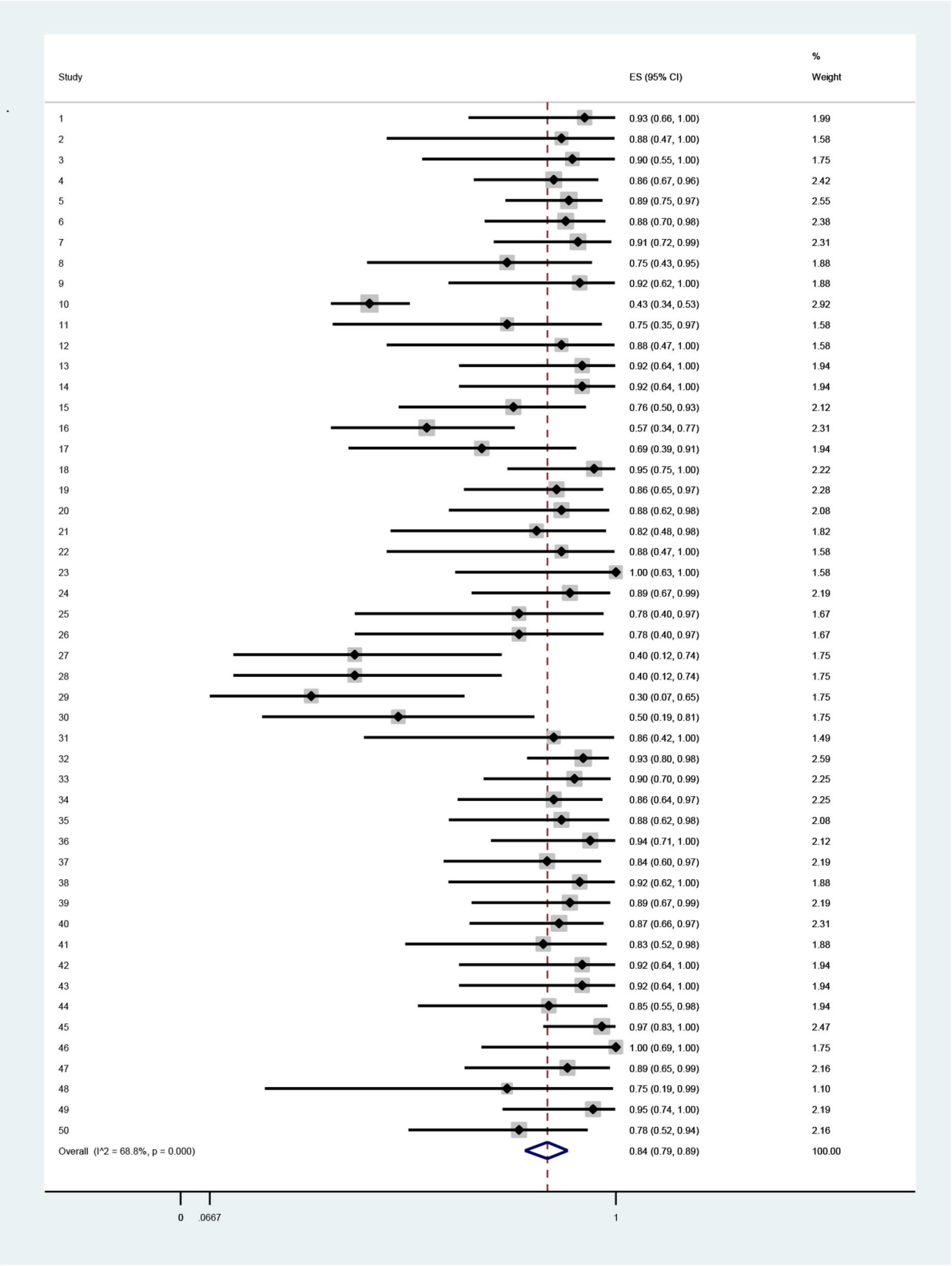

FIGURE 2 | Forest plot of estimates and 95% confidence intervals of the technical efficiency among Iranian hospitals.

articles remained. Also, the assessment of references of these articles resulted in two new articles. In this way, 49 articles were entered into the final step of the systematic review and meta-analysis. Twelve articles (24.48%) of these were in the Persian language and the remaining were in the English language. A PRISMA flow chart of the study retrieval and selection process with reasons for exclusion at each stage is provided in **Figure 1**.

By attention that some studies have reported efficiency in several forms or in different scenarios and different inputs were used in them, we considered them as separated studies. In this regard, studies of Hatam et al., Karimi et al., Salehzade et al., Raeisian et al., Firouzi Jahantigh et al., and Sheikhzadeh et al. were each considered as two separated studies. Studies of Joshan et al. and Asadi et al. were each considered as three studies and lastly studies of Fazeli et al. and Mahfoozpor et al. were each considered as four studies. The average number of hospitals entered into the studies was 17.59 hospitals. The lowest and the highest number of hospitals belonged to Rezapour et al. with 4 hospitals and Aboulhalaj et al. with 122 hospitals, respectively.

As mentioned before, each type of efficiency was entered into the meta-analysis separately, so that 50 studies for technical efficiency, 36 studies for managerial efficiency, and 41 studies for scale efficiency had entry requirements to the analysis.

The studies were performed from 1996 to 2016. After performing all the steps of study selection, 49 articles were entered into the final phase of the study. The number of hospitals assessed in these articles ranged from 4 to 122. The inputs considered in the studies included number of beds, number of operation rooms, physicians, nurses, support forces and other human resources, costs, education hours, and working days. The outputs were number of surgeries, outpatients, occupancy rate, bed days, admission, inpatients, surgeries, emergencies, bed turnover, mean patient stay, hospital income, bed occupancy rate, SERVQUAL score, number of clinical, paraclinical, and outpatient services, number of discharged patients, number of contracted insurance companies, access to emergency room, confronted with hospital infections, anesthesia problems, employee consent, active to fixed bed ratio, number of deaths, and patient-day. Two studies assessed charity hospitals, four studies assessed private hospitals, and five studies assessed Social Security Organization (SSO) hospitals. The remaining studies assessed hospitals affiliated with universities of medical sciences belonging to the Iran Ministry of Health (**Table 1**).

The results indicated that there was heterogeneity in studies related to technical efficiency (heterogeneity $chi^2 = 156$, $p < 0.001$), managerial efficiency (heterogeneity $chi^2 = 79.58$, $p < 0.001$), and scale efficiency (heterogeneity $chi^2 = 67.22$, $p < 0.001$). I^2 index in technical and managerial efficiency was higher than 50%, which indicates high heterogeneity between the studies. This index was lower than 50% for scale efficiency.

Study publication error using Egger's test indicated that there was publication bias in technical and managerial efficiencies ($P < 0.001$), but there was no publication bias in scale efficiency ($p = 0.19$). **Table 2** displays the results of Egger's testing for the three types of efficiencies. Begg's test indicated that there was no publication bias in the three types of efficiencies ($P < 0.001$).

The results indicated that technical efficiency of Iranian hospitals had high variation, so that it ranged from 0.34 in the Mahfoozpor et al. study to 1 in Raeisian et al. and Najafi et al. On the basis of random effects modeling, random pooled estimation of hospital technical efficiency was 0.84 (95% CI = 0.52, 0.78) (**Table 3**, **Figure 2**). The managerial efficiency of Iranian hospitals was between 0.59 in the Aboulhalaj et al. study and 1 in studies

TABLE 4 | The results of pool estimation for managerial efficiency among Iranian hospitals.

Study	Authors	Estimation	95% confidence intervals	Weight
1	Joshan et al. (24)	0.93	(0.66, 1)	2.71
2	Joshan et al. (24)	1	(0.63, 1)	2.01
3	Joshan et al. (24)	0.9	(0.55, 1)	2.28
4	Sepehrdost et al. (25)	0.93	(0.76, 0.99)	3.59
5	Sepehrdost et al. (25)	0.95	(0.82, 0.99)	3.9
6	Ghaderi et al. (26)	0.88	(0.7, 0.98)	3.5
7	Karimi et al. (28)	0.96	(0.78, 1)	3.35
8	Pourreza et al. (29)	0.92	(0.62, 1)	2.51
9	Aboulhalaj et al. (30)	0.59	(0.5, 0.68)	4.83
10	Askari et al. (33)	0.92	(0.64, 1)	2.61
11	Ilbeigi et al. (34)	0.88	(0.64, 0.99)	2.96
12	Rahimi et al. (35)	0.74	(0.52, 0.9)	3.35
13	Najarzadeh et al. (36)	0.85	(0.55, 0.98)	2.61
14	Akbari et al. (37)	0.95	(0.75, 1)	3.17
15	Safi Aryan et al. (39)	0.94	(0.7, 1)	2.88
16	Kazemi et al. (40)	0.91	(0.59, 1)	2.4
17	Raeisian et al. (41)	1	(0.63, 1)	2.01
18	Raeisian et al. (41)	1	(0.63, 1)	2.01
19	Mohebifar et al. (42)	0.95	(0.74, 1)	3.11
20	Mahfoozpor et al. (44)	0.6	(0.26, 0.88)	2.28
21	Mahfoozpor et al. (44)	0.6	(0.26, 0.88)	2.28
22	Mahfoozpor et al. (44)	0.7	(0.35, 0.93)	2.28
23	Mahfoozpor et al. (44)	0.8	(0.44, 0.97)	2.28
24	Nabilou et al. (9)	0.94	(0.71, 1)	2.96
25	Rezapour et al. (50)	0.95	(0.74, 1)	3.11
26	Torabipour et al. (51)	0.92	(0.62, 1)	2.51
27	Ahmad Kiadaliri et al. (19)	0.89	(0.67, 0.99)	3.11
28	Nabilou et al. (52)	0.91	(0.72, 0.99)	3.35
29	Rezaei et al. (53)	0.83	(0.52, 0.98)	2.51
30	Goudarzi et al. (54)	0.92	(0.64, 1)	2.61
31	Askari et al. (55)	0.92	(0.64, 1)	2.61
32	Sabermahani et al. (56)	0.92	(0.64, 1)	2.61
33	Najafi et al. (57)	1	(0.69, 1)	2.28
34	Rezapour et al. (50)	0.75	(0.19, 0.99)	1.29
35	Hadian et al. (59)	0.95	(0.74, 1)	3.11
36	Mehrtak et al. (60)	0.94	(0.73, 1)	3.04
Random pooled estimation		0.9	(0.85, 0.94)	100

Heterogeneity $chi^2 = 79.58$ (d.f. = 35), $p = 0.000$.
I^2 (variation in ES attributable to heterogeneity) = 56.02%.
Estimate of between-study variance $Tau^2 = 0.07$.
Test of ES = 0: $z = 34.96$, $p = 0.00$.

of Joshan et al., Raeisian et al., and Najafi et al. Random pooled estimation of managerial efficiency of Iranian hospitals was 0.90 (95% CI = 0.85,0.94) (**Table 4**, **Figure 3**). The lowest amount of scale efficiency (0.52) was in the Mahfoozpor et al. study and the highest (1) was in the Raeisian et al. and Torabipour et al. studies. Random pool estimation of scale efficiency for Iranian hospitals

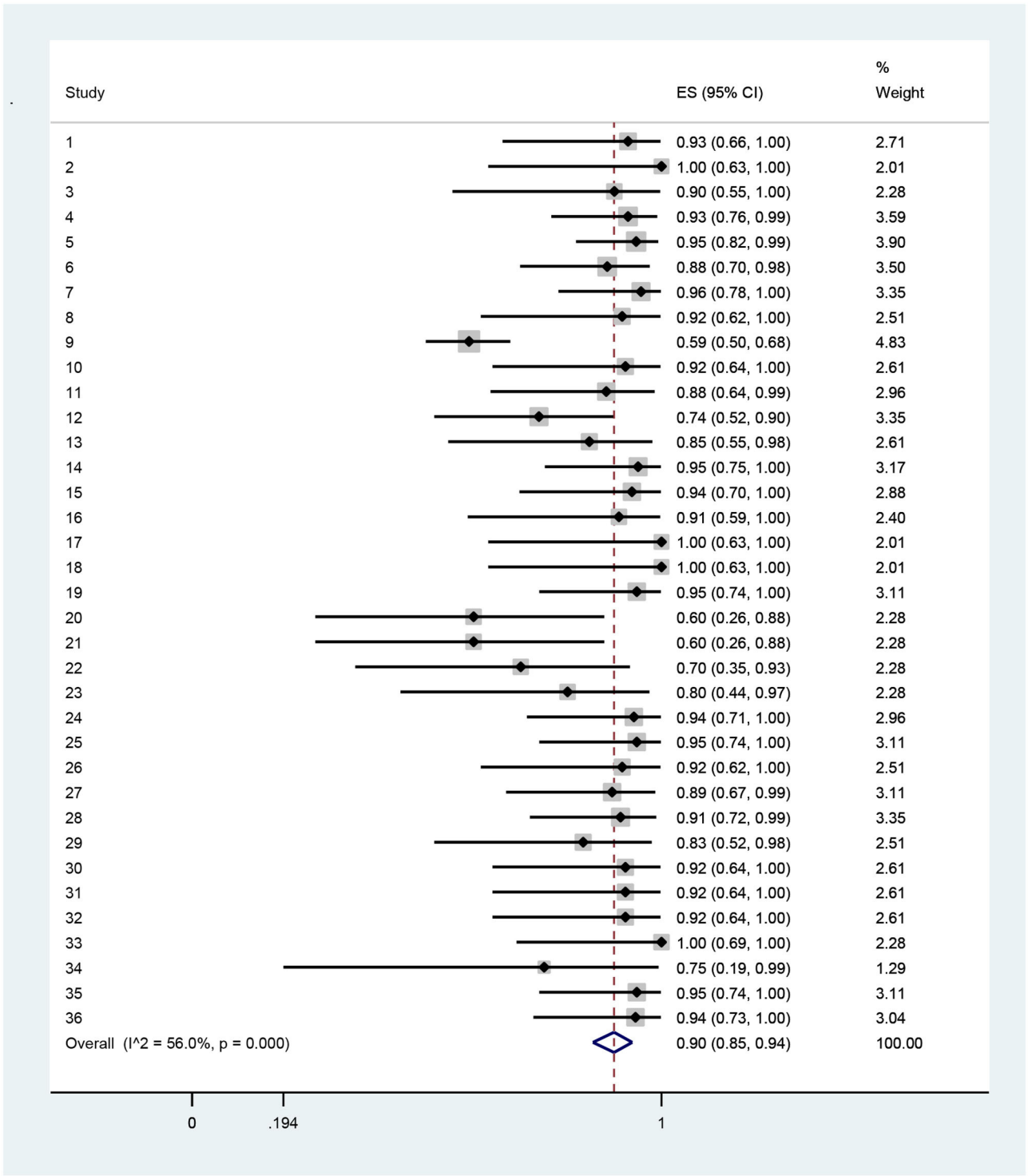

FIGURE 3 | Forest plot of estimates and 95% confidence intervals of the managerial efficiency among Iranian hospitals.

was 0.88 (95%CI = 0.84, 0.91) (**Table 5**, **Figure 4**). The results of technical, managerial, and scale efficiencies are presented in **Tables 2**, **4**, **5**, respectively. In addition, the forest plots for technical, managerial, and scale efficiencies are presented in **Figures 1**–**3**, respectively.

Sub-group analysis based on study year indicated that random pool estimation of technical efficiency of Iranian hospitals for 2011 and before and after 2011 was 0.86 (95% CI = 0.80, 0.91) and 0.78 (95%CI = 0.64, 0.89), respectively. The status of managerial efficiency for 2011 and before was better than after 2011 (random pool estimation equal to 0.91, compared to 0.86). Random pool estimation of scale efficiency for 2011 and before was 0.90 (95%CI = 0.86, 0.93), while random pool estimation of scale efficiency for after 2011 was 0.74 which is lower (95%CI = 0.61, 0.86) (**Table 6**).

TABLE 5 | The results of pool estimation for economies of scale efficiency among Iranian hospitals.

Study	Authors	Estimation	95%confidence Intervals	Weight
1	Joshan et al. (24)	0.93	(0.66, 1)	2.38
2	Joshan et al. (24)	0.88	(0.47, 1)	1.64
3	Joshan et al. (24)	0.9	(0.55, 1)	1.91
4	Sepehrdost et al. (25)	0.93	(0.76, 0.99)	3.46
5	Sepehrdost et al. (25)	0.95	(0.82, 0.99)	3.91
6	Ghaderi et al. (26)	0.96	(0.8, 1)	3.34
7	Karimi et al. (27)	0.96	(0.78, 1)	3.15
8	Poureza et al. (29)	0.92	(0.62, 1)	2.16
9	Aboulhalaj et al. (30)	0.75	(0.66, 0.82)	5.44
10	Salehzade et al. (31)	0.88	(0.47, 1)	1.64
11	Salehzade et al. (31)	0.88	(0.47, 1)	1.64
12	Askari et al. (33)	0.92	(0.64, 1)	2.27
13	Ilbeigi et al. (34)	0.82	(0.57, 0.96)	2.67
14	Rahimi et al. (35)	0.74	(0.52, 0.9)	3.15
15	Najarzadeh et al. (36)	0.77	(0.46, 0.95)	2.27
16	Akbari et al. (37)	0.95	(0.75, 1)	2.92
17	Safi Aryan et al. (39)	0.94	(0.7, 1)	2.58
18	Kazemi et al. (40)	0.91	(0.59, 1)	2.04
19	Raeisian et al. (41)	0.88	(0.47, 1)	1.64
20	Raeisian et al. (41)	1	(0.63, 1)	1.64
21	Mohebifar et al. (42)	0.95	(0.74, 1)	2.84
22	Fazeli et al. (43)	0.67	(0.3, 0.93)	1.78
23	Fazeli et al. (43)	0.91	(0.59, 1)	2.04
24	Mahfoozpor et al. (44)	0.5	(0.19, 0.81)	1.91
25	Mahfoozpor et al. (44)	0.7	(0.35, 0.93)	1.91
26	Mahfoozpor et al. (44)	0.5	(0.19, 0.81)	1.91
27	Mahfoozpor et al. (44)	0.6	(0.26, 0.88)	1.91
28	Nabilou et al. (9)	0.94	(0.71, 1)	2.67
29	Rezapour et al. (50)	0.89	(0.67, 0.99)	2.84
30	Torabipour et al. (51)	1	(0.74, 1)	2.16
31	Ahmad Kiadaliri et al. (19)	0.95	(0.74, 1)	2.84
32	Nabilou et al. (9)	0.91	(0.72, 0.99)	3.15
33	Rezaei et al. (53)	0.92	(0.62, 1)	2.16
34	Goudarzi et al. (54)	0.92	(0.64, 1)	2.27
35	Askari et al. (55)	0.92	(0.64, 1)	2.27
36	Sabermahani et al. (56)	0.85	(0.55, 0.98)	2.27
37	Najafi et al. (57)	0.9	(0.55, 1)	1.91
38	Hatam et al. (58)	0.5	(0.26, 0.74)	2.76
39	Rezapour et al. (50)	0.75	(0.19, 0.99)	0.98
40	Hadian et al. (59)	0.95	(0.74, 1)	2.84
41	Mehrtak et al. (60)	0.78	(0.52, 0.94)	2.76
Random pool estimation		0.88	(0.84, 0.91)	100

Heterogeneity $chi^2 = 67.22$ *(d.f. = 40)* $p = 0.00$.
I^2 *(variation in ES attributable to heterogeneity) = 40.5%.*
Test of ES = 0: $z = 40.93$ $p = 0.00$.

DISCUSSION

The assessment of hospital efficiency provides the groundwork to assess their performance and increase productivity when using limited resources. One of the ways of assessing allocated resources to obtain specified goals is efficiency studies. In summary, efficiency means using resources to their maximum to produce goods and services (61).

This is the first systematic review and meta-analysis study regarding assessment of the efficiency of Iranian hospitals in terms of its subcategories namely technical, managerial, and scale efficiencies. Different methods have been used to assess Iranian hospital efficiency such as DEA, Pabon-Lasso, and Stochastic Frontier Analysis (SFA) in past decades (21). In this regard, as this study indicates, the DEA method is the most widely applied method to assess hospital efficiency (19).

Our findings showed that the random pool estimations of technical, managerial, and economics of scale efficiency were 0.87, 0.9, and 0.88, respectively. This finding indicates that the resources of the studied hospitals in Iran have been used in an inefficient way. One idea about hospital efficiency is that the expectation from hospitals to work efficiently is far from reality. The reasoning for this claim is the economic theory of firms that declare the hospitals cannot work at full efficiency because of uncertainty in costs and prices of services that they provide. In summary, lack of information on costs and prices is one of the main factors that has a negative effect on hospital efficiency (6, 62).

Most of the studies were implemented in Tehran province (13 studies). Four studies investigated the efficiency of hospitals across all provinces of Iran. However, some provinces such as Sistan and Baluchistan had no individual reports about the efficiency of their hospitals. Therefore, there is an information gap for health policymakers and hospital managers in this field.

As the results indicated, most of the researchers tended to perform analyzes through the input-oriented method, because inputs are in the control of hospital managers, so that by creating changes in the inputs can change the rate of outputs to the desired extent. However, it is suggested that private and for-profit hospitals are excluded from this rule, because the managers of these type of hospitals want to maximize outputs and, as a result, hospital profits (63).

Human and capital resources such as the number of nurses and physicians and the number of beds were the main inputs in

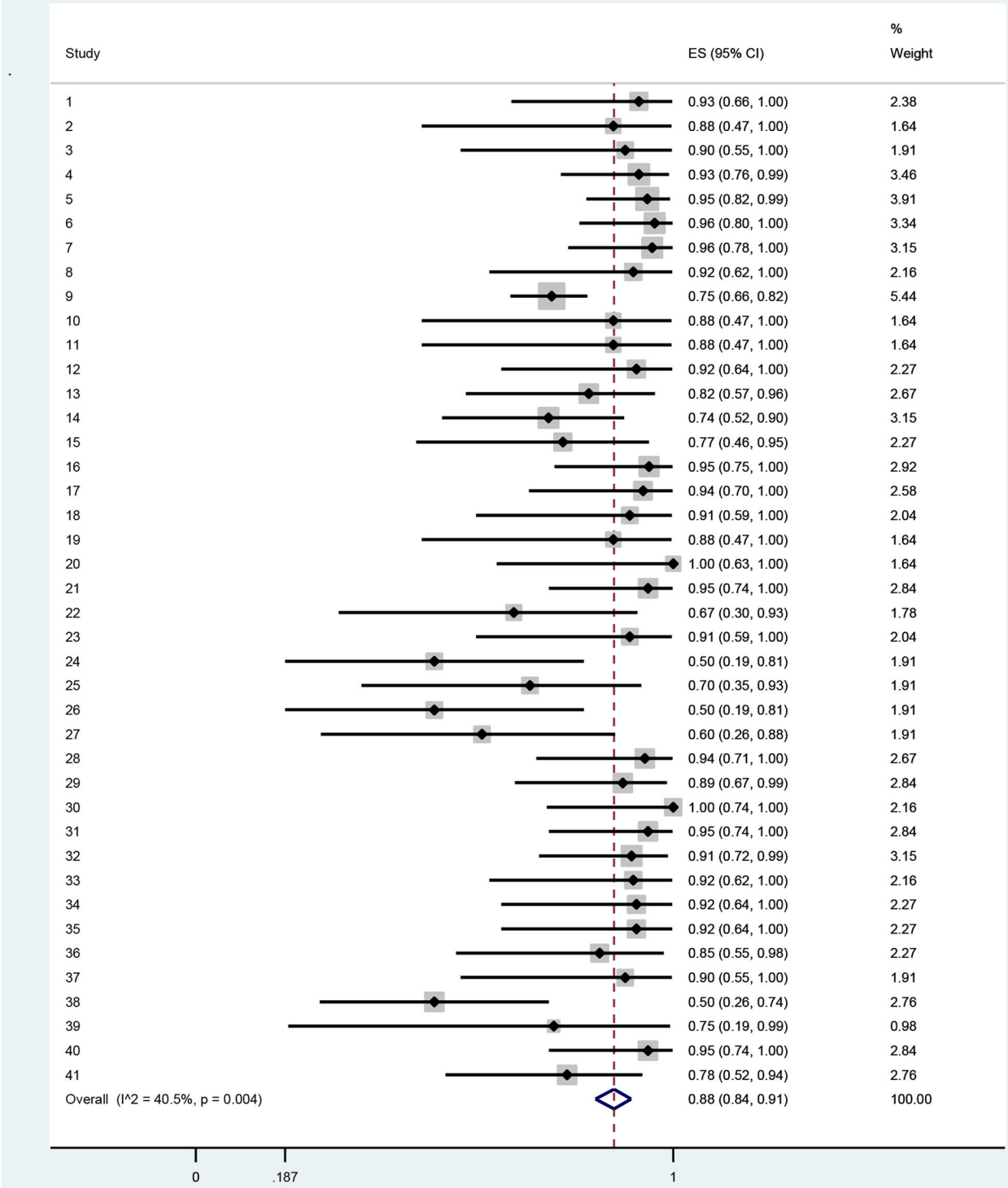

FIGURE 4 | Forest plot of estimates and 95% confidence intervals of the economics of scale efficiency among Iranian hospitals.

TABLE 6 | The random pool estimation of technical, managerial, and economics of scale efficiencies among Iranian hospitals by time of studies.

	Subgroup	Estimation	95% confidence intervals	Weight	Test(s) of heterogeneity		Random Test for heterogeneity between sub-groups:
					I^2**	*P*-value	*P*-value
Technical efficiency	2011 and before	0.86	(0.80, 0.91)	75.97	66.80%	0.000	0.23
	After 2011	0.78	(0.64, 0.89)	24.03	75.55%	0.000	
Managerial efficiency	2011 and before	0.91	(0.86, 0.95)	77.89	60.74%	0.000	0.27
	After 2011	0.86	(0.75, 0.94)	22.11	39.89%	0.100	
Economics of scale efficiency	2011 and before	0.90	(0.86, 0.93)	83.67	32.79%	0.040	0.01
	After 2011	0.74	(0.61, 0.86)	16.33	35.26%	0.150	

** *I^2, the variation in ES attributable to heterogeneity.*

all included studies. Number of surgeries, outpatient admissions, inpatient admissions, bed days, and bed occupancy rate were the most frequent outputs considered in the studies to estimate the efficiency of hospitals. Today, the management of all resources, especially human resources in the health care industry is recognized as a vital issue for all healthcare organizations (64). Furthermore, better management of human resources is associated with higher patient outcomes without significant increases in the cost of hospitals (65).

The results indicated that most hospital efficiency studies suffer from some weak points. Therefore, the selection of inputs has been performed on the basis of resource review (e.g., previous published articles) not consideration of each hospital situation. Also, the inputs were not weighted, so that the resources with high specialty and expenditure receive the same weight as others. Hospital case mix has not been considered in this hospital efficiency assessment. This leads to low efficiency assessment in hospitals which have the most complicated cases. Lastly, some studies have not considered the data validity and the appropriate ratio of inputs and outputs with the number of hospitals precisely.

The study of Contor VJM and Poh Kl also provides some theoretical and methodological limitations of the DEA method in capturing the full view of efficiency of healthcare centers (66). However, with a suitable study design, the DEA method is among the most important and most applicable methods in the assessment of health system efficiency, especially hospitals (67).

The results indicated that technical, managerial, and scale efficiency of Iranian hospitals after performing HSEP decreased in comparison with before it. A study on Turkey hospitals from 2001 to 2006, which measured the effect of Turkey health sector reform on hospital efficiency to provide policy implications for policymakers, indicated that this reform had increased the efficiency of public hospitals but the efficiency of private hospitals had decreased (68).

As there was no hospital with full efficiency in the study and the increasing trend of health system costs and scarce resources, it is proposed to design and implement a system to monitor efficiency and consumption of resources especially in hospitals. This can help to identify inefficient hospitals and the causes of it. Health policymakers through cost management planning and increasing outputs can pave the way in this regard.

Strengths and Limitations

This is the first comprehensive systematic review and meta-analysis evaluating efficiency of Iranian hospitals which is applicable for comparison of the efficiency of hospitals before and after HSEP. The methodology adhered to the PRISMA statement (22).

The strength of the study is in performing meta-analysis after the systematic review which has specified the exact amount of technical, managerial, and scale efficiencies of Iranian hospitals. The Cochrane Consumers and Communication Review Group's data extraction template (69) was used to obtain the needed information about the studies included. Nevertheless, the retrieved studies were mainly administered in some easily accessible areas rather than in a balanced distribution all over the country. This limits the generalizability of the results.

CONCLUSION

This study indicated that many hospitals are inefficient. This implies that there is considerable room for efficiency improvement in hospitals. Hospital management has a unique role in this regard. Health systems have reformed in spite of increasing access and utilization of patients to the services, but have not considered efficiency improvement in hospitals. So, health policymakers and hospital managers should design and implement some related programs in order to monitor and improve the efficiency of hospitals.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

This systematic review and meta-analysis study was approved by the Ethical Committee of Arak University of Medical Sciences (Ethical Code Number: IR.ARAKMU.REC.1398.044).

AUTHOR CONTRIBUTIONS

SA, BK, AK-K, and MD: conception and design of study/review/case series. SA, MA, YS, and AA: acquisition of data. YS, AA, MD, BK, and MA: analysis of collected data. SA, AK-K, YS, MD, AA, and YS: interpretation of data and drafting of paper and/or critical revision. All the authors have read and approved the manuscript to be submitted to BMC Health Services Research.

FUNDING

This paper is retrieved from an approved research project. The Deputy of Research of Arak University of Medical Sciences has financially supported this study in different parts of the study including design, data collection, analysis, interpretation, and writing the manuscript (Grant Number: 3382).

ACKNOWLEDGMENTS

The authors would like to acknowledge the Deputy of Research of Arak University of Medical Sciences for financial support of this study and providing the background to perform this study.

REFERENCES

1. Rezaei S, Fallah R, Kazemi Karyani A, Daroudi R, Zandiyan H, Hajizadeh M. Determinants of healthcare expenditures in Iran: evidence from a time series analysis. *Med J Islam Repub Iran*. (2016) 30:313.
2. Johnson G, Frakt A. Hospital markets in the United States, 2007–2017. *Healthcare*. (2020) 8:100445. doi: 10.1016/j.hjdsi.2020.100445
3. Li Y, Lei X, Morton A. Performance evaluation of nonhomogeneous hospitals: the case of Hong Kong Hospitals. *Health Care Manag Sci.* (2018) 18:1–14. doi: 10.1007/s10729-018-9433-y
4. Barouni M, Amiresmaieli MR, Shahravan A, Amini S. The efficiency assessment of dental units using data envelopment analysis approach: the case of Iran. *Iran J Public Health.* (2017) 46:552.
5. Kontodimopoulos N, Nanos P, Niakas D. Balancing efficiency of health services and equity of access in remote areas in Greece. *Health Policy.* (2006) 76:49–57. doi: 10.1016/j.healthpol.2005.04.006
6. Hollingsworth B, Dawson P, Maniadakis N. Efficiency measurement of health care: a review of non-parametric methods and applications. *Health Care Manag Sci.* (1999) 2:161–72. doi: 10.1023/A:1019087828488
7. Banker RD, Charnes A, Cooper WW. some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag Sci.* (1984) 30:1078–92. doi: 10.1287/mnsc.30.9.1078
8. Kalhor R, Amini S, Sokhanvar M, Lotfi F, Sharifi M, Kakemam E. Factors affecting the technical efficiency of general hospitals in iran: data envelopment analysis. *J Egypt Public Health Assoc.* (2016) 91:20–5. doi: 10.1097/01.EPX.0000480717.13696.3c
9. Nabilou B, Yusefzadeh H, Rezapour A, Azar FEF, Safi PS, Asiabar AS, et al. The productivity and its barriers in public hospitals: case study of Iran. *Med J Islam Rep Iran.* (2016) 30:316.
10. Sheikhzadeh Y, Roudsari AV, Vahidi RG, Emrouznejad A, Dastgiri S. Public and private hospital services reform using data envelopment analysis to measure technical, scale, allocative, and cost efficiencies. *Health Promot Perspect.* (2012) 2:28. doi: 10.5681/hpp.2012.004
11. Jandaghi G, Matin HZ, Doremami M, Aghaziyarati M. Efficiency evaluation of qom public and private hospitals using data envelopment analysis. *Eur J Econ Finance Adm Sci.* (2010) 22:83–91.
12. Hajialiafzali H, Moss J, Mahmood M. Efficiency measurement for hospitals owned by the iranian social security organisation. *J Med Syst.* (2007) 31:166–72. doi: 10.1007/s10916-007-9051-6
13. Hatam N, Pourmohammadi K, Keshtkaran A, Javanbakht M, Askarian M. Factors affecting efficiency of social security hospitals in iran: data envelopment analysis. *HealthMED.* (2012) 6:1961–8.
14. Abolghasemi K, Teymourzadeh E. Performance evaluation and ranking of selective wards in a military hospital using DEA and promethee Method. *J Mil Med.* (2017) 18:325–34.
15. Keshtkaran A, Barouni M, Ravangard R, Yandrani M. Economic efficiency of radiology wards using data envelopment analysis: case study of Iran. *Health.* (2014) 6:311. doi: 10.4236/health.2014.65045
16. Bahrami MA, Rafiei S, Abedi M, Askari R. Data envelopment analysis for estimating efficiency of intensive care units: a case study in Iran. *Int J Health Care Qual Assur.* (2018) 31:276–82. doi: 10.1108/IJHCQA-12-2016-0181
17. Gharahighehi A, Kheirkhah AS, Bagheri A, Rashidi E. Improving performances of the emergency department using discrete event simulation, DEA and the MADM methods. *Digit Health.* (2016) 2:2055207616664619. doi: 10.1177/2055207616664619
18. Jahangiri A. Application of data envelopment analysis technique in iranian hospitals (a systematic review). *J Hosp.* (2016) 15:103–24. doi: 10.22105/DMOR.2019.82229
19. Kiadaliri AA, Jafari M, Gerdtham U-G. Frontier-based techniques in measuring hospital efficiency in Iran: A systematic review and meta-regression analysis. *BMC Health Serv Res.* (2013) 13:312. doi: 10.1186/1472-6963-13-312
20. Bahadori M, Izadi AR, Ghardashi F, Ravangard R, Hosseini SM. The evaluation of hospital performance in Iran: a systematic review article. *Iran J Public Health.* (2016) 45:855. doi: 10.5812/traumamon.31382
21. Mosadeghrad AM, Esfahani P, Nikafshar M. Hospitals' Efficiency in Iran: a systematic review and meta-analysis of two decades of research. *J Payavard Salamat.* (2017) 11:318–31. doi: 10.4103/jehp.jehp_393_18
22. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* (2010) 6:E1000097. doi: 10.1016/j.ijsu.2010.02.007
23. Moradi-Lakeh M, Vosoogh-Moghaddam A. Health sector evolution plan in iran; equity and sustainability concerns. *Int J Health Policy Manag.* (2015) 4:637–40. doi: 10.15171/ijhpm.2015.160
24. Joshan S, Shah Hoseini R, Fetros MH. Technical Efficiency analyses in hospitals of tabriz university of medical sciences. *Teb and Tazkieh.* (2016) 1:37–48.
25. Sepehrdoust H, Rajabi E. Factors Effective on Technical Efficiency and Ranking of Efficient Units in Social Security Hospitals. *Health Info Manag.* (2012) 9:887–94.
26. Ghaderi H, Goudarzi G, Gohari M. Determination technical efficiency of hospitals affiliated with Iran University of Medical Science by Data Envelopment Analysis (2000–2004). *Jha.* (2007) 9:39–44. Available online at: http://jha.iums.ac.ir/article-1-42-fa.html
27. Karimi S, Sajadi H, Karami M, Torkzad L, Bidram R. efficiency estimation in general hospitals of isfahan university of medical sciences

during 2005-2006 by data envelopment analysis. *J Health Adm.* (2009) 12:39–46.
28. Mohammadi Ardakani A, Mirghafouri SH, Mirfakhreddini SH, Damki AM, Momeni H. Assessment of Relative Efficiency of Governmental Hospitals of Yazd Province Using DEA. *J Yazd Univ Med Sci.* (2009) 17: 200–8.
29. Pourreza A, Goudarzi G, Azadi H. Determination of technical efficiency of hospitals affiliated with tehran university of medical science by the data envelopment analysis method: 1996–2006. *J School Public Health Ins Public Health Res.* (2010) 7:4.
30. Abolhalaj M, Ahmad Kiadaliri AA, Najafi B. Measuring technical efficiency of hospitals affiliated with universities of medical sciences. *Teb va Tazkiye.* (2010) 19:49–61.
31. Salehzadeh R, Ketabi S. Measuring the Efficiency of Qom Hospitals With Data Envelopment Analysis and Analytic Hierarchy Process. *Health Info Manag.* (2011) 8:479–89.
32. Sadeghi Arani Z, Asadi M, Mirghafouri SH, Khosravanian H. Qualitative performance evaluation of hospitals using DEA, balanced scorecard and servqual; a case study of general hospitals of Yazd. *J Shaheed Sadoughi Univ Medical Sci.* (2011) 18:559–69.
33. Askari R, Goudarzi R, Fallahzadeh H, Zarei B, Dehqani Tafti A. Efficiency appraisal of yazd university of medical science hospitals by quantitative approach data envelopment Analysis(DEA). *Payavard Salamat.* (2012) 6:215–24.
34. Ilbeigi A, Kazemi M, Peivandi MT. Relationship Between Official Performance Measurement Scores and Relative Efficiency in General Hospitals. *Hosp.* (2012) 11:31–44.
35. Rahimi B, Yusefzade H, Khalesi N, Valinejadi A, Gozali A, Akbari S, et al. Analysis of the efficiency and optimal consumption of resources in selected hospitals in urmia province through data envelopment analysis. *J Health Adm.* (2012) 15:91–102.
36. Najjar Zadeh M, Ghasem Zadeh R, Salehi R, Torabi Pour A. Assessment of hospitals efficiency by data envelopment Analysis in Ahvaz in 2006-2010. *Jundishapur J Health Sci.* (2012) 4:77–87.
37. Arab M, Akbari F, Dadashi A, Keshavarz K. Technical efficiency analyses in hospitals of Tabriz University of medical sciences. *J Hosp.* (2012) 11:2011.
38. Azar A, Valipour Khatir M, Moghbel Baerz A, Hasas Yeganeh Y. evaluation of hospital efficiency by data envelopment analysis: Tehran University of medical sciences: 2009-2011. *J Health Adm.* (2013) 16:36–46.
39. Safi Arian R, Shahhoseini R. Assessment of technical efficiency of hospitals under hamadan university of medical sciences on performance indicators and data envelopment analysis model in 2010. *Pajouhan Sci J.* (2013) 11:27–34.
40. Kazemi Z, Ahmad Kiadaliri AA. Haghparast h. Estimating efficiency and optimal resource utilization in selected hospitals in East of Iran: applying the data envelopment analysis. *J Iran Inst Health Sci.* (2013) 12:449–58.
41. Raeesian S, Eslamian M, Lotfi F, Yousefzadeh H, Moradi N, Kalhor R. Comparison of productivity between university and non-university hospitals in Ahvaz: Using MALEMQUIST'S index and DEA. *Jundishapur Sci Med J.* (2014) 13:689–98.
42. Mohebbi far R, Goudarzi G, Rahimi H. Technical efficiency in hospitals of guilan university of medical sciences. *Guilan Univ Med Sci.* (2014) 22:73–9.
43. Fazeli E, Vafaei F, Jamshidi Navid B. Investigation on efficacy of the hospitals affiliated to Ilam university of medical sciences by DEA method. *J Ilam Univ Med Sci.* (2015) 23:89–97.
44. Mahfoozpour S, Pouragh B, Abedi Z, Satarivand S. Assessing efficiency in hospitals of Shahid Beheshti university of medical sciences using data envelopment analysis method. *J Health Promotion Management.* (2015) 5:11–20.
45. Goudarzi G, Soufi M, Rajabi Guilan N, Ghasemi SR. Measuring the technical efficiency of kermanshah university of medical sciences hospitals, using data envelopment analysis method (DEA): 2005-2011 (Nonparametric Approach). *J Rafsanjan Univ Med Sci.* (2015) 13:739–48.
46. Firouzi Jahantigh F, Ostovare M. Performance Evaluation of hospitals affiliated to tehran university of medical sciences using a hybrid model of data envelopment analysis and PROMETHEE Method. *Iran Occup Health Journal.* (2017) 14:140–52.
47. Amoozadeh Doghikola A, Shahverdi R, Rezaei Balf F. Comparison of relative efficiency of hospitals affiliated to babol and mazandaran universities of medical sciences before and after health development plan with data envelopment analysis method. *J Babol Univ Med Sci.* (2018) 20:68–4.
48. Yozi H, Maleki S, Mohammadian F. The Effect of Various management styles on hospital's efficiency of Tehran University of medical science in 2016. *Zanko Med J Kurd Univ Med Sci.* (2018) 4:1–15.
49. Lotfi F, Kalhor R, Bastani P, Shaarbafchi Zadeh N, Eslamian M, Dehghani MR, et al. Various Indicators for the assessment of hospitals' performance status: differences and similarities. *Iran Red Crescent Med J.* (2014) 16:E12950. doi: 10.5812/ircmj.12950
50. Rezapour A, Ebadifard Azar F, Yousef Zadeh N, Roumiani Y, Bagheri Faradonbeh S. Technical efficiency and resources allocation in university hospitals in Tehran, 2009-2012. *Med J Islam Repub Iran.* (2015) 29:266.
51. Torabipour A, Najarzadeh M, Arab M, Farzianpour F, Ghasemzadeh R. Hospitals productivity measurement using data envelopment analysis technique. *Iran J Public Health.* (2014) 43: 1576–81.
52. Nabilou B, Yusefzadeh H, Salem Safi P. Performance assessment of health system reform plan in the hospitals affiliated with urmia university of medical sciences. *J Urmia Nurs Midwifery Fac.* (2017) 14:8 96–905.
53. Rezaei S, Dopeykar N, Barouni M, Jafari M, Gharibi F. Do hospitals affiliated with the kurdistan university of medical sciences perform efficiently? non-parametric data envelopment analysis. *Hosp Pract Res.* (2016) 1:135–40. doi: 10.21859/hpr-0104135
54. Goudarzi G, Imani NMH, Jahanmehr N, Rostami K, Omidfar R, Mahooti F. Hospital performance assessment of lorestan university of medical sciences. *Health Monitor.* (2012) 11:3309–315.
55. Askari R, Farzianpour F, Goudarzi R, Shafii M, Sojaei S. Efficiency evaluation of hospitals affiliated with yazd university of medical sciences using quantitative approach of data envelopment analysis in the year 2001 to 2011. *Pensee J.* (2014) 76:416–25.
56. Sabermahani A, Goudarzi G, Barouni M, Khakian M. Estimation of technical efficiency of general hospitals of kerman university of medical sciences by data envelopment analysis (dea) method in 2007. *J Kerman Univ Med Sci.* (2010) 17:59–67.
57. Najafi B, Beheshti Dehkordi A, Emami Meibodi A. The productivity of general hospitals of Ardebil Province (1999–2006). *J Qazvin Univ Med Sci.* (2011) 14:64–70.
58. Hatam N, Pourmohammadi K, Bastani P, Javanbakht M. Using parametric methods in hospital efficiency measurement: a case study in hospitals affiliated to social security and welfare ministry. *J Razi J Med Sci.* (2013) 20:29–36.
59. Hadian M, Rahimi B, Valinejadi A, Yusefzadeh H, Faghisolouk F. An analysis of productivity in hospitals affiliated to Iran and Tehran universities of medical sciences: 2006-2011. *J Health Admin.* (2015) 17:72–84.
60. Mehrtak M, Yusefzadeh H, Jaafaripooyan E. Pabon lasso and data envelopment analysis: a complementary approach to hospital performance measurement. *Glob J Health Sci.* (2014) 6:107–16. doi: 10.5539/gjhs.v6n 4p107
61. Nayar P, Ozcan YA. Data envelopment analysis comparison of hospital efficiency and quality. *J Med Syst.* (2008) 32:193–9. doi: 10.1007/s10916-007-9122-8
62. Evans RG. "Behavioural" cost functions for hospitals. *Can J Econ.* (1971) 4:198215. doi: 10.2307/133526
63. Emamrezaei A, Barouni M. Efficiency analysis of hospitals in iran: a systematic review. *PAYESH.* (2017) 16:7–16.
64. Kabene SM, Orchard C, Howard JM, Soriano MA. Leduc RJ. The importance of human resources management in health care: a global context. *Hum Resour Health.* (2006) 4:20. doi: 10.1186/1478-449 1-4-20

65. Stock GN, McDermott C. McDermott MJH. The effects of capital and human resource investments on hospital performance. *Hosp Top.* (2014) 92:14–9. doi: 10.1080/00185868.2014.875316
66. Cantor VJM, Poh KL, Joms J. Integrated analysis of healthcare efficiency: a systematic review. *J Med Syst.* (2018) 42:8. doi: 10.1007/s10916-017-0848-7
67. Kohl S, Schoenfelder J, Fügener A, Brunner JOJH. The use of data envelopment analysis (DEA) in healthcare with a focus on hospitals. *Health Care Manag Sci.* (2018) 18:1–42. doi: 10.1007/s10729-018-9443-9
68. Gok MS, Sezen B. Analyzing the efficiencies of hospitals: an application of data envelopment analysis. *J Glob Strateg Manag.* (2011) 10:137–46. doi: 10.20460/JGSM.2011515804
69. Ryan R, Synnot A, Prictor M, Hill S. "Cochrane consumers and communication review group. data extraction template for included studies," In: *Standard Protocol Text and Additional Guidance for Review Authors Edited by Group CaCR.* Available online at: https://CccrgCochraneorg/Authorresources Melbourne: La Trobe University. (2016).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

ORIGINAL RESEARCH
published: 10 December 2021
doi: 10.3389/fpubh.2021.799529

Myocardial Infarction Complications After Surgery in Vietnam: Estimates of Incremental Cost, Readmission Risk, and Length of Hospital Stay

My Hanh Bui[1,2], Quynh Long Khuong[3], Phuoc Thang Dao[4], Cao Phuong Duy Le[5], The Anh Nguyen[6], Binh Giang Tran[7], Duc Hung Duong[8], Tuan Duc Duong[9], Tien Hung Tran[9], Hoang Ha Pham[7], Xuan Thanh Dao[10] and Quang Cuong Le[11]*

[1] Department of Tuberculosis and Lung Diseases, Hanoi Medical University, Hanoi, Vietnam, [2] Department of Functional Exploration, Hanoi Medical University Hospital, Hanoi, Vietnam, [3] Center for Population Health Science, Hanoi University of Public Health, Hanoi, Vietnam, [4] Department of Monitoring and Evaluation, Interactive and Research Development, Ho Chi Minh City, Vietnam, [5] Department of Interventional Cardiology, Nguyen Tri Phuong Hospital, Ho Chi Minh City, Vietnam, [6] Department of Intensive Care, Huu Nghi Hospital, Hanoi, Vietnam, [7] Department of Gastroenterology Surgery, Viet Duc Hospital, Hanoi, Vietnam, [8] Department of Cardiovascular Surgery, Bach Mai Hospital, Hanoi, Vietnam, [9] Vietnam Social Insurance, Hanoi, Vietnam, [10] Department of Orthopedic, Hanoi Medical University Hospital, Hanoi, Vietnam, [11] Department of Neurology, Hanoi Medical University, Hanoi, Vietnam

OPEN ACCESS

Edited by:
Mihajlo Jakovljevic, Hosei University, Japan

Reviewed by:
Huy Nguyen, Federation University Australia, Australia
Li-Nien Chien, Taipei Medical University, Taiwan

***Correspondence:**
My Hanh Bui
buimyhanh@hmu.edu.vn

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *21 October 2021*
Accepted: *18 November 2021*
Published: *10 December 2021*

Citation:
Bui MH, Khuong QL, Dao PT, Le CPD, Nguyen TA, Tran BG, Duong DH, Duong TD, Tran TH, Pham HH, Dao XT and Le QC (2021) Myocardial Infarction Complications After Surgery in Vietnam: Estimates of Incremental Cost, Readmission Risk, and Length of Hospital Stay. Front. Public Health 9:799529. doi: 10.3389/fpubh.2021.799529

Myocardial infarction is a considerable burden on public health. However, there is a lack of information about its economic impact on both the individual and national levels. This study aims to estimate the incremental cost, readmission risk, and length of hospital stay due to myocardial infarction as a post-operative complication. We used data from a standardized national system managed by the Vietnam Social Insurance database. The original sample size was 1,241,893 surgical patients who had undergone one of seven types of surgery. A propensity score matching method was applied to create a matched sample for cost analysis. A generalized linear model was used to estimate direct treatment costs, the length of stay, and the effect of the complication on the readmission of surgical patients. Myocardial infarction occurs most frequently after vascular surgery. Patients with a myocardial infarction complication were more likely to experience readmission within 30 and 90 days, with an OR of 3.45 (95%CI: 2.92–4.08) and 4.39 (95%CI: 3.78–5.10), respectively. The increments of total costs at 30 and 90 days due to post-operative myocardial infarction were 4,490.9 USD (95%CI: 3882.3–5099.5) and 4,724.6 USD (95%CI: 4111.5–5337.8) per case, while the increases in length of stay were 4.9 (95%CI: 3.6–6.2) and 5.7 (95%CI: 4.2–7.2) per case, respectively. Perioperative myocardial infarction contributes significantly to medical costs for the individual and the national economy. Patients with perioperative myocardial infarction are more likely to be readmitted and face a longer treatment duration.

Keywords: coronary artery disease, perioperative, economic burden, re-hospitalization, length of stay

INTRODUCTION

Myocardial infarction (MI) is a common cause of coronary artery disease (CAD) and is a cause of quality-of-life decline and in-hospital mortality (1). MI occurs when the blood flow that supplies oxygen to the myocardium is reduced, leading to fatal myocardial ischemia and necrosis (2, 3). The incidence of MI complications during and after surgery ranges from 0.3 to 33%, according to age group, patient cardiac history, and type of intervention (4–6), while myocardial injury rates are from 9.9 to 17.9% (7). Perioperative myocardial infarction (PMI) occurs most frequently within 1–3 days of surgery, with 57.4% reported 1 day after surgery and nearly 20% more by the third day (8, 9). Although the mortality rate for PMI is on a downward trend (10, 11), in-hospital numbers was 11.6% among patients who are 45 years old or older in 2011 and reached 75.6% over the period 2003–2015 (12, 13). Further, Dennis et al. recorded a rate of 30-day post-discharge mortality at 4.4% among patients who are 65 years old or older in 2017 (10). Alongside the risk of immediate onset, PMI might occur asymptomatically, and so close post-operative follow-up and periodical clinical evaluation is needed.

PMI is not only life-threatening to patients but also results in long-term loss of quality of life. A systematic review indicated the 30-day readmission rate at 12% and was found to be up to 19.1% by Smilowitz (14, 15). Umesh et al. showed a high risk of readmission due to PMI at least once within 12 months of discharge, at 21.3% (16). A higher possibility of admission could lead to longer treatment duration, as has been shown by the correlation between MI and the longest length of stay (11). The burden of treatment cost has been highlighted in a number of studies (17–19), and hospitalization costs increased considerably among patients who undergo intervention (17), with an accumulated high burden on the national economy (19). The cost of MI treatment cost in Vietnam in 2013 was estimated to be US$ 2,503 per hospitalization, which was significantly higher than contemporary GDP per capita (20).

Vietnam is a low-middle income country (LMIC) locating in the South-East region of Asia, with an estimated population in 2019 of ∼96.5 million people (21). Over the same period of 2019, Vietnam's GDP showed an increase of 7.02%, and per capita income was reported at 4.3 million dong (21). Even though treatment therapies for CHD are constantly evolving, CHD negatively contributed 13.22% of all-cause mortality in 2018 (22). Hoa et al. estimated that the prevalence of ST-segment elevation MI in patients who had undergone percutaneous coronary therapy was 14.5% (23). Nonetheless, little is known about the burden of PMI in terms of the risk of readmission, length of hospital stay (LOS), or burden of hospitalization costs. Additional studies are needed to understand the burden of PMI in LMICs and to provide valuable findings to promote protective strategies against impoverishing expenditure as part of the Global Surgery 2030 (24).

This study was conducted to explore the epidemiological features of PMI in Vietnam and evaluate the burden on the national economy, readmission risk, and LOS using data collected from the Vietnam Social Insurance database.

MATERIALS AND METHODS

Data Source

Our study analyzed data exported from a standardized national system, which is the electronic payment portal database under the management of the Vietnam Social Insurance (VSI) agency using an Oracle© database. The system was implemented nationwide on January 1, 2017, and required all medical data from domestic hospitals to be entered and uploaded onto the VSI server to enable reimbursement (25). Data on care for all administrative levels and medical procedures are reported through the VSI database. The administrative levels in Vietnam are organized hierarchically from primary to tertiary care, while medical procedures include preventive and treatment care, medical consultations, maternity care, recovery, and prescribed medications. The 10th revision of International Classification of Diseases (ICD-10) codes was used to code VSI diagnosis data, while a domestic coding system adapted from the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) classification system is used for medical procedures (26). The data for this study were exported from January 1, 2017, to September 30, 2018.

Study Participants

Patients aged 18 years old or above who had undergone at least one of seven procedures were eligible for the study: (1) spinal-neurological, (2) cardiothoracic, (3) vascular, (4) gastrointestinal, (5) urological, (6) orthopedic, and (7) plastic. Patients were excluded from the study if they had undergone another type of surgery within 30 days before being recruited into the study.

Measurements

Myocardial Infarction

MI describes a clinical phenomenon where the bloodstream that supplies oxygen to cardiac muscle tissues is interrupted, with a consequent risk of myocardial ischemia and necrosis (1, 2). One of the most common reasons for myocardial infarction is a thrombosis that forms gradually and asymptomatically inside the coronary arteries (1, 2). Myocardial infarction is diagnosed through clinical symptoms of myocardial ischemia, electrocardiogram (ECG), the presence of relevant biomarkers, and imaging.

MI is diagnosed when the evidence of myocardial necrosis is available accompanied with clinical evidence of myocardial ischemia and the increase and/or decrease of cardiac biomarkers (troponin) with at least one value exceeding the 99th percentile upper reference limit, followed by one of the following criteria as follows (1, 3): (1) symptoms of ischemia; (2) changes in ECG indicating new ischemia as new changes of ST segment and T wave and left bundle branch block; (3) development of pathological Q wave recorded in ECG; (4) evidence of new loss of viable myocardium or regional wall motion abnormality detected by imaging; (5) evidence of thrombosis detected by coronary angiography or autopsy.

Readmission

We examined the impact of PMI on the readmission of surgical patients. Readmission refers to any overnight stay at a hospital. Readmission was assessed within 30 and 90 days after surgery.

Cost

A bottom-up approach was used to capture the total amount of cost reimbursed by VSI and any costs incurred by the patient that were ineligible for reimbursement by VSI. In these cases, direct medical costs were directly assessed. Direct medical costs were separated into three indicators, the total index treatment costs, costs post-operatively within 30 days of surgery, and costs within 90 days. The total index treatment cost is the sum of costs that patients paid for full episodes of initial treatment, while post-operative costs are the total index treatment cost and all extra costs incurred with 30 and 90 days after surgery. A full episode of initial treatment was considered as all stages that patients had completed during their hospital visit, from admission to discharge, and the corresponding costs relate to registration, consolation, testing, medications, and hospitalization. We calculated all the costs in USD with the 2018 rate at 1 USD = 23,255 Vietnamese dong.

Length of Stay

LOS is defined as the number of days of treatment during the index hospital admission, from the date of surgery to discharge. LOS at 30 and 90 days included both the stay for the index hospital admission and stays for any corresponding readmission.

Statistical Analysis

To detect differences in patients' characteristics between the two groups (with and without PMI), we used Chi-squared tests for categorical variables and *t*-tests for continuous variables.

Effects of PMI on Readmission

Multivariable logistic regressions were carried out to evaluate the effect of PMI on readmission within 30 and 90 days after surgery. The models were adjusted for patient socio-demographic, emergency hospitalization status, and pre-operative comorbidities.

Matching

The propensity score matching (PSM) method was used with a matched ratio of one-to-one using nearest neighbor algorithm, to eliminate the risk of potential bias in PMI-related cost estimation as a result of distinction in socio-demographic and hospital characteristics between the two groups of participants. Each individual was assigned a propensity score, which was the output of the logistic regression. Subsequently, an individual in the first group was matched with another individual with the same propensity score in a second group for the purpose of promoting the characteristic balance in the analytic sample (27, 28). The covariates for the logistic regression included socio-demographic characteristics, hospital information, emergency status, and a list of chronic comorbidities provided by Elixhauser et al. (29).

Incremental Cost Estimation

We excluded patients with multiple complications to avoid the compounding effect and to achieve a more precise estimate of the burden attributable to PMI. We used the matched sample to estimate the incremental cost of PMI and LOS. A generalized linear model with a log link function was used for all the models. The Gamma distribution was chosen to reflect the right-skewed nature of the costs, while Poisson distribution was used to estimate the incremental LOS, reflecting the nature of the count data. The 95% confidence intervals (95% CI) were estimated based on the bootstrapping method with 1,000 replications.

A significance level of 0.05 was used for all statistical tests. Data were managed and analyzed using Stata v16 (StataCorp, College Station, TX, USA).

Ethical Approval

All procedures performed in this study were in accordance with the ethical standards of the Ethical Review Board of Hanoi Medical University (IRB approval No. 67/HDDDDHYHN; Dated: March 24, 2017). All patient information was anonymous.

RESULTS

Study Sample

Figure 1 describes the sample size of our study. The original study sample was 1,241,893 surgical patients, of which 1,466 patients had PMI (0.12%). After removing 190,453 patients with other post-operative complications and 25,445 samples with missing values for covariates, the complete case sample included 1,025,985 patients. In the matched sample for costs analysis, 764 patients with PMI were matched with 764 patients without PMI.

Incidence of PMI

The incidence of PMI for each of seven types of surgery is shown in **Figure 2**. The incidence of PMI was highest with vascular surgery (2.30%), followed by cardiothoracic surgeries (0.38%). PMI was <0.1% in other types of surgery.

Patient Characteristics

Table 1 summarizes the pre-operative characteristics of surgical patients with PMI before and after matching. In the unmatched sample, the patient's characteristics were significantly different between groups of patients with PMI (n = 764) and without PMI (n = 1,025,221). After matching 764 patients who had PMI with 764 patients with no post-operative complications using the PSM method, the distribution of patient characteristics was similar in the two groups and no significant differences were found.

Impact of PMI on Readmission

Table 2 shows the results of multivariable logistic regression for the impact of PMI on readmission within 30 and 90 days. After controlling for socio-demographic, emergency hospitalization status, and pre-operative comorbidities, patients with PMI had higher odds of readmission within 30 and 90 days as compared to

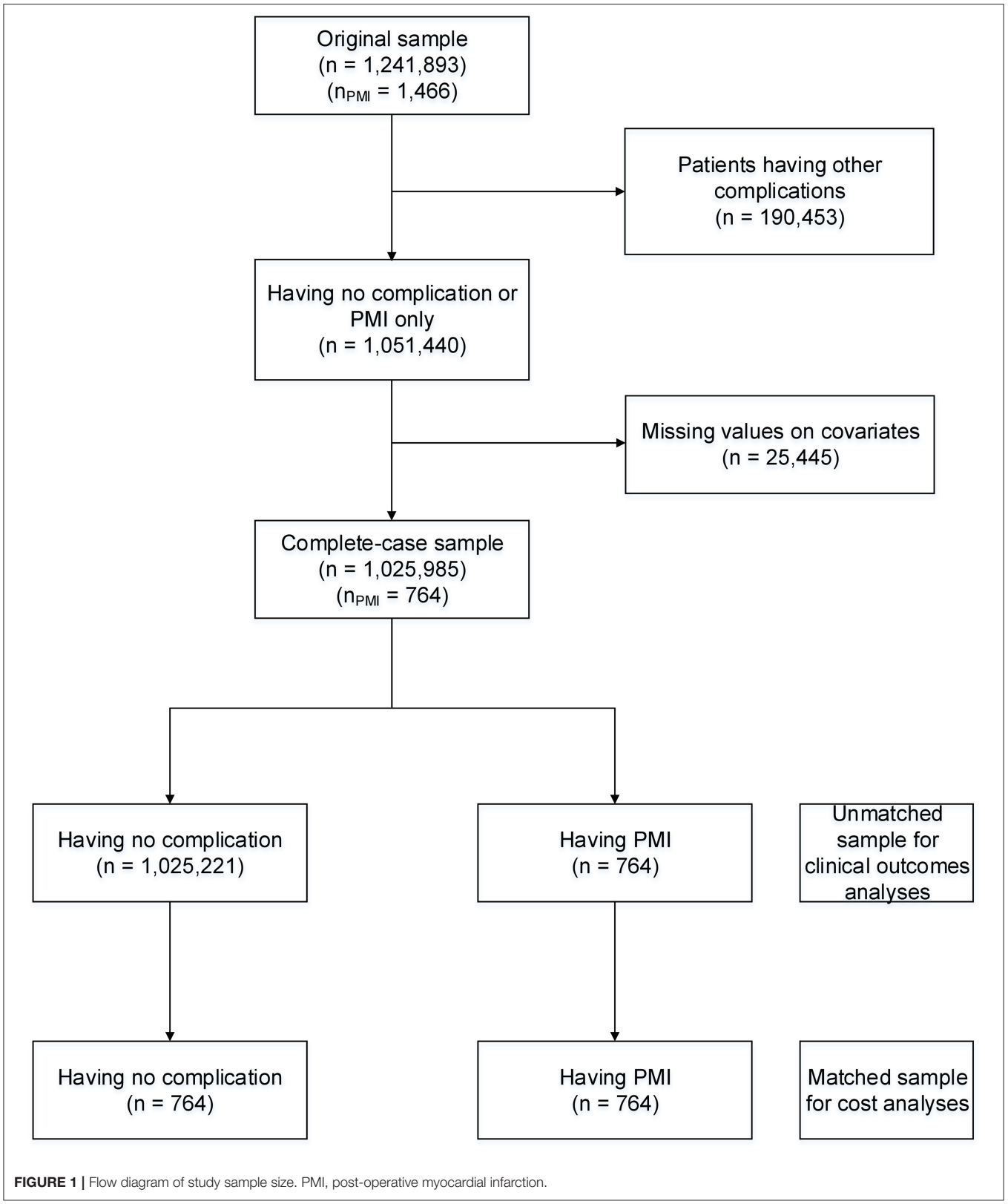

FIGURE 1 | Flow diagram of study sample size. PMI, post-operative myocardial infarction.

those who did not have a post-operative complication (OR = 3.45; 95%CI: 2.92–4.08; and OR = 4.39; 95%CI: 3.78–5.10, respectively).

Incremental Cost and LOS Due to PMI

The incremental costs and LOS due to PMI are shown in **Table 3**. A significant increase was found in most types of cost, except

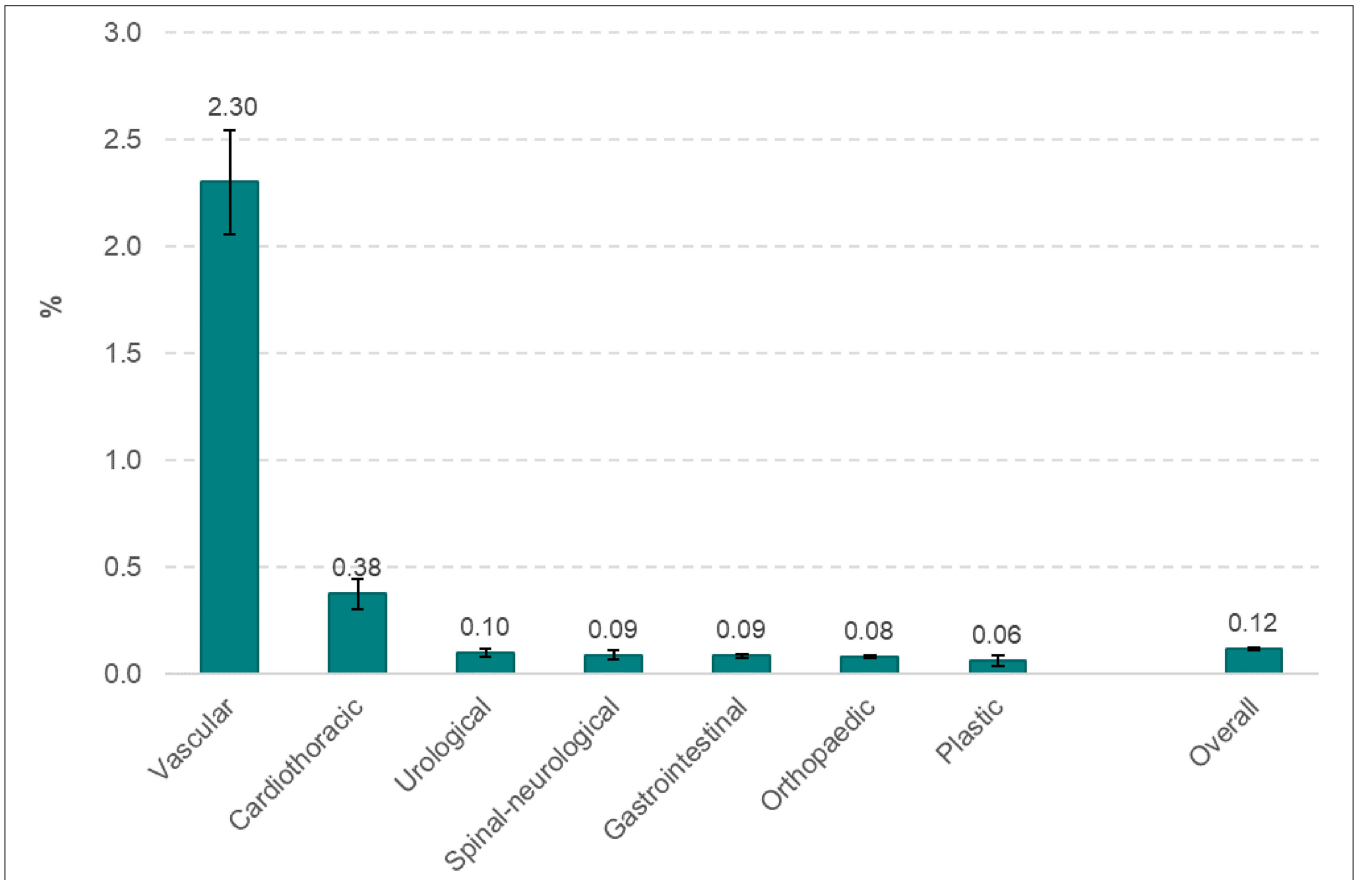

FIGURE 2 | Incidence of post-operative myocardial infarction by seven types of surgery.

for the cost of outpatient visits and the drug cost associated with outpatient visits. Specifically, the incremental cost associated with PMI of the index treatment was 2,176.7 USD (95% CI: 1902.7–2450.8). The increments of total costs within 30 days after surgery were 4,490.9 USD (95% CI: 3882.3–5099.5) and within 90 days were 4,724.6 USD (95% CI: 4111.5–5337.8), which is equivalent to 175% and 184% of the GDP per capita in Vietnam in 2018, respectively. The service cost related to readmission contributed the largest part of the total costs.

LOS was also higher in patients with PMI, with an average increase of 3.5 days for the index treatment, 4.9 days within 30 days, and 5.7 days within 90 days after surgery. All the increments in LOS were significant with a p-value < 0.001.

DISCUSSION

Cardiovascular complications have become a significant public health issue in recent years. Alongside putting persons at risk of poor prognosis, PMI exposes patients to invasive treatment and diminishes their post-operative quality of life. Any poor outcome due to cardiovascular complications, especially PMI, increases the relevant medical costs, and so generates an economic burden that negatively affects patients, especially those living in LMICs. Nevertheless, understanding this impact requires a great deal of effort on data collection at the national level. For these reasons, this present study set out to describe the epidemiological characteristics of PMI in Vietnam and estimate PMI-related medical cost at the individual level, as well as the correlation between PMI and readmission. The data analyzed in our study came from a standardized national system under the management of the VSI.

We found that 1,466 (0.12%) participants were diagnosed with PMI in the initial sample. PMI predominantly occurred after vascular surgery, with 2.3% of patients, and was less frequently recorded among the other types of surgery. Patients with PMI were shown to have higher odds of 30- and 90-day readmission compared to patients without PMI. The total costs of the PMI-related treatment were higher in the PMI group than in the non-PMI groups, and the same trend was found for LOS days. The estimated incremental cost was US$2,176.70 for the index treatment, which accounted for 84.8% of GDP per capita in the contemporary year.

The incidence of PMI in our study was calculated at 0.12%, which is lower than the range of incidence of 0.3–33.3% reported in the literature review by Gualandro et al. (9). The figure of Poldermans et al. showed that the rate of PMI occurring within 30 days after surgery was up to 32.7%; however, the sample size was only 101 patients with extensive

TABLE 1 | Baseline characteristics of surgical patients by myocardial infarction before and after matching.

Factors	No PMI		PMI *n* (%)	*P*-value	
	Unmatched sample *n* (%)	Matched sample *n* (%)		Unmatched sample	Matched sample
N	1,025,221	764	764		
Age, mean (SD)	46.9 (17.3)	65.4 (16.2)	65.0 (15.9)	<0.001	0.60
Gender					
Male	613,861 (59.9)	457 (59.8)	462 (60.5)	0.74	0.79
Female	411,360 (40.1)	307 (40.2)	302 (39.5)		
Region					
Northern Midlands and Mountainous	134,976 (13.2)	49 (6.4)	50 (6.5)	<0.001	0.99
Red River Delta	229,401 (22.4)	169 (22.1)	170 (22.3)		
North Central and South Central Coast	259,894 (25.4)	136 (17.8)	129 (16.9)		
Central Highland	60,444 (5.9)	21 (2.7)	23 (3.0)		
Southeast	182,402 (17.8)	213 (27.9)	207 (27.1)		
Mekong River Delta	158,104 (15.4)	176 (23.0)	185 (24.2)		
Emergency hospitalization	196,872 (19.2)	217 (28.4)	207 (27.1)	<0.001	0.57
Pre-operative concomitant diseases					
Heart failure	7,055 (0.7)	44 (5.8)	49 (6.4)	<0.001	0.59
Valvular heart disease	3,940 (0.4)	9 (1.2)	9 (1.2)	<0.001	1.00
Peripheral vascular disease	2,109 (0.2)	8 (1.0)	10 (1.3)	<0.001	0.64
Hypertension	9,1907 (9.0)	252 (33.0)	252 (33.0)	<0.001	1.00
Paralysis	986 (0.1)	2 (0.3)	1 (0.1)	0.76	0.56
Chronic lung disease	15,907 (1.6)	32 (4.2)	28 (3.7)	<0.001	0.60
Diabetes	36,439 (3.6)	119 (15.6)	124 (16.2)	<0.001	0.73
Complicated diabetes	583 (0.1)	1 (0.1)	3 (0.4)	<0.001	0.32
Hypothyroidism	1,231 (0.1)	2 (0.3)	2 (0.3)	0.26	1.00
Chronic renal failure	6,069 (0.6)	20 (2.6)	27 (3.5)	<0.001	0.30
Liver diseases	24,108 (2.4)	13 (1.7)	16 (2.1)	0.64	0.57
Metastatic cancer	4,797 (0.5)	2 (0.3)	2 (0.3)	0.40	1.00
Cancer	44,609 (4.4)	24 (3.1)	21 (2.7)	0.030	0.65
Joint disease	11,616 (1.1)	6 (0.8)	7 (0.9)	0.57	0.78
Weight loss	8,453 (0.8)	18 (2.4)	18 (2.4)	<0.001	1.00
Fluid and electrolyte disorders	1,476 (0.1)	2 (0.3)	2 (0.3)	0.39	1.00
Anemia	3,483 (0.3)	2 (0.3)	2 (0.3)	0.71	1.00
Depression/addiction	585 (0.1)	0 (0.0)	0 (0.0)	0.51	–

PMI, post-operative myocardial infarction; SD, standard deviation.

ischemia (5), compared to more than one million surgical patients in our study. It should be noted that the incidence might vary across studies due to differences in either socio-demographic or clinical characteristics or disease definitions (9). Kertai et al. found a PMI incidence of 8.9% among patients experiencing abdominal aortic aneurysm surgery (30), while Durazzo et al. found a rate of 11% for patients undergoing arterial surgery (31), which is higher than the rate of 2.3% in our study. However, the sample size of the two studies was significantly smaller than the VSI sample and focused on specific groups.

We found that participants with PMI had an adjusted OR of 30-day readmission 3.45-times higher than ones without PMI, and the effect was even stronger for 90-day readmission. Smilowitz et al. revealed that, in 2014, patients in the US diagnosed with perioperative acute MI were more likely to undergo readmission within 30 days after discharge (15). The finding by Smilowitz is consistent with our results, and the analysis of Smilowitz was based on a national database (15). The same trend was displayed in a systematic review of 14 studies on PMI topics and MI was recorded as a potential risk factor for 30-day readmission with the pool OR at 2.26 (14).

Although national social care packages have been widely developed in the attempt to support patients in reducing health-related burdens, the accumulative medical care contributes to a significant expense for patients. Furthermore, even though medical costs are partly reimbursed by VSI, the high level of cost may prevent PMI patients from seeking treatment, leading to a higher risk of poor prognosis due to the urgent nature

TABLE 2 | Impact of post-operative myocardial infarction on readmission within 30 and 90 days of surgical patients.

Factors	30-day readmission		90-day readmission	
	OR	95%CI	OR	95%CI
PMI	3.45***	2.92–4.08	4.39***	3.78–5.10
Age	1.01***	1.01–1.01	1.02***	1.01–1.02
Gender (Ref: Male)				
Female	0.85***	0.84–0.86	0.89***	0.88–0.90
Region (Ref: Northern Midlands and Mountainous)				
Red River Delta	0.86***	0.84–0.88	0.85***	0.83–0.87
North Central and South Central Coast	0.82***	0.80–0.84	0.86***	0.84–0.87
Central Highland	0.83***	0.80–0.86	0.84***	0.82–0.87
Southeast	0.60***	0.58–0.61	0.60***	0.59–0.61
Mekong River Delta	0.64***	0.62–0.66	0.63***	0.61–0.65
Emergency hospitalization	1.12***	1.10–1.14	1.09***	1.07–1.11
Hospital classification (Ref: Level 2 or lower)				
Pre-operative concomitant disease				
Heart failure	1.03	0.95–1.11	1.02	0.95–1.09
Valvular heart disease	1.17**	1.06–1.30	1.16**	1.06–1.28
Peripheral vascular disease	1.27***	1.11–1.44	1.18**	1.05–1.33
Hypertension	1.13***	1.10–1.16	1.14***	1.12–1.17
Paralysis	2.13***	1.81–2.51	1.97***	1.69–2.30
Chronic lung disease	1.22***	1.16–1.28	1.29***	1.24–1.35
Diabetes	1.22***	1.18–1.26	1.23***	1.19–1.26
Complicated diabetes	1.11	0.88–1.42	1.16	0.94–1.44
Hypothyroidism	0.97	0.80–1.17	1.04	0.89–1.23
Chronic renal failure	3.27***	3.08–3.48	3.09***	2.91–3.27
Liver disease	1.07***	1.03–1.12	1.08***	1.05–1.13
Metastatic cancer	2.35***	2.19–2.53	2.26***	2.11–2.42
Cancer	5.52***	5.40–5.65	5.11***	5.00–5.22
Joint disease	0.95	0.89–1.01	0.99	0.93–1.04
Weight loss	1.07*	1.00–1.15	1.04	0.98–1.11
Fluid and electrolyte disorders	1.13	0.97–1.32	1.10	0.95–1.26
Anemia	1.42***	1.29–1.56	1.43***	1.31–1.55
Depression/addiction	1.30	0.98–1.70	1.21	0.94–1.55

*All estimates were calculated using unmatched sample, ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.*
OR, odds ratio; 95%CI, 95% confidence interval; PMI, post-operative myocardial infarction; Ref, reference.

of the disease. We found that patients with PMI had to pay three times more than patients without PMI. As a result, the incremental cost equated to roughly 84.8% of the 2018 year GDP per capita. This high impact is in line with results published in previous studies (20, 32, 33). Nguyen et al. showed that acute MI treatment costed Vietnamese patients US$ 2,503 per hospitalization in 2013, which is lower than our estimated cost in 2018 (20). This difference may be because our sample size was larger than the sample of only 89 patients analyzed by Nguyen (20). In addition, we collected data from the VSI database that contains medical cost information across regions in Vietnam, while the data range recorded by Nguyen focused on one regional hospital (20). The variation in currency exchange rates across years may also contribute to the differences between the two studies. Our finding was equivalent to that of Soekhlal et al. who found a mean cost of acute MI treatment of US$3,094 for 2012 in Netherlands (32). The data in a study by Soekhlal was collected from nationwide at the same administrative levels as our database (32). In 2005, Tiemann et al. calculated total acute MI-related cost per case was ∼US$3,027 using a European database (34). Moreover, the estimated cost increased across European countries based on the availability of modern invasive procedures (34). The differences in treatment cost remained unchanged between two groups after summing all the costs incurring within 30 and 90 days after surgery. In addition, PMI generated the risk of suffering a longer hospital LOS. This negative point is reflected in the values of the PMI group; which were significantly higher than the non-PMI group for both the index treatment and post-operative duration.

Our study applied a standardized national system under VSI management that provided the individual data across all regions in Vietnam. The large scale of the database

TABLE 3 | Incremental costs and length of hospital stay due to post-operative myocardial infarction complication.

	Mean cost (USD)		Incremental cost of PMI, USD (% GDP[a])	95% Bootstrap CI	*P*-value
	No PMI	PMI			
N	764	764			
Treatment cost					
Total cost of the index treatment	1019.9	3196.6	2176.7 (84.8)	1902.7–2450.8	<0.001
30-day costs					
Service cost for 30-day readmission	858.8	2773.8	1914.9 (74.6)	1649.0–2180.8	<0.001
Drug cost for 30-day readmission	177.4	547.5	370.2 (14.4)	292.6–447.7	<0.001
Service cost for 30-day outpatient visits	79.8	89.7	9.9 (0.4)	−19.7–39.5	0.511
Drug cost for 30-day outpatient visits	37.6	56.8	19.2 (0.7)	−0.8–39.1	0.060
Total 30-day cost	2173.4	6664.3	4490.9 (175.0)	3882.3–5099.5	<0.001
90-day costs					
Service cost for 90-day readmission	882.7	2982.3	2099.6 (81.8)	1834.4–2364.8	<0.001
Drug cost for 90-day readmission	187.9	572.3	384.4 (15.0)	305.4–463.5	<0.001
Service cost for 90-day outpatient visits	96.6	126.9	30.3 (1.2)	−2.9–63.5	0.074
Drug cost for 90-day outpatient visits	49.7	83.3	33.6 (1.3)	10.9–56.3	0.004
Total 90-day cost	2236.7	6961.4	4724.6 (184.1)	4111.5–5337.8	<0.001
			Difference in LOS		
Length of stay (days)					
Length of hospital stay of the index treatment	11.8	15.3	3.5	2.4–4.6	<0.001
Total length of treatment within 30 days[b]	13.0	17.9	4.9	3.6–6.2	<0.001
Total length of treatment within 90 days[c]	14.0	19.7	5.7	4.2–7.2	<0.001

[a] *Compared to GDP per capita in Vietnam in 2018 (2,566 USD).*

[b] *Includes length of hospital stay in the treatment and 30-day re-hospitalization periods.*

[c] *Includes length of hospital stay in the treatment and 90-day re-hospitalization periods. All estimates were calculated using matched sample.*

CI, confidence interval; GDP, gross domestic product; LOS, length of stay; PMI, post-operative myocardial infarction; Ref, reference.

strengthened the costing analysis that was central to this study. The propensity scores matching was executed to reduce the imbalance in to characteristics between two groups. Hence, the estimated costs were adjusted for the socio-demographic characteristics, hospital information, emergency status, and history of comorbidities.

There were limitations to our study. The analysis did not include indirect costs that possibly exist as a result of the negative effects of PMI treatment. The indirect costs may emerge due to the expenses of either additional functional foods or drugs outside regular hospital medications. Also, the VSI database does not cover the data on care for cases seeking treatment in the private sector. The estimated medical costs are not adjusted for the different thresholds of reimbursement that align with VSI regulations and are not stratified according to different types of surgery. The retrospective data potentially contains a certain number of missing records as well as the issues of data uncertainty, which could potentially affect the precision of the estimates.

CONCLUSION

Even though the incidence of PMI has decreased in recent years, medical costs affect the lives of patients and create a heavy burden on the national economy. The other impact of PMI on LOS and risk of readmission, if not properly considered, harm the patient's quality of life and prompt the need for medical care. Additional research is needed to explore the economic burdens of PMI on individual and national levels across various types of surgery and

social status and to estimate the medical costs relating to the private sector.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

MB, QK, and PD conceived the study, performed the official statistical analyses, interpreted the results, and wrote the manuscript. MB and TT collected and cleaned data. CL, TN, BT, DD, TD, TT, HP, XD, and QL provided the critical revision of the manuscript for important intellectual content. All authors have read and approved the final manuscript.

REFERENCES

1. Thygesen K, Alpert Joseph S, Jaffe Allan S, Chaitman Bernard R, Bax Jeroen J, Morrow David A, et al. Fourth universal definition of myocardial infarction. *J Am Coll Cardiol.* (2018) 72:2231–64. doi: 10.1016/j.jacc.2018.08.1038
2. Saleh M, Ambrose JA. Understanding myocardial infarction. *F1000Res.* (2018) 7:F1000 Faculty Rev-378. doi: 10.12688/f1000research.15096.1
3. Thygesen K, Alpert Joseph S, White Harvey D. Universal definition of myocardial infarction. *J Am Coll Cardiol.* (2007) 50:2173–95. doi: 10.1016/j.jacc.2007.09.011
4. Fleischmann KE, Goldman L, Young B, Lee TH. Association between cardiac and noncardiac complications in patients undergoing noncardiac surgery: outcomes and effects on length of stay. *Am J Med.* (2003) 115:515–20. doi: 10.1016/S0002-9343(03)00474-1
5. Poldermans D, Schouten O, Vidakovic R, Bax JJ, Thomson IR, Hoeks SE, et al. A clinical randomized trial to evaluate the safety of a noninvasive approach in high-risk patients undergoing major vascular surgery: the DECREASE-V Pilot Study. *J Am Coll Cardiol.* (2007) 49:1763–9. doi: 10.1016/j.jacc.2006.11.052
6. Mangano Dennis T, Hollenberg M, Fegert G, Meyer ML, London Martin J, Tubau Julio F, et al. Perioperative myocardial ischemia in patients undergoing noncardiac surgery—I: incidence and severity during the 4 Day perioperative period. *J Am Coll Cardiol.* (1991) 17:843–50. doi: 10.1016/0735-1097(91)90863-5
7. Smilowitz NR, Redel-Traub G, Hausvater A, Armanious A, Nicholson J, Puelacher C, et al. Myocardial injury after noncardiac surgery: a systematic review and meta-analysis. *Cardiol Rev.* (2019) 27:267–73. doi: 10.1097/CRD.0000000000000254
8. Kikura M, Oikawa F, Yamamoto K, Iwamoto T, Tanaka KA, Sato S, et al. Myocardial infarction and cerebrovascular accident following non-cardiac surgery: differences in post-operative temporal distribution and risk factors. *J Thromb Haemost.* (2008) 6:742–8. doi: 10.1111/j.1538-7836.2008.02948.x
9. Gualandro DM, Campos CA, Calderaro D, Yu PC, Marques AC, Pastana AF, et al. Coronary plaque rupture in patients with myocardial infarction after noncardiac surgery: frequent and dangerous. *Atherosclerosis.* (2012) 222:191–5. doi: 10.1016/j.atherosclerosis.2012.02.021
10. Ko DT, Khera R, Lau G, Qiu F, Wang Y, Austin PC, et al. Readmission and mortality after hospitalization for myocardial infarction and heart failure. *J Am Coll Cardiol.* (2020) 75:736–46. doi: 10.1016/j.jacc.2019.12.026
11. Baek H, Cho M, Kim S, Hwang H, Song M, Yoo S. Analysis of length of hospital stay using electronic health records: a statistical and data mining approach. *PLoS ONE.* (2018) 13:e0195901. doi: 10.1371/journal.pone.0195901
12. Devereaux PJ, Xavier D, Pogue J, Guyatt G, Sigamani A, Garutti I, et al. Characteristics and short-term prognosis of perioperative myocardial infarction in patients undergoing noncardiac surgery. *Ann Intern Med.* (2011) 154:523–8. doi: 10.7326/0003-4819-154-8-201104190-00003
13. Zhou X, Chen L, Su Z, Li Y, Tu M, Xiao J, et al. Perioperative acute myocardial infarction in patients after non-cardiac surgery in China: characteristics and risk factors. *Medicine.* (2019) 98:e16929. doi: 10.1097/MD.0000000000016929
14. Wang H, Zhao T, Wei X, Lu H, Lin X. The prevalence of 30-day readmission after acute myocardial infarction: a systematic review and meta-analysis. *Clin Cardiol.* (2019) 42:889–98. doi: 10.1002/clc.23238
15. Smilowitz NR, Beckman JA, Sherman SE, Berger JS. Hospital readmission after perioperative acute myocardial infarction associated with noncardiac surgery. *Circulation.* (2018) 137:2332–9. doi: 10.1161/CIRCULATIONAHA.117.032086
16. Khot UN, Johnson MJ, Wiggins NB, Lowry AM, Rajeswaran J, Kapadia S, et al. Long-term time-varying risk of readmission after acute myocardial infarction. *J Am Heart Assoc.* (2018) 7:e009650. doi: 10.1161/JAHA.118.009650
17. Sugiyama T, Hasegawa K, Kobayashi Y, Takahashi O, Fukui T, Tsugawa Y. Differential time trends of outcomes and costs of care for acute myocardial infarction hospitalizations by st elevation and type of intervention in the United States, 2001–2011. *J Am Heart Assoc.* 4:e001445. doi: 10.1161/JAHA.114.001445
18. Tran DT, Ohinmaa A, Thanh NX, Welsh RC, Kaul P. The healthcare cost burden of acute myocardial infarction in Alberta, Canada. *Pharmacoecon Open.* (2018) 2:433–42. doi: 10.1007/s41669-017-0061-0
19. Seo H, Yoon S-J, Yoon J, Kim D, Gong Y, Kim AR, et al. recent trends in economic burden of acute myocardial infarction in South Korea. *PLoS ONE.* (2015) 10:e0117446. doi: 10.1371/journal.pone.0117446
20. Nguyen TP, Nguyen T, Postma M. Economic burden of acute myocardial infarction in Vietnam. *Value Health.* (2015) 18:389. doi: 10.1016/j.jval.2015.09.857
21. Puelacher C, Lurati-Buse G, Singeisen H, Dang M, Cuculi F, Müller C. Perioperative myocardial infarction/injury after noncardiac surgery. *Swiss Med Wkly.* (2015) 145:w14219. doi: 10.4414/smw.2015.14219
22. World Health Rankings. *Viet Nam: Coronary Heart Disease.* (2018). Available online at: https://www.worldlifeexpectancy.com/viet-nam-coronary-heart-disease (accessed May 01, 2021).
23. Vu HTT, Pham HM, Nguyen HTT, Nguyen QN, Do LD, Pham NM, et al. Novel insights into clinical characteristics and in-hospital outcomes of patients undergoing percutaneous coronary intervention in Vietnam. *Int J Cardiol Heart Vasculat.* (2020) 31:100626. doi: 10.1016/j.ijcha.2020.100626
24. Meara JG, Leather AJ, Hagander L, Alkire BC, Alonso N, Ameh EA, et al. Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. *Surgery.* (2015) 158:3–6. doi: 10.1016/j.surg.2015.04.011
25. Bank TW. *Current Health Expenditure Per Capita - Vietnam*. Available online at: https://data.worldbank.org/indicator/SH.XPD.CHEX.PC.CD?locations=VN (accessed May 01, 2020).
26. Ng JYS, Ramadani RV, Hendrawan D, Duc DT, Kiet PHT. National health insurance databases in Indonesia, Vietnam and the Philippines. *Pharmacoecon Open.* (2019) 3:517–26. doi: 10.1007/s41669-019-0127-2
27. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res.* (2011) 46:399–424. doi: 10.1080/00273171.2011.568786
28. Dehejia RH, Wahba S. Propensity score-matching methods for nonexperimental causal studies. *Rev Econ Stat.* (2002) 84:151–61. doi: 10.1162/003465302317331982
29. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care.* (1998) 36:8–27. doi: 10.1097/00005650-199801000-00004
30. Kertai MD, Boersma E, Westerhout CM, Klein J, van Urk H, Bax JJ, et al. A combination of statins and beta-blockers is independently associated with a reduction in the incidence of perioperative mortality and nonfatal myocardial infarction in patients undergoing abdominal aortic aneurysm surgery. *Eur J Vasc Endovasc Surg.* (2004) 28:343–52. doi: 10.1016/j.ejvs.2004.07.008
31. Durazzo AES, Machado FS, Ikeoka DT, De Bernoche C, Monachini MC, Puech-Leão P, et al. Reduction in cardiovascular events after vascular surgery with atorvastatin: a randomized trial. *J Vasc Surg.* (2004) 39:967–75. doi: 10.1016/j.jvs.2004.01.004

32. Soekhlal RR, Burgers LT, Redekop WK, Tan SS. Treatment costs of acute myocardial infarction in the Netherlands. *Netherlands Heart J.* (2013) 21:230–5. doi: 10.1007/s12471-013-0386-y
33. Tran DT, Welsh RC, Ohinmaa A, Thanh NX, Kaul P. Resource use and burden of hospitalization, outpatient, physician, and drug costs in short- and long-term care after acute myocardial infarction. *Can J Cardiol.* (2018) 34:1298–306. doi: 10.1016/j.cjca.2018.05.022
34. Tiemann O. Variations in hospitalisation costs for acute myocardial infarction – a comparison across Europe. *Health Econ.* (2008) 17:S33–45. doi: 10.1002/hec.1322

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Public Health

ORIGINAL RESEARCH
published: 04 April 2022
doi: 10.3389/fpubh.2022.794177

Spatial Effect Analysis of Health Expenditure and Health Output in China From 2011 to 2018

Penghui Xu[1,2], *Xicang Zhao*[1]*, *Haili Li*[2] *and Shi Guo*[2]

[1] School of Management, Jiangsu University, Zhenjiang, China, [2] School of Humanities and Management, Wannan Medical College, Wuhu, China

Objective: The objective of this study is to study the spatial effects of health expenditure and health output in China.

Methods: Using the spatial panel data of 31 provinces in China from 2011 to 2018, the spatial weight matrix was introduced to analyze the spatial correlation, and the spatial Durbin model (SDM) was used to investigate the health output effect of health expenditure.

Results: Excluding the number of doctors per thousand, the provincial health expenditure, the number of beds per thousand population, and per capita education level had a positive impact on the regional health output. The health effect of China's health inputs showed a spatial spillover effect.

Conclusion: Due to the significant spatial effect, the health output of 31 provinces in China benefits not only from the local health inputs, but also from the health inputs of neighboring provinces.

Suggestions: This article puts forward some suggestions based on the conclusion: China should strengthen the health cooperation among neighboring provinces, promote the free flow of various health factors among provinces, make full use of the spillover and interdependence of health investment among provinces, and improve the medical policy environment in China.

Keywords: health expenditure, health output, spatial lag model, spatial Durbin model, spatial spillover effect

OPEN ACCESS

Edited by:
Mihajlo Jakovljevic, Hosei University, Japan

Reviewed by:
P. Wilner Jeanty, OhioHealth, United States
Mahyudin Ahmad, MARA University of Technology, Malaysia

***Correspondence:**
Xicang Zhao
549612136@qq.com

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *13 October 2021*
Accepted: *14 January 2022*
Published: *04 April 2022*

Citation:
Xu P, Zhao X, Li H and Guo S (2022) Spatial Effect Analysis of Health Expenditure and Health Output in China From 2011 to 2018. Front. Public Health 10:794177. doi: 10.3389/fpubh.2022.794177

INTRODUCTION

Under the concept of the neoclassical growth model, human capital, particularly good human health, is an imperative factor for attaining the desired economic growth and development of any country (1). Besides, based on Grossman's human capital model, health increases the human capital by making more time for working and increasing utility (2). The promotion of health is substantial because health is vital, a fundamental right (3). At present, health is the center of sustainable development, and increasing the individuals' health level is one of the most important policies of countries. Therefore, investment in health is very important.

Health expenditures contain all the expenditures which are used for preparing and improving the individuals' health (4). The concept of health expenditures is different from one country to another. In China, health expenditure refers to the total amount of funds consumed by the whole society for medical and health services in a certain period of time (usually 1 year) in a

country or region, which includes the total monetary amount of living labor and materialized labor consumed in providing healthcare services (5). Health expenditure can not only comprehensively reflect the basic situation of a country and region, such as total health investment, resource allocation efficiency, and population medical burden, but also an important macroindicator of health economy to scientifically evaluate whether the investment is sufficient, whether the financing is fair, and whether the allocation is effective (6). From 2011 to 2018, China's health expenditure maintained a rapid growth, with an increase of about 242.84% in 8 years, from 2,434.591 billion Yuan in 2011 to 5,912.191 billion Yuan in 2018; the infant mortality rate decreased from 12.1% in 2011 to 3.9% in 2018, a decrease of 67.76%.

Over the past few decades, scholars and researchers have paid more attention to evaluate the link between health expenditure and health outcomes. Relevant studies can be mainly summarized in two aspects. Some studies showed that health expenditure leads to better health outcomes, whereas others have reported an insignificant effect. For instance, Musgrove (7), Filmer and Pritchett (8), and Fayissa and Gutema (9) both concluded that health expenditure is not a crucial determinant of health outcomes. Furthermore, Fayissa and Gutema (9) reported a strong negative effect of increases in health expenditure on life expectancy at birth, and Gupta et al. (10) reported that the effect of public health expenditure on health outcomes is weak. Kiross et al. (11) used panel data from World Bank Development Indictors (WDI) from 2000 to 2015 covering 46 countries in sub-Saharan Africa, used the random effects model, and found that public and external healthcare expenditure was significantly negatively correlated with health output. Based on dynamic panel system GMM model, Shen et al. (12) suggest that private health expenditure had no significant impact on health output.

On the other hand, studies by Anyanwu and Ehijakpor (13), Kamiya (14), and Novignon et al. (15) have reported a positive effect of health expenditure on health outcomes. Whereas these studies have reported a positive effect of health expenditure on health outcomes, the conclusions from these studies have differed. Bein et al. (16) and Raeesi et al. (17) suggest that health outcomes improved with the increase of public and private healthcare expenditure. Zhang et al. (18) and Mao et al. (19) both propose an increase in public health expenditure to improve health outcomes. Novignon et al. (15) suggest that the effect of public health expenditure is stronger than that of private. However, Rahman et al. (20) suggest that the effect of private health expenditure is greater than that of public health expenditure. Zhao et al. (21) adopted the vector error correction model and found that the total health expenditure in China had a significant improvement effect on health output.

Based on our review of the literature, first, we identify that there is no consistent conclusion on the effects of health expenditure on health outcomes. Second, most scholars have used the random and fixed effect models, cointegration test, and the vector error correction model to investigate the effects of health expenditure on health outcomes, which is likely to cause errors or biasness in estimating and analyzing the process, and they fail to consider the spatial correlation of health outputs and whether there are spatial spillover effects of health expenditure. Based on the spatial panel data of 31 Chinese provinces obtained from 2011 to 2018, we employ the spatial Durbin model (SDM) to analyze the effects of health expenditure on health outcomes and finally puts forward some conclusions and suggestions based on the deepening of cognition of healthy China.

RESEARCH DESIGN

Model Setting and Variable Selection

Based on the aim of this study, we included the relationship between health expenditure and health development into the system subframework. Based on the health production function Grossman (2) and the definition of "social determinants of health" by the World Health Organization (22), a macrohealth production function model is established as follows:

$$lnHealth_{it} = \alpha_0 + \beta_1 lnHI_{it} + \beta_2 lnEdu_{it} + \beta_3 lnBed_{it} + \beta_4 lnDoc_{it} + \varepsilon_{it} \quad (1)$$

where i is region, t represents year, α denotes constant, ε represents the random error term, and $\beta_1 - \beta_4$ represent the coefficient matrix of each variable, respectively, and measure the influence degree of each variable on the explained variable. As the dependent variable, we use health to express the health output. Other variables are independent variables. The specific description and sources of the dependent and independent variables are as follows:

(1) Health output (health): Based on our review of the literature, we found that the measurement indicators of health output generally focus on life expectancy per capita, infant mortality, maternal mortality, mortality, mortality of children under the age of 5 years, etc. Considering the availability of data and the authority of indicators, according to the research of Yang and Lu (23), we adopt the provincial annual population mortality (%) indicator to measure health, which is a negative indicator, that is, the higher the value is, the lower the regional health level is. Failure to forward the reverse index will lead to the deviation in the reliability of the analysis results, which will have an adverse impact on the correctness of the decision (24). Therefore, we use the reciprocal method to forward the data of the reverse index based on SPSS20.0 software.
(2) The provincial health expenditure (HI) denotes health expenditure and health input of each province. The health expenditure refers to the total amount of funds consumed by the whole society for medical and health services in 1 year in each province, which includes the total monetary amount of living labor and materialized labor consumed in providing healthcare services (5).
(3) The number of beds per thousand (Bed) refers to the number of medical beds per thousand people, which indicates the medical facilities and service level of each region. The number of doctors per thousand has a significant positive impact on health output (13), and the increase in the number of beds will promote residents' demand for health, so as to improve their health status and improve their health level (25).

TABLE 1 | Descriptive statistics of variables.

Variable	Description	Mean	Std. dev.	Min	Max
Health	The provincial annual population mortality	0.5407	0.2392	0	1
HI	The provincial health expenditure	1286.31	979.3121	63.97	5198.69
Doc	The number of doctors per thousand	5.9034	1.6025	2.68	15.46
Bed	The number of beds per thousand	4.9939	0.9736	2.77	7.55
Edu	Per capita education level	9.0162	1.1308	4.2219	12.6754

(4) Per capita education level (Edu) refers to the average level of education in a region. According to Grossman's research, good education plays an important role in obtaining job opportunities, obtaining nutrition, forming a good lifestyle, and efficient use of drugs (2). It is of great significance to improve the quality of life and health level. That is, per capita education level has a significant positive impact on health output (26).

(5) The number of doctors per thousand (Doc) refers to the number of health technical personnel per thousand people, representing the medical technical personnel in each region, reflecting the level of health services in a region. In the context of shortage of health human resources, increasing human input can obtain higher health output (27). Establishing dynamic panel data from the macrolevel and using GMM estimation method, Shen et al. (28) found that the number of doctors per thousand people can significantly improve the health level.

Data Sources

In accordance with the consistency, validity, and availability of data, statistical data are collected from 31 provinces during 2011 and 2018, and the data sources include the China Statistics Yearbook (2012–2020), the China Health Statistics Yearbook (2012), China Health and Family Planning Statistics Yearbook (2013–2017), the China Health Statistics Yearbook (2018–2020), and wind database. All variables are adopted as their natural logarithm for processing potential heteroscedasticity, as listed in **Table 1**.

METHOD

Spatial Econometric Models

To examine and measure the possible spatial effects, the spatial lag model (SLM), the spatial error model (SEM), and the SDM have been utilized, including that SLM includes the regression model of spatial dependence through the addition of a lagged dependent variable; SEM believes that the spatial correlation of variables may ignore the error term of the regression model through the addition of the dependent variable, whereas SDM not only reveals spatial spillover effect of dependent variable in adjacent regions, but also captures influence of independent variables in adjacent regions on their own dependent variables. It is a common model for empirical test of spatial spillover effect (29). These three spatial econometrics models were constructed as follows:

SLM:

$$InHealth_{it} = \rho WInHealth_{it} + \beta_1 InHI_{it} + \beta_2 InDoc_{it} + \beta_3 InBed_{it} + \beta_4 InEdu_{it} + \mu_i + \lambda_t + \varepsilon_{it}, \varepsilon_{it} \sim N(0, \sigma^2 I_n) \quad (2)$$

SEM:

$$Inhealth_{it} = \beta_1 InHI_{it} + \beta_2 InDoc_{it} + \beta_3 InBed_{it} + \beta_4 InEdu_{it} + \mu_i + \lambda_t + \phi_{it}$$
$$\phi_{it} = \eta \Sigma_{j=1}^{N} W_{ij}\phi_{it} + \varepsilon_{it}, \varepsilon_{it} \sim N(0, \sigma^2 I_n) \quad (3)$$

SDM:

$$InHealth_{it} = \rho WInHealth_{it} + \beta_1 InHI_{it} + \beta_2 InDoc_{it} + \beta_3 InBed_{it} + \beta_4 InEdu_{it} + W(\theta_1 InHI_{it} + \theta_2 InDoc_{it} + \theta_3 InBed_{it} + \theta_4 InEdu_{it}) + \mu_i + \lambda_t + \varepsilon_{it}$$
$$\varepsilon_{it} \sim N(0, \sigma^2 I_n) \quad (4)$$

Where β denotes the direct coefficient of the independent variable, which indicates the impact exerted by the independent variable on the dependent variable. θ denotes the space lag coefficient of the independent variable, which indicates the impact exerted by the independent variable in surrounding cities on the dependent variable in the local city. ρ denotes the spatial autoregressive coefficient, which indicates the degree of dependent variable's spatial dependence. λ_t stands for the time fixed effect, μ_i refers to the space fixed effect. W denotes the economic distance weight matrix (W). Wy, Wx denote the spatial lag terms of dependent variable and independent variable, which allows us to analyze the spillover effects of independent variables. ε_{it} denotes a random error vector, satisfying $\varepsilon_{it} \sim N(0, \sigma_{it}^2)$. η measures the impact intensity of the error of the dependent variable in neighborhood cities. i is the 31 province units in China, and t represents year.

To judge which spatial econometric model is more appropriate, Elhorst (29) proposed a test method: to test and sum the two hypotheses of SDM panel model through Wald test and LR test based on $H_0 : \theta = 0$ and $H_0 : \theta + \rho\beta = 0$. If both assumptions are rejected, SDM panel model should be used; If $\theta = 0$, and LM test and robust LM test show that the dependent variables have spatial correlation, SDM converts into a SLM; If $\theta + \rho\beta = 0$, and LM test and robust LM test show that the residual has spatial autocorrelation, SDM simplifies into a SEM. The specific econometric model is Equations (2)–(4). All variables are presented in logarithmic forms to eliminate possible heteroscedasticity.

Spatial Weight Matrix

Generally, the spatial weight matrix based on the economic correlation has also been widely utilized in spatial econometrics

(30). Therefore, we construct and employ the economic distance weight matrix (W) to investigate the effects of health expenditure and health output. The spatial weight matrix was constructed as follows:

$$W = \begin{cases} \frac{1}{|\overline{GDP_i} - \overline{GDP_j}|}, & if \ i \neq j \\ 0, & if \ i = j \end{cases} \quad (5)$$

where $\overline{GDP_i}$, $\overline{GDP_j}$ refer to the annual average GDP of province i and j during the research period 2011–2018, respectively. Here, we adopt the row normalization to ensure the rows sum to 1, and their diagonal elements are set to 0.

Data Description of Space Metering Panel

The global Moran's I index is used to examine and measure the spatial autocorrelation and spatial heterogeneity. The equation for calculating global Moran'I is defined as follows:

$$Moran'I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}(Y_i - \bar{Y})(Y_j - \bar{Y})}{S^2 \sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}} \quad (6)$$

among them, $S^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})$, $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, Y_i represents the observed value of the region, n denotes the total number of provinces, and W_{ij} denotes the economic distance weight matrix.

The values of global Moran's I range from −1 to 1. The value of global Moran's I is greater than zero that means there has positive spatial correlation, with value less than zero that means there has negative spatial correlation, with value equal to zero that means there has no spatial correlation.

As shown in **Table 2**, the Morans'I index of each variable was significantly positive, which indicates that health, HI, Bed, Edu, and Doc in China's 31 provinces have spatial autocorrelation. Therefore, it is appropriate for us to adopt the empirical method of spatial measurement.

Model Test

First, the use of two Lagrange multiplier tests (i.e., LM_test for no spatial lag and Robust LM_test for no spatial lag, and LM_test for no spatial error and Robust LM_test for no spatial error) determines whether the spatial lag effect or the spatial error effect is significant (31). If one LM test shows a significant effect whereas the other effect is not significant, this study should adopt the significant form spatial effect model. If the LM test results show that the two effects are significant or are not significant simultaneously, this study should adopt the SDM, and by the Wald or likelihood ratio (LR) test could determine whether the SDM can be simplified into the SLM or SEM. Through the Hausman test and LR joint significance test (i.e., space fixed effect or time fixed effect), this study could determine whether the spatial econometric model should adopt pool fixed effect, space fixed effect, time fixed effect, or space-and-time fixed effect (32).

Because the SDM has jointly captured the influence of spatial lag dependent variable and spatial lag explanatory variables, Equations 2–4 may demonstrate the endogenous problem, which violates the classical assumptions of ordinary least square (OLS) method. Hence, LeSage et al. (33) put forward the maximum likelihood (ML) method to solve the endogenous problem effectively and provided the theoretical framework to estimate or analyze the spatial lag values of both dependent and independent variables (direct and indirect effects). The specific derivation processes are rewritten as:

$$y = (I - \rho W)^{-1}(X\beta + \theta WX + \mu_i + \varepsilon_{it}) \quad (7)$$

$$(I - \rho W)^{-1} = I + \rho W + \rho^2 W^2 + +\rho^3 W^3 + \cdots \quad (8)$$

$$\partial y_i / \partial x_{ir} = (I - \rho W)^{-1}(I\beta_r + (W)_{ii}\theta_r), \quad (9)$$

for all i and for all r,

$$\partial y_i / \partial x_{jr} = (I - \rho W)^{-1}(I\beta_r + (W)_{ii}\theta_r), \quad (10)$$

for all $i \neq j$ and for all r,

where y is the health, ρ denotes the spatial autocorrelation coefficient, W stands for non-negative spatial weight matrix, X represents the independent variables, and β and θ are the spatial regressive coefficients. Wy and WX denote the spatial lag terms of dependent variable and independent variable, which allows us to analyze the spillover effects of independent variables. I represents an $N \times 1$ unit matrix, N is the number of the provinces, $(I - \rho W)^{-1}$ stands for spatial Leontief inverse matrix, $\partial y_i / \partial x_{ir}$ denotes the direct effect, $\partial y_i / \partial x_{jr}$ refers to the indirect effect. indicates the coefficient of the rth independent variable, and θ_r denotes the coefficient of the spatial lag of the rth independent variable. Referring to the methods presented by Elhorst (32), we use MATLAB to estimate the magnitude and sign of the direct and indirect effects in the SDM model.

RESULTS

Before performing the spatial econometric regression analysis, it is necessary to perform Lagrange multiplier (LM) test on the model to further determine whether the model has the spatial correlation. The original assumption of the LM test is that there is no spatial lag term and spatial error term. In **Table 3**, both the LM test and the robust LM test reject the null hypothesis at the significance level of 10%. Therefore, the introduction of space weight matrix is necessary and effective.

To make the model more robust, the Hausman test was performed to determine whether the model should adopt fixed or random effects. Then, the Wald and likelihood ratio (LR) tests were conducted to determine whether the optimal SDM model should degenerate into the SLM or SEM models. The Hausman test results reject the random effects model hypothesis at the 1%

TABLE 2 | Moran's I index test results of core variables.

Year	Health		HI		Doc		Bed		HR	
	Moran's I	Z value	Moran's I	Z value	Moran's I	Z value	Moran's I	Z value	Moran's I	Z value
2011	0.151*	1.579	0.128**	1.933	0.260***	4.144	0.144**	1.629	0.302***	3.106
2012	0.171**	1.753	0.131**	1.980	0.155***	2.562	0.138**	1.488	0.307***	3.193
2013	0.128*	1.380	0.137**	2.055	0.266***	4.250	0.179**	1.834	0.287***	3.184
2014	0.201**	2.006	0.145**	2.157	0.114**	1.990	0.215**	2.120	0.255***	2.923
2015	0.123*	1.331	0.150**	2.217	0.105**	1.892	0.244***	2.352	0.331***	3.407
2016	0.172**	1.750	0.153**	2.261	0.097**	1.755	0.252***	2.422	0.328***	3.388
2017	0.197**	1.954	0.156**	2.300	0.079*	1.515	0.306***	2.877	0.301***	3.108
2018	0.181**	1.832	0.160**	2.346	0.072*	1.454	0.328***	3.068	0.304***	3.141

***, **, * *Represent significance levels of 1, 5, and 10%, respectively.*

TABLE 3 | The result of LM test and Hansman test.

Variable	Statistics	*p*
LM_lag	3.0409*	0.071
LM_error	4.1388**	0.040
Robust LM_lag	3.4522*	0.063
Robust LM_error	4.2861**	0.038
Hausman	60.790***	0.001

***, **, * *Represent significance levels of 1, 5, and 10%, respectively.*

TABLE 4 | Wald test and LR test of spatial panel model.

Model / Test variables	Space fixed effect	Time fixed effect	Space time fixed effect
Wald_spatial_lag	2.1770	18.3616***	19.1156***
	$P = 0.7032$	$P = 0.0048$	$P = 0.0045$
LR_spatial_lag	2.4562	15.4859***	17.1050***
	$P = 0.6525$	$P = 0.0098$	0.0046
Wald_spatial_error	18.8924***	17.2244***	22.0234***
	$P = 0.0028$	$P = 0.0051$	$P = 0.0033$
LR_spatial_error	17.1828***	18.3018***	19.1228***
	$P = 0.0043$	$P = 0.0032$	$P = 0.0034$

*** *Represent the significance levels of 1%.*

significance level. Furthermore, in **Table 4**, both the Wald and LR test results reject the hypothesis that the SDM degenerates at the 1% significance level, which means that both null hypotheses are rejected; Besides, the dynamic SDM model was used to estimate, and the results were not ideal. As a result, the SDM model was selected in this study, and no dynamic models were pursued.

Elhorst believes that the panel data model with spatial lag-dependent variables is more reasonable than the R2 indicator (29). Through comparison of corrected goodness-of-fit (corrected R2) of the models, the values of no fixed effect, spatial fixed effect model, time fixed effect model, and double fixed model are 0.3544, 0.3591, 0.6544, and 0.5307, respectively. The time fixed effect model is obviously larger than those of other three models. The log-likelihood (log-L) values of time fixed effect model are significantly larger than those of other three models. Except LnPGDP variable, the coefficients of LnHI, LnDoc, LnBed, and LnEdu are positive and all significant at 5% significance level. To sum up, this study will select the time fixed effect model as a spatial econometric model to study the impact of health expenditure on health output in China.

From the results of the dynamic SDM regression in column (3) of **Table 5**, we analyze the impact of spatial effects on health. We find that the estimated parameter of LnHI is 0.0660 and significant at 1% significance level. LnHI has played a significant role in promoting health, and it is imperative to increase health expenditure. This also proves that China has been actively engaged in "Healthy China." The coefficient of W*LnHI is 0.0625 and significant, which means that LnHI in local provinces is conducive to improve health in adjacent provinces. At present, LnHI has a strong spatial spillover effect in the geographical adjacent space. The coefficient of LnDoc is negative and significant. The coefficient of W*LnDoc is 0.1324 and significant, which means that LnDoc in local provinces is conducive to improve health in adjacent provinces. It indicates that when emphasis is attached to number of health technicians, per thousand people can achieve the goal of improving health. The coefficient of LnBed is 0.1317 and significant at 1% level. LnBed has played a significant role in promoting health, and it is imperative to increase the number of beds per thousand people. This also proves that China has been actively engaged in "Healthy China." The coefficient of W*LnBed is 0.0966 and significant, which means that LnBed in local provinces is conducive to improve health in adjacent provinces. LnBed has a strong spatial spillover effect in the geographical adjacent space. The coefficient of LnEdu is positive and significant. LnEdu has played a significant role in promoting health in local provinces. The coefficient of W*LnEdu is −0.1152, which is significant at 10% significance level, which means that LnEdu in local provinces is not conducive to improve health in adjacent provinces. To further explore the spatial effects of regression coefficients in the spatial econometric model, we estimate the direct and indirect effects caused by changes in health inputs and their sum effects. **Table 6** shows the decomposition results of the SDM.

TABLE 5 | Setting results of four effects in SDM panel model.

Model / Test variables	1	2	3	4
	No fixed effect	Space fixed effect	Time fixed effect	Double-fixed
LnHI	0.0260	0.0100	0.0660***	0.0620***
	(0.5348)	(0.5317)	(4.1186)	(3.9044)
LnDoc	−0.0229*	−0.0226*	−0.1057***	−0.1048***
	(−1.7325)	(−1.6013)	(−8.6248)	(−8.1047)
LnBed	0.0343*	0.0524*	0.1317***	0.1338**
	(1.9193)	(1.7128)	(7.0003)	(6.9497)
LnEdu	−0.0533*	−0.0369*	0.0260**	0.0318*
	(−1.8437)	(−1.9213)	(1.6946)	(1.9351)
W* LnHI	0.0048	−0.0160	0.0625*	0.0390
	(1.0419)	(−0.9570)	(1.3677)	(0.6509)
W* LnDoc	0.0376	0.0863	0.1324*	0.1027**
	(0.9938)	(1.9092)	(1.8386)	(2.4543)
W* LnBed	−0.0683*	−0.0387	0.0966**	−0.0936
	(1.3488)	(−0.9411)	(1.9616)	(−1.4904)
W* LnEdu	−0.0386	0.1489	−0.1152*	−0.1270**
	(−0.4205)	(1.3573)	(−1.9342)	(−2.0066)
W*dep.var.	0.1199*	0.1339*	0.1340**	0.1099
	(1.5016)	(1.2130)	(1.9292)	(0.9944)
	0.0047	0.0056	0.0366	0.0375
R^2	0.3577	0.9139	0.6577	0.6361
Corrected.R^2	0.3544	0.3591	0.6544	0.5307
LogL	57.8541	303.5458	310.4060	58.9300

***, **, * *Represent significance levels of 1, 5, and 10%, respectively.*

TABLE 6 | Descriptive statistical table of cumulative effect scalars.

Variable	Direct effect	T value	Indirect effect	T value	Total effect	T value
LnHI	0.0660***	4.2275	0.0112*	1.5572	0.0772***	1.9983
LnDoc	−0.1058***	−8.4654	−0.0143*	−1.0119	−0.1201***	−6.0495
LnBed	0.1312***	7.0924	0.0576*	2.5033	0.1888***	8.5785
LnEdu	0.0262*	1.6706	0.0038	0.7846	0.0301*	1.6046

***, * *Represent significance levels of 1% and 10%, respectively.*

LeSage et al. (33) deduced the direct effect of measuring the influence of the change of independent variable on the dependent variable of the neighborhood and the indirect effect of measuring the influence of the change of independent variable on the dependent variable of the neighborhood based on the model's own partial derivative and crosspartial derivative. Elhorst applied the partial differential method to measure direct effect, indirect effect, and total effect (34), with the results shown in **Table 6**.

Overall, the direct effect coefficient of LnHI is 0.0660, which is significant at 1% significance level. The indirect effect coefficient is 0.0112 and significant. Improving LnHI in local regions can improve their health, and it may improve health improvement in adjacent regions. The total effect is positive and significant, which indicates that LnHI does improve health. This shows that health input improves the health level of residents in the region (14–16). Besides, it has a positive demonstration effect on neighboring provinces to increase health input and health output.

The direct and indirect effects of LnDoc on health are −0.1058 and −0.0143, respectively, which are significant at 1 and 10% confidence level. The significance test reflects that LnDoc has significant spatial spillover effects. Direct effect of LnDoc is greater than indirect effect. The total effect is negative and significant, which is consistent with some scholars' conclusion (13, 25). This shows that increasing the number of health personnel in this region does not promote the health output of this region, which may be because economically developed regions will attract the population of surrounding regions to transfer to areas with rich health resources and economically developed regions, which leads to the uneven regional distribution of health human capital in various regions and seriously affects the exertion of the spatial spillover effect of labor force.

The direct effect coefficient of LnBed is 0.1312, which is significant at 1% confidence level. The indirect effect coefficient

is 0.0576, which is significant at 10% confidence level. Improving LnBed promotes health in local and neighboring provinces. The total effect is positive and significant. This phenomenon is not difficult to understand. According to Grossman's model (2), this may be increasing the number of beds which will promote residents' demand for health, so as to improve their health status and improve their health level; besides, it has a positive demonstration effect on neighboring provinces to increase LnBed input.

The direct effect coefficient of LnEdu is 0.0262, which is significant at 10% confidence level. The indirect effect coefficient is 0.0038, which is not significant. Improving LnEdu promotes health in local regions (2, 26). This shows that improving the education level of local residents will play an important role in promoting people to obtain job opportunities, improve the quality of life, form a good lifestyle, and obtain good medical conditions, and the better their health level will be. The total effect is positive and significant. This may be because on the one hand, China's per capita education level is still relatively low compared with developed countries such as Europe and the United States; on the other hand, Chinese college students prefer to work in economically developed cities and coastal areas with high wage levels after graduation, resulting in huge regional differences in education and urban–rural differences.

The spatial autoregressive coefficient on W*dep.var is 0.1340, which is significant at 5% confidence level, which indicates that the spatial lag variable has a positive and significant effect on health. That is, the interaction and radiation effect of health output in the 31 provinces are significant, and the positive spillover effect of places with high health level will spread to the surrounding areas and play a positive role in promoting it. Therefore, Chinese provinces should continue to maintain the momentum of increasing health investment, improve the health level of the people, and then promote the common prosperity of other regions. In addition, health policies should not be formulated in isolation, but should take into account the overall situation of the country and develop in coordination with other provinces.

CONCLUSIONS, RECOMMENDATIONS, AND LIMITATIONS

The above results show that the health output of the 31 provinces in China not only benefits from the local health input, but also benefits from the health input of neighboring provinces. At the same time, due to the significant spatial dependence among different regions, the interaction effect of health output between each province and its neighboring provinces is obvious. Therefore, to give full play to the spatial spillover effect of health input, rationally allocate health resources, and further promote regional economic development, the following suggestions are put forward:

(1) Chinese government should strengthen health collaboration between neighboring provinces and promote the free flow of resources, health professionals, and other elements between provinces. China should strengthen the spillover effect between provinces, give play to the benign interaction of health output, and promote the coordinated development of China's overall health output. Local governments should break the standard ideology, establish the overall awareness, and actively establish the health cooperation system. In the process of implementing health policies, China government should pay attention to overall planning, avoid waste of resources, and take improving China's health output as the primary task of building a healthy China, so as to realize the coordinated development of regional health.

(2) China government should take advantage of the spillover and interprovincial interdependence of health inputs at the provincial level. China should encourage exchanges and cooperation among different regions, actively encourage the flow of health technical personnel, and improve the level of knowledge spillover. At present, China's health resources and health technicians are concentrated in the economically developed coastal areas, whereas there are relatively few in the northwest and border areas. Therefore, attention should be paid to building a central area with diffusion effect, and transregional talent exchange should be used as a link to drive the surrounding areas, so as to realize the balanced development of health levels in different regions. For the coastal economically developed areas, enhance their own radiation effect, the less developed areas should make full use of the spillover effect to make up for the lack of medical technology capacity, save the cost of health research and development, give full play to the advantages of late-comers, and actively connect with the developed areas.

(3) Another important factor that influences knowledge spillover effect is the health policy, so it is necessary to improve the China's health policy environment. It is conducive to the spread of spillover effect of health investment to improve the hard environment of basic supporting service facilities such as hospitals, health centers, and clinics and to establish efficient and reasonable modern medical and health system to narrow the gap of medical technology level between provinces and regions. The most important is to continuously improve the soft environment for health development, accelerate the perfection of relevant laws, improve the open market mechanism, standardize the hospital management system and supervision mechanism, and form a good medical atmosphere.

There are two limitations in the study. First, the models we built are based on the spatial panel data of 31 provinces. But most health management activities take place in cities, whereas the utilization of provincial-level data makes it difficult to capture the spatial heterogeneity of different cities, so the utilization of city-level data is necessary to increase the reliability of estimations. Second, this article simply uses a province's overall health input to measure the impact of health input indicators (HI) on health output, without considering the impact on health output from the perspective of the structural characteristics of health input. In the future, we can measure the impact of health input indicators on health output in terms of government health expenditures, social health expenditures, and personal health expenditures.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This study was funded by the grants from the Ministry of Education of China (Grant 20YJCZH193) and Anhui Provincial Department of Education (Grant SK2018A0203).

ACKNOWLEDGMENTS

The authors thank all the participants who made an effort in this study.

REFERENCES

1. Romer D. *Advanced Macroeconomics*. Singapore: McMGraw MHill Companies Inc. (1996).
2. Grossman M. On the concept of health capital and the demand for health. *J Polit Econ.* (1972) 80:223–55. doi: 10.1086/259880
3. Aydan S, Bayin Donar G, Arikan C. The effects of public health and social spendings on health outcomes: a panel data analysis of OECD countries. *Hacettepe Saglik Idaresi Dergisi.* (2021) 24: 469–482. Available online at: https://dergipark.org.tr/en/pub/hacettepesid/issue/65158/867836
4. Wolfe BL. Health status and medical expenditures: is there a link? *Soc Sci Med.* (1986) 22:993–9. doi: 10.1016/0277-9536(86)90199-1
5. Meng QY. *Health Economics*. Beijing: People's Medical Publishing House (2013).
6. Wang CY, Xiao H. Study on the prediction and influencing factors of the total health expenditure in Hubei Province. *Med Soc.* (2020) 33:11–5. doi: 10.13723/j.yxysh.2020.10.003
7. Musgrove P. *Public and Private Roles in Health: Theory and Financing Patterns*. Washington, DC: World Bank (1996).
8. Filmer D, Pritchett L. *Child Mortality and Public Spending on Health: How Much Does Money Matter?* Washington, DC: World Bank Publications (1997).
9. Fayissa B, Gutema P. The determinants of health status in Sub-Saharan Africa (SSA). *Am Econ.* (2005) 49:60–6. doi: 10.1177/056943450504900207
10. Gupta S, Verhoeven M, Tiongson ER. The effectiveness of government spending on education and health care in developing and transition economies. *Eur J Polit Econ.* (2002) 18:717–37. doi: 10.1016/S0176-2680(02)00116-7
11. Kiross G T, Chojenta C, Barker D, Loxton D. The effects of health expenditure on infant mortality in sub-Saharan Africa: evidence from panel data analysis. *Health Econ Rev.* (2020) 10:1–9. doi: 10.1186/s13561-020-00262-3
12. Shen YQ, Sheng HQ, Ma GF, Ma AN. Health inputs and health outputs in OECD countries: an econometric analysis based on the dynamic panel system GMM model. *Chin J Health Econ.* (2019) 38:64–6.
13. Anyanwu JC, Andrew EO. Health expenditures and health outcomes in Africa. *Afric Dev Rev.* (2009) 21:400–33. doi: 10.1111/j.1467-8268.2009.00215.x
14. Kamiya Y. *Determinants of Health in Developing Countries: Cross-Country Evidence*. Discussion Paper. Osaka: Osaka School of International Public Policy (OSIPP) (2010).
15. Novignon J, Olakojo S A, Nonvignon J. The effects of public and private health care expenditure on health status in sub-Saharan Africa: new evidence from panel data analysis. *Health Econ Rev.* (2012) 2:1–8. doi: 10.1186/2191-1991-2-22
16. Bein MA, Unlucan D, Olowu G, Kalifa W. Healthcare spending and health outcomes: evidence from selected East African countries. *Afric Health Sci.* (2017) 17:247–254. doi: 10.4314/ahs.v17i1.30
17. Raeesi P, Harati-Khalilabad T, Rezapour A, Azari S, Javan-Noughabi J. Effects of private and public health expenditure on health outcomes among countries with different health care systems: 2000 and 2014. *Med J Islamic Rep Iran.* (2018) 32:35. doi: 10.14196/mjiri.32.35
18. Zhang N. Dynamic econometric analysis of the impact of health resource investment on residents' health in Beijing. *Chin J Hosp Manag.* (2018) 34:1051–4.
19. Mao WL, Wei LB. Research on the impact of government health expenditure scale on residents' health. *Zhejiang Soc Sci.* (2020) 03:11–20. doi: 10.14167/j.zjss.2020.03.002
20. Rahman MM, Khanam R, Rahman M. Health care expenditure and health outcome nexus: new evidence from the SAARC-ASEAN region. *Global Health.* (2018) 14:1–11. doi: 10.1186/s12992-018-0430-1
21. Zhao TL, Pang J, Zhao JM. The relationship between total health expenditure and health output in China: an empirical analysis based on vector error correction model. *Chin J Health Manage.* (2020) 37:906–9.
22. World Health Organization. *A Conceptual Framework for Action on the Social Determinants of Health*. Geneva: WHO (2010).
23. Yang M, Lu B. An assessment of mortality and life expectancy for china's provinces: based on the 2010 provincial census data. *Popul Res.* (2019) 43:18–35.
24. Liu XH. Necessity of data forward processing in factor analysis and its software implementation. *J Chongqing Inst Technol.* (2009) 23:152–5.
25. Retzlaff-Roberts D, Chang CF, Rubin RM. Technical efficiency in the use of health care resources: a comparison of OECD countries. *Health Policy.* (2004) 69:55–72. doi: 10.1016/j.healthpol.2003.12.002
26. Shen YQ, Sheng HQ, Ma GF, Ma AN. OECD countries' health input and health output: Econometric Analysis Based on Dynamic Panel System GMM model. *China's Health Econ.* (2019) 07:64–6.
27. Li HM, Wang YQ. Health resource investment and health China Construction: a value-based health system perspective. *China Admin.* (2018) 08:65–9.
28. Shen SG, Zheng QY. Research on health productivity and its influencing factors in China. *J Sun Yat sen Univ.* (2017) 57:153–166. doi: 10.13471/j.cnki.jsysusse
29. Elhorst JP. Matlab software for spatial panels. *Int Regional Sci Rev.* (2012) 37:389–405. doi: 10.1177/0160017612452429
30. Tian Y, Sun CW. Comprehensive carrying capacity, economic growth and the sustainable development of urban areas: a case study of the Yangtze River Economic Belt. *J Clean Prod.* (2018) 195:486–96. doi: 10.1016/j.jclepro.2018.05.262
31. Anselin L. *Spatial Econometrics: Methods and Models* (Vol. 4). Berlin: Springer (1988).
32. Zhao LS, Sun CZ, Liu FC. Interprovincial two-stage water resource utilization efficiency under environmental constraint and spatial spillover effects in China. *J Clean Prod.* (2017) 164:715–25. doi: 10.1016/j.jclepro.2017.06.252
33. LeSage J, Kelley P, Pace R. Spatial econometric models. In: *Handbook of Applied Spatial Analysis*. Berlin, Heidelberg: Springer (2010), 355–376.
34. Li S, Liu Y, Li WJ. Non-performing loan, spatial spillover and regional economic growth. *Rev Econ Manage.* (2018) 34:26–41. doi: 10.13962/j.cnki.37-1486/f.2018.03.003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Public Health

ORIGINAL RESEARCH
published: 06 May 2022
doi: 10.3389/fpubh.2022.847822

The Impact of the Long-Term Care Insurance on the Medical Expenses and Health Status in China

Yao Tang[1], *Tianran Chen*[1], *Yuan Zhao*[2*] and *Farhad Taghizadeh-Hesary*[3,4*]

[1] School of Public Administration, Zhejiang University of Finance and Economics, Hangzhou, China, [2] School of International Business, Southwestern University of Finance and Economics, Chengdu, China, [3] School of Global Studies, Tokai University, Hiratsuka, Japan, [4] TOKAI Research Institute for Environment and Sustainability (TRIES), Hiratsuka, Japan

Based on the panel data of China Health and Retirement Longitudinal Study (CHARLS) in 2011, 2015, and 2018, this paper used the difference-in-difference (DID) method to evaluate the implementation effect how the Long-Term Care Insurance (LTCI) policy impacted on the medical expenses and the health status of the middle-aged and elder population. The empirical results show that LTCI has reduced the outpatient and inpatient quantity by 0.1689 and 0.1093 per year, and cut the outpatient and inpatient expenses by 23.9% and 19.8% per year. Moreover, the implementation of LTCI has improved the self-rated health, the activity of daily living (ADL), as well as the mental health. These conclusions verify the implementation value of LTCI system and provide policy implications for the medical reform and the further LTCI implementation in a larger scale.

Keywords: Long-Term Care Insurance, medical expenses, health status, difference-in-difference method, medical reform

OPEN ACCESS

Edited by:
Mihajlo Jakovljevic,
Hosei University, Japan

Reviewed by:
Narimasa Kumagai,
Seinan Gakuin University, Japan
Krzysztof Kaczmarek,
Medical University of Silesia, Poland

***Correspondence:**
Yuan Zhao
zhaoyuan@swufe.edu.cn
Farhad Taghizadeh-Hesary
farhad@tsc.u-tokai.ac.jp

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *03 January 2022*
Accepted: *21 March 2022*
Published: *06 May 2022*

Citation:
Tang Y, Chen T, Zhao Y and Taghizadeh-Hesary F (2022) The Impact of the Long-Term Care Insurance on the Medical Expenses and Health Status in China. Front. Public Health 10:847822. doi: 10.3389/fpubh.2022.847822

INTRODUCTION

According to the latest data of the seventh national census released by the National Bureau of Statistics of China in May 2021, the proportion of the population aged 60 and over reached 18.70%, and the proportion of those aged 65 and above reached 13.50%. The proportion changes of the elderly demonstrate that the population aging in China has further deepened. In addition, as the average life span of the population increases, China's elderly population is significantly characterized by aging and longevity, correspondingly the number of the disabled elderly is constantly going up as well. According to the *2018–2019 Survey Report on Long-term Care in China*, the proportion of the severely disabled elderly population reached nearly 5%, with a disability rate of 11.8%. More than 40 million disabled, elderly people are in sore need of medical care and at the same time, the rapidly expanding medical expenses have brought heavy burdens not only to elder individuals, but also their families. Given the severe challenges that may come with the gradual increase of the disabled population, Long-Term Care Insurance (LTCI) has been gradually initiated by the governments. From a worldwide perspective, the long-term care insurance development can be traced back to the Social Insurance Law enacted by the Austrian government in 1956, which legislates the right to health care covering a small part of long-term care liability. Later, the LTCI has been implemented gradually in other countries, until Germany has passed the Long-term Care Act in 1995, when the LTCI system was formally proposed as an independent pillar of the welfare system in Germany.

Learning the experience from other countries, the local government of Qingdao City, Shandong Province of China, took the lead in implementing the LTCI system in the cities and towns under

its jurisdiction in 2012, and further expanded the implementation scope of the LTCI system to its rural areas in 2015. The trial LTCI system can be implemented firstly in Qingdao was because the elder population aged above 65 has reached 920 thousand at the end of 2011, accounting 12% and the disable and semi-disable elderly has reached 250 thousand. At that time, the medical security system in Qingdao has already been unable to meet the increasing care need for the aging elderly. Therefore, the nursing services have been separated from the medical services and constituted a part of the trial LTCI system in Qingdao. As to the insured targets, those who have basic medical insurance, no matter the medical insurance for the employed or the general residents in the urban area (extending the scope to the rural area after 2015), were automatically insured with the trial LTCI system. They can make use of the care services of the LTCI system since they have been disabled and in bed for more than 6 months, or completely unable to take care of themselves, or their assessed score of the "Daily Living Ability Assessment" were <60. The LTCI system provides economic compensation for the care services generated during the hospitalization, and provides the integrated care services such as health management, maintenance treatments, basic living care, and functional maintenance (rehabilitation training) etc. through the medical care institutions and nursing centers.

The LTCI treatment and nursing types have been determined according to the disable status. The care services of LTCI system in Qingdao can be mainly categorized as four kinds from the severe disable status to the mild disable status: (1) Special medical care, those who are graded II or above are provided 24-h continuous care services in the designated medical institution; (2) Nursing home care, the institutions integrated treatment and convalesce take the obligations to provide 24-h continuous care services; (3) Home care, door-to-door services provided by the nursing institutions; (4) Patrol care, provided by the nursing institutions (including village clinics) of which the duration and categories are less than the third one. The reimbursement rate of LTCI can be as high as 90% for the insured urban employees (including the retired). For those who have basic resident medical insurance (including rural residents since 2015), the rates are about 70–80% according to the specific care they have accepted. As to the patrol care, the cost such as visiting cost, medical consumable material cost etc. that happens during patrol care are paid by LTCI fund, while the cost such as medicine, examination etc. that happens during hospitalization are still paid by the medical insurance.

After Qingdao being the first city to conduct the trial implementation of the LTCI policy in China, the Ministry of Human Resources and Social Security of the PRC issued the *Guiding Opinions on Piloting the Long-term Care Insurance System* on June 27, 2016, which marks the determination and endeavor of the central government to solve the care problems for the elderly. Since then, another 15 cities across the country has started their trial work of the LTCI system. Several years past, what is the policy effect of the LTCI system? Can it effectively reduce medical expenses of the elderly and their family? Will it have negative effects on the health of the elderly with the reduced medical expenses or have positive effects on their health because the insured elderly patients can access to care services in a better way? In order to answer a series of questions related to the effects of the LTCI policy implementation, a few quantitative studies have been carried out.

By analyzing the long-term care insurance system in the trial cities, scholars concluded that LTCI has lowered the medical expenses and controlled the scale of expenses of the elderly without harming their health (1–3). Using the synthetic control method, Yu et al. (4) found that after implementing the LTCI system, the per capita medical expenses decreased briefly, but later increased. As these studies used the macro data at the municipal level, it was difficult to accurately investigate the long-term impact of LTCI at the micro perspective, for example the medical expenses of elder individuals. Scholars found that the LTCI subsidies for home care significantly reduced their expenses of medical services (5, 6). About every 1 yuan invested in the elderly care system can save about 8.6 yuan of the medical insurance fund (7). These studies contribute the literature on this issues, still there are rare quantitative studies on the effect of trial LTCI policy in China, and the conclusions have been no closer to consensuses. Therefore, based on the 2011, 2015 and 2018 panel data of China Health and Retirement Longitudinal Study (CHARLS), this paper uses the difference-in-difference (DID) method to empirically analyze the effect of the LTCI policy in the experimental cities, exploring whether the original intention of formulating the policy has been achieved and whether it should be implemented nationwide as one of the formal pillars of the social security system of China.

The rest of this paper is structured as follows: As a literature review on LTCI, the Section Literature Review analyzes the impact of LTCI on the medical expenses and the health status of the elderly and proposes the corresponding theoretical hypotheses; Section Data and Recognition Strategies describes the data, samples, model and empirical methods used in this paper; in Section Empirical Results and Interpretation, this paper analyzes the empirical results, conducts the robust test and discusses the mechanism; the Section Conclusions and Policy implications summarizes the whole paper and presents the policy implications for the future implement and medical reform.

LITERATURE REVIEW

Governments hold the cautious opinions on the LTCI system, worrying the possible increasing care expenses which may be not able to afford. In addition, no consistent conclusions have been reached on the impact of the LTCI policy according to currently available researches, especially its impact on the medical expenses. This paper reviews relevant literature from the following two perspectives.

The Impact of LTCI on the Medical Expenses of the Elderly

On the one hand, the implementation of LTCI can reduce medical expenses because of its "substitution effect". Although the development of LTCI are different in different areas, there are many similarities in essence, which basically fall in between

the commercial insurance and social insurance. There are neither pure commercial LTCI nor pure state-borne LTCI. Theoretically, with the long-term care services provided, no matter which design of LTCI, the unnecessary utilization of medical services will reduce (8–10). As the LTCI system responds to the much demand for the nursing and care, the burden caused by excessive medical expenses can be eased (11–13). Gade et al. (14) found the medical care providing to the critically ill patients can reduce their use of medical resources, such as the intensive care units, which help to control the medical cost. With the quasi-experimental and randomized controlled trial evidence, researchers have verified that LTCI in U.S can relieve the medical expense burden of elderly patients to some extent (15, 16). Similar research conclusions were found in Japan and South Korea (17, 18). Jae et al. (19) found that patients with LTCI spent less time in hospitals, thus paid significantly reduced medical fees than those patients without LTCI.

On the other hand, some researchers believe that the implementation of LTCI will increase medical expenses because it can release the demand for medical services, namely the "release effect" (20). Jone et al. (21) compared the effect among the different earning groups and concluded the LTCI lead to the higher initial spending on the medical care among middle-income individuals. It is found that the more convenient provision of the medical apparatus and service, the more care services will be utilized for the elderly people (22). Wooldridge and Schore (23) analyzed how channeling impact on the use of hospital, nursing home, and other medical services, and found that the reductions in hospital use among the treatment group were neither large nor statistically significant. Comparing the medical expenses at different ages of the Czech Republic and the Netherlands, the proportion of medical expenses of the elderly who are 85 years old and above is about 20% in Netherlands. In contrast, due to the lack of LTCI in the Czech Republic, the medical expenses of the population over 85 years old did not increase rapidly (4). All the above researches on LTCI indicate that LTCI can release medical needs, thereby increasing the medical expenses of the elderly. However, it would be easy to control the medical expenses simply by closing hospitals, but that would be worthless (24). Only in the way of cutting medical costs without harming the health status of the population, is consistent with the idea of "value-based health care" (25).

Impact of LTCI on the Health Status of the Elderly

From the macro social perspective, the implementation of LTCI can strengthen the social health welfare because the care services at various levels construct a more complete social security system (26–28). From the micro perspective of elder individuals, research has shown that the degree of care accepted by the elderly is directly proportional to the quality of their longevity and health (29). Based on the experimental results obtained from Spiers et al. (30), cognitive training for the elderly significantly improves their daily activity abilities, which can last for about 5 years. Chen et al. (31) found that the nursing management of the medical institutions and the service attitude of care-givers has been improved with the implementation of LTCI, thus improved the health status of the elderly. Yu and Tseng (32) found LTCI effectively satisfies the nursing demands of the aging society, improves the health status of the elderly, as well as the household caregivers In addition, Fu et al. (33) found the LTCI has a positive effect on labor supply, which had increased after the LTCI system implemented, and had decreased after the LTCI system had reformed. Researchers have realized the critical role of LTCI in improving the health status of elder individuals and their family, members, as well as its positive role in other social field.

DATA AND RECOGNITION STRATEGIES

Data Source

The data used in this paper is extracted from the panel data of CHARLS in 2011, 2015, and 2018, including the micro-data representing families with middle-aged and elderly people over 45 years old. This data is mainly used to analyze the problems of population aging faced by China, which provides a high-quality basic guarantee for the data needed for interdisciplinary research. This paper chooses the CHARLS database for the following reasons: firstly, it is a survey data specifically for the middle-aged and elderly people over 45 years old, which is the target group of the LTCI system, who are potentially taking the care service. Secondly, the LTCI system was first implemented in Qingdao in 2012, and implemented in other 15 cities in 2016. In these 15 cities which implemented the trial LTCI, the data of twelves cities were included in the CHARLS database which are Chengde, Qiqihar, Shanghai, Suzhou, Ningbo, Anqing, Shangrao, Jingmen, Guangzhou, Chongqing, Chengdu, as well as Qingdao. The 2011, 2015, and 2018 CHARLS data contribute to the observation of the situation before and after the implementation of the LTCI policy at different time nodes, which is consistent with the model setting of the DID method. Thirdly, the data covers a series of key variables about medical expenses and the health status of the middle-aged and the elderly, which satisfies the study objects of this article.

Model and Methods

With the pilot programs of LTCI in different cities as the "natural experiments", this paper establishes a DID model, with a specific model setting as follows:

$$Y_{ijt} = \alpha + \beta Trial_{ij} \times Treat_{it} + \delta X_{ijt} + \tau_t + \omega_i + \varepsilon_{ijt} \quad (1)$$

As in formula (1), Y_{ijt} represents the explained variables which this paper emphasized, namely the medical consumption variable (medical expenses and treatment frequency) and health level (self-rated health status and the ability of activities of daily living etc.). i represents the individual, j represents the cities of trial LTCI, t represents time, alternatively the year. $Trial_{ij} \times Treat_{it}$ represents the interaction term of the city variable whether it had implemented the LTCI and the time variable whether the LTCI had been implemented at the year t. The coefficient β is the focus of this study, which is used to observe the policy effect of LTCI. X_{ijt}represents the characteristics of individuals. τ_t and

ω_i represent the time and individual fixed effect respectively. ε_{ijt} represents the random disturbance term.

Variable Descriptions

Firstly, this paper discusses whether the implementation of LTCI can reduce the medical expenses of the elderly. The dependent variables related to medical expenses, are represented by the outpatient expense, the outpatient frequency, the inpatient expense, and the inpatient frequency in the past year (the logarithms of the variables mentioned above are taken). The main explanatory variable is the impact of LTCI policy, which is the interaction term of dummy variables whether it is the trial cities and the year before and after the implementation of the LTCI system, namely $Trial_{ij} \times Treat_{it}$. In this paper, the year 2016 is set as the time benchmark. As to the non-trial cities of LTCI, $Trial_{ij} = 0$; As to the trial cities of LTCI, $\text{Treat}_{it} = 1$ if the samples are extracted from the CHARLS in 2018, because the LTCI has been implemented; otherwise $\text{Treat}_{it} = 0$ (when the samples are extracted from CHARLS in 2011 and 2015). To the special city Qingdao, which has implemented the LTCI system in 2012, $\text{Treat}_{it} = 1$ if the samples are extracted from the urban area of Qingdao in 2015 and 2018 of the CHARLS database. Otherwise $\text{Treat}_{it} = 0$. (Note: Qingdao has further expanded the implementation scope of the LTCI system to its rural areas in 2015, the rural area is not regarded as the treatment group in 2015, because the interval is about only half a year between the data collection and the LTCI implementation).

Secondly, this paper explores whether the implementation of LTCI will improve the health status of the elderly. The dependent variable, which is the health status of the elderly, is represented by the variables, such as the health self-evaluation score, the ADL score and the mental health score (MHS). The health self-evaluation score ranges from 1 to 5, with 1 point indicating extremely unhealthy and 5 points indicating extremely healthy. Covering six activities of daily life, the ADL score ranges from 0 to 40 points. Similarly, the range of the MHS is also from 0 to 40 points, which is calculated through a set of questions testing the depression symptom of the individuals. For both of the ADL score and MHS, the higher the scores, the worse the conditions. In addition, all variables are standardized before putting in the DID regression analysis.

The control variables are selected based on Anderson's health service utilization model and Grossman's theory of health care demand (34, 35). Grossman's theory of health care demand emphasizes that health investment is necessary to stay healthy, which is impacted by the price of the medical resources. The relatively lower cost of the care services compared with their earnings, provoke individuals to obtain more health resources. In addition, according to Anderson (34), the impacting factors of health service utilization mainly include the personal characteristic factors composed of age, gender, education, earnings, subjective and objective health status etc. Therefore, this paper controls the following variables whose definitions are shown in **Table 1**.

Descriptive Statistics

As the paper focused on the changes of the medical services utilization and expenses, and the health condition, the mean difference of the related variables before and after the LTCI implementation have been analyzed. The results are shown in **Tables 1**, **2** respectively.

As demonstrated in the upper part of the **Table 2**, for the trial cities, the medical services utilization and consumption has changed greatly after the implementation of LTCI policy. The outpatient expenses and frequency, and the inpatient expenses and frequency all decreased with the reform. To those cities which have not implemented the LTCI system, as shown in the lower part of the **Table 2**, statistical results demonstrate that the medical expenses have increased, no matter the outpatient or the inpatient expenses, and the frequency of the medical services utilization, especially the inpatient frequency also have increased.

The upper part of the **Table 3** demonstrated the physical and mental health changes with the implementation of the LTCI system. For the trail cities, the ADL and the MHS of the individuals has improved, even not much. The self-rated health has decreased a little, but not significant. As to those cities which have not implemented the LTCI system, the self-rated health has firstly decreased in 2015, and increase later in 2018. In contrast, the scores of the mental health and the ADL have increased, which means the worse physical and mental circumstance as time goes by.

TABLE 1 | Definitions of variables.

Variable	Definition	N	Mean	S.D.
Age	Age of the elderly	42,591	62.3	13.524
Gender	Dummy variable (male = 1; female = 0)	42,591	0.353	0.496
Marriage	Dummy variable (widowed/divorced = 1; others = 0)	42,591	0.219	0.483
Education	Uneducated = 1; primary school = 2; junior high school = 3; senior high school = 4; college and above = 5	42,591	2.885	2.231
Individual-earnings	Earnings over the past year	42,591	3,211	17,435
Self-rated health	Extremely unhealthy = 1; relatively unhealthy = 2; healthy = 3; relatively healthy = 4; extremely healthy = 5	42,591	3.974	1.213
Disease quantity	The total quantity of chronic diseases suffered by the elderly	42,591	18.95	14.72
Physical pain	Yes = 1; no = 0	42,591	0.402	0.527

Source: the charls 2011, 2015, and 2018.

TABLE 2 | The comparison of medical service utilization, 2011, 2015, and 2018.

	Trial cities			
	Before		After	
	Mean	S.D.	Mean	S.D.
Outpatient expense (Yuan)	273.5726	898.6157	189.4393***	1,126.5433
Outpatient frequency	0.3218	1.0914	0.2278**	0.8348
Inpatient expense (Yuan)	1,612.3099	4,893.1747	1,277.6557***	4,099.8214
Inpatient frequency	0.1361	0.4733	0.1026**	0.3744
N	1,327		653	

	Non-trial cities					
	2011		2015		2018	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Outpatient expense (Yuan)	219.0321**	963.7452	247.7403**	1,253.9243	265.6823	1,306.4376
Outpatient frequency	0.3974*	1.0857	0.4283**	1.1695	0.4041	1.0264
Inpatient expense (Yuan)	1,032.8137**	6,866.2960	1,384.0247*	8,857.2057	1,471.0475	5,798.36769
Inpatient Frequency	0.1385*	0.5372	0.1522*	0.4795	0.1563	0.4092
N	14,286		15,130		14,127	

Source: CHARLS 2011, 2015, and 2018.
The trial cities were compared before and after the LTCI system implemented; and the trial cities are Qingdao, Chengde, Qiqihar, Shanghai, Suzhou, Ningbo, Anqing, Shangrao, Jingmen, Guangzhou, Chongqing and Chengdu.
The non-trial cities are compared in adjacent years and marked on the first ones.
, **, and * represent significance at the level of 10, 5, and 1%, respectively.*

TABLE 3 | The comparison of the health status, 2011, 2015, and 2018.

	Trial cities			
	Before		After	
	Mean	S.D.	Mean	S.D.
Self-rated health	3.5313	7.7615	3.5308	6.1343
ADL score	4.3654	11.2351	4.3058*	12.7523
Mental health score	17.7856	90.3986	15.7183***	75.4348
N.	1,327		653	

	Non-trial cities					
	2011		2015		2018	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Self-rated health	3.7452	7.2358	3.5034*	8.243	3.6762	8.7743
ADL score	4.2376	9.4209	4.3458*	11.4875	4.3987	12.7541
Mental health score	16.7052**	73.0985	17.875**	93.7835	18.2423	93.6573
N.	14,286		15,130		14,127	

Source: CHARLS 2011, 2015, and 2018.
The trial cities were compared before and after the LTCI system implemented; and the trial cities are Qingdao, Chengde, Qiqihar, Shanghai, Suzhou, Ningbo, Anqing, Shangrao, Jingmen, Guangzhou, Chongqing and Chengdu.
The non-trial cities are compared in adjacent years and marked on the first ones.
, **, and * represent significance at the level of 10, 5, and 1%, respectively.*

Parallel Trend Test

Before taking the DID regress model, it tested whether the trends before the LTCI implementation is parallel. The core related variables of the medical services utilization and the health conditions in 2011, 2015 and 2018 have been analyzed. The results are

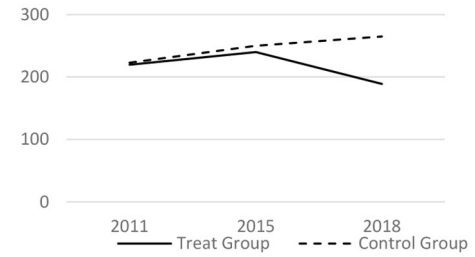

FIGURE 1 | Parallel trend of outpatient expenses (Yuan).

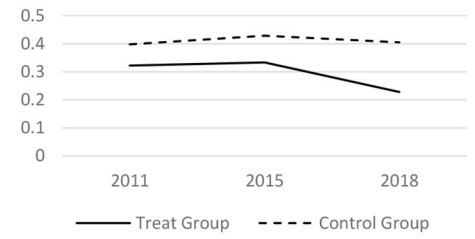

FIGURE 2 | Parallel trend of outexpatient frequency.

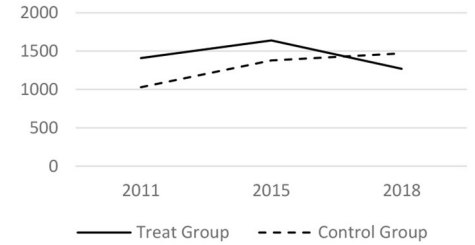

FIGURE 3 | Parallel trend of inexpatient expenses (Yuan).

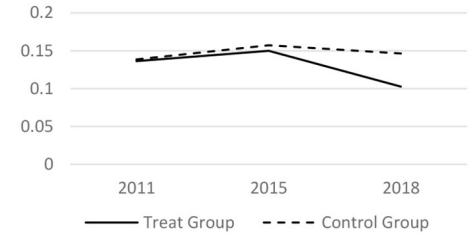

FIGURE 4 | Parallel trend of inexpatient frequency.

shown in **Figures 1**–**6** respectively, which also visually demonstrated how the medical service utilization and health conditions have changed before and after the LTCI implementation.

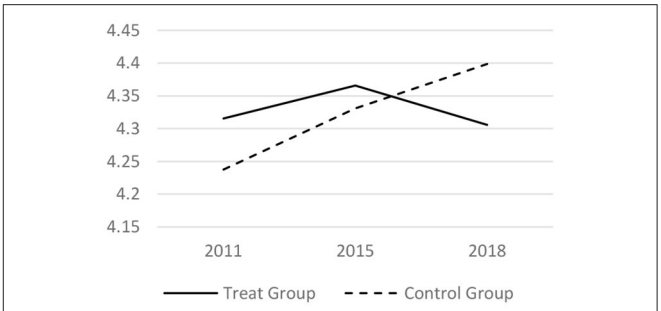

FIGURE 5 | Parallel trend of ADL score.

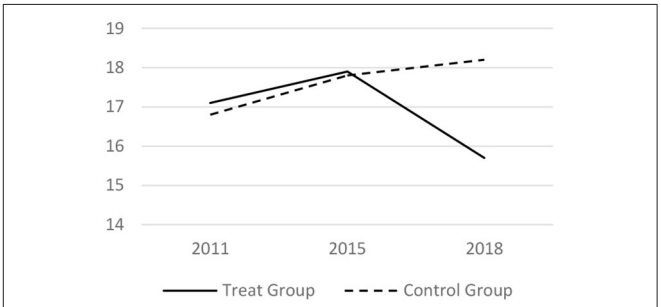

FIGURE 6 | Parallel trend of mental health score.

EMPIRICAL RESULTS AND INTERPRETATION

The Impact of LTCI on Medical Expenses

Regression Results of the DID Model

According to the empirical method introduced above, this paper mixes the CHARLS data in three periods (2011, 2015, and 2018) and identifies whether the implementation of the LTCI policy reduces the expenses and frequency of the medical treatment. **Table 4** shows the empirical results:

As shown in Columns (1) and (2) of the **Table 4**, compared with those in non-trial cities of the LTCI system, the implementation of LTCI has reduced the outpatient expenses by 22.82% and decreased the outpatient quantity by 0.1689 per year. It can be seen from Columns (3) and (4) that the inpatient expenses and quantity in the trial city also have reduced (19.8% and 0.1093 respectively). Moreover, among the controlled variables, the age, education and disease quantity are positively correlated with the outpatient and inpatient medical services. Alternatively, those who are elder and with stronger educational backgrounds, higher earnings and more chronic diseases spend significantly more in outpatient medical services, as the coefficients are 0.0162, 0.0078, 0.0083, and 0.0473 respectively. This conclusion can also be applied to the inpatient expenses. In addition, compared with males, the female have significantly higher inpatient expenses and quantity.

TABLE 4 | DID regression results of the expenses and frequency of the medical services.

Variable	(1) Inoutpatient_expense	(2) Outpatient_frequency	(3) Ininpatient expense	(4) Inpatient_frequency
Did	−0.2282**	−0.1689**	−0.198*	−0.1093**
	(0.1132)	(0.0786)	(0.1057)	(0.0431)
Age	0.0162*	0.0628**	0.0693**	0.0541*
	(0.0097)	(0.0271)	(0.0285)	(0.0312)
Gender	−0.0314	−0.152**	−0.0672*	−0.253***
	(0.0302)	(0.0651)	(0.0399)	(0.0883)
Marriage	−0.0253*	−0.0332*	−0.0376*	−0.0292
	(0.0142)	(0.0181)	(0.0212)	(0.0192)
Education	0.0078*	0.0064	0.0077**	0.0792
	(0.0045)	(0.0042)	(0.0031)	(0.0727)
Income	0.0083**	0.0054**	0.0061*	0.0045*
	(0.0041)	(0.0023)	(0.0035)	(0.0027)
Self–rated health	−0.169*	−0.421***	−0.147*	−0.532***
	(0.1012)	(0.1473)	(0.0821)	(0.1609)
Disease_quantity	0.0473***	0.0242**	0.0524***	0.0367**
	(0.0182)	(0.0113)	(0.0171)	(0.0154)
_Cons	0.402*	−0.0592**	0.0221*	−0.723**
	(0.0653)	(0.0191)	(0.0046)	(0.0924)
N	38728	38728	38728	38728

Source: charls 2011, 2015, and 2018.
The number in the first row of each column is the coefficient and inside the brackets is the standard error clustering to the community level.
, **, and * represent significance at the level of 10, 5, and 1%, respectively.*
The regression has controlled the entity fixed effects and time fixed effects.

TABLE 5 | DID results with reselection of the dependent variable of reimbursement.

Variable	(1) Inout–patient reimbursement	(2) Inin–patient reimbursement
Did	−0.2461**	−0.2617**
	(0.0301)	(0.0265)
Age	−0.0819*	0.0474*
	(0.0423)	(0.0252)
Gender	0.1926*	0.1865
	(0.0726)	(0.0843)
Marriage	−0.1206**	−0.0224
	(0.0419)	(0.0479)
Education	0.0187*	0.0068*
	(0.0098)	(0.0037)
Income	−0.0059*	−0.0074***
	(0.0031)	(0.0028)
Self–rated health	−0.1843***	−0.3082***
	(0.0538)	(0.1164)
Disease_quantity	−0.0642***	0.0442*
	(0.0236)	(0.0243)
_Cons	2.1057***	1.4133***
	(0.0606)	(0.0148)
N	34,948	34,948

Source: charls 2011, 2015, and 2018.
The number in the first row of each column is the coefficient and inside the brackets is the standard error clustering to the community level.
, **, and * represent significance at the level of 10, 5, and 1%, respectively.*
The regression has controlled the entity fixed effects and time fixed effects.

Robustness Test

Robust Test With Reselection of the Dependent Variable

The medical expenses and treatment quantities were used as explanatory variables to study the impact of LTCI system on the medical utilization in trial cities. To verify the robustness of previous conclusions, the related variables such as the reimbursement for the inpatient and outpatient expenses have been taken to replace the explanatory variables mentioned above. As the inpatient and outpatient expenses are not fully covered by the medical insurance, the reimbursement can be used to the check whether the burden of medical consumption has changed before and after the LTCI implementation. The results are demonstrated in **Table 5**, where the regression coefficients of DID in Column (1) and Column (2) are significantly negative (−0.2461 and −0.2617), indicating that the implementation of LTCI makes the reimbursement of the medical insurance decreased. This is basically consistent with the previous regression results, implying the robustness of the above model. (Note: considering the outliers, sample tail reduction treatment (at top 1% level) have been done and then redo the DID with the same control variables, the basic conclusions did not change).

Robust Test With Reselection of the Control Group

The control group analyzed above is composed of all cities that have not implemented LTCI. To further verify the robustness of previous conclusions, we select the second trial cities of LTCI system to form a new control group. Alternatively, Chengde, Guangzhou, Shanghai, Suzhou, Anqing, Shangrao, Jingmen, Chongqing, Ningbo, Qiqihar and Chengdu, were selected from

TABLE 6 | DID results with the second trial tier 11 cities as the control group.

Variable	(1) lnoutpatient_expense	(2) Outpatient_frequency	(3) lninpatient expense	(4) lnpatient_frequency
Did	−0.1935***	−0.0169***	−0.2297***	−0.0594***
	(0.0272)	(0.0058)	(0.0339)	(0.0158)
Age	0.0313***	0.0353***	0.0468**	0.0311
	(0.0119)	(0.0132)	(0.0238)	(0.0218)
Gender	−0.0316**	−0.0725**	0.0823*	−0.142**
	(0.0126)	(0.0302)	(0.0497)	(0.0649)
Marriage	−0.0283*	−0.0373*	−0.0411*	−0.295**
	(0.0149)	(0.0206)	(0.0231)	(0.1114)
Education	0.0057	0.0027	0.0056	0.0643
	(0.0083)	(0.0021)	(0.0054)	(0.0398)
Income	−0.0023	−0.0015	−0.0057*	−0.0036
	(0.0016)	(0.0029)	(0.0032)	(0.0053)
Self–rated health	−0.141***	−0.4038***	−0.1295***	−0.5873***
	(0.0105)	(0.0247)	(0.0322)	(0.0357)
Disease_quantity	0.0023*	−0.0035	0.0019*	−0.0026**
	(0.0014)	(0.0026)	(0.0011)	(0.0012)
_Cons	0.0328**	−0.0654***	0.0416***	−0.0877
	(0.0176)	(0.0124)	(0.0031)	(0.1093)
N	2,367	2,367	2,367	2,367

Source: charls 2011, 2015, and 2018.
The number in the first row of each column is the coefficient and inside the brackets is the standard error clustering to the community level.
, **, and * represent significance at the level of 10, 5, and 1%, respectively;*
The regression has controlled the entity fixed effects and time fixed effects.

the original control group to constitute a new control group. These 11 cities did not implement LTCI from 2011 to 2015, but they were selected as the trial cities after 2016. These cities may have similar characteristics with Qingdao in terms of the medical service utilization etc., so that they can be chosen to implement the LTCI policy as the second trial tier. Hence, this paper screens the control group to further test the impact of LTCI on the medical expenses and medical treatment quantity, the results of which are shown in **Table 6**.

As shown in Columns (1)–(4) of **Table 6**, the regression coefficients of DID are significantly negative (−0.1935, −0.0169, −0.229, and −0.0594 respectively). The regression results are basically consistent with conclusions obtained above which further indicates the robustness of DID model and strengthens the reliability of the conclusion that LTCI can relieve the pressure of the medical services and resources.

The Impact of LTCI on the Health Status

The health status is represented by the variables of the health self-evaluation score, the ADL score and the MHS. The health self-evaluation score ranges from 1 to 5, the higher the scores, the better the health status. For both of the ADL and MHS, their scores range from 0 to 40 points, which are calculated through a set of related questions. The higher the scores, the worse the conditions. With the DID model, the specific empirical results are shown in **Table 7**.

As shown in Column (1) of **Table 7**, the regression coefficient of DID is positive (0.1369), which indicates that the self-rated health scores in trial cities are higher due to the implementation of LTCI. When it comes to the ADL score and MHS, both coefficients are negative, −0.0477 and −0.0673 respectively. As the higher the scores, the poorer the circumstances, the negative coefficients indicate that the mental health and ability of daily life have improved with implementation of LTCI. Among the control variables, the elder population, the lower health self-rated and ADL scores. Those who are well educated, married and have higher income tend to have significantly better physical and mental health. Alternatively, their scores of the ADL and mental health are lower.

Robust Test

Robust Test With Reselection of the Dependent Variable

To verify the robustness of the conclusions achieved above, the related variables such as "whether the body is in pain" as the proxy variable to replace the explanatory variables mentioned above. The results are demonstrated in **Table 8**, where the regression coefficient of DID is significantly negative (−0.1242), indicating that the implementation of LTCI help relieve the physical pain in trial cities. This is basically consistent with the previous regression results, implying the robustness of the above model.

Robust Test With Reselection of the Control Group

To verify the robustness of previous conclusions that LTCI system has improved the health status of the population, we select the second trial cities of LTCI to constitute a new control group.

TABLE 7 | DID regression results of the health status.

Variable	(1) Self–rated health	(2) ADL score	(3) MHS
Did	0.1369**	−0.0477***	−0.0673***
	(0.0581)	(0.0249)	(0.0239)
Age	−0.0056	−0.0137***	0.1584**
	(0.0045)	(0.0041)	(0.0668)
Gender	−0.0302	−0.2156	−0.2436
	(0.0268)	(0.1544)	(0.1736)
Marriage	0.0536***	−0.0274***	−0.3529***
	(0.0143)	(0.0094)	(0.0931)
Education	0.0029	−0.0049**	−0.0135*
	(0.0023)	(0.0021)	(0.0075)
Income	0.0262	−0.0639***	−0.3184**
	(0.0492)	(0.0226)	(0.1555)
Cons	0.3173***	−0.292***	−0.383***
	(0.0528)	(0.0936)	(0.1173)
N	38,728	38,728	38,728

Source: charls 2011, 2015, and 2018.
The number in the first row of each column is the coefficient and inside the brackets is the standard error clustering to the community level.
*, **, and *** represent significance at the level of 10, 5, and 1%, respectively.*
The regression has controlled the entity fixed effects and time fixed effects.

TABLE 8 | DID results with reselection of the dependent variable of pain.

Variable	(1) Pain
Did	−0.1242***
	(0.0248)
Age	0.0217*
	(0.0112)
Gender	−0.171**
	(0.0763)
Marriage	−0.0391***
	(0.0132)
Education	0.0011
	(0.0035)
Income	0.0156
	(0.0501)
Cons	−1.825***
	(0.1772)
N	32,985

Source: CHARLS 2011, 2015 and 2018.
The number in the first row of each column is the coefficient and inside the brackets is the standard error clustering to the community level.
*, **, and *** represent significance at the level of 10, 5, and 1%, respectively.*
The regression has controlled the entity fixed effects and time fixed effects.

Alternatively, the sample of 11 trial cities in 2011 and 2015 is the control group. The method is similar with the test on the medical service utilization and the results are shown in **Table 9**.

The basic conclusions keep the same as seen from the first line of the column (1), (2) and (3) in **Table 9**. The coefficient of DID is 0.473, which is positive with the dependent variable of self-rated health, indicating that the health status has improved with the implementation of LTCI system in the trial cities. As

TABLE 9 | DID results with the second trial tier 11 cities as the control group.

Variable	(1) Self–rated health	(2) ADL score	(3) MHS
Did	0.3473***	−0.2012***	−0.2361***
	(0.1275)	(0.0719)	(0.0634)
Age	−0.0064**	0.0338	0.0153
	(0.0031)	(0.0797)	(0.0832)
Gender	−0.0289	−0.0297	−0.0328
	(0.0225)	(0.0165)	(0.0732)
Marriage	0.0643***	−0.0217*	−0.0593*
	(0.0132)	(0.0129)	(0.0355)
Education	0.0081	−0.0049	−0.0084
	(0.0055)	(0.0042)	(0.0057)
Income	0.0226	−0.0542**	−0.0628
	(0.0692)	(0.0183)	(0.0846)
Cons	4.613***	−0.3362***	−0.5817***
	(0.4896)	(0.0927)	(0.1137)
N	2,335	2,335	2,335

Source:charls 2011, 2015, and 2018.
The number in the first row of each column is the coefficient and inside the brackets is the standard error clustering to the community level.
*, **, and *** represent significance at the level of 10, 5, and 1%, respectively.*
The regression has controlled the entity fixed effects and time fixed effects.

to the ADL score and MHS, both coefficients are negative (-0.0231 and−0.0226). The negative coefficients indicate that the mental health and ability of daily life have heightened with implementation of LTCI. All these conclusions are consistent with the previous DID regression results.

CONCLUSIONS AND POLICY IMPLICATIONS

China has entered the aging society since 2000 and has gradually become the country which has the largest aging population in the world. The health care demand for the elderly is estimated to a high level which will break out massive, crises by 2030 as the number of disabled elderly people keep increasing. Against such a backdrop, the central government of China has tentatively implemented the Long-Term Care Insurance System in 15 cities trying to solve the elderly care problems. This paper evaluate the implementation effect of the LTCI policy, basing on the panel data of CHARLS in 2011, 2015, and 2018. Statistics analysis demonstrated that the policy intentions on the medical utilization and the health status have been basically achieved. The empirical results from the DID analysis showed that: on the one hand, as to the medical expenses and utilization, the outpatient and inpatient frequency and expenses has decreased with the implementation of LTCI; on the other hand, the self-rated health, the ADL, especially the MHS has significantly improved in the trial cities, and these conclusions are valid under a series of robust tests.

The implementation of LTCI system separated the nursing services from the medical services, making the medical funds play a bigger role and better health outcomes. These findings

provide a useful reference for further deepening medical reform. Firstly, the medical service utilization, such as the hospitalization, is one of the determinant factors on the health and longevity of the population. China should continue to focus on improving the construction of the LTCI system, and at the same time, to reduce the unnecessary excessive medical treatments. In addition, it is necessary to integrate the graded diagnosis and treatment with the LTCI system, making the medical and nursing system mutually matched, as well as in line with China's national condition.

Secondly, the traditional Chinese medical philosophy thinks highly of that the best way to cure a disease is to cure it before it become a disease. Since keeping health is the fundamental goal of the medical system reform, and the regularly patrol care and health examinations can help maintain their physical and mental health, the medical examination services and health education activities needed to be expanded, especially to those patients with chronic diseases. Last but not least, in order to strengthen the health management of the middle-aged and elder population, professional care service practitioners are indispensable. Faced with a huge shortfall in care service practitioners, training mechanism of the care service team to improve their professional service skills are also important. In sum, to construct the LTCI system is just the first step to deal with the huge need of elderly care services with the increasing expanding aging population.

DATA AVAILABILITY STATEMENT

The CHARLS databases are public and open to researchers who has applied for them. The data can be applied at: http://charls.pku.edu.cn/, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

YT and TC: conceptualization, methodology, and writing—original draft. YT and YZ: formal analysis. FT-H: investigation and validation. YZ and TC: software. TC and FT-H: writing–review and editing. All authors have read and agreed to the published version of the manuscript.

FUNDING

The research was funded by the National Social Science Foundation of China, titled: The Cost Estimation and the Policy Research on the Informal Care for the Disabled Elderly in Rural China (No: 19CSH071) and Zhejiang Provincial Natural Science Foundation Project, titled: Study on the Demand and the Supply of Elderly Care Services and its Multiple Effects in Zhejiang Province (No: LQ20G030021), and the Grant in Aid for Excellent Young Researcher of the Ministry of Education of Japan.

REFERENCES

1. Zhang C, Giles J and Zhao Y. Evaluation of the effect of new rural social endowment insurance policy: income, poverty, consumption, subjective welfare and labor supply. *Econ Q.* (2015) 14:203–30.
2. Feng J, Yu Y and Lou P. Medical demand and the growth of medical cost in china: based on the difference of medical cost between urban and rural elderly. *Soc Sci China.* (2015) 85–103.
3. Gong X, Zhou W. Government subsidies, insurance payments and end-of-life care costs: an analysis based on CLHLS data from 2002 to 2014. *South Econ J.* (2018) 2018:68–85.
4. Yu X, Liu H, Yan W. The impact of long-term care insurance on medical costs. *Insur Res.* (2019) 2019:114–27.
5. Li L, Yang Y. the impact of population aging on health care costs: a case study of Beijing. *Soc Secur Res.* (2017) 27–39.
6. Yao H. Comparison reflection of china's long-term care insurance system under the background of aging crisis. *Soc Secur Res.* (2020) 2020:48–56.
7. Wang Z, Feng J. The Effect of long-term care insurance on medical cost substitution and the comparison of different compensation models. *Econ Q.* (2021) 21:557–76.
8. Bei L, Yan MH, Lim GJ, Feng G. Substitution effect of long-term care to hospital inpatient care?. *China Econ Rev.* (2020) 2209–26.
9. Costa-Font J, Jiménez-Martín S, Cristina V. Does Long-term Care Subsidization Reduce Hospital Admissions and Utilization? *J Health Econ.* (2018) 58:43–66. doi: 10.1016/j.jhealeco.2018.01.002
10. Dong J, Smieliauskas F, Konetzka RT. *Effects of Long-Term Care Insurance on Financial Well-Being.* (2015). Available online at: https://ssrn.com/abstract=2719776 (accessed December 21, 2015).
11. Johri M, Beland F, Bergman H. International experiments in integrated care for the elderly: a synthesis of the evidence. *Int J Geriatr Psychiatry.* (2003) 18:108–21. doi: 10.1002/gps.819
12. Kattenberg M, Bakx P. Substitute services: a barrier to controlling long-term care expenditures. *Eur J Ageing.* (2020) 2020:85–97. doi: 10.1007/s10433-020-00570-x
13. Forte PE. The American long term care insurance program: a solution to reduce cost and provide stability. *Generations.* (2017) 40:40–53. Available online at: https://chinesesites.library.ingentaconnect.com/content/asag/gen/2017/00000040/00000004/art00008
14. Gade G, Venohr I, Conner D, McGrady K, Beane J, Richardson R H, et al. Impact of an inpatient palliative care team: a randomized control trial. *J Palliat Med.* (2018) 11:180–90. doi: 10.1089/jpm.2007.0055
15. Penrod JD, Deb P, Dellenbaugh C, Burgess JF, Zhu CWL, Christiansen CL, et al. Hospital-based palliative care consultation: effects on hospital cost. *J Palliat Med.* (2010) 13:210–23. doi: 10.1089/jpm.2010.0038
16. McCarthy IM, Robinson C, Huq S, Philastre M, Robert FL. Cost savings from palliative care teams and guidance for a financially viable palliative care program. *Health Serv Res.* (2015) 50:217–36. doi: 10.1111/1475-6773.12203
17. Zhao C, Sun J. The reform of long-term care insurance system in japan and its enlightenment. *Populat J.* (2018) 40:79–89.
18. Hashimoto HH. Micro data analysis of medical and long-term care utilization among the elderly in Japan. *Int J Environ Res Public Health.* (2010) 7:3022–37. doi: 10.3390/ijerph7083022
19. Jae CW, Park E, Lee SGP. Does long-term care insurance reduce the burden of medical costs? a retrospective elderly cohort study. *Geriatr Gerontol Int.* (2018) 18:207–19. doi: 10.1111/ggi.13536
20. Arrizumi H. Effect of public long-term care insurance on consumption, medical care demand, and welfare. *J Health Econ.* (2008) 27:1423–35. doi: 10.1016/j.jhealeco.2008.07.008
21. Jone CR, Done N, Gerard F. Anderson. Considering Long-term Care Insurance for Middle-income Countries: Comparing South Korea with Japan and Germany. *Health Policy.* (2015) 119:1319–29. doi: 10.1016/j.healthpol.2015.06.001
22. Motel-Klingebiel A, Tesch-Roemer C, Von Kondratowitz HJ. Welfare states do not crowd out the family: evidence for mixed responsibility from comparative analyses. *Aging and Society.* (2015) 25:863–82. doi: 10.1017/S0144686X05003971
23. Wooldridge J, Schore J. The effect of channeling on the use of nursing homes, hospitals, and other medical services. *Health Serv Res.* (2018) 23:78—91. Available online at: https://pubmed.ncbi.nlm.nih.gov/3130323/

24. Einav L, Finkelstein A, Mahoney N. Provider incentives and healthcare costs: evidence from long-term care hospitals. *Econometrica*. (2018) 86:2161–219. doi: 10.3982/ECTA15022
25. Porter E. What is value in health care? *N Engl J Med*. (2010) 363:2477–81. doi: 10.1056/NEJMp1011024
26. Leena F, Mari A, Jutta P, Jani R, Pekka R, Marja J. Long-term care is increasingly concentrated in the last years of life: a change from 2000 to 2011. *Eur J Public Health*. (2017) 49–61. doi: 10.1093/eurpub/ckw260
27. Li F, Otani J. Financing elderly people's long-term care needs: evidence from China. *Int J Health Plan Manag*. (2018) 33:479–88. doi: 10.1002/hpm.2488
28. Zhu M. The analysis of demand for long term care and its insurance system constructing in China. *Chinese Journal of Health Policy*. (2009) 35–44. Available online at: http://en.cnki.com.cn/Article_en/CJFDTOTAL-ZGWZ200907009.htm
29. Klimaviciute J, Pestieau P. Long term care social insurance: how to avoid big losses? *Int Tax Publ Fin*. (2018) 99–139. doi: 10.1007/s10797-017-9445-4
30. Spiers G, Matthew FE, Moffatt S, Barker RO. Impact of social care supply on healthcare utilisation by older adults: a systematic review and meta-analysis. *Age Ageing*. (2018) 2018:107–21. https://doi.org/10.1093/ageing/afy14710.1093/ageing/afy147
31. Chen L, Zhang X, Xu X. Health insurance and long-term care services for the disabled elderly in China: based on CHARLS data. *Risk Healthc Policy*. (2020) 2020:155–62. doi: 10.2147/RMHP.S233949
32. Tsu-Wei Y, Tseng L. Examining the role of commercial long-term care insurance in long-term care services. *Int J Bank Market*. (2019) 37:565–78. doi: 10.1108/IJBM-01-2018-0007
33. Fu RN. Spillover effect of japanese long-term care insurance as an employment promotion policy for family caregivers. *J Health Econ*. (2017) 56:103–12. doi: 10.1016/j.jhealeco.2017.09.011
34. Anderson R. *A Behavioral Model of Families' Use of Health Services. Research Series No. 25*. Chicago, IL: Center for Health Administration Studies; University of Chicago (1968).
35. Grosman M. *The Demand for Health: A Theoretical and Empirical Investigation*. New York, NY: National Bureau of Economic Research. (1972).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Public Health

ORIGINAL RESEARCH
published: 18 May 2022
doi: 10.3389/fpubh.2022.773728

Fiscal Decentralization, Public Health Expenditure and Public Health–Evidence From China

Wangzi Xu[1*†] *and Jia Lin*[2†]

[1] School of Public Health, Xiamen University, Xiamen, China, [2] International School, Jinan University, Guangzhou, China

Since the beginning of the COVID-19 outbreak and the launch of the "Healthy China 2030" strategy in 2019, public health has become a relevant topic of discussion both within and outside China. The provision of public health services, which is determined by public health expenditure, is critical to the regional public health sector. Fiscal decentralization provides local governments with more financial freedom, which may result in changes to public health spending; thus, fiscal decentralization may influence public health at the regional level. In order to study the effects of fiscal decentralization on local public health expenditure and local public health levels, we applied a two-way fixed effect model as well as threshold regression and intermediate effect models to 2008–2019 panel data from China's 30 mainland provinces as well as from four municipalities and autonomous regions to study the effects of fiscal decentralization on public health. The study found that fiscal decentralization has a positive effect on increasing public health expenditure. Moreover, fiscal decentralization can promote improvements in regional public health by increasing public health expenditure and by improving the availability of regional medical public service resources. In addition, fiscal decentralization has a non-linear effect on public health.

Keywords: public health expenditure, public health, fiscal decentralization, intermediary effect test, threshold regression analysis

OPEN ACCESS

Edited by:
Mihajlo Jakovljevic,
Hosei University, Japan

Reviewed by:
Mariela Deliverska,
Medical University Sofia, Bulgaria
Naiefa Rashied,
University of Johannesburg, South Africa
Morro Touray,
University of Surrey, United Kingdom

***Correspondence:**
Wangzi Xu
xuwangzi1016@163.com

†These authors share first authorship

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *10 September 2021*
Accepted: *25 April 2022*
Published: *18 May 2022*

Citation:
Xu W and Lin J (2022) Fiscal Decentralization, Public Health Expenditure and Public Health–Evidence From China. Front. Public Health 10:773728. doi: 10.3389/fpubh.2022.773728

INTRODUCTION

Since the beginning of the COVID-19 outbreak, China's public health system has become an essential contributor to the protection of public health and to the promotion of economic and social development. Public health is valued by society and the government because it is underpinned by individual health. In 2020, the Health Committee of China received an 11.3% increment in its budget allocation from state finances relative to what it received in 2019. China's "Healthy China 2030" strategy, which was launched in 2016, emphasized health as an inevitable requirement for human survival and development as well as the foundation for economic development (1). It also regards health as a significant symbol of national prosperity and economic rejuvenation. China's president, Xi Jinping, also attaches great importance to public health and places the protection of individual health as a development priority. He promoted the reform of China's health system, increased the amount of reimbursement received from health insurance, and ensured access to affordable and effective medical treatment (2). Improving the public health care system, increasing government health expenditure, and promoting public health are essential to ensuring stable and sustainable economic development. In this public health and improvement drive, China introduced

fiscal decentralization programs to foster resources that are efficiently managed to ensure the delivery of effective public health care.

A fiscal decentralization system involves shifting revenue and expenditure to lower levels of government, something that is currently applied in the Chinese context (3) China's fiscal decentralization began in the 1980s (4). Lin and Zhou (5), Akai and Sakata (6), and Zhang and Zou (7) have studied the effects of different aspects of fiscal decentralization on economic growth and environmental pollution (8). However, studies on fiscal decentralization in the context of public health care are limited. The classical theory of fiscal decentralization is based on the tenet that local governments know more about their local situation than the central government does (9). Therefore, under a decentralized system, local governments can better provide public goods in an effective, efficient, and timely manner. However, some studies present different views (10). For example, Fu and Zhang (11) believe that Chinese decentralization distorts the public expenditure structure of local governments and that as fiscal decentralization increases, most local governments focus on effective capital construction investment while ignoring investment in education, culture and health. In addition, local corruption has increased government investment expenditure and reduced expenditure targeting livelihoods such as education and health care (12). Thus, in the fiscal expenditure structure, economic public goods crowd out non-economic livelihood-related public goods, such as public health services.

Considering the possible non-linear relationship, the introduction of a quadratic term and the use of threshold effect regression and other non-linear regression models can capture the possible structural changes that take place in the model and the non-linear relationship between variables more appropriately. Therefore, this paper adopts a non-linear regression model when analyzing the impacts of fiscal decentralization on public health. The main aim of this paper is to analyze the impacts of fiscal decentralization on public health. We employed a two-way fixed effect model as well as threshold regression and intermediate effect models for the analyses. First, using 2008–2019 panel data from 30 provinces in China, the paper examines the effects of fiscal decentralization on public health expenditure and public health outcomes. Second, after establishing that the impact of fiscal decentralization on public health is non-linear through introduction of a quadratic term [as in Wang and Wang (13)], this paper further utilizes a threshold regression model [as in Li et al. (14)] to study the non-linear relationship between fiscal decentralization and public health. Finally, with respect to policy recommendations for the government, the paper also discusses the influence mechanism through an intermediary effect test.

MATERIALS AND METHODS

Variable Selection

Explanatory and Response Variables

Scholars around the world have different opinions on how to measure fiscal decentralization (15, 16). This paper uses fiscal revenue decentralization (FD) and fiscal expenditure decentralization (FED) to represent the level of local fiscal decentralization (17). The calculation formula is presented below in **Table 1**.

In addition, the core of Chinese fiscal decentralization is the distribution of tax revenue between the central and local governments, which is essentially fiscal revenue decentralization. Therefore, this paper takes FD as the core explanatory variable and uses FED as a robustness test to ensure the robustness of the regression results. **Figure 1** illustrates the annual changes in the degree of average FD in China in the period of 2008–2019.

The response variable for public health is a multi-dimensional and macro concept. Some literature and government reports use life expectancy per capita, infant mortality, and maternal mortality to measure public health interventions (18), but the most frequently used indicator in scientific research is population mortality. Therefore, this paper uses population mortality as the measure of public health interventions. Thus, the lower the population mortality, the higher the level of public health interventions. **Figure 2** shows the annual average population mortality in China from 2008 to 2019.

Public health expenditure is another response variable as well as a vital intermediary variable. The data were obtained directly from each province's statistical year book. **Figure 3** demonstrates the trend of China's average annual total health expenditure from 2008 to 2019. In fact, due to the increase in outpatient costs, rising drug prices, and more individualized medical care, health spending has continued to grow at an alarming rate worldwide (19). Hence, it can be seen from **Figure 3** that China's average health expenditure is continuously increasing.

The trend observed in the three figures above suggests that there may be a correlation between the degree of fiscal decentralization, public health, and public health expenditure.

Variable Definition

Table 1 provides the definitions and calculation methods of all variables in this paper.

The control variables are variables other than explanatory variables that may affect population health. Economic development, scientific and technological advancement, degree of urbanization, international trade and air pollution are closely related to public health improvements (20–22). The number of hospital beds and the density and distribution health workers are important embodiments of the resource level of the public healthcare system (23) and are confounding factors of public health. Public health spending also directly affects the public health level (24); thus, the mediating variables are Bed, Tech, and PS.

Data Sources

Due to the lack of some data, the regions of Taiwan and Tibet have been excluded. All data are sourced from the China Statistical Year book, the Finance Year book of China and the China Statistical Year book on Environment. The descriptive statistics of all of the variables are shown in **Table 2**.

The data set covers the period 2008 to 2019 (inclusive). Since this paper includes 30 Chinese provinces, the use of panel data can address endogeneity problems caused by unobservable

TABLE 1 | Variables of interest.

Classification	Variables	Variable definitions	Outcome measure /Computing method
Response	PH	Public health	Population mortality
variables	PS	Public health expenditure	Logarithm of public health expenditure in each province
Explanatory	FD	Fiscal Revenue Decentralization	Per capita provincial fiscal revenue/(per capita provincial fiscal revenue + per capita central fiscal revenue)
variables	FED	Fiscal expenditure decentralization	Per capita provincial fiscal expenditure/(per capita provincial fiscal expenditure + per capita central fiscal expenditure)
	lnGDPave	Real per capita GDP	The logarithm of nominal per capita GDP multiplied by GDP index divided by 100
	Industry	Industrial structure	Added value of secondary industry/added value of tertiary industry
Control	Patent	Scientific/technological level	Regional authorized patents
Variables	Market	Marketization index	2008–2016 from the report, 2017–2019 forecasted by trends
	Trade	Import and export trade	Total imports and exports/nominal GDP
	Pop	Resident population	The logarithm of permanent residents in each region
	Urban Pollution	Urbanization rate aAir pollution	Urban population/resident population Industrial SO2 emissions take logarithm
Mediating	Bed	Number of beds	Number of beds in regional medical institutions
Variables	Tech	Hygienic personnel	Number of health workers per thousand people

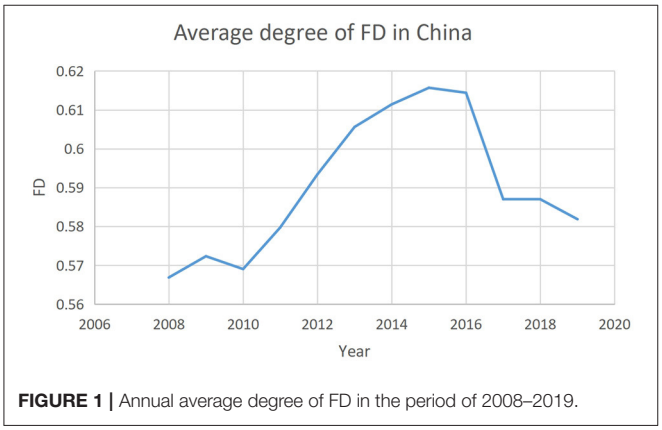

FIGURE 1 | Annual average degree of FD in the period of 2008–2019.

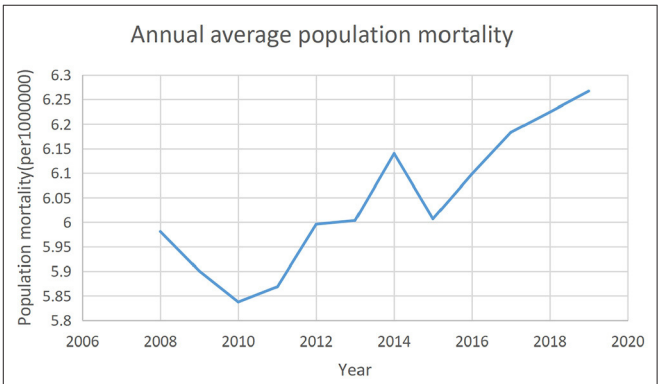

FIGURE 2 | Annual average population mortality in China in the period of 2008–2019.

heterogeneity, because the fixed effect model using panel data can control the time-invariant missing variables that change with individuals. Moreover, panel data provide the characteristics of each province's respective regional dynamic behaviors, making the results more buoyant (25).

Model Development

Benchmark Regression Models

Three models were developed for the benchmark regression analysis. Model 1 is used to test the impact of fiscal decentralization on public health expenditure:

$$PS_{it} = \beta_0 + \beta_1 FD_{it} + \beta_2 X_{it} + \mu_i + \delta_t + \varepsilon_{it} \quad (1)$$

where PS_{it} is public health expenditure; FD_{it} is fiscal revenue decentralization and can be replaced by fiscal expenditure decentralization; and X_{it} represents a series of control variables, including the regional economic level, industrial structure, degree of urbanization degree, imports and exports, level of scientific and technological development, and degree of marketization. μ_i is the provincial fixed effect, δ_t is the time fixed effect, and ε_{it} is a random error term. The Chi^2 value of Hausman's test is 8.28 ($P = 0.0159$). Therefore, the original hypothesis indicating the use of the random effect model is rejected, and Model 1 can be analyzed using a two-way fixed effect model.

Model 2 was used to study the impact of fiscal decentralization on public health. The quadratic term of the explanatory variable is added to test whether there is a non-linear relationship between the explanatory variables and response variables, that is, whether there is a U-shaped or n-shaped relationship. Even if the quadratic term of the explanatory variable is added,

the relationship between the explanatory variable and each parameter is still linear. Since fiscal decentralization has a non-linear relationship with public health, the quadratic term for the FD variable was introduced to test the non-linear relationship between FD and PH [as in Jin et al. (26)]. If the results show that the coefficients of FD^2 and FD are both significant, then the relationship between FD and PH is non-linear.

$$PH_{it} = \alpha_0 + \alpha_1 FD_{it} + \alpha_2 FD_{it}^2 + \alpha_3 X_{it} + \mu_i + \delta_t + \varepsilon_{it} \quad (2)$$

where FD^2 is the squared term of FD; PH_{it} represents public health measured by population mortality; X_{it} is the control variable with confounding effects on PH (as in Model 1); μ_i is the provincial fixed effect; δ_t is the time-fixed effect; and ε_{it} is a random error term and is independently identically distributed.

Model 3 was used to further analyze whether PS affects the effect of FD on PH; in this model, we selected public health expenditure as an intermediary variable and introduced the interactive terms FD × PS to verify whether fiscal decentralization influences public health by affecting public health expenditure or otherwise.

$$PH_{it} = \beta_0 + \alpha_1 FD_{it} + \alpha_2 PS_{it} + \alpha_3 FD \times PS_{it} + \alpha_4 X_{it} + \mu_i + \delta_t + \varepsilon_{it} \quad (3)$$

where $FD \times PS_{it}$ is the interaction between fiscal decentralization and public health expenditure. In order to eliminate the influence of multicollinearity and make the coefficients comparable, the variables FD and PS are centralized (27). The explanations of other variables are the same as in previous equations.

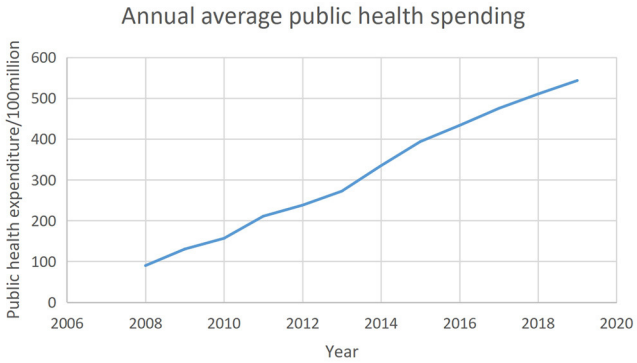

FIGURE 3 | Average annual public health spending in the period of 2008–2019.

Threshold Regression Models

The quadratic term is introduced into Model (2) to preliminarily study the non-linear relationship between FD and PH, but Model (2) cannot confirm whether the relationship is U-shaped or n-shaped, so threshold regression is needed to further explore the non-linear relationship between fiscal decentralization and public health. In reference to Wang and Wang (13), this paper uses the threshold effect model for further analysis. First, taking fiscal decentralization itself as the threshold variable, a double threshold model, Model (4), is established (28).

$$\begin{aligned} PH_{it} = {} & \beta_0 + \beta_1 FD_{it} \times I(FD_{it} \le \gamma_1) + \beta_2 FD_{it} \times I(\gamma_1 < FD_{it} \le \gamma_2) \\ & + \beta_3 FD_{it} \times I(\gamma_2 < FD_{it}) + \beta_4 X_{it} + \mu_{it} + \delta_t + \varepsilon_{it} \end{aligned} \quad (4)$$

where $I(\cdot)$ is a characteristic function, and the value of $I(\cdot)$ is 1 if the corresponding condition is true and is 0 otherwise. γ_1 and γ_2 are threshold values. β_1, β_2, and β_3 represent the impact of fiscal decentralization on public health under different levels of threshold variables. The other variables are as they were in the previous models.

However, the non-linear relationship between fiscal decentralization and public health may also be a single threshold

TABLE 2 | Descriptive statistics.

Variable	Sample size	Average value	Standard deviation	Minimum value	Maximum value
FD	360	0.590	0.469	0.199	2.727
FED	360	0.9996	0.437	0.515	2.938
PS	360	5.472	0.802	2.84	7.365
PH	360	6.042	0.763	4.21	7.57
Tech	360	5.737	1.874	1	15.46
Pollution	360	12.5932	1.2142	6.7799	14.3033
Pop	360	4521.672	2711.412	554	11521
Bed	360	21.205	13.785	1.735	64.01
lnGDPave	360	10.57	0.512	9.196	11.77
Market	360	6.44	1.948	2.33	11.518
Industry	360	0.956	0.309	0.191	1.897
Patent	360	43371.06	68928.77	228	5,27,390
Urban	360	55.801	13.044	29.11	89.6
Trade	360	0.282	0.329	0.0127	1.698

TABLE 3 | Benchmark regression results.

	PS		PH		
	Column (1) Model (1)	**Column (2) Model (1)**	**Column (3) Model (2)**	**Column (4) Model (2)**	**Column (5) Model (3)**
FD	204.7639*** (4.36)	175.1637** (2.03)	−1.775053*** (−2.97)	−1.635604*** (−2.63)	0.254947 (0.85)
lnGDPave		118.4329 (0.67)		1.387735* (2.32)	0.9316115 (1.53)
Industry		−111.4985** (−2.36)		−0.13067 (−0.83)	−0.1142591 (−0.66)
Urban		15.9217*** (4.22)		−0.0386917*** (−2.43)	−0.060338*** (−3.30)
Market		44.14284*** (3.31)		−0.031523 (−0.68)	−0.017638 (-0.36)
Trade		−296.9658*** (−2.89)		0.0585157 (0.27)	−0.066019 (−0.32)
Patent		−57.82005** (−2.48)		−0.1169984 (−1.33)	−0.1752133** (−1.94)
FD^2			0.6677871*** (4.04)	0.5963655*** (3.38)	
PS					−0.0000269 (−1.38)
FD×PS					−0.001067*** (−2.06)
Constant term	194.7079*** (6.99)	−1460.97 (−0.85)	6.710703*** (25.13)	−4.390332 (−0.75)	4.979367 (0.82)
Samples	360	360	360	360	360
Control variables	Not controlled	Controlled	Not controlled	Controlled	Controlled
Year and province effect	Controlled	Controlled	Controlled	Controlled	Controlled
R^2	0.86	0.88	0.85	0.86	0.85

*The columns report n(z-values); n refers to the coefficient of each term, z-values are in parentheses, and ***p < 0.01, **p < 0.05, *p < 0.1. Control variables include lnGDPave, Industry, Urban, Market, Trade, and Patent.*

TABLE 4 | Robustness check results.

	PS		PH		
	Column (1) Model (1)	**Column (2)Model (1)**	**Column (3) Model (2)**	**Column (4)Model (2)**	**Column (5) Model (3)**
FED	154.0399*** (3.09)	119.0458** (1.57)	−2.081253*** (−3.14)	−1.975881*** (−2.92)	0.0554069 (0.22)
FED^2			0.652546*** (4.07)	0.621215*** (3.74)	
FED×PS					−0.001783*** (−2.80)
Control variables	Not controlled	Controlled	Not controlled	Controlled	Controlled
Year and province effect	Controlled	Controlled	Controlled	Controlled	Controlled
R^2	0.88	0.88	0.85	0.86	0.85

*The columns report n (z-values); n refers to coefficient of each term, z-values are in parentheses, and ***p < 0.01, **p < 0.05, *p < 0.1. Control variables include lnGDPave, Industry, Urban, Market, Trade, and Patent.*

effect, so the estimated model can be modified, such as in Model (5) below.

$$PH_{it} = \beta_0+\beta_1 FD_{it}\times I(FD_{it}\leq \gamma)+\beta_2 FD_{it}\times I(\gamma < FD_{it}) +\beta_3 X_{it}+\mu_{it}+\delta_t+\varepsilon_{it} \quad (5)$$

Similarly, Model (6) is obtained by taking the variable PS as the threshold variable, and the interpretation of each variable is the same as that in Model (5).

$$PH_{it} = \beta_0+\beta_1 FD_{it}\times I(PS_{it}\leq \gamma)+\beta_2 FD_{it}\times I(\gamma < PS_{it}) +\beta_3 X_{it}+\mu_{it}+\delta_t+\varepsilon_{it} \quad (6)$$

Mechanism Analysis Models

In Equation (3) above, we used the interaction term between public health expenditure and fiscal decentralization as the intermediary variable to obtain results determining whether or not it has a positive impact on public health. To achieve this, mediating effect analysis is adopted to analyze the mechanism. Intermediary effect analysis is a method that tests whether a variable becomes an intermediary variable and to what extent it plays an intermediary role in the process. Generally speaking, if variable X affects variable Y by affecting variable M, then variable M is an intermediary variable (29). Testing the product of coefficients is the core principal of the intermediary effect tests and assumes that coefficient "a" is the effect of the independent variable on the intermediate variable and that coefficient "b" is the impact of the intermediate variable on the dependent variable after controlling the influence of the independent variable. The Sobel test is a coefficient product test. It checks the existence of the mediating effect by calculating Sobel statistics "z". "z" is computed by "ab" divided by the standard error of "ab," which examines whether the product of "a, b" is significantly or not (30). Therefore, we constructed models based on the concept of the intermediary effect and utilized the Sobel test to verify whether public health as the intermediary variable produces an intermediary effect. At the same time, the number of beds in regional medical institutions and the number of medical

TABLE 5 | More control variable results using FGLS.

	PS		PH			
	Column (1) Model (1)	**Column (2)Model (1)**	**Column (3) Model (2)**	**Column (4) Model (3)**	**Column (5) Model (2)**	**Column (6) Model (3)**
FD	17.65036** (1.88)	174.2401*** (2.74)	−1.08911*** (−3.78)	0.27909 (1.79)	−1.168549** (−1.97)	0.3031356 (1.02)
FD^2			0.413459*** (3.82)		0.451718*** (2.51)	
FD×PS				−0.00209*** (−6.21)		−0.00108** (−2.19)
Control variables	Controlled	Controlled	Controlled	Controlled	Controlled	Controlled

*The columns report n(z-values); n refers to coefficient of each term, z-values are in parentheses, and ***p < 0.01, **p < 0.05, *p < 0.1. Control variables in Columns (1), (3), and (4) include lnGDPave, Industry, Urban, Market, Trade, and Patent, and the control variables in (2), (5), and (6) include lnGDPave, Industry, Urban, Market, Trade, Patent, Pop, and Pollution.*

TABLE 6 | Sobel test results.

Variable	**Project**	**coefficient**	**standard deviation**	***Z*-value/*T*-value**	***P*>|*Z*|/*P*>|*t*|**
PS	Sobel value	−0.276338	0.03416027	−8.089	6.661e-16
	Direct effect	−0.28588	0.1533915	−1.86	0.063
	Total effect	−0.56222	0.1569531	−3.58	0.000
	Proportion of intermediary effect		49.15%		
Bed	Sobel value	−0.338111	0.01692059	−19.98	0.000
	Direct effect	−0.493609	0.1192244	−4.14	0.000
	Total effect	−0.83172	0.1202742	−6.92	0.000
	Proportion of intermediary effect		40.65%		
Tech	Sobel value	−0.017398	0.00677149	−2.569	0.010
	Direct effect	−0.544824	0.1573193	−3.46	0.001
	Total effect	−0.562221	0.1569531	−3.58	0.000
	Proportion of intermediary effect		3.09%		

practitioners per 1,000 people are introduced as intermediary variables, lnGDPave, Patent, and Urban are adopted as control variables to determine how the fiscal decentralization mechanism affects public health. Based on the theory and benchmark models above, the models for the mediating effect test are constructed below (31):

$$M_{it} = \beta_0 + \beta_1 FD_{it} + \beta_2 X_{it} + \mu_i + \varphi_t + \varepsilon_{it} \quad (7)$$

$$PH_{it} = \beta_0 + \beta_3 M_{it} + \beta_4 X_{it} + \mu_i + \varphi_t + \varepsilon_{it} \quad (8)$$

$$PH_{it} = \beta_0 + \beta_5 M_{it} + \beta_6 FD_{it} + \beta_7 X_{it} + \mu_i + \varphi_t + \varepsilon_{it} \quad (9)$$

where M is the intermediary variables PS, Tech, and Bed, X_{it} represents a series of control variables consistent with those in the benchmark regression, FD_{it} is fiscal revenue decentralization degree in province *i* and year *t* and PH_{it} is the population mortality in province *i* and year *t*. If β_1, β_3, and β_5 are all significant, it suggests that fiscal decentralization promotes public health by affecting those intermediary variables.

Stata16 (of StataCorp, 4905 Lakeway Dr. College Station, TX77845, United States) was the statistical software used for all the analyses in this paper.

RESULTS

Benchmark Regression

Table 3 illustrates the benchmark regression results of Equation (1) for public health expenditure (PS) and the results of Equations (2) to (3), which measured public health according to population mortality (PH).

As presented in **Table 3**, regardless of whether the control variable is introduced or not, the coefficient of the effects of fiscal decentralization on public health expenditure is significantly positive at the 1% level (which means that there is a probability that the coefficient is false and is <1%), indicating that fiscal decentralization promotes the government public health expenditure. Additionally, it can be seen from Column (3) and (4) that fiscal decentralization has a significant negative effect on population mortality (at the level of 1%), which further proves the promotion effect that fiscal decentralization has on public health. The coefficient of FD is negative, but in contrast to the FD coefficient, the coefficient of the quadratic term of FD is significantly positive (at the 1% level). And the absolute value of the coefficient of FD^2 is smaller than that of FD, demonstrating the non-linear feature of the impact of fiscal decentralization on public health. In Column (5), the negative coefficients of the interaction terms in Equation (3) are significant, illustrating

that fiscal decentralization can advance public health through increasing public health expenditure.

Robustness Check

In order to verify the robustness of the above regression results, the following robustness tests were conducted. Firstly, FED was used as an explanatory variable to replace FD to estimate Equations (1), (2), and (3). The results are demonstrated in **Table 4**. Secondly, we used the feasible generalized least squares (FGLS) estimation to deal with the problem of heteroscedasticity for the robustness test. The results are shown in columns (1), (3), and (4) of **Table 5**. Other studies emphasize the negative impact of air pollution on public health (32, 33). However, environmental pollution can be represented by many existing control variables, such as industry and because population may also be affected by lnGDPave. However, in order to make the results more robust, we introduced air pollution and population as control variables into all of the models, and the results are illustrated in Columns (2), (5), and (6) in **Table 5**.

As indicated in **Tables 4**, **5**, no matter which method is used to test the robustness of benchmark regression, the magnitude and sign of the coefficients of the main explanatory variables are the same as those of the benchmark regression, showing that fiscal decentralization has positive effect on public health expenditure, that fiscal decentralization also exerts a non-linear favorable effect on public health. Fiscal decentralization can affect public health by increasing public health expenditure. The benchmark regression passed the robustness test and thus, the results are considered valid.

Mechanism Analysis

The results of the Sobel test are presented in **Table 6**.

As shown in **Table 6**, the three intermediary variables passed the intermediary effect test at the 1% level, with the intermediary variables having a negative impact on population mortality. The intermediary effect of public health expenditure accounts for its largest proportion, up to 49.15%, which is significantly mutually confirmed with the previous interaction term in Equation (3). Similarly, the number of medical beds account for 40.65%, and the number of medical practitioners also have a significant impact, accounting for 3.09%. **Figure 4** illustrates the flow of the mechanism effects.

Heterogeneity Analysis

Figures 5, **6** show regional distribution differences in terms of provincial population mortality and the fiscal decentralization level in China in 2008 and 2019, respectively. No data were available for Tibet and Taiwan. The darker the color, the higher the population mortality, and the worse the public health level.

Figures 5, **6** suggest that there are differences in the population mortality and degree of fiscal decentralization between coastal and inland cities. Therefore, we categorized the 30 provinces into coastal and inland areas. Coastal areas include Hebei, Liaoning, Tianjin, Shandong, Zhejiang, Shanghai, Fujian, Jiangsu, Guangdong, Hainan, and Guangxi. In addition, according to the traditional east–west division method, we also divided them into east (Tianjin, Beijing, Hebei, Shanghai, Liaoning, Jiangsu, Fujian, Zhejiang, Shandong, Hainan and Guangdong), west (Sichuan, Gansu, Qinghai, Yunnan, Guizhou, Chongqing, Shaanxi, Xinjiang, Ningxia, Guangxi, Inner Mongolia) and central (Shanxi, Heilongjiang, Anhui, Henan, Hunan, Hubei, Hebei, Jiangxi) regions (34). Using these categorizations, we applied Equation 2 for regression. The regression results are shown in **Table 7**.

The results illustrate that fiscal decentralization in inland cities plays a larger role in improving the level of public health compared to coastal areas. However, no significant differences were observed in the impacts of fiscal decentralization on the level of public health between eastern and western cities.

Unexpectedly, the fiscal decentralization of inland provinces better inhibits population mortality compared to coastal regions. The reasons for this may be that the economies of coastal cities are relatively more developed and that the fiscal decentralization level is better established. During the threshold effect analysis (further explained below), we found that when fiscal decentralization reaches a certain degree, its promotion effect on local public health will be weakened. Another important factor to highlight is that the populations of inland Chinese regions are generally smaller than those found in other parts of the country, so the population mortality statistics may be biased (35). For the central and western regions, the coefficient is not significant. This is because eastern regions and coastal areas were at the forefront of the reform and opening up (which has been one of China's national policies since 1978), resulting in a stronger economic foundation. The central and western cities have a lower population density and lower industrial aggregation, so the role that fiscal decentralization plays in public health may be weaker.

Threshold Effect Analysis

In addition to the verification of a non-linear relationship between fiscal decentralization and public health in the previous sections, a threshold effect analysis was conducted on FD and PS to further understand the non-linear effect of fiscal decentralization on public health. The results of the threshold effect test are presented in **Table 8**.

As is illustrated in **Table 8**, there is a single threshold effect with both variables FD and PS acting as threshold variables. Therefore, threshold effect regression was applied on Models (5) and (6), as shown in **Table 9**.

Table 9 indicates that regardless of whether FD or PS is adopted as the threshold variable, after crossing a certain value, that is, when FD is >2.0361 or PS is >135.1796, the absolute value of the coefficient of core explanatory variable FD becomes larger, implying that FD can inhibit population mortality. On the other hand, the results of the threshold effect regression test also indicate that regardless of the level of FD and PS, FD always shows a negative correlation with population mortality, which again verifies that providing further fiscal decentralization to local governments will promote the improvement of local health.

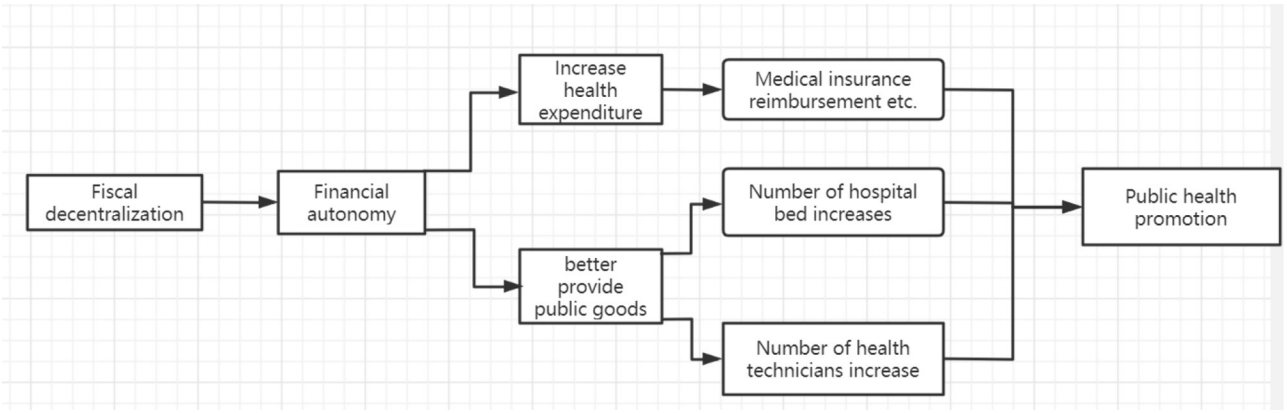

FIGURE 4 | Mechanism flow chart.

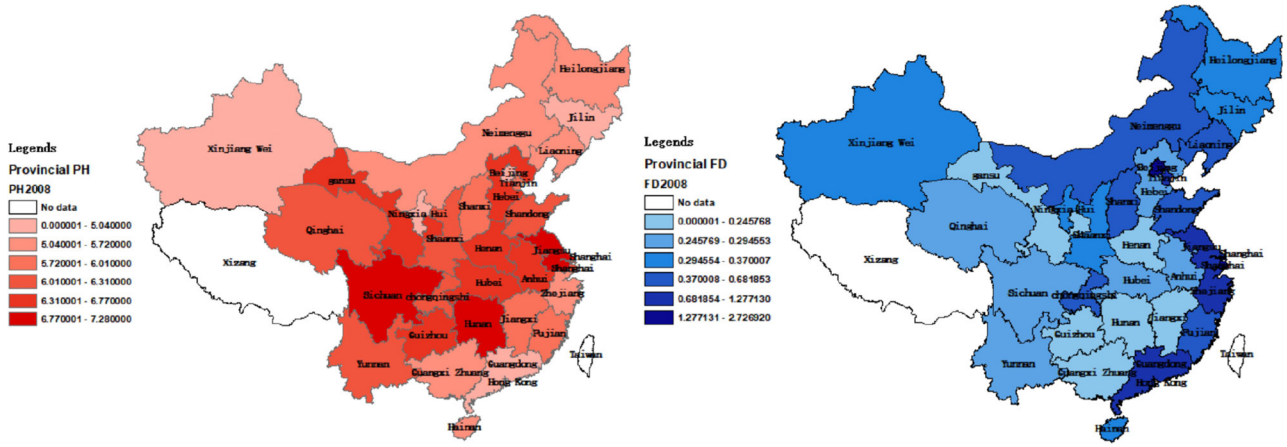

FIGURE 5 | Provincial differences in PH and FD in 2008.

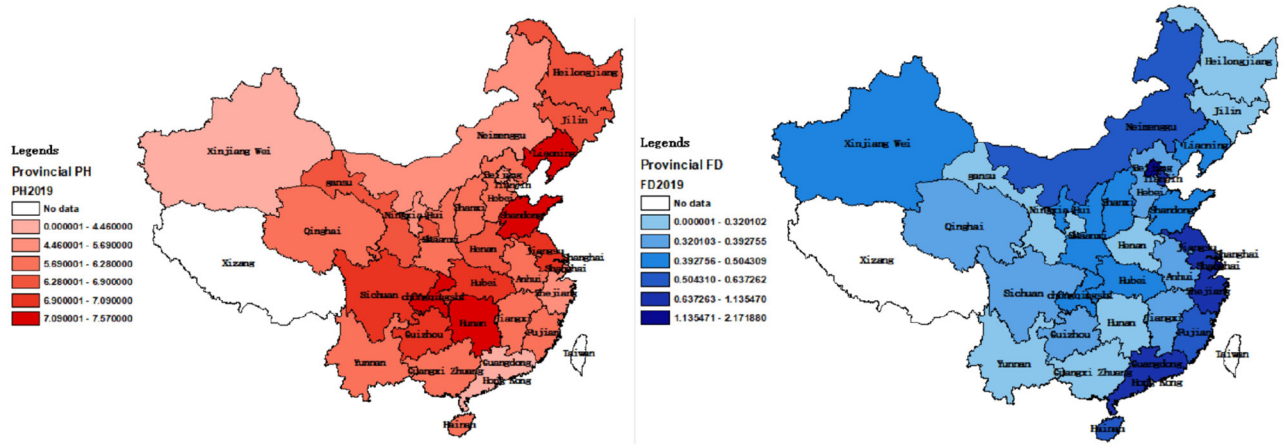

FIGURE 6 | Provincial differences in PH and FD in 2019.

TABLE 7 | Regional regression results.

	Coastal areas	Inland areas	Eastern regions	Central/western regions
FD	−0.4044268*** (−2.89)	−0.579207*** (−3.86)	−1.056453*** (−67.82)	−0.1241495 (−0.42)
Control variable	Controlled	Controlled	Controlled	Controlled
Province and time effect	Controlled	Controlled	Controlled	Controlled
Number of samples	132	228	132	228

*z-values are in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1.*

TABLE 8 | Threshold existence test.

Threshold variables	Threshold number	Threshold estimator	Bootstrap times	F-stat	prob	Critical value		
						1%	5%	10%
FD	Single	2.0361	500	30.35	0.004	36.0683	27.5463	22.7376
	Double	1.1932	500	18.46	0.170	71.3974	44.3315	27.9452
PS	Single	135.1796	500	41.37	0.008	38.8815	28.5789	24.2140
	Double	140.4184	500	−0.32	1.000	37.2183	27.2670	22.1377

DISCUSSION

In recent years, scholars have conducted a lot of research on the impacts of fiscal decentralization on economic development. However, there is little research on the impact of fiscal decentralization on public health. The impact of fiscal decentralization on public goods has been a contested topic since it was first proposed. It was first proposed by Tiebout (36) and Oates (37), who formulated "voting by foot" theory, where residents can freely choose to settle in places where their public goods and tax burden needs are better provided by local governments. These scholars believed that local governments understand the preferences of local residents better than the central government, making them better positioned to provide public goods and services to their populations. However, scholars such as Weingast (38) believe that fiscal decentralization has an adverse impact on the supply of public goods (39). That is, fiscal decentralization enables local governments to compete for scarce capital and labor, increase the number of economic public goods, and then squeezes out the expenditure of non-economic public goods (40). Such decentralization skeptics worry that there would be destructive interregional competition among local governments, which may reduce expenditure or relax regulations; hence, they might not be able to fully provide basic public services (41). Despite these arguments, other scholars such as Khaleghian (42), Cantarero and Pascual (43), and Uchimura and Jütting (44) argue that fiscal decentralization could improve access to health services. In recent years, some studies in the literature have examined the impact of fiscal decentralization on various population health indicators, such as infant mortality, life expectancy, or immunization coverage, and found that fiscal decentralization had a beneficial impact on public health (45). In addition, Uchimura and Jütting (44) found that county-level fiscal decentralization in China was significantly associated with low infant mortality from 1995 to 2001.

The main argument surrounding fiscal decentralization theory is whether fiscal decentralization can improve public welfare. The first generation of decentralization theories hold that local governments are more efficient in supplying certain public goods, improving public welfare levels (46). Starting from public choice theory, the second generation of fiscal decentralization theories hold that the government does not start with the goal of maximizing resident welfare (47, 48), and that especially in countries with an imperfect democratic constitutional system, the competition between local governments under fiscal decentralization cannot improve the level of public welfare and may instead lead to a decline in or even the deterioration of the level of public welfare (49).

Jiménez-Rubio et al. (45) used infant mortality as the public health index and drew the conclusion that fiscal decentralization can only have a significant positive impact on public health if local governments are provided with a large amount of financial autonomy. Hao et al. (50) used panel data from 23 provinces in China from 2002 to 2012 and using a simultaneous equation model to analyze the data, found that fiscal decentralization has both direct and indirect negative impacts on public health. However, the 2002–2012 panel data used by Hao et al. (50) was based on the economic development that took place when local Chinese governments were placing emphasis on GDP and ignoring expenditure on public goods.

The primary aim of this study was to investigate the (non-linear) effect of fiscal decentralization on public health. The empirical results verified that fiscal decentralization plays a significant role in promoting public health expenditure. With greater financial autonomy, local governments have the capacity to directly increase expenditure in the public health areas that are relevant to the needs of the people in each

TABLE 9 | Results of the panel threshold effect regression test.

Variables	Model (5)	Variables	Model (6)
FD (FD ≤2.0361)	−2.240428*** (−4.89)	FD (PS ≤135.1796)	−1.252698*** (−2.66)
FD (2.0361< FD)	−2.740643*** (−5.63)	FD (135.1796< PS)	−1.608206*** (−3.52)
Control variable	Controlled	Control variable	Controlled

*z-values are in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1.*

region since local governments have a better understanding of local issues. Additionally, when the interaction terms of fiscal decentralization and public health spending are introduced, the results show that the coefficient of the interaction term is negative, even though the coefficient of FD is always insignificant in the results. Therefore, what is important is the coefficient of the interaction term. We also established that fiscal decentralization does promote public health by increasing public health expenditure. Based on these results, the intermediary test and threshold effect analysis outcomes show that fiscal decentralization promotes public health through public health expenditure and improves local medical services. At the same time, the effect of fiscal decentralization on public health presents non-linear changes with the changes in the threshold variables of fiscal decentralization and public health expenditure.

The results of this study are different from those obtained in previous studies. Other literature suggests that fiscal decentralization has a negative effect on public health outcomes (51). No previous studies have examined the non-linear relationship between fiscal decentralization and public health. The awareness that Chinese people have about public health and public health measures has improved significantly (52). The performance criteria of local governments are now not only focused on simple economic indicators such as GDP, but also on other indicators such as environmental pollution (53). The Central Committee and the central government of China now pay more attention to people's livelihoods, especially the health of the people in the country (54). Local governments have also paid more attention to public health issues and have created interventions with stricter monitoring measures (55). Since fiscal decentralization has brought greater economic autonomy to local governments, the local governments have consequently paid more attention to public health intervention programs that have been largely developed in response to the demands of local residents for a better quality of health.

A number of research limitations have been identified in this study. First, in the robustness test, indicators such as maternal mortality [which is also a widely used indicator to measure public health (56)] were not used to replace population mortality as a measurement of public health. Second, other models, such as panel smooth transition regression (PSTR), were not used to study the non-linear relationship, so there may be bias in terms of model selection. Thirdly, using the commonly used intermediary effect test in mechanism analysis means that many of the key mechanisms of fiscal decentralization affecting public health are likely to be missed. Fourthly, compared to the city-level or county-level panel data, the sample used in the present study is not representative of the population. Micro-level panel data in the regions would have more accuracy. Finally, the results obtained from the heterogeneity analysis cannot be explained. Since various robustness tests were conducted, these shortcomings have essentially been attenuated if not eliminated. In future research, we can further study the impact of fiscal decentralization on public health at the city or at more micro levels or the impact of fiscal decentralization on specific types of diseases, such as chronic conditions, particularly non-communicable diseases. The control and prevention of chronic diseases may be closely related to local health expenditure. In addition, we can examine the impact of fiscal decentralization on the differences in the public health level between urban and rural areas. It will also be interesting to determine whether fiscal decentralization aggravates the unequal distribution of healthcare resources to study the connection between fiscal decentralization and the supply of public health and local health resources.

CONCLUSIONS AND POLICY RECOMMENDATIONS

Based on the above results, this paper draws the following conclusions: (1) fiscal decentralization can increase the level of regional fiscal health expenditure; (2) fiscal decentralization promotes public health by increasing public health expenditure and improving regional medical resources, such as beds in medical institutions and medical personnel; and (3) the positive effect of fiscal decentralization on public health is non-linear and is affected by fiscal decentralization and public expenditure.

There are several policy implications for the findings of this study: China's public health expenditure still does not comply with the requirements of the WHO (57). The WHO health expenditure targets require an upper-middle-income country to ensure financial health protection at an adequate level, meaning that at least 6.7% of GDP should be allocated to public health spending (58). However, based on the WHO's Global Health Expenditure Database (59), during 2008–2019, China, an upper-middle-income country, does not exceed 5% of GDP. The government should therefore continue to increase its public health expenditure, and, when necessary, legislate to clarify the proportion of GDP required for public health investment, as is the case for education. Post Covid-19, the public health system still needs to be continuously improved to adequately respond to all major challenges. The central government should continue to increase and protect fiscal decentralization. For example, it

should improve the degree of decentralization through fiscal and taxation tools. When promoting fiscal decentralization, the government should refine the scope of its central and local expenditure responsibilities, create an environment for fairness and openness, and refine the responsibilities of local government in terms of public service expenditure in the process of fiscal decentralization. In the process of deepening fiscal decentralization and expanding local fiscal autonomy, local governments should be thoroughly invested in serving the people and should listen to the real needs of the people. Additionally, financial resources should be used to continuously improve people's living conditions as well as to improve people's health.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

WX conceived the concept, designed and conducted empirical analysis, and wrote the paper. JL collected data and helped conducting the empirical research revising and editing the paper as well as making the graphs and tables. All authors read, revised, and approved to the submitted version of the manuscript.

REFERENCES

1. The Central People's Government of the people's Republic of China: the CPC Central Committee and the State Council issued the outline of the "healthy China 2030" plan. Available online at: http://www.gov.cn/zhengce/2016-10/25/content_5124174.htm
2. Xi stresses safeguarding people's health, building quality basic public education (chinatoday.com.cn). Available online at: http://www.chinatoday.com.cn/ctenglish/2018/tpxw/202103/t20210307_800238941.html
3. Ding Y, McQuoid A, Karayalcin C. Fiscal decentralization, fiscal reform, and economic growth in china. *China Eco Rev*. (2019) 53:152–67. doi: 10.1016/j.chieco.2018.08.00
4. Jia J, Guo Q, Zhang J. Fiscal decentralization and local expenditure policy in China. *China Eco Rev*. (2014) 28:107–22. doi: 10.1016/j.chieco.2014.01.002
5. Lin B, Zhou Y. Does fiscal decentralization improve energy and environmental performance? New perspective on vertical fiscal imbalance. *App Ene*. (2021) 302:117495. doi: 10.1016/j.apenergy.2021.117495
6. Akai N, Sakata M. Fiscal decentralization contributes to economic growth: evidence from state-level cross-section data for the United States. *J Urban Eco*. (2002) 52:93–108. doi: 10.1016/S0094-1190(02)00018-9
7. Zhang T, Zou HF. Fiscal decentralization, public spending, and economic growth in China. *J Pub Eco*. (1998) 67:221–40. doi: 10.1016/S0047-2727(97)00057-1
8. Aslim EG, Neyapti B. Optimal fiscal decentralization: redistribution and welfare implications. *Eco Mod*. (2017) 61:224–34. doi: 10.1016/j.econmod.2016.12.008
9. Marlow ML. Fiscal decentralization and government size. *Pub Choice*. (1988) 56:259–69. doi: 10.1007/BF00130275
10. West LA, Wong CP. Fiscal decentralization and growing regional disparities in rural china: some evidence in the provision of social services. *Oxford Rev Econ Pol*. (1995) 11:70–84. doi: 10.1093/oxrep/11.4.70
11. Fu Y, Zhang Y. Chinese decentralization and fiscal expenditure structure bias: the cost of competition for growth. *Manag World*. (2007) 03:4–12. doi: 10.19744/j.cnki.11-1235/f.2007.03.002
12. Mauro P. Corruption and the composition of government expenditure. *J Pub Eco*. (1998) 69:263–79. doi: 10.1016/S0047-2727(98)00025-5
13. Wang Q, Wang L. The non-linear effects of population aging, industrial structure, and urbanization on carbon emissions: a panel threshold regression analysis of 137 countries. *J Clean Prod*. (2022) 2871:125381. doi: 10.1016/j.jclepro.2020.125381
14. Li G, Guo F, Di D. Regional competition, environmental decentralization, and target selection of local governments. *Sci Tot Environ*. (2021) 755:142536. doi: 10.1016/j.scitotenv.2020.142536
15. Du J, Sun Y. The non-linear impact of fiscal decentralization on carbon emissions: from the perspective of biased technological progress. *Environ Sci Pollut Res*. (2021) 28:29890–9. doi: 10.1007/s11356-021-12833-w
16. Adam A, Delis MD, Kammas, P. Fiscal decentralization and public sector efficiency: evidence from OECD countries. *Econ Gov*. (2014) 15:17201. doi: 10.1007/s10101-013-0131-4
17. Su CW, Umar M, Khan Z. Does fiscal decentralization and eco-innovation promote renewable energy consumption? Analyzing the role of political risk. *Sci Tot Environ*. (2021) 751:142220. doi: 10.1016/j.scitotenv.2020.142220
18. The Central People's Government of the People's Republic of China. Available online at: http://www.gov.cn/guoqing/2012-04/18/content_2584164.htm
19. Christopoulos K, Eleftheriou K. The fiscal impact of health care expenditure: evidence from the OECD countries. *Eco Ana Pol*. (2020) 67:195–202. doi: 10.1016/j.eap.2020.07.010
20. Van Schalkwyk MC, Barlow P, Siles-Brügge G, Jarman H, Hervey T, McKee M. Brexit and trade policy: an analysis of the governance of UK trade policy and what it means for health and social justice. *Global Health*. (2021) 17:61. doi: 10.1186/s12992-021-00697-1
21. Alan I, Barreca, Matthew Neidell, Nicholas J, Sanders. Long-run pollution exposure and mortality: evidence from the acid rain program. *J Pub Eco*. (2021) 200:104440. doi: 10.1016/j.jpubeco.2021.104440
22. Hartig T, Mitchell R, De Vries S, Frumkin H. Nature and health. *Annual Rev Pub Health*. (2014) 35:1:207–28.doi: 10.1146/annurev-publhealth-032013-182443
23. Seitio-Kgokgwe O, Gauld RD, Hill PC, Barnett P. Assessing performance of Botswana's public hospital system: the use of the world health organization health system performance assessment framework. *Int J Health Pol Manag*. (2014) 3:179–89. doi: 10.15171/ijhpm.2014.85
24. Gallet CA, Doucouliagos H. The impact of healthcare spending on health outcomes: a meta-regression analysis. *Soc Sci Med*. (2017) 179:9–17. doi: 10.1016/j.socscimed.2017.02.024
25. Hsiao C. Panel data analysis—advantages and challenges. *Test*. (2007) 16:1–22. doi: 10.1007/s11749-007-0046-x
26. Jin G, Yu B, Shen K. Domestic trade and energy productivity in China: an inverted U-shaped relationship. *Energy Eco*. (2021) 97:105234. doi: 10.1016/j.eneco.2021.105234
27. Hayes AF, Rockwood NJ. Rockwood. Regression-based statistical mediation and moderation analysis in clinical research: observations, recommendations, and implementation. *Behav Res Therapy*. (2017) 98:39–57. doi: 10.1016/j.brat.2016.11.001
28. Hansen BE. Threshold effects in non-dynamic panels: estimation, testing, and inference. *J Econ*. (1999) 93:345–68. doi: 10.1016/S0304-4076(99)00025-1
29. Baron RM, Kenny DA. The moderatorsis of 137 countriesure, and urbanization on carbon emissionsor: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. (1986) 51:1173–173:6olhollc1037/0022-3514.51.6.1173
30. Koo N, Leite WL, Algina J. Mediated effects with the parallel process latent growth model: an evaluation of methods for testing mediation in the presence of non-normal data, structural equation modeling. *Multidiscip J*. (2016) 1:32–44. doi: 10.1080/10705511.2014.959419

31. Xu W. The impact and influencing path of the pilot carbon emission trading market–evidence from China. *Front Environ Sci.* (2021) 9:787655. doi: 10.3389/fenvs.2021.787655
32. Zhao S, Liu S, Hou X, Sun Y, Beazley R. Air pollution and cause-specific mortality: a comparative study of urban and rural areas in China, Chemosphere. (2021) 262:127884. doi: 10.1016/j.chemosphere.2020.127884
33. Sun Z, Cheng L. Air pollution and procyclical mortality: evidence from China. *China Eco Rev*. (2021) 68:101625. doi: 10.1016/j.chieco.2021.101625
34. Zhao N. The impact of China's new-type urbanization on individual sustainable development: a comparison between coastal and inland areas. *Marine Pol*. (2022) 136:104938. doi: 10.1016/j.marpol.2021.104938
35. Xu Z, Ouyang A. The factors influencing China's population distribution and spatial heterogeneity: a prefectural-level analysis using geographically weighted regression. *Appl Spatial Ana*. (2018) 11:465–80. doi: 10.1007/s12061-017-9224-8
36. Tiebout CM. A pure theory of local expenditures. *J Pol Eco*. (1956) 64:416–24. doi: 10.1086/257839
37. Oates WE. An essay on fiscal federalism. *J Eco Lit*. (1999) 37:1120–49. doi: 10.1257/jel.37.3.1120
38. Weingast BR. Second generation fiscal federalism: the implications of fiscal incentives. *J Urban Eco*. (2009) 65:279–93. doi: 10.1016/j.jue.2008.12.005
39. Oto-Peralías D, Romero-Ávila D, Usabiaga C. Does fiscal decentralization mitigate the adverse effects of corruption on public deficits?. *Europ J Pol Eco*. (2013) 32:205–31. doi: 10.1016/j.ejpoleco.2013.07.005
40. Keen M, Marchand M. Fiscal competition and the pattern of public spending. *J Pub Eco*. (1997) 66:33–53 doi: 10.1016/S0047-2727(97)00035-2
41. Arends H. The dangers of fiscal decentralization and public service delivery: a review of arguments. *Polit Vierteljahresschr*. (2020) 61:59920el. doi: 10.1007/s11615-020-00233-7
42. Khaleghian P. Decentralization and public services: the case of immunization. *Soc Sci Medi*. (2004) 59:163–83. doi: 10.1016/j.socscimed.2003.10.013
43. Cantarero D, Pascual M. Analysing the impact of fiscal decentralization on health outcomes: empirical evidence from Spain. *Appl Econ Lett*. (2008) 15:109–11 doi: 10.1080/13504850600770913
44. Uchimura H, Jütting JP. Fiscal decentralization, Chinese style: good for health outcomes? *World Develop*. (2009) 37:1924–36 doi: 10.1016/j.worlddev.2009.06.007
45. Jiménez-Rubio D. The impact of fiscal decentralization on infant mortality rates: evidence from OECD countries. *Soc Sci Med*. (2011) 73:1401–7 doi: 10.1016/j.socscimed.2011.07.029
46. Oates WE. Toward a second-generation theory of fiscal federalism. *Int Tax Public Finan*. (2005) 12:349–73. doi: 10.1007/s10797-005-1619-9
47. Qian Y, Weingast BR. Federalism as a commitment to reserving market incentives. *J Eco Perspect*. (1997) 11:83–92. doi: 10.1257/jep.11.4.83
48. Weingast BR. The economic role of political institutions: market-preserving federalism and economic development. *J Law Eco Org*. (1995) 11:1–31.
49. Litvack JI, Ahmad J, Bird RM. *Rethinking Decentralization in Developing Countries*. Washington: The World (1998). Available online at: Bank.https://digitallibrary.un.org/record/276289
50. Hao Y, Liu J, Lu ZN, Shi R, Wu H. Impact of income inequality and fiscal decentralization on public health: evidence from China. *Eco Model*. (2021) 94:34. doi: 10.1016/j.econmod.2020.02.034
51. Brock G, Jin Y, Zeng T. Fiscal decentralization and China's regional infant mortality. *J Pol Model*. (2015) 37:175–88. doi: 10.1016/j.jpolmod.2015.03.001
52. Tan X, Zhang Y, Shao H. Healthy China 2030, a breakthrough for improving health. *Glob Health Promot*. (2019) 26:96. doi: 10.1177/1757975917743533
53. Hong T, Yu N, Mao Z. Does environment centralization prevent local governments from racing to the bottom? –Evidence from China. *J Clean Prod*. (2019) 231:649–59. doi: 10.1016/j.jclepro.2019.05.181
54. Xiao Y, Chen X, Li Q, Jia P, Li L, Chen Z. Towards healthy China 2030: modeling health care accessibility with patient referral. *Soc Sci Med*. (2021) 276:113834. doi: 10.1016/j.socscimed.2021.113834
55. Chai KC, Yang Y, Cui ZX, Ou YL, Chang KC. Threshold effect of the government intervention in the relationship between business cycle and population health: evidence from China. *Front Pub Health*. (2021) 9:689870. doi: 10.3389/fpubh.2021.689870
56. Shen C, Williamson JB. Maternal mortality, women's status, and economic dependency in less developed countries: a cross-national analysis. *Soc Sci Med*. 1999 49:197–214. doi: 10.1016/S0277-9536(99)00112-4
57. WHO Global Spending on Health: A World in Transition (2019). Available online at: https://www.who.int/health_financing/documents/health-expenditure-report-2019.~pdf?ua$=$1
58. Jowett M, Brunal MP, Flores G, Cylus J. *Spending Targets for Health: No Magic Numbe*r. Geneva: World Health Organization (2016) (WHO/HIS/HGF/HFWorkingPaper/16.1; Health Financing WorkingPaperNo.1); Available online at: http://apps.who.int/iris/bitstream/10665/250048/1/WHO-HIS-HGFHFWorkingPaper-16.1-eng.pdf
59. WHO. Global Health Expenditure Database. Available online at: https://apps.who.int/nha/database/

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Public Health

ORIGINAL RESEARCH
published: 11 July 2022
doi: 10.3389/fpubh.2022.847938

Analysis of Multiple Health Risky Behaviours and Associated Disease Outcomes Using Scottish Linked Hospitalisation Data

Damilola Olajide[1]*, *Barbara Eberth*[2] *and Anne Ludbrook*[3]

[1] *Research Design Services East Midlands, School of Medicine, Queens Medical Centre, University of Nottingham, Nottingham, United Kingdom,* [2] *Newcastle University Business School, Newcastle upon Tyne, United Kingdom,* [3] *Health Economics Research Unit, University of Aberdeen, Aberdeen, United Kingdom*

Background: Disease incidence and premature deaths tend to be influenced by multiple health risky behaviours, including smoking, excessive alcohol consumption and unhealthy diet. Risky behaviours tend not to be independent and may have a multiplicative effect on disease incidence and healthcare cost. Thus, understanding the interrelationship between health behaviours and their effect on health outcomes is crucial in designing behavioural intervention programmes.

Objective: To examine the interrelationship between health risky behaviours and associated disease outcomes amongst Scottish adults.

Methods: We use hospitalisation episode data from the Scottish Morbidity Records, (SMR), that have been administratively linked to Scottish Health Surveys (SHeS) respondents with target disease defined by relevant ICD9 and 10 codes. We apply a recursive multivariate probit model to jointly estimate the health risky behaviours and disease incidence to adequately control for unobserved heterogeneity. The model is estimated separately by gender.

Results: Modelling health risk behaviours and disease incidence equations independently rather than jointly may be misleading. We find a clear socioeconomic gradient predicting health risky behaviours and the results differ by gender. Specifically, smoking appears to be a key driver of other health risky behaviours. Current smokers are more likely to be drinking above the recommended limit, physically inactive, and eating inadequate diet.

Conclusions: Interventions targeting current smokers to quit could spillover to other behaviours by reducing excessive drinking, improve physical activity and adequate diet. Thus, improvements in one behaviour may increase the likelihood of adopting other healthier lifestyle behaviours.

Keywords: multiple health behaviours, Scottish adults, linked Scottish health survey, recursive multivariate probit model, lifestyle specific diseases

OPEN ACCESS

Edited by:
Mihajlo Jakovljevic,
Hosei University, Japan

Reviewed by:
Enver Envi Roshi,
University of Medicine, Albania
Alexander Slade,
Stony Brook Medicine, United States

***Correspondence:**
Damilola Olajide
damilola.olajide@nottingham.ac.uk

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *03 January 2022*
Accepted: *09 May 2022*
Published: *11 July 2022*

Citation:
Olajide D, Eberth B and Ludbrook A (2022) Analysis of Multiple Health Risky Behaviours and Associated Disease Outcomes Using Scottish Linked Hospitalisation Data. Front. Public Health 10:847938. doi: 10.3389/fpubh.2022.847938

INTRODUCTION

Health risky behaviours, including smoking, excessive alcohol consumption, physical inactivity and inadequate diet have a direct impact on hospital care demand. These risky behaviours are major factors associated with chronic disease incidence, including heart diseases, stroke, type 2 diabetes, obesity, and certain types of cancer; and a major cause of death. For example, 1.7 million or 2.8% of total number of deaths worldwide are attributable to low fruit and vegetable consumption, and considered amongst the top 10 risk factors for global mortality. Low fruit and vegetable intake was associated with around 14% of cancer and gastrointestinal related deaths and about 9% of stroke related deaths (1). In 2009, smoking attributable admissions in Scotland amounted to 56,153 and 13,044 deaths (2), whilst the prevalence rate was 24% on average between 2009/2010 amongst 16–64 year olds based on self-reported smoking status (3).

Chronic disease incidence resulting from risky behaviours presents a significant economic burden to health care systems and greater costs to the society. Scarborough et al. (4) estimated the economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity individually in the UK. The study found that in 2006/07, the cost of ill health associated with poor diet amounted to £5.8 billion, physical inactivity £0.9b, smoking £3.3b and alcohol-related £3.3b, whilst overweight and obesity was £5.1b.

Previous studies examining links between these behavioural risk factors and subsequent hospital admissions and death tend to assume an independence relationship between the health risky behaviours. Using a unique Scottish data resource (linked Scottish health survey), Hanlon et al. (5) used the 1998 version of the linked data to analyse the independent association between each behavioural factor and subsequent hospital admissions. In Lawder et al. (6), the same authors retrospectively extended their analysis by including the 1995 linked data to also predict deaths. Both studies found that each health behaviour was independently associated (positively or negatively) with the probability of subsequent hospital admissions and death. For example, not meeting the daily fruit and vegetable five-a-day recommendation was associated with increased risk of admissions and death. Smoking was found to have the highest hazard ratio of all the behavioural factors. In their study limitations however, the authors suggested that future research should investigate associations between risk factors and hospital admissions that are disease-specific diagnoses.

Multiple health risky behaviours tend to occur concurrently or complement each other [e.g., (7, 8)], implying a multiplication of associated disease burdens on health and productivity, pressure on health care systems both in terms of disease incidence and associated costs of hospitalisation (9–11). Moreover, cardiovascular diseases tend to be associated with more than a single health behaviour, suggesting potential intercorrelation between these risky behaviours. However, few studies have been undertaken to understand the patterns that such interactions between health risky behaviours and associated health outcomes might take.

This study examined the interrelationships between risky health behaviours and associated disease outcomes, using large linked health care data on Scottish adults (the linked Scottish health survey). The linked Scottish health survey (linked SheS) links the baseline information on risky behaviours and other characteristics in three waves of the Scottish Health Survey [i.e., (12–14)] to subsequent hospital admissions that were behaviour-specific diagnoses up to 2013. Linking behaviour-specific diagnoses provided a way to approximate causality between health behaviours and disease incidence.

To consistently examine the interrelationships, a recursive multivariate probability (probit) model was adopted. This model jointly estimated a disease incidence equation and a reduced form equation for each of the health risky behaviours; smoking, excess alcohol consumption, physical inactivity, and inadequate diet. The model enabled us to assess the association of demographic, socioeconomic, and other factors with risky lifestyle choices, and investigated their influence on the probability of having a related disease incidence. The modelling approach also allowed us to account for (i) the correlation of unobserved characteristics that may impact both disease risk and risky health behaviours, and (ii) the potential endogeneity of the risky health behaviours to disease incidence. The model was estimated separately for female and male samples.

The model statistic showed that modelling health risk behaviours and disease incidence equations independently rather than jointly may be misleading. The estimation results showed a clear socioeconomic gradient predicting health risky behaviours. Specifically, smoking appeared to be a key driver of other health risky behaviours. Current smokers were more likely to be drinking above the recommended limit, physically inactive, and eating inadequate diet. These results differ by age and gender. Undertaking health risky behaviours decreases with age. Also, males were more likely than females to undertake health risky behaviours: smoking, excessive drinking and inadequate diet, longer in their lives.

The findings suggest that targeting current smokers to quit in interventions to promote multiple healthier lifestyles may spill over to other lifestyles by reducing excessive drinking, improve physical activity and adequate diet amongst individuals with multiple health risky behaviours. This implies that, improvements in one behaviour may increase the likelihood of adopting other healthier lifestyles.

The remainder of the paper is organised as follows. Section on Materials and Methods provides a data description and outlines the estimation model. Results are presented next followed by the Discussion.

MATERIALS AND METHODS

Data

The Linked Scottish Health Survey Dataset

The hospitalisation episodes data used in this study were sourced from the Scottish Morbidity Records (SMR) that were administratively linked to respondents to Scottish Health Surveys (SHeS) by the Information Services Division (ISD) of the National Health Services (NHS) Scotland. The SHeS is a

national representative survey of individuals living in private households in Scotland. It collects a wide range of information on individual respondents, including demographic, socioeconomic, environmental, self-reported health, and health behaviours.

The SMR system other the other hand, records details of all hospitalisation episodes in Scottish NHS hospitals for each respondent in the SHeS who consented to having their information passed onto the ISD. Each episode record contains primary and secondary diagnoses identified by disease-specific International Classification of Diseases ICD9 and ICD10 codes, date of admission and discharge. The SMR is routinely linked to cancer and death registrations, covering the period 1981 to the present. Permission from the Privacy Advisory Committee of the ISD was granted to access the linked SHeS datasets.

Health Risky Behaviours and Associated Disease Outcome Variables

The health risky behaviours considered were smoking, alcohol consumption over the limit, physical inactivity, and inadequate diet. All were measured as binary variables in this study. The surveys collected a comprehensive range of data pertaining to alcohol consumption from respondents aged 16 and older, allowing the derivation of weekly units of alcohol consumption. From this, we derived an indicator variable taking a value of one if the individual reported drinking regularly over the moderate limit, and zero otherwise. The drinking limits at the time of the survey were 1 to 14 units for women and 21 units for men (15)[1].

For smoking, individual respondents aged 16 and older were asked if they "smoke cigarettes regularly nowadays". Respondents answered either "Yes" or "No". Thus, the binary "Smoking" variable took a value of one if respondents answered "Yes"; and zero otherwise. For physical inactivity we generated the binary variable "Inactive" taking a value of one if the individual indicated to have spent no time in sporting activities 28 days prior to interview; and zero otherwise.

We proxy diet by use of the number of portions of fruit and vegetables consumed by each respondent over the 24 h period prior to interview. Information on fruit and vegetables include the following food types: fresh, frozen and canned vegetables; vegetables in composites (vegetable lasagne); pulses; salads; fresh, frozen or canned fruit; dried fruit; and fruit in composites (apple pie). Portion size was defined as 80 g and visualised to respondents in tablespoons, cereal bowls and slices. The number of portions of fruits and vegetables were then summed for each respondent to estimate the number of portions of fruits and vegetables consumed in the day prior to interview. The policy recommendation at the time of the survey was five fruit or vegetable portions for adults per day (16) and this has not changed. We derived a binary indicator variable taking the value of one if the individual did not meet the recommended portion size per day; and zero otherwise.

Diseases associated with these health behaviours include cancers, heart diseases, stroke, hypertension, asthma, obesity as a disease, peripheral vascular diseases (PVC), and others identified by the relevant ICD9 and ICD10 codes defining the obesity-, smoking-, and alcohol-related diseases.

Previous evidence showed obesity-related diseases to be associated with diet and physical inactivity [e.g., (16–18)]. We observed the ICD9 and ICD10 codes identifying important diseases such as cancers, stroke and heart diseases tend to overlap, as the same codes identify different behaviour-related diseases. Therefore, disease incidence was defined as the presence of any disease-specific ICD9 and ICD10 codes in any of the diagnosis codes recorded in the data, either as primary diagnosis or cause of death. We used the survey interview date, hospital admission date for an episode, and discharge date to determine when a behaviour-related disease incidence occurred, either pre-survey, post-survey, or both.

A *pre-survey* incidence occurred if the individual was discharged on or before the interview date, whilst an incidence occurred *post-survey* if the individual was admitted on or after the survey interview date. We focus on the first incidence of a disease post-survey to derive the binary disease outcome variable (SMRpost), taking the value one if the individual had the first incidence of a behaviour-related disease *post-survey*; and zero otherwise. We expected each of the risk behaviours to be positively associated with post survey disease incidence.

Other Variables

Other variables included in the model as determinants of risky health behaviour and disease incidence are individual demographic, socioeconomic, environment, and self-reported health-related factors and disease history as measured by the presence of a disease incidence pre-survey (SMRpre). Demographic factors include gender, age and marital status. Age enters the model non-linearly via a squared age term in addition to age. Marital status was categorised into three binary variables: Single; Married or co-habiting; and Divorced, widowed, or separated.

The socio-economic characteristics were educational attainment and occupational status. Educational attainment was categorised into 5 groups; no educational qualifications; O-level or equivalent; A-level or equivalent; Further education or equivalent; and University degree or equivalent professional qualifications. Occupational status was categorised into Unskilled; Partly-skilled manual; Skilled manual; Skilled non-manual; Intermediate/managerial; and professional. We used area deprivation to proxy the individual's living environment, measured by the Scottish Index of Multiple Deprivation (SIMD) which is the Scottish Executive's official measure for identifying small area concentrations of multiple deprivations across Scotland (19). SIMD was categorised into ordered quintiles where SIMD51 = 1 indicated the least deprived and SIMD55 = 1 indicated the most deprived quintile.

Self-reported general health was categorised into three binaries: "Very good" (reference group), "Good", "Fair"; and "Bad/Very bad" health. The remaining self-reported health related factors were binary variables capturing "precondition" including; presence of a medical pre-condition (e.g., IHD, Type 2 diabetes, treatment for high blood pressure, limiting and

[1]The limit for men has since been reduced to the same level as for women.

non-limiting long-standing illness, hospital attendance for a cardiovascular disease (CVD), etc.,), and the presence of the first incidence of a behaviour-related disease incidence pre-survey. Both variables were intended to capture the effect of medical history including pre-survey hospitalisations.

Patterns of Risky Behaviours, Disease Incidence and Individual Characteristics

Restricting attention to adults aged 16 years and older, and excluding respondents whose status within the SMR was unknown at the end of the study period, our final sample consisted of $N = 20{,}751$ individuals. Descriptive statistics in **Table 1** showed health risk behaviours patterns and disease incidence. Figures were presented for the whole sample (ALL) and each risky behaviour separately:

At the time of interview, 50% of the total sample were physically inactive ($N = 10{,}157$), 36.5% reported to be current smokers ($N = 7{,}581$), 90.7% were failing the recommended five a day (inadequate diet) ($N = 18{,}829$), and 24.1% were drinking above moderate alcohol limits ($N = 4{,}991$). Cross-tabulations across risky behaviours (not in **Table 1** but available from the authors) show that 9.6% of respondents had at least one other risky behaviour ($N = 1{,}997$), whilst 4.7% had all the four health risky behaviours considered in this study ($N = 980$). Moreover, 28.5% of the sample experienced a behaviour-related disease incidence post-survey ($N = 5{,}903$). Within this sub-sample, 65% were physically inactive, 42% were smokers, 93% had an inadequate diet, and 21% consumed alcohol regularly above the moderate limit.

The average age in the sample was 46 years. However, current smokers, excessive drinkers, and individuals taking inadequate diet were slightly younger whilst the physically inactive were on average older. Those who experienced a disease incidence were on average 56 years old. Thus, we expected age to impact differentially across risk behaviours and disease incidence. In terms of gender, 46% percent of the whole sample were men. The proportion of males undertaking at least one risky behaviour was generally above 42% in the risk behaviour sub-groups, with the highest being drinking above the recommended limit (63%). Within the groups of individuals with a disease incidence 47% were men.

Regarding marital status, 58.5% of the sample were married or cohabiting. Across the health risk behaviours, married and cohabiting individuals were in the majority, and similarly for those with a disease incidence. Over a third (36%) of the sample had no formal educational qualification. Within the risk behaviour sub-groups, respondents with no educational qualification were more likely to undertake each of the risky behaviours: smoking (40.5%), excessive drinking (26.4%), physically inactive (50.6%), and inadequate diet (36.7%), and to have a disease incidence (54.2%). Thus, we expected the probability of undertaking a risky behaviour to decrease with education level.

With respect to occupational social status, a greater proportion of the total sample were in skilled manual and intermediate/managerial occupations, whilst individuals in the unskilled and professional categories were generally $<10\%$. We observe this pattern across all risk behaviours and disease incidence. Regarding deprivation quintiles, we expected the probability of undertaking each of the risky behaviours to decrease with lower levels of deprivation. This was largely because, individuals in the most deprived group were generally lower than a third of the total sample. In the most deprived group were; almost a third (30.3%) of smokers, almost a quarter (24.5%) of physically inactive, more than one-in-five (22%) of those taking inadequate diet, and over a quarter (27.1%) of those with a disease incidence.

For health-related characteristics, 37.9% reported a health-related pre-condition. Around a third of current smokers (33.4%) and excessive drinkers (34%) reported a pre-condition. respectively Almost two-in-five (37%) as well as over half (52%) of the total sample of those with a pre-condition were undertaking inadequate diet, and physically inactive, respectively. Furthermore, a majority of respondents with a pre-condition (60%) experienced a disease incidence. Therefore, we expected pre-condition to increase the probability of disease incidence but might decrease the probability of undertaking risky behaviours. The majority of respondents in the total sample reported their general health to be "Very good" (33.8%) or "Good" (39.3%), and a similar pattern prevailed across the risky behaviour and disease incidence sub-samples.

Finally, 7.2% of the total sample experienced a disease incidence *pre-survey*; 6% of current smokers; 5% of excessive drinkers; 11% of the inactive; and 7% of those with inadequate diet. Of those with a disease incidence *post-survey*, 19% also reported a pre-survey disease incidence. Thus, we expected the probability of a disease incidence post-survey to increase with health-related respondent characteristics.

Model

The multivariate probit model with endogenous variables belongs to the general class of simultaneous equations system and a generalisation of the bivariate probit model [e.g. (20, 21)]. Previous studies that applied the recursive multivariate probit model in health economics include Contoyannis and Jones (22), Balia and Jones (23), Di Novi (24) and Schneider and Schneider (25, 26). We also considered this model as appropriate to address our study objective. Firstly, the model explicitly accounts for correlations between unobserved characteristics that may determine both health risky behaviours and disease incidence. That is, the random components in the health risky behaviour equations were allowed to be freely correlated with the random component in the structural disease incidence equation. Secondly, the approach accounted for the potential endogeneity bias between health risk behaviours and disease incidence.

In our application, the recursive structure of the model was built around four reduced-form risk behaviour equations and a disease incidence structural equation with risk behaviour variables as predictors. For each latent dependent variable y^* for the ith individual and kth health risk behaviours [$k =$

TABLE 1 | Descriptive statistics of risky behaviours, disease outcome, and other characteristics ($N = 20,751$).

	All		Smoking[a]		Drinking[b]		Inactivity[c]		Diet[d]		SMRpost[e]	
Variable	**Mean**	**sd**	**Mean**	**sd**	**Mean**	**Sd**	**Mean**	**sd**	**Mean**	**sd**	**Mean**	**Sd**
Risky behaviours and SMRpost:												
Smoking	0.365	0.482			0.433	0.495	0.378	0.485	0.386	0.487	0.418	0.493
Drinking above limit	0.241	0.427	0.285	0.451			0.214	0.410	0.236	0.425	0.205	0.404
Inactive	0.500	0.500	0.513	0.500	0.450	0.498			0.506	0.500	0.649	0.477
Inadequate diet	0.907	0.290	0.958	0.201	0.891	0.311	0.921	0.269			0.930	0.256
SMRpost	0.285	0.451	0.326	0.469	0.242	0.429	0.371	0.483	0.291	0.454		
Characteristics:												
Gender (male)	0.455	0.498	0.461	0.499	0.628	0.483	0.429	0.495	0.461	0.499	0.470	0.499
Age	46.15	16.11	43.00	14.76	43.17	15.11	51.61	15.89	45.754	16.079	55.81	14.91
Age squared	2,389.3	1,555.1	2,067.3	1,338.7	2,091.6	1,383.2	2,916.5	1,640.7	2,351.97	1,545.18	3,337.6	1,604.9
Marital status:												
Single	0.220	0.414	0.264	0.441	0.273	0.445	0.173	0.378	0.225	0.418	0.152	0.359
Married or cohabiting	0.585	0.493	0.502	0.500	0.570	0.495	0.583	0.493	0.578	0.494	0.541	0.498
Divorced/wid./separated	0.195	0.396	0.234	0.423	0.157	0.364	0.244	0.429	0.197	0.398	0.307	0.461
Educational qual.:												
None [f]	0.358	0.479	0.405	0.491	0.264	0.441	0.506	0.500	0.367	0.482	0.542	0.498
education2 [g]	0.215	0.411	0.255	0.436	0.223	0.416	0.175	0.380	0.223	0.416	0.183	0.387
education3 [h]	0.126	0.331	0.108	0.310	0.155	0.362	0.103	0.304	0.126	0.331	0.086	0.281
education4 [i]	0.136	0.342	0.126	0.332	0.155	0.362	0.101	0.301	0.135	0.342	0.081	0.273
Degree /professional	0.150	0.357	0.084	0.277	0.192	0.394	0.106	0.307	0.132	0.339	0.088	0.284
education6 (missing)	0.016	0.125	0.022	0.148	0.012	0.108	0.010	0.100	0.017	0.129	0.019	0.136
Occ. social status:												
Unskilled/others	0.063	0.242	0.089	0.284	0.044	0.206	0.068	0.252	0.067	0.250	0.094	0.292
Partly skilled manual	0.141	0.348	0.188	0.391	0.128	0.335	0.155	0.362	0.148	0.355	0.168	0.374
Skilled manual	0.249	0.433	0.282	0.450	0.270	0.444	0.287	0.452	0.256	0.436	0.275	0.446
Skilled non-manual	0.187	0.390	0.172	0.378	0.157	0.364	0.176	0.381	0.188	0.391	0.164	0.370
Intermediate/managerial	0.271	0.445	0.202	0.401	0.305	0.461	0.250	0.433	0.257	0.437	0.235	0.424
Professional	0.056	0.229	0.029	0.167	0.072	0.258	0.038	0.190	0.049	0.217	0.035	0.184
Skill (missing)	0.033	0.179	0.038	0.191	0.024	0.152	0.026	0.160	0.035	0.184	0.028	0.166
Deprivation:												
Least deprived	0.172	0.377	0.111	0.315	0.204	0.403	0.131	0.338	0.165	0.371	0.116	0.320
Deprived = 2	0.196	0.397	0.145	0.352	0.197	0.398	0.178	0.383	0.189	0.392	0.162	0.369
Deprived = 3	0.211	0.408	0.200	0.400	0.197	0.398	0.219	0.414	0.209	0.407	0.214	0.410
Deprived = 4	0.211	0.408	0.241	0.428	0.206	0.405	0.227	0.419	0.216	0.412	0.237	0.425
Most deprived	0.210	0.408	0.303	0.460	0.196	0.397	0.245	0.430	0.220	0.414	0.271	0.444
Self-Assessed Health:												
Very good	0.338	0.473	0.262	0.440	0.350	0.477	0.266	0.442	0.332	0.471	0.192	0.394
Good	0.393	0.489	0.404	0.491	0.418	0.493	0.367	0.482	0.395	0.489	0.342	0.474
Fair	0.197	0.398	0.237	0.426	0.180	0.384	0.249	0.433	0.202	0.401	0.300	0.458
Bad/very bad	0.071	0.257	0.097	0.296	0.052	0.223	0.118	0.323	0.072	0.259	0.167	0.373
Pre-conditions [j]	0.379	0.485	0.334	0.472	0.336	0.472	0.518	0.500	0.370	0.483	0.599	0.490
Disease pre-survey (SMRpre) [k]	0.072	0.259	0.058	0.233	0.051	0.220	0.111	0.314	0.070	0.256	0.188	0.391

[a] *Smoking, current smoker.*
[b] *Drinking, alcohol consumption above the recommended limit (underlimit is 1–14 units for females and 21 units for male=).*
[c] *Inactivity, undertaking no physical activity.*
[d] *Diet, Inadequate diet—not meeting the recommended portions of fruits and vegetable per day.*
[e] *SMRpost,1 if the individual has a behaviour related diseases incidence post-survey.*
[f] *No education qualification.*
[g] *Low level, School leaving, CSE, etc.*
[h] *Low mid. level, SVQ, "A" levels, OND, etc.*
[i] *Upmided, C&G, HNC, HND, etc.*
[j] *Preconditions include IHD, Type 2 diabetes, treatment for high blood pressure, limiting and non-limiting long-standing illness, hospital attendance for a cardiovascular disease, etc.*
[k] *SMRpre, 1 if the individual has a behaviour related diseases incidence post-survey.*

smoking (*smk*), excessive drinking (*drk*), physical inactivity (*inact*), inadequate diet (*diet*) and disease incidence (*dis*)];

$$y^*_{ik} = \beta'_k X_{ik} + \varepsilon_{ik} \quad (1)$$

Let $y_{ik} = 1$ if $y^*_{ik} > 0$; and 0 otherwise, the corresponding binary observed outcome y for each of the kth health risk behaviours and disease incidence can be stated as:

$$\begin{aligned} y_{i,smk} &= \beta'_{smk} X_{i,smk} + \varepsilon_{i,smk}, \\ y_{i,drk} &= \beta'_{drk} X_{i,drk} + \varepsilon_{i,drk}, \\ y_{i,inact} &= \beta'_{inact} X_{i,inact} + \varepsilon_{i,inact}, \\ y_{i,diet} &= \beta'_{diet} X_{i,diet} + \varepsilon_{i,diet}, \\ \text{and} & \\ y_{i,dis} &= \theta_{smk} y_{i,smk} + \theta_{drk} y_{i,drk} + \theta_{inact} y_{i,inact} + \theta_{diet} y_{diet} \\ &\quad + \beta'_{dis} Z_{i,dis} + \varepsilon_{i,dis} \end{aligned} \quad (2)$$

where ε_{ik}were the error terms distributed as multivariate normal, each with a mean of zero and a variance-covariance matrix of the cross-equation error terms, V, which has values of 1 on the leading diagonal; and correlations $\rho_{jk} = \rho_{kj}$ off-diagonal elements for j, $k = 1, \ldots, M$ and $j \neq k$. The parameter ρ_{jk} to be estimated, provides the correlations between the error terms of equations j and k, measuring the extent to which unobserved characteristics of individuals influence their risky behaviour and associated disease incidence. X_{ik} and $Z_{i,dis}$ were vectors of explanatory variables in each of the risky behaviour equations and the disease incidence equation, respectively, whilst β_k and θ_k were parameter vectors to be estimated.

In application, the vector of explanatory variables in X differed from those in vector Z (i.e., $X \neq Z$) due to exclusion restrictions on the structural equation. Equation (2) has the structure of a Zellner (27) Seemingly Unrelated Regression model, except that the dependent variables y_{ik}, are binaries and the explanatory variables need not be the same in all equations. The parameter vectors, θ_{smk}, θ_{drk}, θ_{inact}, θ_{diet}, β_{dis} and ρ were estimated using the MVPROBIT programme developed by Capellari and Jenkins (28) and implemented in STATA.

Finally, in addition to the non-linear functional form of the model, identification of the parameters was achieved by imposing excluded restrictions on the structural equations. Madalla (21) suggested omission of at least one explanatory variable in the reduced-form equations from the structural equation as a predictor variable. Previous applications of this model to health, commonly excluded socioeconomic indicators from the structural equation [e.g., (22–24, 26)]. We follow this approach, assuming that socioeconomic characteristics only indirectly affect disease incidence through lifestyles.

RESULTS

Tables 2–4 present estimation results from the multivariate probit specifications of the full recursive model in Equation (2) for the whole sample, female and male sub-samples, respectively. We tested the null hypothesis H_o: $\rho = 0$ via likelihood ratio test, that the error terms of the equations were jointly statistically equal to zero. This test rejected the null hypothesis of exogeneity, suggesting that estimation of the equations independently of one another would be inefficient.

The first four columns in **Tables 2–4** present the partial effects for the reduced form risk behaviour equations. In **Table 2** for the whole sample, males were strongly more likely than females to undertake risky behaviours; smoking, excessive drinking and inadequate diet. These results agree with the findings from German data by Schneider and Schneider (26). However, males were less likely than females to be physically inactive.

The age effect differed across health risky behaviours and between the estimated models. Generally, age had a quadratic effect on the probability of smoking and excessive drinking. Here, the function was inverted U-shaped, suggesting that the probability of smoking and drinking excessively was increasing with age initially, and then decreased. The critical age (in years) at which this change occurred differed across the estimated models, but was statistically significant at 5% or better. For the whole sample, the critical age at which the probability of smoking began to diminish was 34.5 years, compared to 30.8 years for excessive drinking. However, the diminishing smoking and excessive drinking probabilities occurred at a relatively higher age for males (**Table 4**) than for females (**Table 3**). For males, the diminishing effect occurred at around aged 35 years on average, compared to aged 33 years and ∼22 years for females, respectively.

The age effect for physical inactivity and inadequate diet differed between the whole sample and by gender. For the whole sample, the probability of inactivity and inadequate diet was increasing monotonically and the critical ages were statistically insignificant (**Table 2**). The functional form of the effect of age on physical inactivity and inadequate diet differed markedly between females (**Table 3**) and males (**Table 4**). For the female sample, the functional forms for physical inactivity and inadequate diet were U-shaped, respectively. This suggests an increasing effect of age on these health risky behaviours. Thus, the critical age at which physical inactivity increased amongst females was strongly statistically significant at ∼18 years, but insignificant for inadequate diet. For the male sample however, age has a diminishing effect on inadequate diet with an inverted U-shaped function and the critical age at which inadequate diet decreased amongst males was statistically significant at ∼37 years, but insignificant for physical inactivity.

Regarding marital status, married or cohabiting individuals were less likely than singles to undertake all the health risky behaviours across all models, except physical inactivity and inadequate diet, in which the negative effect was statistically insignificant (**Table 2**). Divorced, separated or widowed women were more likely to be smokers than singles and less likely to drink excessively (**Table 3**). For men however, being divorced, separated or widowed was positively associated with

TABLE 2 | Recursive multivariate probit regression: Whole sample.

	Smoking	Drinking	Inactivity	Diet	SMRpost[a]
Variables	**Coefficient [b]**	**Coefficient**	**Coefficient**	**Coefficient**	**Coefficient**
Characteristics:					
Gender (male)	0.073***	0.523***	−0.093***	0.212***	0.048
Age	0.040***	0.013***	0.0065*	0.0027	0.016***
Age squared	−0.00058***	−0.00021***	0.00015***	−0.00011**	0.00016***
Critical age (Years)[c]	34.5***	30.8***	21.8	13.1	51.4*
Marital status:[d]					
Married/cohabiting	−0.208***	−0.162***	−0.022	−0.052	
Divorced widowed/separated	0.145***	−0.101***	−0.051	0.081*	
Education:[e]					
Low level	−0.051*	0.123***	−0.390***	−0.029	
Low middle level	−0.278***	0.225***	−0.336***	−0.200***	
Upper middle level	−0.271***	0.115***	−0.416***	−0.249***	
Degree/professional qual.	−0.451***	0.231***	−0.509***	−0.516***	
Education (missing)	0.133*	−0.107	−0.657***	−0.297***	
Occupational social status:[f]					
Partly skilled manual	−0.054	0.114**	0.173**	−0.087**	
Skilled manual	−0.176***	0.166***	0.237***	−0.339***	
Skilled non-manual	−0.264***	0.110**	0.106**	−0.343***	
Intermediate/managerial	−0.260***	0.216***	0.151***	−0.470***	
Professional	−0.395***	0.180***	0.040	−0.544***	
Skill (missing)	−0.238***	−0.156**	0.064	−0.117	
Deprivation:[g]					
Deprived = 2	0.036	−0.094***	0.094***	−0.030	
Deprived = 3	0.166***	−0.122***	0.185***	0.009	
Deprived = 4	0.274***	−0.059*	0.171***	0.092**	
Most deprived	0.445***	−0.066*	0.248***	0189***	
SAH and precondition:[h]					
SAH (Good)	0.260***	0.055**	0.058***	107***	0.079***
SAH (Fair)	0.467***	0.029	0.198***	0.225***	0.208***
SAH (bad/very bad)	0.656***	−0.0647	0.580***	0.172***	0.515***
Precondition[i]	−0.263***	−0.0025	0.350***	−0.229***	0.264***
SMRpre[j]	−0.110***	−101**	0.055	−0.064	0.648***
Endogenous risky behaviours:					
Smoking					1.086***
Drinking					0.034
Inactivity					0.017
Inadequate diet [k]					0.373***
Constant	−0.763***	−1.168***	−0.831***	1.884***	−3.067***
LR test of Rho's [l]: $\chi^2_{(10)}$	545.5, Prob > $\chi^2_{(10)} = 0$				
Log partial likelihood			−49,521.4		
Observations	20,332	20,332	20,332	20,332	20,332

[a] *SMRpost, first incidence of a risk related behaviour post survey.*
[b] *The asterisks: ***, **, and * indicate statistical significance at the $p < 0.01$, $p < 0.05$, $p < 0.1$, respectively.*
[c] *Critical age, ($\hat{\beta}_{age}/2*(\hat{\beta}_{agesq})$)in absolute value terms.*
[d] *Reference, Single.*
[e] *Reference, No educational qualification, Low level, School leaving, CSE, etc., Low middle level, SVQ, "A" levels, OND, etc., Upper middle level, City and Guilds, Higher National Diploma, etc.*
[f] *Reference, Unskilled/others.*
[g] *Reference, Least deprived (Deprived = 1).*
[h] *SAH, Self-assessed health, reference, "Very good".*
[i] *Preconditions include IHD, Type 2 diabetes, treatment for high blood pressure, limiting and non-limiting long-standing illness, hospital attendance for a cardiovascular disease, etc.*
[j] *SMRpre, 1 if the individual has a behaviour related diseases incidence post-survey.*
[k] *Inadequate diet is defined in terms of not meeting the recommended five portions a day fruit and vegetable consumption.*
[l] *Likelihood Ratio test of the null hypothesis that the covariance parameters (ρ_{jk}) are jointly statistically not different from zero.*

TABLE 3 | Recursive multivariate probit regression: Females.

	Smoking	Drinking	Inactivity	Diet	SMRpost[a]
Variables	**Coefficient[b]**	**Coefficient**	**Coefficient**	**Coefficient**	**Coefficient**
Characteristics:					
Age	0.038***	0.0091	−0.0102**	−0.010	0.014***
Age squared	−0.00057***	−0.00021***	0.0003***	1.00e-05	0.00014**
Critical age (Years)[c]	33.0***	21.9***	17.7***	49.5	60.0
Marital status:[d]					
Married/cohabiting	−0.221***	−0.160***	−0.004	−0.059	
Divorced/ widowed/separated	0.138***	−0.241***	−0.026	0.017	
Education:[e]					
Low level	−0.053	0.117***	−0.345***	−0.039	
Low middle level	−0.370***	0.221***	−0.393***	−0.242***	
Upper middle level	−0.318***	0.030	−0.518***	−0.162***	
Degree/professional qual.	−0.535***	0.346***	−0.521***	−0.537***	
Education (missing)	−0.0222	−0.463***	−0.785***	−0.269	
Occupational social status:[f]					
Partly skilled manual	−0.036	0.172**	0.169***	−0.148	
Skilled manual	−0.169***	0.144*	0.308***	−0.385***	
Skilled non-manual	−0.285***	0.234***	0.079	−0.300***	
Intermediate/managerial	−0.216***	0.332***	0.162***	−0.496***	
Professional	−0.433***	0.290***	0.133	−0.629***	
Skill (missing)	−0.128*	0.088	0.057	−0.124	
Deprivation: [g]					
Deprived = 2	0.0017	−0.107**	0.080*	−0.023	
Deprived = 3	0.207***	−0.193***	0.188***	−0.029	
Deprived = 4	0.280***	−0.187***	0.173***	0.095*	
Most deprived	0.436***	−0.170***	0.242***	0.144**	
SAH and precondition:[h]					
SAH (Good)	0.232***	0.039	0.017	0.069*	0.130***
SAH (Fair)	0.426***	−0.088*	0.125***	0.176***	0.227***
SAH (bad/very bad)	0.543***	−0.163**	0.541***	0.140*	0.568***
Precondition[i]	−0.226***	0.073**	0.320***	−0.187***	0.280***
SMRpre[j]	−0.071	−0.108	0.118*	−0.052	0.664***
Endogenous risky behaviours:					
Smoking					1.007***
Drinking					−0.463**
Physically Inactive					−0.031
Inadequate diet[k]					0.414*
Constant	−0.658***	−0.997***	−0.359***	2.245***	−2.541***
LR test of Rho's [l]: $\chi^2_{(10)}$			chi2(10) = 326.98, Prob > $\chi^2_{(10)}$ = 0		
Log partial likelihood			−26,519.76		
Observations	11,169	11,169	11,169	11,169	11,169

[a] *SMRpost, first incidence of a risk related behaviour post survey.*
[b] *The asterisks: ***, **, and * indicate statistical significance at the $p < 0.01$, $p < 0.05$, $p < 0.1$, respectively.*
[c] *Critical age, ($\hat{\beta}_{age}/2^*(\hat{\beta}_{agesq})$)in absolute value terms.*
[d] *Reference, Single.*
[e] *Reference, No educational qualification, Low level, School leaving, CSE, etc., Low middle level, SVQ, "A" levels, OND, etc., Upper middle level, City and Guilds, Higher National Diploma, etc.*
[f] *Reference, Unskilled/others.*
[g] *Reference, Least deprived (Deprived = 1).*
[h] *SAH = Self-assessed health, reference, "Very good".*
[i] *Preconditions include IHD, Type 2 diabetes, treatment for high blood pressure, limiting and non-limiting long-standing illness, hospital attendance for a cardiovascular disease, etc.*
[j] *SMRpre, 1 if the individual has a behaviour related diseases incidence post-survey.*
[k] *Inadequate diet is defined in terms of not meeting the recommended five portions a day fruit and vegetable consumption.*
[l] *Likelihood Ratio test of the null hypothesis that the covariance parameters (ρ_{jk}) are jointly statistically not different from zero.*

TABLE 4 | Recursive multivariate probit regression: Males.

	Smoking	Drinking	Inactivity	Diet	SMRpost[a]
Variables	**Coefficient [b]**	**Coefficient**	**Coefficient**	**Coefficient**	**Coefficient**
Characteristics:					
Age	0.041***	0.014**	0.026***	0.019**	0.017***
Age squared	−0.00058***	−0.0002***	−2.52e-05	−0.00026***	0.00019***
Critical age (Years)[c]	35.4***	35.3***	50.7	37.1***	44.2
Marital status:[d]					
Married/cohabiting	−0.214***	−0.134***	−0.054	−0.052	
Divorced/ widowed/separated	0.135***	0.110**	−0.061	0.228***	
Education:[e]					
Low level	−0.052	0.103**	−0.465***	−0.039	
Low middle level	−0.177***	0.198***	−0.294***	−0.171**	
Upper middle level	−0.223***	0.180***	−0.328**	−0.351***	
Degree/professional qual.	−0.365***	0.118**	−0.505***	−0.514***	
Education (missing)	0.309***	0.091	−0.501***	0.324	
Occupational social status:[f]					
Partly skilled manual	−0.086	0.0541	0.165**	−0.281*	
Skilled manual	−0.202***	0.111	0.179**	−0.335**	
Skilled non-manual	−0.266***	−0.0413	0.166**	−0.507***	
Intermediate/managerial	−0.313***	0.103	0.130*	−0.466***	
Professional	−0.407***	0.113	−0.053	−0.506***	
Skill (missing)	−0.424***	−0.439***	−0.0015	−0.130	
Deprivation:[g]					
Deprived = 2	0.0728	−0.089**	0.110**	−0.038	
Deprived = 3	0.110***	−0.064	0.186***	0.066	
Deprived = 4	0.260***	0.0432	0.167***	0.073	
Most deprived	0.446***	0.016	0.246***	0.242***	
SAH and precondition:[h]					
SAH (Good)	0.290***	0.081**	0.109***	0.156***	0.021*
SAH (Fair)	0.510***	0.134***	0.294***	0.295***	0.179***
SAH (bad/very bad)	0.793***	0.0068	0.624***	0.215**	0.404***
Precondition[i]	−0.303***	−0.057*	0.393***	−0.279***	0.293***
SMRpre[j]	−0.167***	−0.108*	−0.046	0.075	0.596***
Endogenous risky behaviours:					
Smoking					1.290***
Drinking					−0.261
Physically Inactive					−0.181
Inadequate diet[k]					0.861***
Constant	−0.746***	−0.722***	−1.451***	1.701***	−3.139***
LR test of Rho's [l]: $\chi^2_{(10)}$			Chi2(10) = 218.51, Prob > $\chi^2_{(10)}$ = 0		
Log partial likelihood			−22,749.8		
Observations	9,163				9,163

[a] *SMRpost, first incidence of a risk related behaviour post survey.*
[b] *The asterisks: ***, **, and * indicate statistical significance at the $p < 0.01$, $p < 0.05$, $p < 0.1$, respectively.*
[c] *Critical age, $(\hat{\beta}_{age}/2 * (\hat{\beta}_{agesq}))$in absolute value terms.*
[d] *Reference, Single.*
[e] *Reference, No educational qualification, Low level, School leaving, CSE, etc., Low middle level, SVQ, "A" levels, OND, etc., Upper middle level, City and Guilds, Higher National Diploma, etc.*
[f] *Reference, Unskilled/others.*
[g] *Reference, Least deprived (Deprived = 1).*
[h] *SAH, Self-assessed health, reference, "Very good".*
[i] *Preconditions include IHD, Type 2 diabetes, treatment for high blood pressure, limiting and non-limiting long-standing illness, hospital attendance for a cardiovascular disease, etc.*
[j] *SMRpre, 1 if the individual has a behaviour related diseases incidence post-survey.*
[k] *Inadequate diet is defined in terms of not meeting the recommended five portions a day fruit and vegetable consumption.*
[l] *Likelihood Ratio test of the null hypothesis that the covariance parameters (ρ_{jk}) are jointly statistically not different from zero.*

smoking, excessive drinking and inadequate diet (**Table 4**). For both females and males, marital status had no effect on physical inactivity.

For education, the results suggested that whilst the more educated were less likely than those with no education to be smoking, physically inactive and taking inadequate diet, they were also more likely to drink excessively. The result on drinking is consistent with findings by Schneider and Schneider (26), in which the result was associated with social acceptance of alcohol consumption amongst the more educated. The results for education attainment were consistent across all estimated models.

The effect of occupational social status also differed across the health risky behaviours and by gender. Generally, and particularly for the whole and female sample, smoking and inadequate diet probability decreased progressively with higher occupation levels, whilst the probability of excessive drinking and physical inactivity increased generally with higher occupational status (**Tables 2**, **3**).[2]

Similarly for males (**Table 4**), the probability of smoking and inadequate diet decreased progressively with higher occupational status, whilst the probability of being physical inactive increased. In contrast to females, the effect of occupational status on the probability of excessive drinking was not statistically different from zero. Overall, these results suggested that whilst the more skilled individuals were less likely to smoke or have an inadequate diet, they were more likely than those with no skills to be physically inactive, particular females. Results for excessive drinking were less clear-cut.

The deprivation effect also differed across health risky behaviours and by gender. The most deprived females (SIMD55) were more likely than females in the least deprived quintile (SIMD51) to be smoking, physically inactive, and had an inadequate diet, but were less likely to drink excessively.[3] Males in the most deprived quintile were also more likely than males in the least deprived quintile to be smoking, physically inactive, and taking inadequate diet. There was no statistical evidence supporting the influence of deprivation on excessive drinking.[4, 5]

Generally, the probability of smoking, physical inactivity, and inadequate diet was positively associated with worse self-assessed general health (**Table 2**). However, the relationship between health risky behaviours and self-assessed general health also differed by gender. For females, the probability of smoking, physical inactivity, and inadequate diet was positively associated with worse self-assessed general health, whilst the probability of excessive drinking was negatively associated with worse self-assessed general health.

Further, having a self-reported precondition decreased the probability of smoking and inadequate diet, but increased the probability of physical inactivity in all samples. However, females who reported a precondition were also more likely to drink excessively whilst for males, a precondition reduced the likelihood of excess drinking. Additionally, a pre-survey disease incidence (SMRpre) decreased smoking and drinking probability in the whole sample and for males. For females however, a pre-survey disease incidence was only statistically associated with physical inactivity.

The last columns in **Tables 2–4** present results for the structural disease incidence equation. For the whole sample, there was no statistical evidence of gender differences in the probability of a behaviour related disease incidence. As expected, the probability of a behaviour-related disease incidence increased monotonically with age across all estimated models, but the critical age was only statistically significant for the whole sample.

For the health-related characteristics, those in worse self-assessed health were strongly more likely than those assessing their heath as "Very Good", to suffer a behaviour-related disease incidence post-survey. The probability of a disease incidence increased progressively with worsening self-assessed health. Experiencing a behaviour-related disease incidence pre-survey or self-reported pre-condition was positively associated with experiencing a disease incidence post-survey.

For endogenous risk behaviour variables, it should be noted that the estimates were conditional on the individual characteristics influencing each of the health risk behaviours. Therefore, the estimates were the direct and indirect effects on disease incidence. There was a gender difference in the impact of excessive drinking on the probability of a disease incidence post-survey, which was negatively associated for females and insignificant for males. For males and females, smoking and inadequate diet showed a positive statistically significant effect on the probability of disease incidence post-survey whilst physical inactivity was statistically insignificant.

Finally, we assessed dependency between the estimated equations and tested whether the health risky behaviours were endogenous to disease incidence. **Table 5** presents the results. Each sub-table reports the pairwise correlations between the health behaviours and disease incidence and the correlation coefficients ρ_{jk}.[6] The results showed that most health risk behaviours were positively correlated and coefficients were highly statistically significant, except in the males sample, where the correlation between smoking and inactivity ($\rho_{smoking * inactive}$) and between drinking and inadequate diet were insignificant.

These results suggested that unobserved factors influencing the likelihood of one risky health behaviour also contribute positively to the likelihood of other health risky behaviours and disease incidence. Secondly, only smoking and unhealthy diet were significantly negatively correlated with a disease incidence. This suggested that these healthy risky behaviours

[2]Statistically insignificance was based on joint tests post-estimation, which gave Chi squared estimates and associated p-value.

[3]For inadequate diet, the coefficient were statistically significant only at the 4th and the 5th quintiles. However, they were jointly significant: Chi2(4) = 14.08, p-value = 0.007.

[4]For excessive drinking, only the coefficient on the 2nd quintile is statistically significant and a joint significance test of the quintiles cannot reject the null hypothesis that the effect of deprivation on excessive drinking is not different from zero: Chi2(4) = 13.16, p-value = 0.105.

[5]For inadequate diet, only the coefficient on the 5th quintile is statistically significant and a joint significance test suggests that the quintiles are jointly significant: Chi2(4) = 16.66, p-value= 0.0023.

[6]In estimation, the statistic $z = \hat{\rho}/s_{\hat{\rho}}$ was computed, where was the asymptotic standard error of $\hat{\rho}$.

TABLE 5 | Correlation matrices for the restricted and the full model[a].

	Smoking	Drinking	Inactivity	Diet	Incidence
Whole sample					
Smoking	1.00				
	1.00				
Excessive drinking	0.141***	1.00			
	*0.149****	*1.00*			
Physical inactivity	0.034***	−0.096***	1.00		
	*0.055****	*0.005*	1.00		
Inadequate diet	0.318***	−0.061***	0.086***	1.00	
	*0.211****	*-0.069****	*0.118****	*1.00*	
Disease incidence	0.115***	−0.099***	0.308***	0.115***	1.00
	*-0.525****	*-0.068*	*-0.063*	*-0.287****	*1.00*
Female sample					
Smoking	1.00				
	1.00				
Excessive drinking	0.124***	1.00			
	*0.183****	*1.00*			
Physical inactivity	0.044***	−0.127***	1.00		
	*0.060****	*-0.008*	*1.00*		
Inadequate diet	0.307***	−0.153***	0.083***	1.00	
	*0.195****	*-0.121****	*0.109****	*1.00*	
Disease incidence	0.091***	−0.208***	0.269***	0.108***	1.00
	*-0.480****	*0.091*	*0.108*	*-0.277****	*1.00*
Male sample					
Smoking	1.00				
	1.00				
Excessive drinking	0.158***	1.00			
	*0.120****	*1.00*			
Physical inactivity	0.025	−0.051***	1.00		
	*0.045***	*0.009*	*1.00*		
Inadequate diet	0.333***	−0.004	0.095***	1.00	
	*0.239****	*-0.012*	*0.123****	*1.00*	
Disease incidence	0.146***	−0.033*	0.362***	0.125***	1.00
	*-0.597****	*0.151*	*0.068*	*-0.367****	*1.00*

[a] *The figures for the full models are shown in italics.*
*The asterisks: ***, **, and * indicate statistical significance at $p < 0.01$, $p < 0.05$, $p < 0.1$, respectively.*

were endogenous to disease risk and should be modelled jointly with disease incidence. These general patterns of correlation were observed across all estimated models.

However, this study is without limitations. Firstly, our modelling approach did not account for potential sources of endogeneity bias. Specifically, self-assessed general health included as an explanatory variable may be jointly determined with the health outcomes. Also, the exclusion of year dummies from the model constrained an understanding of whether the year of survey predicted the outcomes. Future research should consider these limitations in modelling approach.

DISCUSSION

This study was motivated by the need to understand the interrelationships between health risky behaviours and their effect on related health outcomes such as disease incidence. This was considered crucial in designing behaviour intervention programmes. Incidence of many of the main diseases and premature deaths tend to be influenced by multiple health risky behaviours, such as those considered in this study. Risky behaviours tend to be dependent and they may have a multiplicative effect on disease incidence and healthcare cost.

We examined interrelationships between health risky behaviours smoking, excessive drinking, physical inactivity and inadequate diet and associated disease outcomes, using linked healthcare data on Scottish adults aged 16 and above. Specifically, we sought to assess the demographic, socioeconomic, and other factors associated with health risky lifestyle choices and to investigate whether or not these risky lifestyle choices have a significant influence on the probability of having a behaviour-specific disease incidence.

The study adds to the literature on the modelling of the determinants of multiple health behaviours, particularly risky behaviours, and how these determinants might differ by gender. This was made possible by use of a unique data set linking individual characteristics directly to their hospitalisation episodes. This avoided biases often associated with wholly self-reported data.

A key advantage of using linked healthcare dataset was that the record linkage provided a way to approximate causality since the health outcome was behaviour-specific. Also, the modelling strategy employed enabled examination of associations between respondent characteristics and each health risky behaviour, and the overall direct and indirect effects they have on disease incidence. Moreover, focusing on health risky behaviours rather than health protective behaviours was premised on the growing evidence that targeting multiple risky behaviours rather than a single risky behaviour at a time, offers the potential for maximising health promotion and intervention benefit, and reducing costs associated with behaviour-related disease outcomes [e.g., (9, 29)].

Understanding the interrelationships between risky health behaviours and associated disease outcomes is useful for intervention design aiming to promote multiple risky behaviour change at a time. This understanding is based on potential spillover effects arising when improvement in one behaviour increases the likelihood of adopting other healthier lifestyle behaviours.[7] That is; a beneficial single behaviour change may incentivize adoption of another healthy behaviour (30).

The results showed clear demographic and socioeconomic gradients in the risky behaviour determinants, particularly smoking, excessive drinking and inadequate diet, but are less clear cut for physical inactivity. Gender differences in the observed characteristics influencing risky behaviours and disease outcome were found. Undertaking risky behaviours increased generally at younger ages and tended to decrease as one becomes older. However, males were more likely than females to undertake risky behaviours longer in their lives, such as smoking, excessive drinking and inadequate diet. The estimate of the critical ages suggested that physical inactivity amongst females could begin as early as $<$13 years of age.

Our findings also support education as a key driver of better health behaviour irrespective of gender. A social gradient in healthy risk behaviours was also found in the analysis given that our findings suggested that a lower social standing, as measured by deprivation, induced health risky behaviour which increased the probability of associated disease incidence.

Moreover, we found that disease incidence was positively correlated with smoking and inadequate diet, whilst these same lifestyles were negatively correlated with having a precondition. This an important result suggesting that having a pre-condition may motivate healthier lifestyles (in terms of reduction in smoking and inadequate diet) in the future. It would be of interest to understand how these results generalise into other populations. In summary, demographic and socioeconomic characteristics such as age, gender, marital status, education, the environment, and having a precondition were some of the key elements influencing health risky behaviours and indirectly behaviour-related health outcome. Hence, designing intervention programmes targeting health risky behaviours may be more beneficial by greater investments in education and improving the environment to be more conducive for the younger age groups. Improvement in the modelling approach in future research will address the limitations.

[7]The idea of spillover effect here is synonymous with the concept of co-variation amongst multiple health behaviours in the literature on preventive medicine [e.g. (29)].

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because permission to access to the Linked SheS dataset used in this study can be granted by application to the Privacy Advisory Committee, Information Services Division of the NHSScotland. Requests to access the datasets should be directed to https://www.isdscotland.org/.

AUTHOR CONTRIBUTIONS

DO: conceptualisation, econometric modelling and analysis, and initial draught preparation and writing. BE: conceptualisation, analytical support, literature review, and review and editing. AL: conceptualisation, literature review, supervision, and review and editing. All authors have read and agreed to the final draft of the manuscript.

FUNDING

The receipt of financial support from the MRC National Preventive Research Initiative Phase 2 Grant G0701874 is greatly acknowledged. The Funding Partners relevant to this award are: British Heart Foundation, Cancer Research UK, Department of Health, Diabetes UK, Economic and Social Research Council, Medical Research Council, Research and Development Office for the Northern Ireland Health and Social Services, Chief Scientist Office, Scottish Government Health Directorates, The Stroke Association, Welsh Assembly Government, and World Cancer Research Fund. The Health Economics Research Unit is funded by the Chief Scientist Office of the Scottish Government Health Directorates. The views expressed in this article are those of the authors.

ACKNOWLEDGMENTS

The team is grateful to the Privacy Advisory Committee of the Information Service Division, NHSScotland for granting the permission to access the linked SHeS SMR datasets. We are also grateful for the identification of disease-specific ICD9 and ICD10 codes. We are grateful for the helpful comments received at national and international conferences and workshop where the findings of the paper were presented.

REFERENCES

1. World Health Organisation. *Diet, Nutrition and the Prevention of Chronic Diseases*. WHO Technical Report Series, 916. Geneva: World Health Assembly, Resolution 55.23 (2003).
2. ScotPHO. *ScotPHO Smoking Ready Reckoner – 2011 Edition.* (2012). Available online at: https://www.scotpho.org.uk/media/1186/scotpho120626_smokingreadyreckoner.pdf
3. Scottish Government. *Scotland's People Annual report: Results from 2009/2010 Scottish Household Survey*. (2011).
4. Scarborough P, Bhatnagar P, Wickramasinghe KK, Foster C, Rayner M. The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to 2006–07 NHS costs. *J Public Health.* (2011) 4:527–35. doi: 10.1093/pubmed/fdr033
5. Hanlon P, Lawder R, Elders A, Clark D, Walsh D, Whyte B, et al. An analysis of the link between behavioural, biological and social risk factors and subsequent hospital admission in Scotland. *J Public Health*. (2007) 29:405–412. doi: 10.1093/pubmed/fdm062
6. Lawder R, Elders A, Clark D. *Using the linked Scottish health survey to predict hospitalisation & death: An analysis of the link between behavioural, biological and social risk factors and subsequent hospital admission and death in Scotland*. Technical Report. Information Services Division (ISD) of NHS Scotland. (2007). Available online at: https://www.scotpho.org.uk/media/1243/scotpho070816_linkedshes_techrep.pdf
7. Pronk NP, Martinson B, Kessler RC, Beck AL, Simon GE, Wang P. The association between work performance and physical activity, cardiorespiratory fitness, and obesity. *J. Occup. Environ*. (2004) 46:19–25. doi: 10.1097/01.jom.0000105910.69449.b7
8. Fine LJ, Philogene GS, Gramling R, Coups EJ, Sinha S. Prevalence of multiple chronic disease risk factors: 2001 national health interview survey. *Am J Prev Med*. (2004) 27:18–24.
9. Prochaska J, Spring B, Nigg CR. Multiple health behaviour change research: an introduction and overview. *Prevent Med.* (2008) 46:181–8. doi: 10.1016/j.ypmed.2008.02.001
10. Edington DW. Emerging research: a view from one research center. *Am J Health Promot.* (2001) 15:341–9. doi: 10.4278/0890-1171-15.5.341
11. USDHHS. *Mental health; a report of the Surgeon General*. Washington, DC: Department of Health and Human Services (1999).
12. Scottish Executive Department of Health. *The Scottish Health Survey 2003*. Edinburgh: HMSO (2005).
13. Scottish Executive Department of Health. *The Scottish Health Survey 1998*. Edinburgh: HMSO (2000).
14. Scottish Office Department of Health. *The Scottish Health Survey 1995*. Edinburgh: HMSO (1997).
15. Beeston C, McAdams R, Craig N, Gordon R, Graham L, MacPherson M, et al. *Monitoring and Evaluating Scotland's Alcohol Strategy, Final Report.* Edinburgh: NHS Health Scotland (2016).
16. Scottish Government. *Healthy Eating, Active Living: An action plan to improve diet, increase physical activity and tackle obesity (2008 – 2011)*. Edinburgh: St Andrew's House (2008).
17. Popkin BM, Kim S, Rusev ER, Du S, Zizza C. Measuring the full economic costs of diet, physical activity and obesity-related chronic diseases. *Obes Rev*. (2006) 7:271–93. doi: 10.1111/j.1467-789X.2006.00230.x
18. WHO. *Global Strategy on Diet, Physical Activity and Health*. WHO (2004). Available online at: https://www.who.int/dietphysicalactivity/strategy/eb11344/en/index.html (accessed September 27, 2012).
19. Scottish Executive. (2006). *Using the Scottish Index of Multiple Deprivation 2004: Guidance*. Edinburgh: St Andrew's House.
20. Cappellari L, Jenkins SP. Multivariate probit regression using simulated maximum likelihood. *Stata J.* (2003) 3:278–94 doi: 10.1177/1536867X0300300305
21. Maddala GS. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press (1983).
22. Contoyannis P, Jones AM. Socio-economic status, health and lifestyle. *J Health Econ*. (2004) 23:965–95. doi: 10.1016/j.jhealeco.2004.02.001
23. Balia S, Jones AM. *Mortality, Lifestyle and Socioeconomic Status*. York: HEDG Working Paper, 05/02, University of York (2004).
24. Di Novi C. "An Economic Evaluation of Lifestyle and Air-pollution-related Damages: Results from the BRFSS∗," in *Cinzia Di Novi† JEPS Working Paper No. 07-001*. (2007). Available online at: http://jeps.repec.org/papers/07-001.pdf
25. Schneider U, Schneider BS. The effects of education and working hours on health: A multivariate probit approach. (2006).
26. Schneider U, Schneider BS. Willing to be healthy? On the health effects of smoking, drinking and an unbalanced diet. A multivariate probit approach. (2009).
27. Zellner A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *J Am Stat Assoc.* (1962) 57:348–68. doi: 10.1080/01621459.1962.10480664
28. Cappellari L, Jenkins S. Multivariate probit regression using simulated maximum likelihood. *Stata J.* (2003) 3:278–94.
29. Prochaska J. Multiple Health Behaviour Research represents the future of preventive medicine. *Prevent Med.* (2008) 46:281–5. doi: 10.1016/j.ypmed.2008.01.015
30. Hyman DJ, Pavlik VN, Taylor WC, Goodrick G, Moye K. Simultaneous vs. sequential counselling for multiple behaviour change. *Arch Int Med.* (2007) 167:1152–8. doi: 10.1001/archinte.167.11.1152

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership