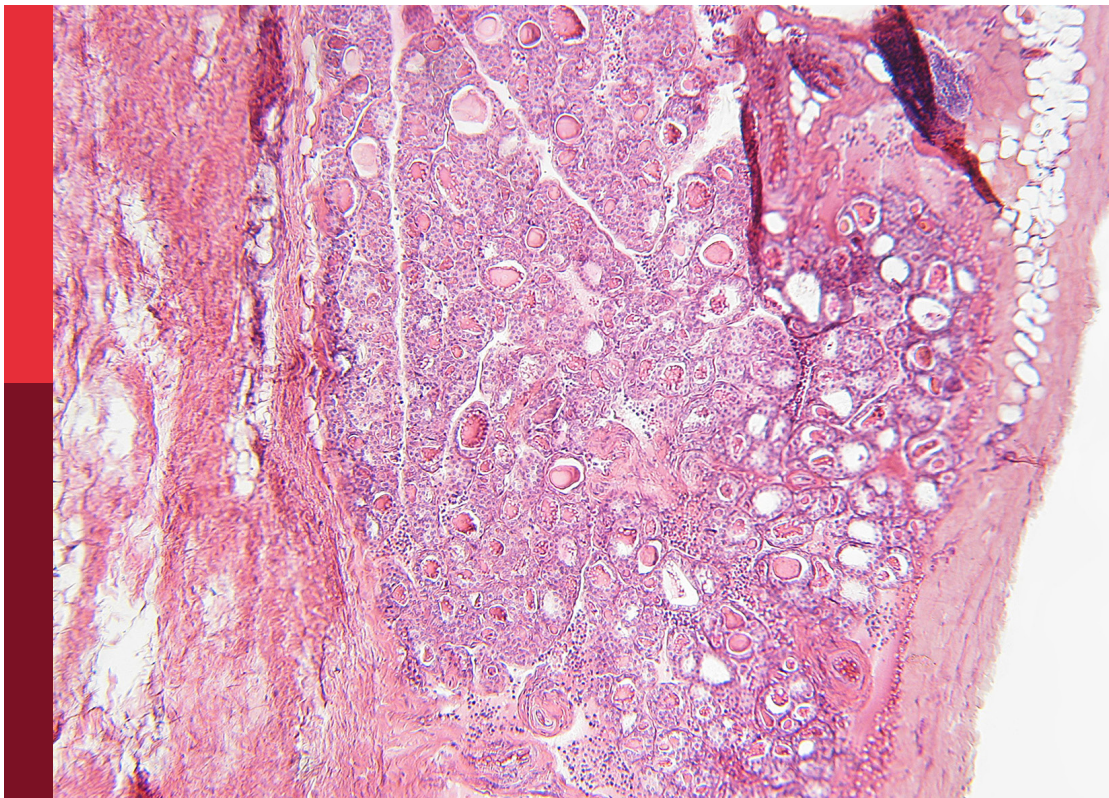
frontiers | Research Topics

COVID-19 and diabetes: Current findings and future perspectives

Edited by
Pranav Kumar Prabhakar

Published in
Frontiers in Endocrinology
Frontiers in Public Health



ISSN 1664-8714
ISBN 978-2-8325-4915-5
DOI 10.3389/978-2-8325-4915-5

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

COVID-19 and diabetes: Current findings and future perspectives

Topic editor
Pranav Kumar Prabhakar — Lovely Professional University, India

Citation
Prabhakar, P. K., ed. (2024). *COVID-19 and diabetes: Current findings and future perspectives*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4915-5

Table of contents

frontiers | Frontiers in Endocrinology

TYPE Editorial
PUBLISHED 09 May 2024
DOI 10.3389/fendo.2024.1421721

OPEN ACCESS

EDITED AND REVIEWED BY
Lixin Li,
Central Michigan University, United States

*CORRESPONDENCE
Pranav Kumar Prabhakar
prabhakar.iitm@gmail.com

RECEIVED 22 April 2024
ACCEPTED 02 May 2024
PUBLISHED 09 May 2024

CITATION
Prabhakar PK (2024) Editorial: COVID-19 and diabetes: Current findings and future perspectives.
Front. Endocrinol. 15:1421721.
doi: 10.3389/fendo.2024.1421721

Editorial: COVID-19 and diabetes: Current findings and future perspectives

Pranav Kumar Prabhakar*

Research and Development Cell, Lovely Professional University, Phagwara Punjab, India

KEYWORDS

diabetes, COVID - 19, complication, SARS-CoV2, cardiovascular

Editorial on the Research Topic
COVID-19 and diabetes: Current findings and future perspectives

Introduction

COVID-19 and diabetes represent a complex intersection in the realm of public health, with profound implications for individuals' health outcomes and healthcare systems worldwide. The COVID-19 pandemic, caused by the novel coronavirus SARS-CoV-2, has rapidly spread across the globe since its emergence in late 2019, resulting in millions of infections and deaths (1). Individuals with diabetes have been identified as a particularly vulnerable population, facing an increased risk of severe outcomes from COVID-19. Conversely, COVID-19 infection can exacerbate glycemic control and increase the risk of diabetic complications, highlighting the bidirectional relationship between these two conditions.

Diabetes is a chronic metabolic disorder characterized by elevated blood sugar levels due to insulin deficiency or resistance. It encompasses several subtypes, including type 1 diabetes, type 2 diabetes, and gestational diabetes, each with its own etiology and management considerations. Diabetes affects millions of people worldwide and is associated with a range of complications, including cardiovascular disease, kidney failure, neuropathy, and retinopathy (2).

Despite these challenges, the COVID-19 pandemic has also catalyzed innovation and adaptation in diabetes care delivery. Telemedicine, remote monitoring, and digital health technologies have emerged as valuable tools for delivering diabetes care and education remotely, enhancing access and convenience for patients. Healthcare providers have embraced virtual consultations, telehealth platforms, and mobile applications to maintain continuity of care and support patients in managing their diabetes during the pandemic (3). These digital health solutions offer opportunities to reach underserved populations, improve patient engagement, and optimize diabetes outcomes in the era of COVID-19 and beyond. This Research Topic aimed at the relationship and coexistence of diabetes and COVID-19. The Research Topic currently includes 16 papers which contain 2 case studies, 2 mini reviews, and 12 original research articles on the various topics.

Interactions between COVID-19 infection and diabetes

The relationship between COVID-19 and diabetes is multifaceted and multifactorial. Individuals with diabetes are more likely to experience severe outcomes from COVID-19, including hospitalization, admission to intensive care units, and death. Several factors contribute to this heightened risk, including impaired immune function, underlying comorbidities, and physiological changes associated with diabetes. Moreover, individuals with diabetes often have other risk factors for severe COVID-19 outcomes, such as obesity, hypertension, and cardiovascular disease, further increasing their vulnerability. Conversely, COVID-19 infection can exacerbate glycemic control and increase the risk of diabetic complications (4). The inflammatory response triggered by COVID-19 can lead to insulin resistance and hyperglycemia, particularly in individuals with pre-existing diabetes. Furthermore, the stress of illness, changes in diet and physical activity, and disruptions to routine diabetes care can all contribute to worsening glycemic control during COVID-19 infection. Yuanyuan et al. analyzed the web of science database for the comprehensive analysis of current publications related to diabetes mellitus (DM) research during the COVID-19 epidemic reveals a growing body of literature addressing various aspects of this intersection.

The COVID-19 pandemic, caused by SARS-CoV-2, has profoundly affected global health and economy since 2020. The virus binds to ACE2 receptors, found in various organs including endocrine glands, impacting multiple endocrine systems. COVID-19 promotes obesity through lifestyle changes, exacerbating diabetes risk. Additionally, it directly affects pancreatic function, worsening type 1 or type 2 diabetes. High adiposity and chronic hyperglycemia increase COVID-19 susceptibility and severity. Bidirectional interactions exist between COVID-19 and diabetes, influencing each other's progression. Healthcare systems have adapted services to manage diabetes amidst the pandemic's challenges. Overall, COVID-19 and diabetes share complex interactions, necessitating tailored healthcare responses. Wolińska et al. discussed the role of various environmental factors that lead to obesity either before COVID-19 or after COVID-19. The rapid rise in overweight and obesity over recent decades has been influenced by various factors, including environmental and novel elements emerging during the COVID-19 pandemic. Lockdown measures during the pandemic led to increased BMI in many countries, driven by reduced physical activity, increased screen time and sleep duration, and elevated consumption of processed foods. Environmental factors such as policy issues, socioeconomic status, lifestyle choices, and neighborhood conditions also contribute to obesity trends. Air pollution's role in obesity remains debated. However, the pandemic's impact extends beyond weight gain, affecting individuals with diabetes disproportionately.

Challenges during hospitalization or post covid-19

In this nationwide retrospective investigation, Kania et al. examined the association between diabetes and in-hospital mortality among COVID-19 patients. Conducted in Poland since the pandemic's onset in 2019, the study revealed diabetes as a significant factor linked to increased hospitalization rates and higher risk of in-hospital mortality, even after adjusting for various factors like age, sex, and comorbidities such as chronic kidney disease and heart conditions. The study also noted variations in relative risk across different age groups and genders, with heightened risks observed in males and patients in their sixties. The research underscores the importance of recognizing diabetes as a crucial risk factor in COVID-19 prognosis, offering valuable insights for healthcare providers.

Meanwhile, Gorchane et al. and Bukara-Radujkovic et al. investigated the impact of the COVID-19 pandemic on new-onset diabetic ketoacidosis (DKA) in Africa, an area with limited prior research on this topic. Their analysis compared DKA incidence trends before and during the pandemic, highlighting an increase in DKA cases alongside rises in both type 1 and type 2 diabetes. These findings suggest that the pandemic may have contributed to the observed uptick in DKA cases, emphasizing the need for further research and attention to COVID-19's effects on diabetes-related complications in Africa (9, 10).

Case study

A case of young pregnant Chinese woman developed sudden hyperglycemia and ketoacidosis in her last trimester, following mild SARS-CoV-2 infection were studies. Despite near-normal glycohemoglobin levels, low C peptide levels indicated severe insulin deficiency, leading to a diagnosis of fulminant type 1 diabetes (FT1D). Insulin therapy swiftly improved ketoacidosis and hyperglycemia, but β cell function remained impaired. The patient transitioned to insulin pump therapy post-discharge, with favorable glucose control at the first follow-up. This case underscores the potential for FT1D onset following SARS-CoV-2 infection and the importance of prompt recognition and management during the COVID-19 pandemic (11). Another case study of a 16-year-old boy developed symptoms of polyuria, polydipsia, and weight loss after receiving the BNT162b2 Comirnaty COVID-19 vaccine, worsening after the second dose has been performed. Diagnostic tests revealed central diabetes insipidus due to neuroinfundibulohypophysitis. Treatment with Desmopressin alleviated symptoms, with ongoing follow-up. This case highlights the need for vigilance in recognizing and reporting potential adverse effects of COVID-19 vaccines, including rare conditions like hypophysitis. Further research is required to determine any causal link between COVID-19 vaccination and the development of central diabetes insipidus (12).

Future prospective

A large number of articles were focused on the prediction of various aspects of post covid complications. This study done by Byeon employed machine learning techniques to identify major risk factors for depression in community diabetic patients and

developed predictive models for high-risk group identification. Analyzing 26,829 adults diagnosed with diabetes, it found a 22.4% prevalence of depression. Utilizing CatBoost, the top nine influential factors included gender, smoking status, COVID-19-related changes in drinking and smoking, subjective health, economic concerns, sleep alterations, economic activity, and social support. Early identification of high-risk individuals is crucial for implementing personalized psychological support at the primary medical level, enhancing mental health outcomes for diabetic patients (13).

Shoaib et al.'s study underscores the complex interplay between COVID-19 and diabetes, recognizing heightened vulnerability and potential post-complications for diabetic individuals. Additionally, it suggests a potential association between cough medicine containing steroids and an increased risk of developing diabetes. The study utilized deep-learning models on chest x-ray images sourced from publicly available datasets, validated by a certified radiologist, to aid diagnosis (14). Another study by Ahmad et al. explored a deep transfer learning approach for COVID-19 detection, achieving a high accuracy of 99.11% with the CIDICXR-Net50 model. This study also investigated the relationship between COVID-19 and diabetes, aiming to enhance diabetes prediction through advanced machine learning techniques. Initial assessment favored the Support Vector Machines (SVM) classifier with 76.62% accuracy (15). Advanced feature engineering revealed hidden patterns, particularly in Glucose levels. Correlation analyses highlighted significant associations, and integrating Decision Trees, Gradient Boosting, and SVM in an ensemble model improved accuracy to 93.2%. This research offers a robust framework for diabetes prediction, crucial for early diagnosis, personalized treatment, and preventive care, addressing global health challenges and enhancing life expectancy (16).

Future prospects regarding COVID-19 and diabetes are multifaceted, encompassing various aspects of research, treatment, preventive measures, healthcare infrastructure, public health policies, and education. Ongoing research endeavors aim to elucidate the complex interaction between COVID-19 and diabetes, seeking to understand why individuals with diabetes face a heightened risk of severe outcomes from the virus. This research is crucial for developing targeted treatments to mitigate complications and improve outcomes for diabetic patients infected with COVID-19. In addition to research, future strategies are likely to prioritize preventive measures tailored to individuals with diabetes. This may include vaccination campaigns aimed at diabetic populations, lifestyle interventions to manage diabetes effectively, and improved diabetes management protocols to reduce the risk of severe COVID-19 outcomes. Furthermore, the pandemic has underscored the importance of robust healthcare infrastructure, particularly for managing chronic conditions like diabetes during public health crises. Investments in telemedicine, remote monitoring technologies, and integrated care models are anticipated to enhance the delivery of healthcare services and improve outcomes for diabetic patients during future outbreaks. Moreover, public health policies may be developed to address the intersection of COVID-19 and diabetes. These policies could involve prioritizing vaccination for diabetic individuals, ensuring equitable access to healthcare services, and implementing measures to reduce the risk of COVID-19 transmission in vulnerable populations. Concurrently, education and awareness efforts will likely intensify, emphasizing the increased risk of COVID-19 complications among individuals with diabetes. Such campaigns may promote vaccination uptake, adherence to preventive measures such as mask-wearing and social distancing, and regular monitoring of blood glucose levels to manage diabetes effectively in the context of the pandemic.

In conclusion, the intersection of COVID-19 and diabetes presents complex challenges and opportunities for healthcare systems, policymakers, and individuals alike. By prioritizing vaccination efforts, optimizing diabetes management, strengthening healthcare infrastructure, investing in research and innovation, and promoting health equity, we can mitigate the impact of COVID-19 on individuals with diabetes and safeguard their health and well-being in the face of future pandemics.

Author contributions

PP: Writing – review & editing, Writing – original draft.

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.

Publisher's note

References

1. Al-Kuraishy HM, Al-Gareeb AI, Mostafa-Hedeab G, Dubey R, Prabhakar PK, Batiha GES. COVID-19 and diabetes: will novel drugs for diabetes help in COVID-19? *Curr Mol Pharmacol.* (2023) 16:494–506. doi: 10.2174/1874467215666220908091604

2. Prabhakar PK, Singh K, Kabra D, Gupta J. Natural SIRT1 modifiers as promising therapeutic agents for improving diabetic wound healing. *Phytomedicine.* (2020) 76:153252. doi: 10.1016/j.phymed.2020.153252

3. Abdi A, Jalilian M, Sarbarzeh PA, Vlaisavljevic Z. Diabetes and COVID-19: A systematic review on the current evidences. *Diabetes Res Clin Pract*. (2020) 166:108347. doi: 10.1016/j.diabres.2020.108347

4. Peric S, Stulnig TM. Diabetes and COVID-19: disease—management—people. *Wiener Klinische Wochenschrift*. (2020) 132:356–61. doi: 10.1007/s00508-020-01672-3

frontiers | Frontiers in Endocrinology

TYPE Case Report
PUBLISHED 04 April 2023
DOI 10.3389/fendo.2023.1168927

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Fabrizio Barbetti,
University of Rome Tor Vergata, Italy
David Baidal,
University of Miami, United States
Antonio Toniolo,
University of Insubria, Italy

*CORRESPONDENCE
Lan Shou
2021011012013@stu.hznu.edu.cn

SPECIALTY SECTION
This article was submitted to Clinical Diabetes, a section of the journal Frontiers in Endocrinology

RECEIVED 18 February 2023
ACCEPTED 20 March 2023
PUBLISHED 04 April 2023

CITATION
Zhou L, Qu H, Zhang Q, Hu J and Shou L (2023) Case report: Fulminant type 1 diabetes following paucisymptomatic SARS-CoV-2 infection during late pregnancy.
Front. Endocrinol. 14:1168927.
doi: 10.3389/fendo.2023.1168927

Case report: Fulminant type 1 diabetes following paucisymptomatic SARS-CoV-2 infection during late pregnancy

Lingling Zhou, Huanjia Qu, Qiuling Zhang, Jinhua Hu and Lan Shou*

Department of Endocrinology and Metabolic Disease, Affiliated Hospital of Hangzhou Normal University, Hangzhou Normal University, Hangzhou, China

Background: Dysregulation of glucose metabolism has been linked to SARS-CoV-2 infection. In addition, the occurrence of new onset diabetes mellitus, including fulminant type 1 diabetes, has been reported after SARS-CoV-2 infection or vaccination.

Methods and results: A young Chinese woman in her last trimester of pregnancy presented with an abrupt progression of hyperglycemia and ketoacidosis, but with a near-normal glycohemoglobin level following paucisymptomatic SARS-CoV-2 infection. The low C peptide levels, both fasting and postprandial, reflected profound insulin deficiency in the setting of negative islet autoantibody testing, consistent with a diagnosis of fulminant type 1 diabetes. Ketoacidosis and hyperglycemia quickly improved following the introduction of insulin therapy, but not the β cell function. The patient received treatment with insulin pump therapy after being discharged, and the first follow-up revealed a well-controlled glucose profile.

Conclusions: New-onset FT1D can occur after SARS-CoV-2 infection. Our report raises awareness of this rare but serious situation, promoting early recognition and management of FT1D during the COVID-19 pandemic.

KEYWORDS

fulminant type 1 diabetes, coronavirus disease 2019 (COVID- 19), severe acute respiratory syndrome coronavirus 2, pregnancy, case report

Highlights

1. It is unknown if the risk of long-term diabetes mellitus incidence increases, but SARS-CoV-2 infection causes aberrant glycometabolism.
2. FT1D often follows a preceding viral infection in a susceptible individual. There have been a few cases of FT1D after SARS-CoV-2 infection or vaccination.

3. SARS-CoV-2-induced islet dysfunction likely occurs not only *via* direct viral entry but also *via* inflammation and oxidative stress systematically or in islet micro-environment.
4. During the COVID-19 pandemic, attention should be paid to identify FT1D.

Introduction

There are mixed epidemiological results on the association between new-onset diabetes and infection by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Studies from some regions have reported an increased type 1 diabetes mellitus (T1DM) incidence during the pandemic, which supports the diabetogenic effect of COVID-19, while other studies have not (1, 2). Fulminant type 1 diabetes (FT1D), a relatively rare subtype of T1DM first reported in Japan, is characterized by a rapid progression of insulin deficiency at disease onset (3). The diagnosis is based on the following: i) occurrence of diabetic ketosis soon after the onset of hyperglycemic symptoms, ii) high level of plasma glucose (≥16.0 mmol/L, ≥288 mg/dL) but a relatively mismatched level of glycohemoglobin (<8.7%), and iii) presence of endogenous insulin deficiency including urinary C peptide excretion <10 μg/day or fasting serum C peptide level <0.3 ng/ml (<0.10 nmol/L) and <0.5 ng/ml (<0.17 nmol/L) after intravenous glucagon or after meal load without verifiable islet-related autoantibodies (4). Though the etiology of FT1D has not been fully elucidated, viral infection is considered to be the most important environmental risk factor for FT1D (4).

Recently, several case reports of FT1D after COVID-19 vaccination or infection were mainly on patients from East Asia (5). Here, we describe a pregnant woman who eventually experienced a rapid clinical course of diabetic ketoacidosis following SARS-CoV-2 infection, drawing attention to FT1D during the COVID-19 pandemic.

Case presentation and diagnosis assessment

A 34-year-old Chinese woman with a singleton pregnancy in her 34th gestational week presented to our hospital due to the abrupt onset of polydipsia, nausea, and vomiting for one day. Five weeks prior to admission, she had a brief fever with a maximum temperature of 37.8°C, accompanied by mild fatigue and muscle pain lasting for two days. She did not seek additional medical attention despite a positive COVID-19 antigen self-test at that time. She had been regularly evaluated in the obstetrics department for close follow-up and her OGTT test result was negative at the gestational age of 25 weeks. She was in good health with a normal weight before pregnancy (53 kg and a BMI of 21 kg/m^2) and was not on any medication or vaccination, and did not smoke or consume alcohol during pregnancy. None of her family members had a history of type 1 or type 2 diabetes mellitus.

At admission, the patient appeared to be dehydrated but was conscious. A quick physical examination revealed a heart rate of 108 beats/minute, blood pressure of 132/72 mmHg, temperature of 36.6°C, respiratory rate of 20 breaths/minute, and oxygen saturation of 98% breathing ambient air. The patient was 159 cm in height and 67 kg in weight, with a BMI of 26.5 kg/m^2. A physical examination of the abdomen revealed no upper abdominal pain, guarding, or rebound tenderness. The symphysis-to-fundal height was 30 cm, and the abdominal circumference was 102 cm. On obstetric evaluation, she was found to have uterine contractions every 20 s per 2 min. The fetus presented as a vertex presentation with an unengaged fetal head. Electronic fetal heart rate monitoring showed a fetal heart rate (FHR) of 114 bpm and an unsatisfactory contraction stress test (CST) result. Emergency obstetric ultrasound showed a biparietal diameter (BPD) of 88 mm, femur length (FL) of 69 mm, FHR of 110 bpm, and a single deepest vertical pocket of 70 mm.

Initial lab work showed an arterial pH of 7.08, actual bicarbonate of 8.6 mmol/L, β-hydroxybutyrate of 5.8 mmol/L, glucose level of 29 mmol/L (522 mg/dL), sodium level of 125 mmol/L, potassium level of 5.8 mmol/L, chloride level of 93 mmol/L, white cell count of 25.6×10^9/L, hemoglobin level of 114 g/L, and mildly elevated pancreatic enzymes (amylase and lipase levels less than 3*upper limit of normal). The lipids profile indicated a severe hypertriglyceridemia level of 10.99 mmol/L. Further laboratory tests revealed a glycohemoglobin level of 5.9%, a glycated serum protein level of 1.45 mmol/L, and an extremely low serum C peptide level of 0.02 ng/ml. Immunological examination yielded an absence of serum islet autoantibodies, including glutamic acid decarboxylase antibody (GADAb), insulinoma associated antigen-2 antibody (IA-2Ab), islet cell antibody (ICAb), and insulin antibody(I-Ab). She tested positive for IgG but not IgM antibodies against SARS-CoV-2. Other potential viral antibodies, including coxsackievirus, cytomegalovirus, parainfluenza virus, human herpes virus, and Epstein–Barr virus, were also tested, but the results returned negative. We ruled out acute pancreatitis from an abdominal computed tomography scan, with no signs of pancreatic edema or exudation soon after the patient received urgent Cesarean section surgery.

Diagnosis

She was diagnosed with diabetic ketoacidosis (DKA) and fulminant type 1 diabetes mellitus.

Treatment

The patient underwent emergency C-section surgery for acute fetal distress and was immediately treated with fluid resuscitation, intravenous insulin infusion, and low-dose sodium bicarbonate. Her main complaints, metabolic acidosis and electrolyte disturbance, were solved within six hours. Two days after admission, β-hydroxybutyrate level was 0.15mmol/L. She was

then switched to a subcutaneous insulin regimen for glycemic control and was initiated with fenofibrate to lower the triglyceride level. Her initial subcutaneous insulin regimen was insulin Glargine 14u SC QHS and insulin aspart 6u SC before meals. On the tenth day after admission the patient started to use an insulin pump and over the next few days the glucose levels stabilized at 4.5–15.6 mmol/L (81–281 mg/dL).

Outcome and follow-up

The newborn infant had a poor Apgar score (4/10) at birth, which was re-evaluated to be 3/10 (at both 5 min and 10 min after birth), and was transferred to a neonatal unit but died afterward. After the C-section, the patient was transferred to intensive care unit and then the endocrine ward. She was hospitalized for another two weeks and then was discharged on continuous subcutaneous insulin infusion therapy. However, three weeks after the disease onset, she still had profound insulin deficiency, as evidenced by low levels of and a slight but unsatisfactory increase of C peptide in the fasting state and after a mixed meal load (shown in Table 1).

Discussion

To the best of our knowledge, this is the first case of newly diagnosed fulminant type 1 diabetes following SARS-CoV-2 infection during pregnancy. The patient reported herein presented with hyperglycemia, ketoacidosis with an extremely rapid course, a near-normal glycohemoglobin level, and exhaustion of endogenous insulin secretion but without evidence of islet-related autoantibodies, which fulfilled the diagnosis of FT1D (4). Although the gestational status itself is a pre-existing risk factor for FT1D, we still suspect that in this case, FT1D occurred as rare organ damage in the pancreas due to a previous COVID-19 infection, according to the timeline of the medical history (shown in Figure 1), since no positive results of other suspicious virus infections reported to be associated with FT1D were obtained.

During the COVID-19 pandemic and the rolling back of strict anti-COVID-19 restrictions in China, cases of autoimmune-mediated disorders, including Graves' disease, type 1 diabetes, Guillain-Bare Syndrome, and systemic lupus erythematosus have been reported following SARS-CoV-2 infection or vaccination. Although the underlying mechanisms and whether SARS-CoV-2 affects pancreatic islets or other endocrine organs remain unknown, SARS-CoV-2 infection does result in aberrant glycometabolic control (6). Existing clinical data suggest that the infection can aggravate insulin resistance, increase hepatic glucose production, and impair peripheral glucose uptake through increased counter-regulatory hormones, release of cytokines and lipids, and also through direct hepatocyte injury. In addition, drugs often used in COVID-19 treatment, such as corticosteroids, also result in metabolic dysregulation and impaired glucose homeostasis.

New-onset diabetes after SARS-CoV-2 infection or vaccination has been reported. Autoantibody-negative, insulin-dependent diabetes was reported following infection (7). Tang et al. also documented a case of FT1D after receiving the first dose of an inactivated COVID-19 vaccine (5). Individuals with HLA genotypes predisposed to T1DM were diagnosed with FT1D several days after COVID-19 mRNA vaccination (8, 9). Typical symptoms of hyperglycemia and ketoacidosis following vaccination were reported in patients who had also been treated with immune checkpoint inhibitors (10, 11).

Studies have revealed that receptors involved in SARS-CoV-2 viral entry, including angiotensin-converting enzyme 2 (ACE2), neurophilin-1 (NRP1), and transmembrane serine proteases 2 (TMPRSS), were detected in human islet β cells, though at low

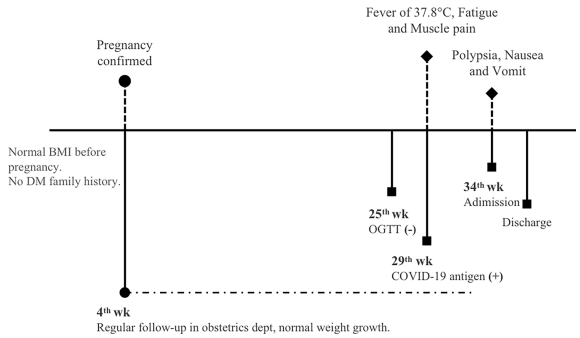


FIGURE 1
Timeline of events and important information of this patient.

TABLE 1 Results of mixed meal tolerance test in the 1st follow-up and OGTT at the 25th gestational week.

	Glucose level (mmol/L)	C peptide level (ng/ml)
Mixed meal load* in the 1st follow-up		
Fasting	6.57 (118 mg/dL)	0.02
Postprandial (2 h)	21.39 (385 mg/dL)	0.05
OGTT at the 25th gestational week		
Fasting	3.64 (66 mg/dL)	Not applicable
Post oral glucose (1h)	8.67 (156 mg/dL)	Not applicable
Post oral glucose (2h)	6.66 (120 mg/dL)	Not applicable

* The meal contained approximately 70 g carbohydrates, 24 g proteins, and 15 g fat.

levels. This may indicate that human islet cells are permissive to SARS-CoV-2 infection. Additionally, clinical data demonstrated that SARS-CoV-2-triggered necroptosis and apoptosis of islets cells were linked to increased glucose levels, a significant viral load, and strong ACE2 expression in β cells (12). Preclinical studies also supported the role of ACE2 in β cell homeostasis: deletion of ACE2 impairs β cell proliferation, decreases β cell mass, and induces β cell oxidative stress and thus decreases insulin secretion. Besides, in high-fat diet mice, de-differentiation of β cells was characterized by a reduction of ACE2 (13). Moreover, molecular mimicry between the SARS-CoV-2 spike protein and human endocrine cells, including pancreatic β cells, has been proposed as a possible element in pathogenesis (14). A recent study using autopsy samples also concluded that the SARS-CoV-2 viral antigen was detected in both endocrine and non-endocrine human pancreas cells, and that the expression of multiple chemokines as well as cytokines were higher in SARS-CoV-2-infected human islets (15). It has also been reported that pancreatic β cells presented with a lower expression of insulin but with a higher expression of glucagon and trypsin1, suggesting that cellular transdifferentiation takes place upon SARS-CoV-2 infection (15). The virus-induced inflammatory cytokines storm, a prothrombotic state, and endothelial derangement *via* ACE2 receptors might also injure β cells function, potentially by affecting the islet microvascular system. Local islet inflammation and systematic oxidative stress after infection might also induce post-translational protein modifications, enhance the generation of neoepitopes, and thus initiate islet autoimmunity (16).

However, the genotype of classical human leukocyte antigen (HLA) alleles was not determined in this case. A direct causal relationship could not be proven in this patient; however, we regard the SARS-CoV-2 infection as a suspicious trigger of FT1D onset in the gestational setting. More studies on the interactions between SARS-CoV-2 and pancreatic cells are warranted.

In conclusion, we presented a case report of a woman who developed FT1D after SARS-CoV-2 infection during late pregnancy. This is a rare and life-threatening situation, with high stillbirth or miscarriage rates. Clinicians should be aware of the possibility of the onset of FT1D within the COVID-19 background, especially in expectant mothers.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

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 was obtained from the patient for the publication of this case report.

Author contributions

All authors participated in treating this patient on the ward and were involved in the paper's conception and analysis as well as interpretation of the lab results. LZ collected and summarized the case data and wrote the first draft of the manuscript; HQ visualized the timeline; and all authors edited, reviewed, and approved the final version of the manuscript. LS is the guarantor of this work and thus has full access to all the data in the study and takes responsibility for its integrity. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by grants from the Medical Health Science and Technology Project of Zhejiang Provincial Health Commission (2023RC234, HQ is the recipient) and the Hangzhou Medical Health Science and Technology Project (A20220027, JH is the recipient).

Acknowledgments

We are grateful that the patient permitted this paper to be published even though she was going through a very difficult time, and the loss of her beloved son also saddens us. We acknowledge the joint efforts of the emergency department, obstetrics department, neonatology department, intensive care unit, operating room, and nursing team regarding the diagnosis, treatment, and care of this patient. We also thank Dr. Muthukumaran Jayachandran from the Shanghai Tenth People's Hospital, affiliated with Tongji University, for editing the English language.

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. Rathmann W, Kuss O, Kostev K. Incidence of newly diagnosed diabetes after covid-19. *Diabetologia* (2022) 65(6):949–54. doi: 10.1007/s00125-022-05670-0

2. Tittel SR, Rosenbauer J, Kamrath C, Ziegler J, Reschke F, Hammersen J, et al. Did the covid-19 lockdown affect the incidence of pediatric type 1 diabetes in Germany? *Diabetes Care* (2020) 43(11):e172–e3. doi: 10.2337/dc20-1633

3. Imagawa A, Hanafusa T, Miyagawa J, Matsuzawa Y. A novel subtype of type 1 diabetes mellitus characterized by a rapid onset and an absence of diabetes-related antibodies. Osaka iddm study group. *N Engl J Med* (2000) 342(5):301–7. doi: 10.1056/NEJM200002033420501

4. Imagawa A, Hanafusa T, Awata T, Ikegami H, Uchigata Y, Osawa H, et al. Report of the committee of the Japan diabetes society on the research of fulminant and acute-onset type 1 diabetes mellitus: New diagnostic criteria of fulminant type 1 diabetes mellitus (2012). *J Diabetes Investig* (2012) 3(6):536–9. doi: 10.1111/jdi.12024

5. Tang X, He B, Liu Z, Zhou Z, Li X. Fulminant type 1 diabetes after covid-19 vaccination. *Diabetes Metab* (2022) 48(2):101324. doi: 10.1016/j.diabet.2022.101324

6. Kazakou P, Lambadiari V, Ikonomidis I, Kountouri A, Panagopoulos G, Athanasopoulos S, et al. Diabetes and covid-19; a bidirectional interplay. *Front Endocrinol (Lausanne)* (2022) 13:780663. doi: 10.3389/fendo.2022.780663

7. Hollstein T, Schulte DM, Schulz J, Glück A, Ziegler AG, Bonifacio E, et al. Autoantibody-negative insulin-dependent diabetes mellitus after sars-Cov-2 infection: A case report. *Nat Metab* (2020) 2(10):1021–4. doi: 10.1038/s42255-020-00281-8

8. Kobayashi T, Yakou F, Saburi M, Hirose A, Akaoka H, Hirota Y, et al. New-onset atypical fulminant type 1 diabetes after covid-19 vaccination: A case report. *Clin Case Rep* (2022) 10(10):e6473. doi: 10.1002/ccr3.6473

9. Sasaki K, Morioka T, Okada N, Natsuki Y, Kakutani Y, Ochi A, et al. New-onset fulminant Type 1 diabetes after severe acute respiratory syndrome Coronavirus 2 vaccination: A case report. *J Diabetes Investig* (2022) 13(7):1286–9. doi: 10.1111/jdi.13771

10. Nishino K, Nakagawa K, Yase E, Terashima M, Murata T. Diabetic ketoacidosis after the second dose of sars-Cov-2 mrna vaccination in a patient with pembrolizumab-induced fulminant type 1 diabetes. *Diabetol Int* (2022) 1–5. doi: 10.1007/s13340-022-00614-w

11. Ohuchi K, Amagai R, Tamabuchi E, Kambayashi Y, Fujimura T. Fulminant type 1 diabetes mellitus triggered by coronavirus disease 2019 vaccination in an advanced melanoma patient given adjuvant nivolumab therapy. *J Dermatol* (2022) 49(5):e167–e8. doi: 10.1111/1346-8138.16304

12. Steenblock C, Richter S, Berger I, Barovic M, Schmid J, Schubert U, et al. Viral infiltration of pancreatic islets in patients with covid-19. *Nat Commun* (2021) 12 (1):3534. doi: 10.1038/s41467-021-23886-3

13. Xuan X, Gao F, Ma X, Huang C, Wang Y, Deng H, et al. Activation of Ace2/Angiotensin (1-7) attenuates pancreatic B cell dedifferentiation in a high-Fat-Diet mouse model. *Metabolism* (2018) 81:83–96. doi: 10.1016/j.metabol.2017.12.003

14. Churilov LP, Normatov MG, Utekhin VJ. Molecular mimicry between sars-Cov-2 and human endocrinocytes: A prerequisite of post-Covid-19 endocrine autoimmunity? *Pathophysiology* (2022) 29(3):486–94. doi: 10.3390/pathophysiology29030039

15. Tang X, Uhl S, Zhang T, Xue D, Li B, Vandana JJ, et al. Sars-Cov-2 infection induces beta cell transdifferentiation. *Cell Metab* (2021) 33(8):1577–91.e7. doi: 10.1016/j.cmet.2021.05.015

16. Strollo R, Vinci C, Man YKS, Bruzzaniti S, Piemonte E, Alhamar G, et al. Autoantibody and T cell responses to oxidative post-translationally modified insulin neoantigenic peptides in type 1 diabetes. *Diabetologia* (2023) 66(1):132–46. doi: 10.1007/s00125-022-05812-4

frontiers | Frontiers in Endocrinology

TYPE Case Report
PUBLISHED 18 April 2023
DOI 10.3389/fendo.2023.1166953

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Liza Das,
Post Graduate Institute of Medical Education and Research (PGIMER), India
Valeria Hasenmajer,
Sapienza University of Rome, Italy

*CORRESPONDENCE
Cristina Partenope
partenope.cristina@gmail.com

SPECIALTY SECTION
This article was submitted to Pediatric Endocrinology, a section of the journal Frontiers in Endocrinology

RECEIVED 15 February 2023
ACCEPTED 31 March 2023
PUBLISHED 18 April 2023

CITATION
Partenope C, Pedranzini Q, Petri A, Rabbone I, Prodam F and Bellone S (2023) AVP deficiency (central diabetes insipidus) following immunization with anti-COVID-19 BNT162b2 Comirnaty vaccine in adolescents: A case report.
Front. Endocrinol. 14:1166953.
doi: 10.3389/fendo.2023.1166953

AVP deficiency (central diabetes insipidus) following immunization with anti-COVID-19 BNT162b2 Comirnaty vaccine in adolescents: A case report

Cristina Partenope[1]*, Quincy Pedranzini[1], Antonella Petri[1], Ivana Rabbone[1], Flavia Prodam[1,2] and Simonetta Bellone[1]

[1]Division of Pediatrics, Department of Health Sciences, University of Piemonte Orientale, Novara, Italy, [2]Endocrinology, Department of Translational Medicine, University of Piemonte Orientale, Novara, Italy

Introduction: The coronavirus disease 19 (COVID-19) pandemic has prompted the development of new vaccines to reduce the morbidity and mortality associated with this disease. Recognition and report of potential adverse effects of these novel vaccines (especially the urgent and life-threatening ones) is therefore essential.

Case presentation: A 16-year-old boy presented to the Paediatric Emergency Department with polyuria, polydipsia and weight loss over the last four months. His past medical history was unremarkable. Onset of symptoms was referred to be few days after first dose of anti-COVID-19 BNT162b2 Comirnaty vaccine and then worsened after the second dose. The physical exam was normal, without neurological abnormalities. Auxological parameters were within normal limits. Daily fluid balance monitoring confirmed polyuria and polydipsia. Biochemistry laboratory analysis and urine culture were normal. Serum osmolality was 297 mOsm/Kg H_2O (285-305), whereas urine osmolality was 80 mOsm/Kg H_2O (100-1100), suggesting diabetes insipidus. Anterior pituitary function was preserved. Since parents refused to give consent to water deprivation test, treatment with Desmopressin was administered and confirmed ex juvantibus diagnosis of AVP deficiency (or central diabetes insipidus). Brain MRI revealed pituitary stalk thickening (4 mm) with contrast enhancement, and loss of posterior pituitary bright spot on T1 weighted imaging. Those signs were consistent with neuroinfundibulohypophysitis. Immunoglobulin levels were normal. Low doses of oral Desmopressin were sufficient to control patient's symptoms, normalizing serum and urinary osmolality values and daily fluid balance at discharge. Brain MRI after 2 months showed stable thicken pituitary stalk and still undetectable posterior pituitary. Due to persistence of polyuria and polydipsia, therapy with Desmopressin was adjusted by increasing dosage and number of daily administrations. Clinical and neuroradiological follow-up is still ongoing.

Conclusion: Hypophysitis is a rare disorder characterized by lymphocytic, granulomatous, plasmacytic, or xanthomatous infiltration of the pituitary gland and stalk. Common manifestations are headache, hypopituitarism, and diabetes

insipidus. To date, only time correlation between SARS-CoV-2 infection and development of hypophysitis and subsequent hypopituitarism has been reported. Further studies will be needed to deepen a possible causal link between anti-COVID-19 vaccine and AVP deficiency.

KEYWORDS

AVP deficiency, diabetes insipidus, COVID-19, SARS-CoV-2 vaccination, vaccine, pituitary stalk thickening, BNT162b2 Comirnaty vaccine, adolescent

Introduction

At the end of 2019 severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2) infection, named COVID-19, outbroke and plagued our healthcare systems, causing more than 6 million deaths worldwide over the last three years (1). We have learned that COVID-19 is a primarily respiratory disease, however it can affect nearly every organ system, including endocrine system (2, 3). Here, the expression of angiotensin-converting enzyme 2 (ACE2) receptors and transmembrane protease serine 2 (TMPRSS2) on many endocrine cells seems to play a crucial role in the direct pathogenetic mechanism by which the virus infects these organs (4).

Since no specific treatment was available, the pandemic urged scientists to develop targeted vaccines to face the high mortality rate of the disease. SARS-CoV-2 vaccines are generally safe and effective in preventing COVID-19 severe symptoms. Injection site reactions, fever, headache, myalgia, and skin rash are the most common vaccine side effects (5). Surprisingly, some cases of endocrinopathies have occurred even after anti-COVID-19 vaccines in adults (6–8), suggesting cytokine release syndrome exacerbated by the vaccine as the possible underlying mechanism for the disease, as a result of endocrine cells susceptibility to elevation of pro-inflammatory molecules (9). To the best of our knowledge, only a few cases of AVP deficiency, also known as central diabetes insipidus (CDI) occurring after anti-COVID-19 vaccines have been cited in literature until now (10–12) and still no one in children. Here we report the first case of new onset CDI in a pediatric patient after BNT162b2 mRNA COVID-19 vaccine.

Case presentation

B.AR., a 16-year-old boy, presented to the Paediatric Emergency Department with polyuria, polydipsia and concomitant weight loss, reporting a urine output of 9 liters in the last 24 hours and a weight loss of nearly 6 kilograms in the previous four months. Symptoms seemed to start a few days after inoculation of the first dose of BNT162b2 mRNA COVID-19 vaccine (at the end of August 2021) and worsened after the second dose, which had been administered 28 days later. Due to the persistence of intense thirst and polyuria the patient had already undergone some medical investigations: no abnormalities were found at urologic assessment and renal ultrasound (Figure 1). Moreover, no significant features emerged from his past medical history, and familial history was unremarkable too. On admission to the Paediatric Emergency Department, the patient's vital parameters included heart rate 110 bpm, temperature 36°C, oxygen saturation 99% on room air. Physical examination revealed a well-being adolescent boy, with auxological parameters within normal limits (height -0.12 standard deviations, weight +0.35 standard deviations), cardiac and pulmonary auscultation without pathological findings and no neurological abnormalities (including intact sense of smell and taste and cognition). Again, head and neck/gastrointestinal/musculoskeletal examinations were all grossly unremarkable. Furthermore, biochemistry laboratory analysis, venous blood gas analysis, and urine analysis were performed (Table 1) and no significant alterations were found. Diabetes mellitus was excluded

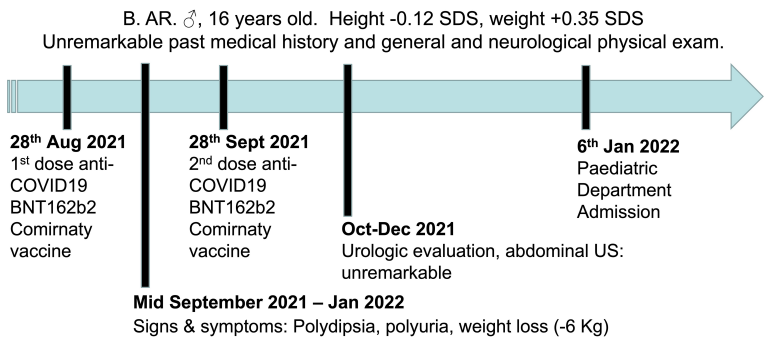


FIGURE 1
Timeline depicting the clinical course of the disease, from the onset of signs and symptoms until admission to the Paediatric Department.

10.3389/fendo.2023.1166953

TABLE 1 Results from blood and urine tests.

Lab test	Patient result	Normal ranges
SERUM CREATININE	0.69 mg/dL	0.9-1.3
BLOOD PH	7.35	7.35-7.45
BICARBONATES	31.8 mmol/L	22-26
SERUM GLUCOSE	88 mg/dl	70-100
PLASMA CALCIUM	9.36 mg/dL	8.5-10.5
PLASMA SODIUM	145 mmol/L	135-145
PLASMA POTASSIUM	4.5 mmol/L	3.5-5
C-REACTIVE PROTEIN	1 mg/dL	0-1
URINE PH	6	5.5-6.5
URINE SPECIFIC GRAVITY	1002	1005-1020
URINE HAEMOGLOBIN, GLUCOSE, PROTEIN	absent	
TSH	1.308 uIU/mL	0.450 - 3.500
FT4	1.35 ng/dL	0.89 -1.76
FT3	3.20 pg/mL	2.30 - 4.20
ACTH	28.2 pg/mL	3.6 - 60
CORTISOL	9 microg/dL	4.5 - 24
PROLACTIN	41.3 mIU/L	44.5 - 375.0
LH	2.50 mIU/mL	1.5 - 34.6
FSH	2.6 mIU/L	1.4 - 18.1
TESTOSTERONE	539.9 ng/dL	144.0 - 842.0
AFP	4.1 IU/mL	0.0 - 7.0
HCG	<2 mIU/mL	0 - 5

TSH, thyroid stimulating hormone; fT4, thyroxine; fT3, free triiodothyronine; ACTH, adrenocorticotropic hormone; LH, luteinizing hormone; FSH, follicle stimulating hormone; AFP, alpha-fetoprotein; hCG, human chorionic gonadotropin.

by detection of normal hematic glucose levels and by the absence of glycosuria. In contrast, diabetes insipidus could be suspected because of plasma sodium level at the upper limit of normal and low urine specific gravity. Given this hypothesis, the patient was tested for SARS-CoV-2 ongoing infection by nasopharyngeal swab – which resulted negative – and was then admitted to the pediatric ward to confirm the diagnosis.

Management and outcome

Primarily, daily fluid balance monitoring confirmed polyuria and polydipsia (IN 6.250 L/OUT 7.100 L). Secondly, serum osmolality (p-Osm) levels were 287 mOsm/Kg H_2O (normal values 285-305), urine osmolality (u-Osm) was 68 mOsm/Kg H_2O (normal values 100-1100) and u-Osm/p-Osm ratio was < 1, so that diabetes insipidus could be suspected. However, parents did not give consent to water deprivation test. A test with desmopressin (the synthetic analog of antidiuretic hormone) was then performed to discern between central and nephrogenic DI. After administration of low dose oral desmopressin (sublingual 60 micrograms) an immediate response was evident in view of both normalization of daily fluid balance (IN 0.520 L/OUT 0.600 L) and rapid increase of u-Osm (119 mOsm/Kg H_2O). Consistent with those findings, ex juvantibus diagnosis of complete CDI could be made. To complete investigations, hormonal tests showed no significant impairment of anterior pituitary function (Table 1), immunological assessment revealed normal immunoglobulin levels (IgA, IgM, and IgG subclasses, including IgG4) and urine culture and Quantiferon were negative as well. Brain contrast-enhanced magnetic resonance imaging (MRI) focused on the study of the pituitary region was carried out. It revealed pituitary stalk thickening (PST) (maximum diameter of 4 mm) with contrast enhancement and loss of posterior pituitary bright spot on T1 weighted imaging (Figure 2). Those signs were consistent with neuroinfundibulohypophysitis. In the view of strict time correlation between COVID-19 vaccine and onset of symptoms and granted that no other differential diagnosis had fit the clinical picture, the case was signaled as an adverse drug reaction to the Italian Pharmacological Agency (AIFA). During hospitalization, B.AR. maintained good general conditions. He was discharged after 7 days with prescription of low doses of oral desmopressin (60 mcg

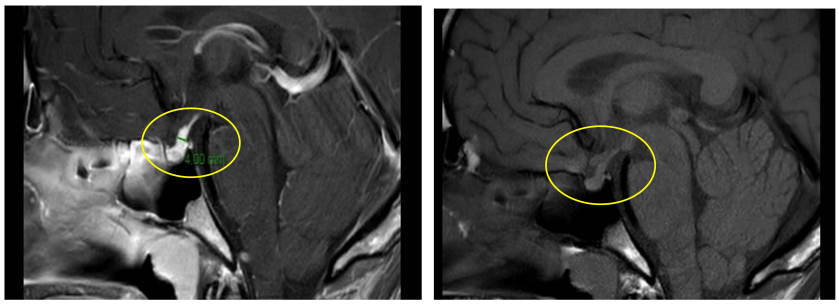


FIGURE 2
Pituitary stalk enlargement with contrast enhancement (left panel, yellow circle); loss of posterior pituitary bright (right panel, yellow circle).

twice daily), as the same dose had been sufficient to control signs and symptoms during hospitalization. After 2 months, follow-up MRI showed stable PST and still undetectable posterior pituitary, consistent with the persistence of inflammation of this cerebral area (Figure 3). An endocrinological follow-up was also planned at our center: three months after discharge, due to persisting polyuria and polydipsia, therapy with desmopressin was adjusted by increasing dosage and number of daily administrations (60 mcg three times a day). Clinical and neuroradiological follow-up is still ongoing.

Discussion

It is now widely accepted that COVID-19 is a multiorgan disease, as many tissues and organs are affected during ongoing or recent infection by SARS-CoV-2 virus, including endocrine system (13). It is becoming well established that in some cases SARS-CoV-2 infection can trigger an inflammation of the hypothalamus-pituitary axis (hypophysitis) both in adults and in children, resulting in a dysfunction that causes diabetes insipidus, either alone (14–17) or associated with anterior hypopituitarism (2, 14) with a latency of 0–16 weeks between recognition of the SARS-CoV-2 infection and development of symptoms (18). Although the exact pathogenetic mechanism has still to be defined, a direct pathway (involving ACE2-mediated hypothalamic viral infection) as well as an indirect, delayed, and immune-mediated pathway have been hypothesized. Moreover, endothelial damage in the blood–neuron interface, thrombotic microangiopathy (pituitary apoplexy), infected leukocyte-mediated transportation and cytokine storming are other accepted theories (19).

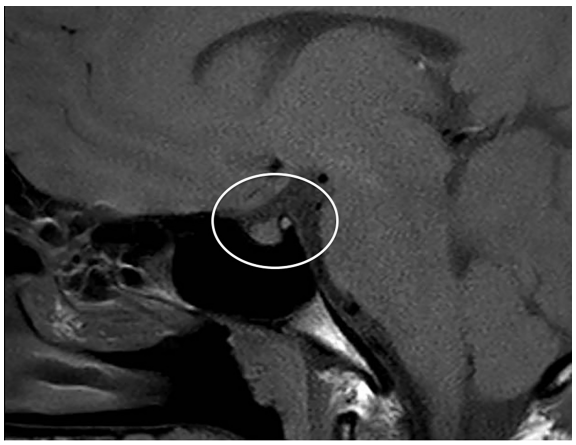


FIGURE 3
Brain MRI after 2 months: stable thickened pituitary stalk and undetectable posterior pituitary (circle).

Interestingly, hypopituitarism might have a bidirectional relationship with COVID-19 since pre-existent hypopituitarism can be per se a potential risk factor for COVID-19 due to its comorbidities (e.g. hypothalamic obesity) and can be worsened by SARS-CoV-2 infection (20). Moreover, COVID-19 vaccine can affect endocrine system, albeit infrequently and with good prognosis. Pezzaioli et al. recently collected and reviewed all published data on potential endocrine adverse effects post-COVID-19 vaccines in adult patients (6). Thyroid disorders are the most common; only eight cases of pituitary dysfunction have been described so far (10–12, 21–25), of which two presented with CDI (11, 12).

Our patient complained of exacerbation of DI symptoms after the second dose of COVID-19 vaccine. Although we were not able to conduct dose-specific analyses, previously published cases also reported onset or worsening of adverse endocrinological events following the second dose of vaccine (5, 22, 23). Mechanisms of increased reactogenicity after the second dose are largely unknown. One possibility is that the vaccine directly might cause pituitary impairment with cumulative effect. Autoimmune/inflammatory syndrome induced by vaccine adjuvants (ASIA syndrome), molecular mimicry, cross-reactivity or a pro-inflammatory state induced by vaccine components and subsequent activation of autoreactive B and T cells are speculated to be involved, as in Guillain-Barrè syndrome or optic neuritis (26–28).

COVID-19 symptoms are generally milder in the pediatric age than in the adult population, but this is not always true when the endocrine system is involved. Lizzi et al. described a pediatric case of acute onset of isolated CDI associated with recent SARS-CoV-2 infection (16), requiring 7 days of hospitalization. Here we described the case of an adolescent who developed CDI after COVID-19 vaccination, in whom posterior pituitary function has not recovered yet. Diabetes insipidus is a tricky disease whose symptoms can be underestimated by both patients and clinicians. CDI is characterized by decreased or absent secretion of antidiuretic

hormone (ADH; also called arginine vasopressin or AVP), resulting in a variable degree of polyuria. Lack of AVP can be caused by disorders or lesions in the hypothalamic osmoreceptor, in the supraoptic or paraventricular nuclei, in the superior portion of the supraopticohypophyseal tract or in the pituitary sella. Recently, a panel of experts from national and international endocrinology and endocrine pediatric societies has proposed to change diabetes insipidus' name to "arginine vasopressin deficiency (AVP-D)" for central etiologies to avoid detrimental confusion with diabetes mellitus for both patients and their caretakers (29).

Finally, pituitary stalk is a funnel-shaped structure that connects the hypothalamus to the pituitary gland. There are several etiologies that give rise to PST (30, 31), which often manifests clinically with CDI (32, 33): autoimmunity/inflammation (neuroinfundibulohypophysitis (34)) – sometimes referred as idiopathic PST, infectious diseases (e.g. tuberculosis (35)) or neoplastic lesions (33, 36). Anti-pituitary, anti-hypothalamus autoantibodies or high IgG4 levels have been detected in some patients with hypophysitis. However, their causal role remains unclear (37). Searching for these antibodies may help to diagnose an autoimmune hypophysitis, especially in cases like ours presenting with non-diagnostic pituitary MRI or hypoprolactinemia (38, 39). Unfortunately, no autoantibodies testing was performed in this case report since the lab kit was not available in our institution. The approach to the neuroradiological finding of PST is still controversial (32, 40): empirical management recommend to conduct follow-up MRI every 3-6 months and proceed with pituitary stalk biopsy only in case of stalk size ≥ 7 mm, progressive infundibular enlargement or worsening of symptoms. Moreover, all patients with pituitary stalk lesions and CDI should be routinely assessed for anterior pituitary hormonal function, which was normal in our patient.

Conclusion

To date, CDI following anti-COVID-19 vaccine remains a rare and only temporally linked occurrence, even though its pathophysiological explanation has been hypothesized. Further studies will be needed to deepen a possible causal link between the two events.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

Author contributions

CP and QP drafted the manuscript. AP, IR, and FP were involved in the clinical management of the patient and critically revised the manuscript. SB supervised the whole process. All 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. Barber RM, Sorensen RJD, Pigott DM, Bisignano C, Carter A, Amlag JO, et al. Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through nov 14, 2021: A statistical analysis. *Lancet* (2022) 399(10344):2351–80. doi: 10.1016/S0140-6736(22)00484-6

2. Mirza SA, Sheikh AAE, Barbera M, Ijaz Z, Javaid MA, Shekhar R, et al. COVID-19 and the endocrine system: A review of the current information and misinformation. *Infect Dis Rep* (2022) 14(2):184–97. doi: 10.3390/idr14020023

3. Thakur V, Ratho RK, Kumar P, Bhatia SK, Bora I, Mohi GK, et al. Multi-organ involvement in covid-19: Beyond pulmonary manifestations. *J Clin Med* (2021) 10 (3):1–19. doi: 10.3390/jcm10030446

4. Lazartigues E, Qadir MMF, Mauvais-Jarvis F. Endocrine significance of SARS-CoV-2's reliance on ACE2. *Endocrinol* (2020) 161(9):1–7. doi: 10.1210/endocr/bqaa108

5. Dighriri IM, Alhusayni KM, Mobarki AY, Aljerary IS, Alqurashi KA, Aljuaid FA, et al. Pfizer-BioNTech COVID-19 vaccine (BNT162b2) side effects: A systematic review. *Cureus* (2022) 14(3). doi: 10.7759/cureus.23526

6. Pezzaioli LC, Gatta E, Bambini F, Facondo P, Gava M, Cavadini M, et al. Endocrine system after 2 years of COVID-19 vaccines: A narrative review of the literature. *Front Endocrinol (Lausanne)* (2022) 13:1–12. doi: 10.3389/fendo.2022.1027047

7. Zhao Y, Wu X. Influence of COVID-19 vaccines on endocrine system. *Endocrine* (2022) 78(2):241–6. doi: 10.1007/s12020-022-03119-3

8. Sheikh AB, Javaid MA, Sheikh AAE, Shekhar R. Central adrenal insufficiency and diabetes insipidus as potential endocrine manifestations of covid-19 infection: A case report. *Pan Afr Med J* (2021) 38:222. doi: 10.11604/pamj.2021.38.222.28243

9. Au L, Fendler A, Shepherd STC, Rzeniewicz K, Cerrone M, Byrne F, et al. Cytokine release syndrome in a patient with colorectal cancer after vaccination with BNT162b2. *Nat Med [Internet]* (2021) 27(8):1362–6. doi: 10.1038/s41591-021-01387-6

10. Murvelashvili N, Tessnow A. A case of hypophysitis following immunization with the mRNA-1273 SARS-CoV-2 vaccine. *J Investig Med High Impact Case Rep* (2021) 9:0–4. doi: 10.1177/23247096211043386

11. Bouça B, Roldão M, Bogalho P, Cerqueira L, Silva-Nunes J. Central diabetes insipidus following immunization with BNT162b2 mRNA COVID-19 vaccine: A case report. *Front Endocrinol (Lausanne)* (2022) 13:889074. doi: 10.3389/fendo.2022.889074

12. Ankireddypalli AR, Chow LS, Radulescu A, Kawakami Y, Araki T. A case of hypophysitis associated with SARS-CoV-2 vaccination. *AACE Clin Case Rep [Internet]* (2022) 8(5):204–9. doi: 10.1016/j.aace.2022.06.001

13. Das L, Dutta P, Walia R, Mukherjee S, Suri V, Puri GD, et al. Spectrum of endocrine dysfunction and association with disease severity in patients with COVID-19: Insights from a cross-sectional, observational study. *Front Endocrinol (Lausanne)* (2021) 12:1–9. doi: 10.3389/fendo.2021.645787

14. Sheikh AB, Javed N, Sheikh AAE, Upadhyay S, Shekhar R. Diabetes insipidus and concomitant myocarditis: A late sequelae of COVID-19 infection. *J Investig Med High Impact Case Rep* (2021) 9:0–3. doi: 10.1177/2324709621999954

15. Yavari A, Sharifan Z, Larijani B, Mosadegh Khah A. Central diabetes insipidus secondary to COVID-19 infection: a case report. *BMC Endocr Disord [Internet]* (2022) 22(1):4–6. doi: 10.1186/s12902-022-01048-w

16. Lizzi M, Aricò M, Carlone G, Anzellotti MT, Trotta D, Palatino V. Central diabetes insipidus: Another rare complication of SARS-CoV-2 infection in children? *Pediatr Infect Dis J* (2022) 41(10):E448. doi: 10.1097/INF.0000000000003632

17. Misgar RA, Rasool A, Wani AI, Bashir MI. Central diabetes insipidus (Infundibuloneuro hypophysitis): A late complication of COVID-19 infection. *J Endocrinol Invest [Internet]* (2021) 44(12):2855–6. doi: 10.1007/s40618-021-01627-z

18. Balmain J, Jarebi M, Al-Salameh A, Toussaint P, Timmerman M, Chenin L, et al. Pituitary apoplexy in the aftermath of a SARS-CoV-2 infection: A case series from amiens university hospital. *Eur J Endocrinol* (2021) 187(3):K19–25. doi: 10.1530/EJE-22-0056

19. Balcom EF, Nath A, Power C. Acute and chronic neurological disorders in COVID-19: potential mechanisms of disease. *Brain* (2021) 144(12):3576–88. doi: 10.1093/brain/awab302

20. Frara S, Loli P, Allora A, Santini C, di Filippo L, Mortini P, et al. COVID-19 and hypopituitarism. *Rev Endocr Metab Disord [Internet]* (2022) 23(2):215–31. doi: 10.1007/s11154-021-09672-y

21. Pinar-Gutiérrez A, Remón-Ruiz P, Soto-Moreno A. Case report: Pituitary apoplexy after COVID-19 vaccination. *Med Clin (Barc)* (2022) 158:496–500. doi: 10.1016/j.medcle.2021.09.021

22. Roncati L, Manenti A. Pituitary apoplexy following adenoviral vector-based COVID-19 vaccination. *Brain Hemorrhages* (2023) 4(1):27–9. doi: 10.1016/j.hest.2022.04.002

23. Morita S, Tsuji T, Kishimoto S, Uraki S, Takeshima K, Iwakura H, et al. Isolated ACTH deficiency following immunization with the BNT162b2 SARS-CoV-2 vaccine: A case report. *BMC Endocr Disord [Internet]* (2022) 22(1):1–6. doi: 10.1186/s12902-022-01095-3

24. Mizuno T, Takahashi R, Kamiyama T, Suzuki A, Suzuki M. Neuroleptic malignant syndrome with adrenal insufficiency after BNT162b2 COVID-19 vaccination in a man taking valproate: A case report. *Am J Case Rep* (2022) 23:4–7. doi: 10.12659/AJCR.936217

25. Lindner G, Ryser B. The syndrome of inappropriate antidiuresis after vaccination against COVID-19: Case report. *BMC Infect Dis [Internet]* (2021) 21 (1):1–3. doi: 10.1186/s12879-021-06690-8

26. Burçin Gİ, Süleyman NŞ, Uğur Ü.. Three cases of subacute thyroiditis following SARS-CoV-2 vaccine: Post- vaccination ASIA syndrome burçin. *J Clin Endocrinol Metab* (2021) 106(9):2600–5. doi: 10.1210/clinem/dgab373

27. Lo Sardo L, Parisi S, Ditto MC, De Giovanni R, Maletta F, Grimaldi S, et al. New onset of giant cell arteritis following ChAdOx1-s (Vaxevria®) vaccine administration. *Vaccines* (2023) 11(2):1–8. doi: 10.3390/vaccines11020434

28. Paras PS, Samuel G, Daniel Z, Amanda W, Rashmi V. Optic neuritis secondary to the pfizer-BioNTech-162b2 COVID-19 vaccine managed with plasmapheresis: A case report and review. *Oman J Ophthalmol* (2022) 15(3):397–402. doi: 10.4103/ojo.ojo_354_21

29. Arima H, Bichet DG, Cheetham T, Christ-Crain M, Drummond J, Gurnell M, et al. Changing the name of diabetes insipidus. *Pituitary* (2022) 0:1–3. doi: 10.1007/s11102-022-01276-2

30. Satogami N, Miki Y, Koyama T, Kataoka M, Togashi K. Normal pituitary stalk: High-resolution MR imaging at 3T. *Am J Neuroradiol* (2010) 31(2):355–9. doi: 10.3174/ajnr.A1836

31. Elisabetta G, Morana G, Di Iorgi N, Pistorio A, Allegri AEM, Napoli F, et al. Role of MRI T2-DRIVE in the assessment of pituitary stalk abnormalities without gadolinium in pituitary diseases. *Eur J Endocrinol* (2018) 178:613-6–22. doi: 10.1530/EJE-18-0094

32. Zhou X, Zhu H, Yao Y, Lian X, Feng F, Wang L, et al. Etiological spectrum and pattern of change in pituitary stalk thickening: Experience in 321 patients. *J Clin Endocrinol Metab* (2019) 104(8):3419–27. doi: 10.1210/jc.2018-02297

33. Moszczyńska E, Kunecka K, Baszyńska-Wilk M, Perek-Polnik M, Majak D, Grajkowska W. Pituitary stalk thickening: Causes and consequences. the children's memorial health institute experience and literature review. *Front Endocrinol (Lausanne)* (2022) 13:868558. doi: 10.3389/fendo.2022.868558

34. Prodam F, Caputo M, Mele C, Marzullo P, Aimaretti G. Insights into non-classic and emerging causes of hypopituitarism. *Nat Rev Endocrinol [Internet]* (2021) 17 (2):114–29. doi: 10.1038/s41574-020-00437-2

35. Kumar T, Nigam JS, Jamal I, Jha VC. Primary pituitary tuberculosis. *Autops Case Rep* (2021) 11:e2020228. doi: 10.4322/acr.2020.228

36. Partenope C, Carceller F, Pozzobon G, Albanese A, Weber G. Implications of deferred diagnosis of paediatric intracranial germ cell tumours. *Pediatr Blood Cancer* (2023) 70(3):e30168. doi: 10.1002/pbc.30168

37. Bellastella G, Maiorino MI, Bizzarro A, Giugliano D, Esposito K, Bellastella A, et al. Revisitation of autoimmune hypophysitis: knowledge and uncertainties on pathophysiological and clinical aspects. *Pituitary* (2016) 19(6):625–42. doi: 10.1007/s11102-016-0736-z

38. Iwama S, Welt CK, Romero CJ, Radovick S, Caturegli P. Isolated prolactin deficiency associated with serum autoantibodies against prolactin-secreting cells. *J Clin Endocrinol Metab* (2013) 98(10):3920–5. doi: 10.1210/jc.2013-2411

39. Das L, Dutta P. Approach to the patient: A case with an unusual cause of hypopituitarism. *J Clin Endocrinol Metab* (2022) dgac747. doi: 10.1210/clinem/dgac747

40. Di Iorgi N, Allegri AEM, Napoli F, Calcagno A, Calandra E, Fratangeli N, et al. Central diabetes insipidus in children and young adults: Etiological diagnosis and long-term outcome of idiopathic cases. *J Clin Endocrinol Metab* (2014) 99(4):1264–72. doi: 10.1210/jc.2013-3724

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 04 May 2023
DOI 10.3389/fendo.2023.1161637

OPEN ACCESS

EDITED BY
Fariba Ahmadizar,
University Medical Center Utrecht,
Netherlands

REVIEWED BY
Soroush Mohammadi Jouabadi,
Erasmus Medical Center, Netherlands
Nicoletta Luxi,
University of Verona, Italy

*CORRESPONDENCE
Maciej T. Malecki
maciej.malecki@uj.edu.pl

SPECIALTY SECTION
This article was submitted to Clinical Diabetes, a section of the journal Frontiers in Endocrinology

RECEIVED 08 February 2023
ACCEPTED 06 April 2023
PUBLISHED 04 May 2023

CITATION
Kania M, Koń B, Kamiński K, Hohendorff J, Witek P, Klupa T and Malecki MT (2023) Diabetes as a risk factor of death in hospitalized COVID-19 patients – an analysis of a National Hospitalization Database from Poland, 2020. *Front. Endocrinol.* 14:1161637. doi: 10.3389/fendo.2023.1161637

Diabetes as a risk factor of death in hospitalized COVID-19 patients – an analysis of a National Hospitalization Database from Poland, 2020

Michal Kania[1,2], Beata Koń[3], Konrad Kamiński[3], Jerzy Hohendorff[1,2], Przemysław Witek[1,2], Tomasz Klupa[1,2] and Maciej T. Malecki[1,2]*

[1]Department of Metabolic Diseases and Diabetology, Jagiellonian University Medical College, Krakow, Poland, [2]Department of Metabolic Diseases and Diabetology, University Hospital, Krakow, Poland, [3]Department of Analysis and Innovation, National Health Fund, Warsaw, Poland

Introduction: Diabetes is one of the comorbidities associated with poor prognosis in hospitalized COVID-19 patients. In this nationwide retrospective study, we evaluated the risk of in-hospital death attributed to diabetes.

Methods: We analyzed data from discharge reports of patients hospitalized with COVID-19 in 2020 as submitted to the Polish National Health Fund. Several multivariate logistic regression models were used. In each model, in-hospital death was estimated with explanatory variables. Models were built either on the whole cohorts or cohorts matched with propensity score matching (PSM). The models examined either the main effects of diabetes itself or the interaction of diabetes with other variables.

Results: We included 174,621 patients with COVID-19 who were hospitalized in the year 2020. Among them, there were 40,168 diabetic patients (DPs), and the proportion of DPs in this group was higher than in the general population (23.0% vs. 9.5%, p<0.001). In this group of COVID-19 hospitalizations, 17,438 in-hospital deaths were recorded, and the mortality was higher among DPs than non-diabetics (16.3% vs. 8.1%, p<0.001). Multivariate logistic regressions showed that diabetes was a risk factor of death, regardless of sex and age. In the main effect analysis, odds of in-hospital death were higher by 28.3% for DPs than for non-diabetic patients. Similarly, PSM analysis including 101,578 patients, of whom 19,050 had diabetes, showed that the risk of death was higher in DPs regardless of sex with odds higher by 34.9%. The impact of diabetes differed among age groups and was the highest for patients aged 60-69.

Conclusions: This nationwide study confirmed that diabetes was an independent risk factor of in-hospital death in the course of COVID-19 infection. However, the relative risk differed across the age groups.

KEYWORDS

COVID-19, diabetes, mortality, epidemiology, modelling, propensity-score matching

1 Introduction

Coronavirus disease-2019 (COVID-19) caused by the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), emerged at the end of 2019 and caused a global pandemic (1). The first case of COVID-19 in Poland was identified on March 3, 2020. This was followed by the first wave of the pandemic and triggered the first national lockdown that lasted approximately until the end of June 2020. The second wave started at the beginning of August and ended in January 2021. According to official data from the Polish government, a total of 68,505 excess deaths were recorded in 2020, which accounts for a 16% increase in total mortality compared to the period 2017-2019 (2). The challenges concerning preparation of the healthcare system to properly respond to this new epidemiological threat, required urgent action. The COVID-19 pandemic profoundly impacted the healthcare system in Poland. The Polish Ministry of Health initially decided to create a network of hospitals dedicated exclusively to serve as a multispecialty reference center for COVID-19 patients (3, 4). This model was modified in October 2020 when temporary COVID-19 hospitals were established and numerous regional hospitals were converted in order to play a supportive role in managing COVID-19 patients (5). Until the end of July 2022, there were 6,050,000 cases of SARS-Cov-2 infections in Poland with deaths totaling around 115,500 (6).

Numerous factors, such as male gender, age, diabetes, obesity, or cardiovascular disease, have been identified as risk factors for hospitalization of COVID-19 patients (7–10). Diabetes, along with hypertension, cardiovascular and respiratory diseases, is one of the most common comorbidities of COVID-19. It has been associated with a higher risk of unfavorable COVID-19 outcomes, such as longer hospitalization, higher in-hospital mortality, ICU admission and requirement for mechanical ventilation (11–14). However, most original studies reporting this association were retrospective single-center reports. There are few analyses of datasets from large healthcare initiatives or from the nationwide level of COVID-19 patients (15, 16).

In this retrospective study, we evaluated diabetes as a risk factor for in-hospital mortality using data from the Polish National Health Fund, a state institution that finances healthcare services from contributions paid by the insured persons (17) and runs a nationwide healthcare database.

2 Methods

For this retrospective analysis, information concerning patients hospitalized between the 1st Jan and 31st Dec 2020 due to COVID-19 was extracted from data reported to the Polish National Health Fund. Analyzed data included hospitalizations that were financed from the governmental COVID-19 Counteracting Fund and included information on both insured and uninsured patients. If the patient was hospitalized more than once in 2020, then the first hospitalization of the patient was included in the study. Only adult patients were included in the study defined as people aged at least 17 years old as of Dec 31st 2019.

Hospital information from discharge charts of COVID-19 patients was merged with data from diabetes databases and data regarding comorbidities from reports delivered to the National Health Fund that health care providers sent to receive reimbursement for health care benefits delivered due to other health problems than COVID-19. They included medical benefits reported with a specific ICD-10 diagnosis code (either as a main or additional diagnosis), or ICD-9 procedure code in the case of hemodialysis, if the patient had it within the last 3 years. The list included diabetes, obesity, loss of weight, cardiac arrhythmias, arterial hypertension, heart failure, peripheral artery disease, dyslipidemia, ischemic heart disease, history of myocardial infarction and stroke, cardiomyopathy, valvular disorders, atrial fibrillation, pulmonary circulation disorders, chronic pulmonary diseases, chronic kidney disease, hypothyroidism and hyperthyroidism, other thyroid diseases, liver disorders, coagulopathies, anemia, electrolyte imbalances, history of neoplasm, rheumatoid diseases, paralysis, other neurological disorders, depression, drug abuse, and psychosis. The full list with corresponding ICD codes is presented in Table S1 in the Appendix.

Separately, for diabetic patients (DPs) and patients without diabetes, characteristics such as mean age, proportion of population older than 65, proportion of population older than 85, sex and in-hospital death were compared. For each of these variables, we examined whether differences between diabetic and non-diabetic populations were statistically significant. For age, Welch two sample t-test was used, and for other characteristics, referring to share of population, a chi-square test was used.

2.1 Logistic regression modelling

Econometric modelling was used to analyze the impact of multiple variables on the risk of in-hospital death of COVID-19 patients. Four different multivariate logistic regression models were estimated. In each model, a dependent variable was binary in nature describing if a patient's hospitalization ended at the same day, when the patient died. Explanatory variables included diagnosis of diabetes, as well as additional confounders: sex, logarithm of age (recorded as difference between 2020 and year of patient's birth) and the diagnosis of the selected comorbidities (named above). The first two models were built for the whole cohort, another two for the matched cohort using propensity score matching (PSM). All models were estimated with the same set of explanatory variables, however in the first and third model only main effects were included. In the second and fourth model, interactions between diabetes and other variables were also included.

PSM was applied, as diabetes can affect probability of being hospitalized due to COVID-19 causing a non-random sample. Each DP was matched with non-diabetic patient(s) according to other 36 variables (sex, age group 18-24, 25-29, 30-34,…, 85-89, 90-94, 95+, comorbidities other than diabetes as shown in Table S1, i.e. arterial hypertension, other neoplasms, health failure, cardiac arrhythmias etc.). One-to-many exact matching was used (18), i.e. non-diabetic patient or patients were matched to each singular DP (according to

above mentioned parameters). The number of matched non-diabetics was not limited. If DP was not matched with non-diabetic patient, their data was discarded from the analysis.

In order to minimize the number of variables and include only those with strong effect on the in-hospital death, the following approach to select variables in the final models for each of four setting was used: the dataset was divided into train (80% of observations) and test dataset (20% of observations). In the first step, a model only with the logarithm of age was built. Then, for each explanatory variable, a model with the logarithm of age and that variable was built. Furthermore, each of those models was compared with the model containing only the logarithm of age and Bayes Factor (BF) was calculated (19). If the maximum value of BF was greater than 30, then the variable with the highest BF was added to the final model. Subsequently, the procedure was repeated, and a model with the logarithm of age and selected variable was built; each of the remaining variables was separately added to the model, and for each model BF was calculated. If the maximum BF was greater than 30, then the variable was added to the model and the procedure was repeated until maximum BF for remaining variables was above 30.

The models' estimates were used to calculate the probability of in-hospital death according to the patient's age, sex and diagnosis of diabetes. The quality of the models was tested on a test dataset and final models were estimated on the whole dataset. To assess goodness of fit, McFadden's pseudo R^2 was used. Furthermore, the quality of the models was verified with area under ROC curve (AUC) calculated on the test set.

All the analyses were conducted with SPSS Pro 6.0 and in R, version 3.6.1 (20, 21). For statistical significance a threshold of 0.05 for p-value was used.

3 Results

The study included a total of 174 621 COVID-19-infected patients who were hospitalized in Poland between the 1st Jan and 31st Dec 2020. The mean age was 60.5 ± 18.9 years and the median age was 63 years (IQR 29). The median age in diabetic patients was 72 (IQR 16) and 59 (IQR 31) in non-diabetic patients. There were 90 628 (51.9%) men in the study group (Table 1). The estimated number of patients with diabetes in Poland in 2018 was 2,864,000 (approximately 9.1% of adult population according to official registries (22)). Of 174 621 patients hospitalized with COVID-19 infection, 40 168 had diabetes (23.0%) and it was more common than in the general population (p<0.001). The prevalence of comorbidities, including but not limited to cardiovascular, metabolic, pulmonary, psychiatric and endocrine disorders, was higher in DP than in non-diabetic COVID-19 patients (Table 2). 17 438 in-hospital deaths were recorded (10.0%), the mortality was higher in DPs, compared to non-diabetics (16.3% vs. 8.1%, p<0.001). It was higher across all age groups regardless of patients gender and age (Table S2).

Results of multivariate logistic regression (Table 3) showed that diabetes and some other comorbidities, including chronic kidney disease, heart failure, chronic ischemic heart disease, cardiac arrythmias and neoplasms, were associated with a higher risk of death, regardless of sex and age. The risk was higher in older patients and in males. For patients with diabetes, odds of in-hospital death were higher by 28.3% than for patients without diabetes. The regression model that included interactions of variable referring to diabetes with other variables, showed that impact of diabetes on COVID-19 in-hospital death was influenced by the patient's age (Table S3). Results of the model are visualized on Figures 1; 2. Figure 1 shows probability of in-hospital death according to patients' sex, age and the fact of being diabetic without other comorbidities. Figure 2 shows probability of in-hospital death according to patients' sex, age and fact of being diabetic for analyzed group of patients. Results show that the highest difference in median value of COVID-19 in-hospital death between diabetic and non-diabetic patients was for patients in age group 65-69 (4.3 percent point), 60-64 and 55-59 (4.2 percent point) (Table S3).

For PSM, data from 101 578 patients was analyzed, of whom 19 050 had diabetes (i.e. 21 118 diabetic were not matched to non-diabetic patients). The characteristics of population matched by PSM is shown in Tables S4—S6 in the Appendix. In the group of patients matched with PSM, the risk of COVID-19 in-hospital death was higher in DPs than for non-diabetic patients regardless of sex and age. Results of the multivariate logistic regression model showed that the odds for in-hospital death were 72.5% higher for men than for women (Table 4). For DPs, odds of death were higher by 34.9% than for non-diabetic patients. After adding interactions to the model, it showed similar results to the model estimated without PSM, i.e. age differentiated impact of the diabetes on the in-

TABLE 1 Basic characteristics of DPs and non-diabetic patients with in-hospital death data.

Characteristics	All	Diabetes	Non-diabetes	P value
Number	174 621	40 168 (23%.0)	134 453 (77.0%)	–
Age [years (SD)]	60.5 (18.92)	70.9 (12.58)	57.4 (19.39)	<0.001
Age > 65 years [n (%)]	113 643 (65.1%)	29 927 (74.5%)	53 789 (40%)	<0.001
Age > 85 years [n (%)]	21 679 (12.4%)	5 094 (12.7%)	11 491 (8.5%)	<0.001
Male [n (%)]	90 628 (51.9%)	21 113 (52.6%)	69 515 (51.7%)	0.002
Endpoint				
In-hospital death [Yes (%)]	17 438 (10.0%)	6 541 (16.3%)	10 897 (8.1%)	<0.001

TABLE 2 The prevalence of comorbidities in the hospitalized population.

Comorbidities	Non-diabetic	Diabetic	P value
Number	134 453 (77.0%)	40 168 (23%.0)	-
Arterial hypertension	60 742 (45.2%)	34 363 (85.6%)	<0.001
Neoplasm. Other	26 445 (19.7%)	9 366 (23.3%)	<0.001
Heart failure	19 843 (14.8%)	15 286 (38.1%)	<0.001
Cardiac arrhythmias	22 871 (17.0%)	12 216 (30.4%)	<0.001
Dyslipidaemia	21 237 (15.8%)	13 283 (33.1%)	<0.001
Ischemic heart disease	18 708 (13.9%)	14 446 (36.00%)	<0.001
Chronic pulmonary disorders	19 826 (14.8%)	8 406 (20.9%)	<0.001
Peripheral artery disease	15 029 (11.2%)	9 887 (24.6%)	<0.001
Atrial fibrillation	13 189 (9.8%)	8 615 (21.5%)	<0.001
Neoplasm. malignant	14 212 (10.6%)	6 278 (15.6%)	<0.001
Anaemia	12 096 (9.0%)	6 207 (15.5%)	<0.001
Hypothyroidism	11 010 (8.2%)	4 712 (11.7%)	<0.001
Chronic kidney disease	7 409 (5.5%)	7 253 (18.1%)	<0.001
Depression	10 104 (7.5%)	3 227 (8.0%)	<0.001
Other neurological disorders	8 481 (6.3%)	2 911 (7.3%)	<0.001
History of stroke	6 500 (4.8%)	4 013 (10.0%)	<0.001
Angina	5 227 (3.9%)	4 002 (10.0%)	<0.001
Electrolyte imbalances	5 552 (4.1%)	3 023 (7.5%)	<0.001
Alcohol abuse	7 104 (5.3%)	1 377 (3.4%)	<0.001
Obesity	3 600 (2.7%)	4 039 (10.1%)	<0.001
Liver disorders	4 361 (3.2%)	2 145 (5.3%)	<0.001
Valvular disorders	4 172 (3.1%)	2 331 (5.8%)	<0.001
Rheumatoid diseases	4 527 (3.4%)	1 627 (4.1%)	<0.001
History of myocardial infarction	2 974 (2.2%)	2 536 (6.3%)	<0.001
Other thyroid disorders	3 874 (2.9%)	1 237 (3.1%)	0.03
Paralysis	2 857 (2.1%)	1 473 (3.7%)	<0.001
Psychosis	3 198 (2.4%)	847 (2.1%)	0.002
Hyperthyroidism	2 350 (1.8%)	1 038 (2.6%)	<0.001
Dialysis	1 680 (1.3%)	1 410 (3.5%)	<0.001
Pulmonary circulation disorders	1 996 (1.5%)	846 (2.1%)	<0.001
Coagulopathies	1 767 (1.3%)	735 (1.8%)	<0.001
Cardiomyopathy	1 425 (1.2%)	966 (2.4%)	<0.001
Drug abuse	1 897 (1.4%)	355 (0.9%)	<0.001
Weight loss	1 746 (1.3%)	491 (1.2%)	0.23

hospital risk of death (Table 4). Results of the models for patients without other comorbidities are presented on Figure 3. Figure 4 presents values of estimated COVID-19 in-hospital death according to patients' age group and fact of being diabetic for the analyzed group of patients. The median values are additionally presented in Table S7. The results show that the median value of the probability of in-hospital death was highest for patients aged 95 and more. The highest difference (in absolute values) between diabetic and non-

TABLE 3 Results of logistic regression modelling for the risk of in-hospital death.

Variable	Model without interactions of variable diabetes with other variables			Model with interactions of variable diabetes with other variables		
	Estimate (SE)	Odds ratio (95% CI)	P-value	Estimate (SE)	Odds ratio (95% CI)	P-value
Intercept	-18.714 (0.208)	0.000 (0.000 – 0.000)	<0.001	-19.845 (0.236)	0.000 (0.000 – 0.000)	<0.001
log(Age)	3.838 (0.049)	46.414 (42.179 - 51.075)	<0.001	4.104 (0.055)	60.560 (54.334 - 67.5)	<0.001
Gender: male	0.425 (0.017)	1.529 (1.478 - 1.581)	<0.001	0.413 (0.017)	1.511 (1.460 - 1.563)	<0.001
Diabetes	0.249 (0.018)	1.283 (1.238 - 1.328)	<0.001	5.787 (0.46)	325.885 (132.16 - 803.578)	<0.001
Chronic kidney disease	0.218 (0.026)	1.243 (1.181 - 1.308)	<0.001	0.228 (0.026)	1.256 (1.193 - 1.322)	<0.001
Other neurological disorders	0.299 (0.029)	1.348 (1.273 - 1.428)	<0.001	0.299 (0.029)	1.348 (1.273 - 1.428)	<0.001
Heart failure	0.245 (0.020)	1.278 (1.228 - 1.329)	<0.001	0.247 (0.02)	1.280 (1.231 - 1.332)	<0.001
Neoplasm, malignant	0.257 (0.024)	1.293 (1.235 - 1.354)	<0.001	0.257 (0.023)	1.293 (1.235 - 1.354)	<0.001
Psychosis	0.391 (0.053)	1.478 (1.332 - 1.640)	<0.001	0.375 (0.053)	1.456 (1.312 - 1.615)	<0.001
Cardiac arrhythmias	-0.126 (0.02)	0.881 (0.847 - 0.917)	<0.001	-0.121 (0.02)	0.886 (0.851 - 0.921)	<0.001
Haemodialysis	0.368 (0.052)	1.445 (1.304 - 1.601)	<0.001	0.337 (0.052)	1.401 (1.264 - 1.552)	<0.001
Neoplasm, other	-0.156 (0.022)	0.856 (0.82 - 0.893)	<0.001	-0.159 (0.022)	0.853 (0.817 - 0.891)	<0.001
Weight loss	0.363 (0.058)	1.438 (1.284 - 1.610)	<0.001	0.357 (0.058)	1.430 (1.277 - 1.601)	<0.001
Coagulopathies	0.331 (0.061)	1.392 (1.234 - 1.570)	<0.001	0.312 (0.061)	1.366 (1.211 - 1.540)	<0.001
Other thyroid disorders	-0.313 (0.061)	0.731 (0.649 - 0.823)	<0.001	-0.321 (0.061)	0.726 (0.644 - 0.817)	<0.001
Chronic ischemic heart disease	-0.092 (0.020)	0.912 (0.877 - 0.948)	<0.001	-0.094 (0.020)	0.910 (0.876 - 0.946)	<0.001
Anaemia	0.105 (0.024)	1.111 (1.059 - 1.164)	<0.001	0.106 (0.024)	1.112 (1.060 - 1.165)	<0.001
D:log(Age)*	–	–	–	-1.258 (0.107)	0.276 (0.224 - 0.341)	<0.001
*D: - interaction with diabetes	Train dataset – McFadden's pseudo-R2: 0.138 Test dataset – AUC (area under ROC curve): 0.770 Complete dataset – McFadden's pseudo-R2: 0.138			Train dataset – McFadden's pseudo-R2: 0.139 Test dataset – AUC (area under ROC curve): 0.771 Complete dataset – McFadden's pseudo-R2: 0.139		

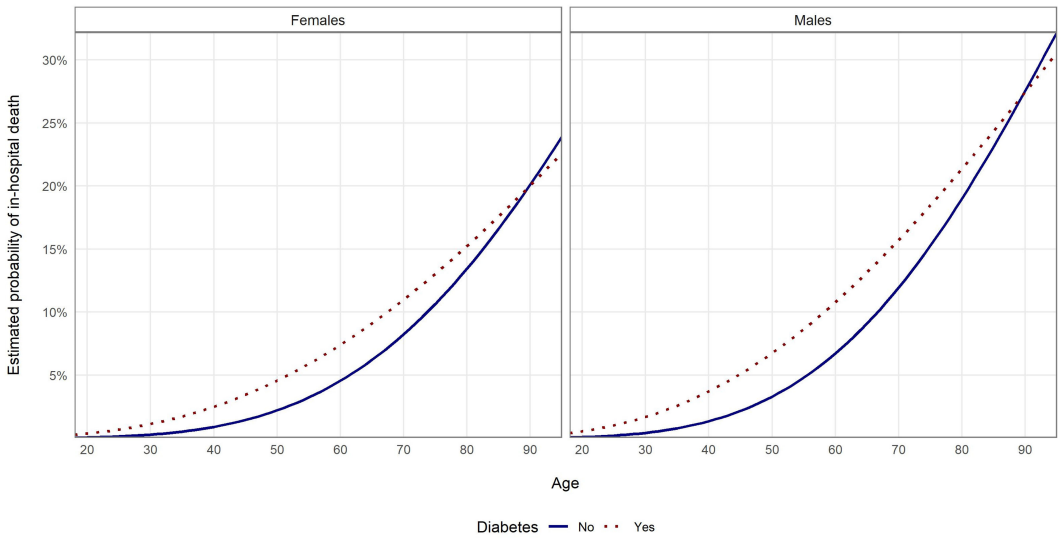


FIGURE 1
Probability of in-hospital death in the examined groups according to age and sex estimated with multivariate logistic regression without PSM and with interaction analysis for patients without other comorbidities.

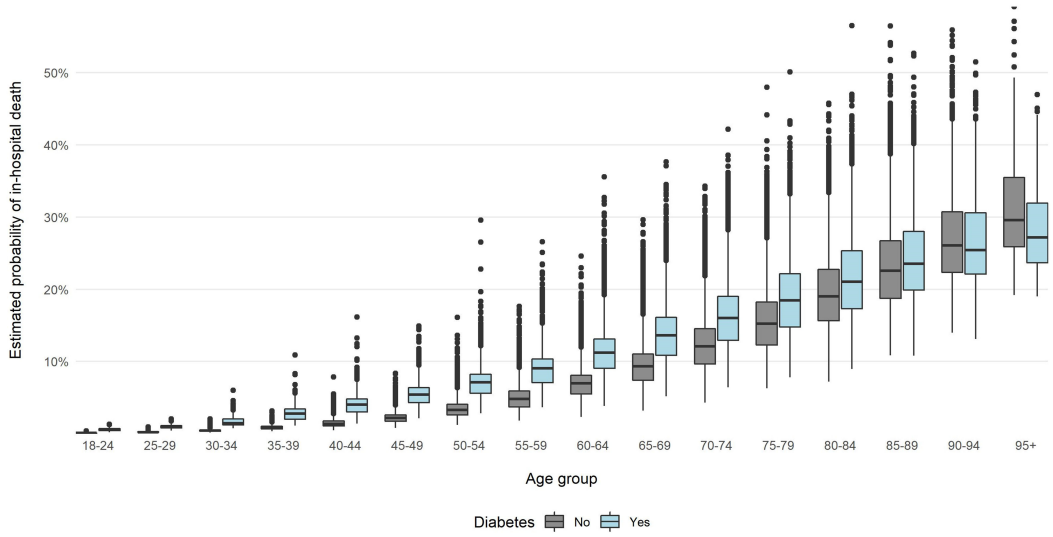


FIGURE 2
Probability of in-hospital death in the examined groups according to age and sex estimated with multivariate logistic regression without PSM and with interaction analysis.

diabetic patients was observed for age groups 60-64 (4.8 percent point) and 65-69 (4.5 percent point).

4 Discussion

In this report we present the analysis of a nationwide cohort of patients hospitalized due to COVID-19 in 2020. Of note, our study is one of the very few assessing the prevalence of COVID-19 hospitalizations and mortality on the nationwide level, with most studies to date focusing on the case series or hospital-based cohort of patients. We confirmed that diabetes was an independent risk of hospitalization and in-hospital death in the course of COVID-19, and provided also some new observations that are discussed below.

Results of our unadjusted analysis showed that the hospitalization rate was almost 2.5 times higher for DP than for non-diabetics. Most studies, to date, mainly up to mid-2021, reported a higher frequency of hospitalization due to COVID-19 (7–10, 23), with the risk being up to 4 times higher in DPs (23). After adjustment for age, gender, and co-existing comorbidities, the risk decreased, but in most studies was

TABLE 4 Results of logistic regression modelling after PSM analysis for the risk of in-hospital death.

Variable	Model without interactions of variable diabetes with other variables			Model with interactions of variable diabetes with other variables		
	Estimate (SE)	Odds ratio (95% CI)	P-value	Estimate (SE)	Odds ratio (95% CI)	P-value
Intercept	-21.926 (0.317)	0.000 (0.000 – 0.000)	<0.001	-23.170 (0.358)	0.000 (0.000 – 0.000)	<0.001
log(Age)	4.561 (0.074)	95.716 (82.78 - 110.675)	<0.001	4.855 (0.084)	128.335 (108.918-151.214)	<0.001
Gender: male	0.545 (0.026)	1.725 (1.638 - 1.817)	<0.001	0.530 (0.026)	1.699 (1.614-1.789)	<0.001
Diabetes	0.300 (0.027)	1.349 (1.28 - 1.422)	<0.001	6.526 (0.694)	682.533 (175.129-2660.04)	<0.001
Chronic kidney disease	0.305 (0.056)	1.357 (1.215 - 1.515)	<0.001	0.319 (0.056)	1.375 (1.232-1.535)	<0.001
Other neurological disorders	0.420 (0.068)	1.523 (1.333 - 1.739)	<0.001	0.424 (0.068)	1.529 (1.338-1.746)	<0.001
Neoplasm, malignant	0.329 (0.045)	1.390 (1.273 - 1.518)	<0.001	0.328 (0.045)	1.389 (1.272-1.516)	<0.001
Neoplasm, other	-0.220 (0.042)	0.803 (0.739 - 0.871)	<0.001	-0.223 (0.042)	0.800 (0.737-0.868)	<0.001
Heart failure	0.233 (0.034)	1.263 (1.181 - 1.351)	<0.001	0.247 (0.034)	1.280 (1.197-1.369)	<0.001
Cardiac arrhythmias	-0.230 (0.038)	0.795 (0.738 - 0.856)	<0.001	-0.221 (0.038)	0.802 (0.745-0.864)	<0.001
D:log(Age)*	–	–	–	-1.450 (0.162)	0.234 (0.171-0.322)	<0.001
*D: - interaction with diabetes	Train dataset – McFadden's pseudo-R2: 0.173 Test dataset – AUC (area under ROC curve): 0.807 Complete dataset – McFadden's pseudo-R2: 0.172			Train dataset – McFadden's pseudo-R2: 0.174 Test dataset – AUC (area under ROC curve): 0.808 Complete dataset – McFadden's pseudo-R2: 0.174		

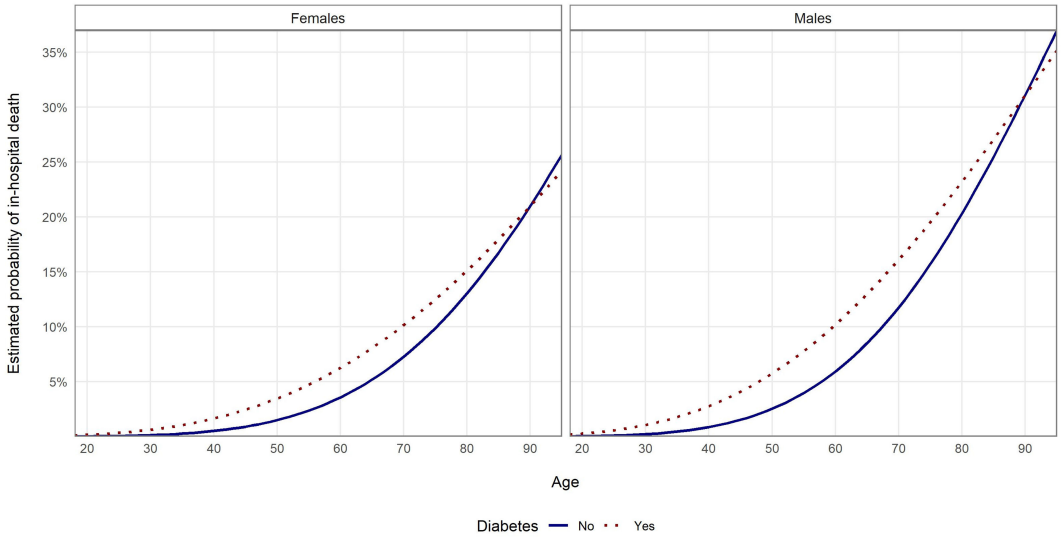


FIGURE 3
Probability of in-hospital death in the examined groups according to age and sex estimated with multivariate logistic regression with PSM and with interaction analysis for patients without other comorbidities.

still statistically significant (7, 8, 10). In cohorts of the European origin, the risk for hospitalization in DPs was ca. twice as high as in population without diabetes (10, 24). Importantly, our data covers the period of the first two waves of COVID-19 pandemic in Poland. These numbers should be seen in the context that initially, due to a mandatory supervision by epidemiological services or hospitalization related to COVID-19, the number of patients admitted to the hospitals was disproportionately higher, as hospitalization of all patients with SARS-CoV-2 infection was obligatory (25). Of note, the proportion of DPs among all hospitalized COVID-19 patients identified on the state level was slightly lower than in our earlier report involving a large single center cohort from the University Hospital in Krakow (26,3%) (26). Still, multiple reasons may contribute to the higher frequency of hospitalizations in patients with comorbidities, including diabetes. DPs were older, with a higher prevalence of comorbidities, including cardiovascular disorders, thus subjecting them to COVID-19 complications. Infections are generally more common in DPs and often occur with increased severity (27).

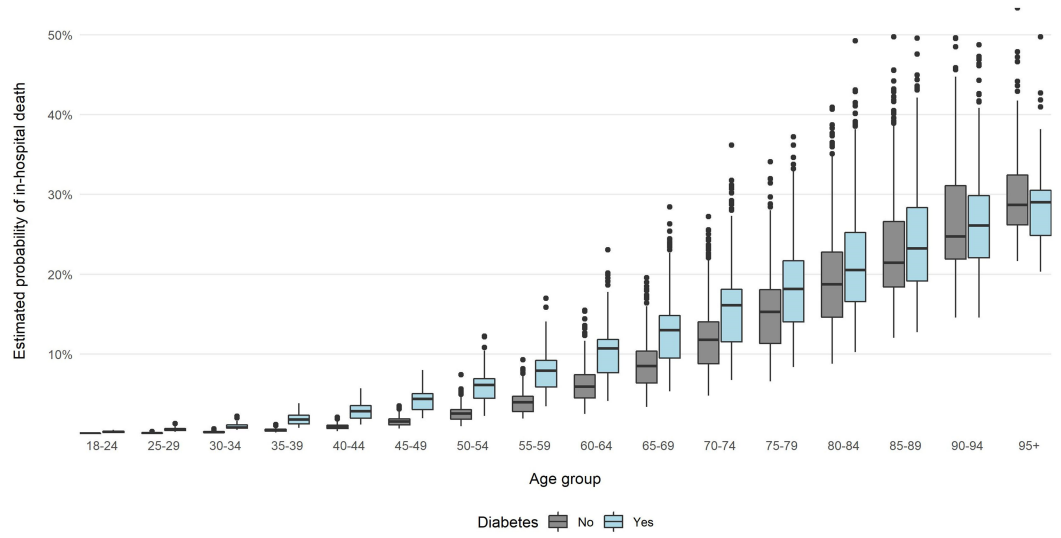


FIGURE 4
Probability of in-hospital death in the examined groups according to age and sex estimated with multivariate logistic regression with PSM and with interaction analysis.

Some aspects of immune response to infections, including lymphocyte response, macrophage and granulocyte function, may be impaired in DPs group (28). Some other potential mechanisms responsible for a higher susceptibility to severe COVID-19 in DPs may include predisposition to hyperinflammatory reaction, higher affinity of SARS-Cov-2 virus to cell membranes and decreased viral clearance (29).

Diabetes was a risk factor for in-hospital death in hospitalized COVID-19 patients in the large cohort examined in our study. This is in line with previous studies, showing that diabetes is an independent risk factor for disease severity and in-hospital death (10, 15, 16, 30–32). It should be noted that the prevalence of numerous comorbidities, including cardiovascular diseases, was higher in DP than in non-diabetic patients. To date, research efforts focused primarily on single-center cohorts of patients, with less common analyses of datasets from large healthcare initiatives, or on nationwide level, thus, making our report noteworthy (15, 16).

Older age has been reported as an important risk factor for COVID-19-related mortality (33), which was also, as expected, observed in our study. Notably, the biggest difference in in-hospital mortality between DPs and non-diabetic patients was observed in 40-64 age groups, which is a result similar to some other studies (10, 34). Our results specifically point to patients aged 60-69, as in this group the impact of diabetes on COVD-19 in-hospital death differed the most between the two examined groups. It is likely that diabetes may not further increase the risk of mortality in elderly patients as the advanced age itself is one of the strongest risk factors (34). Conversely, in the younger population groups, the role played by diabetes is more apparent. Similar observations were reported for male patients in a recent systematic review, with nearly 50% higher risk of death than in males than in females (33). These distinctions were previously attributed to differences in the function of the immune system, including counts of selected lymphocytes, lack of some regulatory genes that are located on the X chromosome (33), and age-related mosaic loss of chromosome Y in elderly males (35). Moreover, a recent study explained these disparities to some extent by social and contextual factors (36).

Uniquely, in our study we developed a model that enabled us to quantify the risk of mortality in the cohort of patients hospitalized due to COVID-19 in Poland. The main strength of this model is the fact that it was supplied with data from a large population.

Finally, limitations of this study should be discussed. First, this is a retrospective, observational study and, thus, prone to many biases related to this study design. For this reason, causative relationships cannot be claimed based on this research. Secondly, we aimed to investigate solely patients admitted to the hospital, thus, our results cannot be automatically extended on the entire population affected by COVID-19. As previous studies showed, around half of the deaths due to COVID-19 occurred in non-hospitalized persons, with the majority of these patients residing in long-term care facilities (16). In addition, our models could have been supplied with deficient data as the data investigated included only that reported by various health providers. We are not able to assure the data's full credibility. As we used billing data, some of comorbidities may have not been reported as individual diagnoses, such as obesity, with surprisingly low frequency reported in our dataset. We also cannot verify the criteria that were used to diagnose reported disorders.

Another limitation is that National Health Fund data includes only insured people. i.e. 90% of citizens of Poland. While our data on COVID-19 considered both insured and uninsured people, this can cause a potential bias on the information concerning other comorbidities for uninsured people, as only health services related to COVID-19 were state-financed for them.

Moreover, we were not able to differentiate between type 1 and type 2 diabetes, as there some inconsistency regarding reporting these with an appropriate ICD-10 code (in Polish translation of ICD-10 classification code E11 refers to non-insulin dependent diabetes and in English version it refers to type 2 diabetes). Nevertheless, the same data gathering methodology is routinely utilized by the Polish National Health Fund to inform the decision making. Thus, we believe that its quality was adequate for this study.

To summarize, in this nationwide retrospective study, diabetes was associated with higher frequency of hospitalization and a higher risk of in-hospital death in the course of COVID-19, regardless of sex, age and some of selected comorbidities, including chronic kidney disease, heart failure, chronic ischemic heart disease, cardiac arrythmias and neoplasms. However, the relative risk attributed to diabetes differed significantly across the age groups and genders. This relative risk was particularly high in males and patients in their sixties. This was one of the largest datasets of hospitalized diabetic patients analyzed since the outbreak of the COVID-19 pandemic in 2019 in Poland. Our findings can inform individual clinicians' decisions and public healthcare providers on the risk associated with COVID-19 for individual populations.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: Raw data analyzed in the article is proprietary to National Health Fund, Warsaw, Poland. Requests to access these datasets should be directed to maciej.malecki@uj.edu.pl.

Ethics statement

The studies involving human participants were reviewed and approved by Jagiellonian University Ethics Committee. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

MK, JH, PW, TK and MTM contributed to conception and design of the study. BK and KK organized the database. BK and KK

performed the statistical analysis. MK wrote the first draft of the manuscript. JH, PW, TK, MTM wrote sections of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This publication was supported by the National Center for Research and Development CRACoV-HHS project (Model of multi-specialist hospital and non-hospital care for patients with SARSCoV-2 infection) through the initiative "Support for specialist hospitals in fighting the spread of SARSCoV-2 infection and in treating COVID-19" (contract number SZPITALE-JEDNOIMIENNE/18/2020). The described research was implemented by consortium of the University Hospital in Krakow and the Jagiellonian University Medical College.

Acknowledgments

The authors acknowledge the editorial assistance of Ms. Joana Mielko, School of Medicine in English, Jagiellonian University Medical College.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2023.1161637/full#supplementary-material

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. Brief report: a novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* (2020) 382:727. doi: 10.1056/NEJMOA2001017

2. Barański K, Brożek G, Kowalska M, Kaleta-pilarska A, Zejda JE. Impact of COVID-19 pandemic on total mortality in Poland. *Int J Environ Res Public Health* (2021) 18:4388. doi: 10.3390/IJERPH18084388

3. Sagan A, Bryndova L, Kowalska-Bobko I, Smatana M, Spranger A, Szerencses V, et al. A reversal of fortune: comparison of health system responses to COVID-19 in the visegrad group during the early phases of the pandemic. *Health Policy* (2021) 126 (5):446–45. doi: 10.1016/J.HEALTHPOL.2021.10.009

4. Król Z, Szymanski P, Bochnia A, Abramowicz E, Płachta A, Rzeplinski R, et al. Transformation of a large multi-speciality hospital into a dedicated COVID-19 centre during the coronavirus pandemic. *Ann Agric Environ Med* (2020) 27:201–6. doi: 10.26444/AAEM/123801

5. Szarpak L, Pruc M, Nadolny K, Smereka J, Ladny JR. Role of a field hospital in COVID-19 pandemic. *Disaster Emergency Med J* (2020) 5:221–3. doi: 10.5603/DEMJ.A2020.0046

6. *Raport zakażeń koronawirusem (SARS-CoV-2) - koronawirus: informacje i zalecenia - portal gov.pl n.d* Available at: https://www.gov.pl/web/koronawirus/wykaz-zarazen-koronawirusem-sars-cov-2 (Accessed January 31, 2022).

7. Carrillo-Vega MF, Salinas-Escudero G, García-Peña C, Gutiérrez-Robledo LM, Parra-Rodríguez L. Early estimation of the risk factors for hospitalization and mortality by COVID-19 in Mexico. *PloS One* (2020) 15:e0238905. doi: 10.1371/JOURNAL.PONE.0238905

8. Vahey GM, McDonald E, Marshall K, Martin SW, Chun H, Herlihy R, et al. Risk factors for hospitalization among persons with COVID-19–Colorado. *PloS One* (2021) 16:e0256917. doi: 10.1371/JOURNAL.PONE.0256917

9. Ko JY, Danielson ML, Town M, Derado G, Greenlund KJ, Kirley PD, et al. Risk factors for coronavirus disease 2019 (COVID-19)–associated hospitalization: COVID-19–associated hospitalization surveillance network and behavioral risk factor surveillance system. *Clin Infect Dis* (2021) 72:e695–703. doi: 10.1093/CID/CIAA1419

10. Bonora E, Fedeli U, Schievano E, Trombetta M, Saia M, Scroccaro G, et al. SARS-CoV-2 and COVID-19 in diabetes mellitus. population-based study on ascertained infections, hospital admissions and mortality in an Italian region with ~5 million inhabitants and ~250,000 diabetic people. *Nutrition Metabolism Cardiovasc Dis* (2021) 31:2612. doi: 10.1016/J.NUMECD.2021.06.010

11. Yang J, Zheng Y, Gou X, Pu K, Chen Z, Guo Q, et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *Int J Infect Dis* (2020) 94:91. doi: 10.1016/J.IJID.2020.03.017

12. Ito H, Omoto T, Abe M, Matsumoto S, Shinozaki M, Nishio S, et al. Relationships between the duration of illness and the current status of diabetes in elderly patients with type 2 diabetes mellitus. *Geriatr Gerontol Int* (2017) 17:24–30. doi: 10.1111/ggi.12654

13. Nandy K, Salunke A, Pathak SK, Pandey A, Doctor C, Puj K, et al. Coronavirus disease (COVID-19): a systematic review and meta-analysis to evaluate the impact of various comorbidities on serious events. *Diabetes Metab Syndrome: Clin Res Rev* (2020) 14:1017–25. doi: 10.1016/j.dsx.2020.06.064

14. Corona G, Pizzocaro A, Vena W, Rastrelli G, Semeraro F, Isidori AM, et al. Diabetes is most important cause for mortality in COVID-19 hospitalized patients: systematic review and meta-analysis. *Rev Endocrine Metab Disord* (2021) 22(2):275–96. doi: 10.1007/S11154-021-09630-8

15. Barron E, Bakhai C, Kar P, Weaver A, Bradley D, Ismail H, et al. Associations of type 1 and type 2 diabetes with COVID-19-related mortality in England: a whole-population study. *Lancet Diabetes Endocrinol* (2020) 8:813–22. doi: 10.1016/S2213-8587(20)30272-2

16. Rawshani A, Kjölhede EA, Rawshani A, Sattar N, Eeg-Olofsson K, Adiels M, et al. Severe COVID-19 in people with type 1 and type 2 diabetes in Sweden: a nationwide retrospective cohort study. *Lancet Regional Health – Europe* (2021) 4:100105. doi: 10.1016/J.LANEPE.2021.100105

17. *Narodowy fundusz zdrowia (NFZ) – finansujemy zdrowie polaków n.d* Available at: https://www.nfz.gov.pl/ (Accessed March 12, 2023).

18. Ho DE, Imai K, King G, Stuart EA, Abadie A, Beck N, et al. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Anal* (2007) 15:199–236. doi: 10.1093/PAN/MPL013

19. Penny WD, Mattout J, Trujillo-Barreto N. Bayesian Model selection and averaging. In: *Statistical parametric mapping: the analysis of functional brain images* (2007) 454–67. doi: 10.1016/B978-012372560-8/50035-8

20. R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing (2021). Available at: https://www.r-project.org/.

21. *PS IMAGO PRO - predictive solutions n.d* Available at: https://en.predictivesolutions.pl/ps-imago-pro,en (Accessed January 2, 2023).

22. *Centrala narodowego funduszu zdrowia. departament analiz i strategii. NFZ o zdrowiu, cukrzyca 2019* Available at: https://ezdrowie.gov.pl/pobierz/nfz_o_zdrowiu_cukrzyca.

23. van Gerwen M, Alsen M, Little C, Barlow J, Genden E, Naymagon L, et al. Risk factors and outcomes of COVID-19 in new York city; a retrospective cohort study. *J Med Virol* (2021) 93:907–15. doi: 10.1002/JMV.26337

24. Rossi PG, Marino M, Formisano D, Venturelli F, Vicentini M, Grilli R. Characteristics and outcomes of a cohort of COVID-19 patients in the province of reggio Emilia, Italy. *PloS One* (2020) 15:e0238281. doi: 10.1371/JOURNAL.PONE.0238281

25. Bociąga-Jasik M, Wojciechowska W, Terlecki M, Wizner B, Rajzer M, Garlicki A, et al. Comparison between COVID-19 outcomes in the first three waves of pandemic: a reference hospital report. *Pol Arch Intern Med* (2022) 132:16286. doi: 10.20452/PAMW.16286

26. Kania M, Mazur K, Terlecki M, Matejko B, Hohendorff J, Chaykivska Z, et al. Characteristics, mortality, and clinical outcomes of hospitalized patients with COVID-19 and diabetes: a reference single-center cohort study from Poland. *Int J Endocrinol* (2023) 2023:1–11. doi: 10.1155/2023/8700302

27. Joshi N, Caputo GM, Weitekamp MR, Karchmer AW. Infections in patients with diabetes mellitus. *N Engl J Med* (1999) 341:1906–12. doi: 10.1056/NEJM199912163412507

28. Pozzilli P, Leslie RDG. Infections and diabetes: mechanisms and prospects for prevention. *Diabetic Med* (1994) 11:935–41. doi: 10.1111/J.1464-5491.1994.TB00250.X

29. Muniyappa R, Gubbi S. COVID-19 pandemic, coronaviruses, and diabetes mellitus. *Am J Physiol Endocrinol Metab* (2020) 318:E736–41. doi: 10.1152/AJPENDO.00124.2020/ASSET/IMAGES/LARGE/ZH10052083420002.JPEG

30. Al-Ozairi E, Brown R, Hamdan Y, Alabdullah L, Voase N, Kandari J, et al. Risk of mortality among inpatients with COVID-19 and type 2 diabetes: national data from Kuwait. *Endocrinol Diabetes Metab* (2021) 4(4):e00287. doi: 10.1002/EDM2.287

31. Sonmez A, Demirci I, Haymana C, Tasci I, Dagdelen S, Salman S, et al. Clinical characteristics and outcomes of COVID-19 in patients with type 2 diabetes in Turkey: a nationwide study (TurCoviDia). *J Diabetes* (2021) 13:585–95. doi: 10.1111/1753-0407.13171

32. Boye KS, Tokar Erdemir E, Zimmerman N, Reddy A, Benneyworth BD, Dabora MC, et al. Risk factors associated with COVID-19 hospitalization and mortality: a Large claims-based analysis among people with type 2 diabetes mellitus in the united states. *Diabetes Ther* (2021) 12:2223–39. doi: 10.1007/s13300-021-01110-1

33. Dessie ZG, Zewotir T. Mortality-related risk factors of COVID-19: a systematic review and meta-analysis of 42 studies and 423,117 patients. *BMC Infect Dis* (2021) 21:1–28. doi: 10.1186/S12879-021-06536-3/FIGURES/10

34. Woolcott OO, Castilla-Bancayán JP. The effect of age on the association between diabetes and mortality in adult patients with COVID-19 in Mexico. *Sci Rep* (2021) 11 (1):1–10. doi: 10.1038/s41598-021-88014-z

35. Bruhn-Olszewska B, Davies H, Sarkisyan D, Juhas U, Rychlicka-Buniowska E, Wójcik M, et al. Loss of y in leukocytes as a risk factor for critical COVID-19 in men. *Genome Med* (2022) 14:1–14. doi: 10.1186/S13073-022-01144-5/FIGURES/5

36. Danielsen AC, Lee KM, Boulicault M, Rushovich T, Gompers A, Tarrant A, et al. Sex disparities in COVID-19 outcomes in the united states: quantifying and contextualizing variation. *Soc Sci Med* (2022) 294:114716. doi: 10.1016/J.SOCSCIMED.2022.114716

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 16 May 2023
DOI 10.3389/fendo.2023.1167087

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Sawai Singh Rathore,
Dr. Sampurnanand Medical College, India
Evgeny Shlyakhto,
Almazov National Medical Research Centre, Russia

*CORRESPONDENCE
Mohammad Abumayyaleh
mohammad.abumayyaleh@medma.uni-heidelberg.de

RECEIVED 15 February 2023
ACCEPTED 27 April 2023
PUBLISHED 16 May 2023

CITATION
Abumayyaleh M, Núñez Gil IJ, Viana-LLamas MC, Raposeiras Roubin S, Romero R, Alfonso-Rodríguez E, Uribarri A, Feltes G, Becerra-Muñoz VM, Santoro F, Pepe M, Castro Mejía AF, Signes-Costa J, Gonzalez A, Marín F, López-País J, Manzone E, Vazquez Cancela O, Paeres CE, Masjuan AL, Velicki L, Weiß C, Chipayo D, Fernandez-Ortiz A, El-Battrawy I, Akin I and HOPE COVID-19 investigators (2023) Post-COVID-19 syndrome and diabetes mellitus: a propensity-matched analysis of the International HOPE-II COVID-19 Registry. *Front. Endocrinol.* 14:1167087. doi: 10.3389/fendo.2023.1167087

Post-COVID-19 syndrome and diabetes mellitus: a propensity-matched analysis of the International HOPE-II COVID-19 Registry

Mohammad Abumayyaleh[1,2]*, Iván J. Núñez Gil[3], María C. Viana-LLamas[4], Sergio Raposeiras Roubin[5], Rodolfo Romero[6], Emilio Alfonso-Rodríguez[7], Aitor Uribarri[8,9], Gisela Feltes[10], Víctor Manuel Becerra-Muñoz[11], Francesco Santoro[12], Martino Pepe[13], Alex Fernando Castro Mejía[14], Jaime Signes-Costa[15], Adelina Gonzalez[16], Francisco Marín[17], Javier López-País[18], Edoardo Manzone[19], Olalla Vazquez Cancela[20], Carolina Espejo Paeres[21], Alvaro López Masjuan[22], Lazar Velicki[23,24], Christel Weiß[25], David Chipayo[3], Antonio Fernandez-Ortiz[3], Ibrahim El-Battrawy[1,2,26], Ibrahim Akin[1,2] and HOPE COVID-19 investigators

[1]Department of Cardiology, Angiology, Haemostaseology and Medical Intensive Care, University Medical Center Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany, [2]European Center for AngioScience (ECAS) and German Center for Cardiovascular Research (DZHK) partner site Heidelberg/Mannheim, Mannheim, Germany, [3]Hospital Clínico San Carlos, Universidad Complutense de Madrid, Instituto de Investigación, Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain, [4]Hospital Universitario Guadalajara, Guadalajara, Spain, [5]University Hospital Álvaro Cunqueiro, Vigo, Spain, [6]Hospital Universitario Getafe, Getafe, Universidad Europea, Madrid, Spain, [7]Hospital University of Bellvitge, Barcelona, Spain, [8]Cardiology Department, Vall d'Hebron University Hospital and Research Institute, Universitat Autonoma de Barcelona, Barcelona, Spain, [9]Centro de Investigacion Biomedica en Red para Enfermedades Cardiovasculares (CIBERCV), Madrid, Spain, [10]Hospital Nuestra Señora de América, Madrid, Spain, [11]Hospital Clinico Universitario Virgen de la Victoria, Malaga, Spain, [12]Department of Medical and Surgical Sciences, University of Foggia, Foggia, Italy, [13]Azienda Ospedaliero-Universitaria Consorziale Policlinico di Bari, Bari, Italy, [14]Hospital General del norte de Guayaquil IESS Los Ceibos, Guayaquil, Ecuador, [15]Hospital Clínico de Valencia, INCLIVA, Valencia, Spain, [16]Hospital Universitario Infanta Sofia, Madrid, Spain, [17]Hospital Clínico Universitario Virgen de la Arrixaca, Murcia, Spain, [18]Complejo Hospitalario Universitario de Ourense, Ourense, Spain, [19]Hospital del Sureste, Madrid, Spain, [20]Complejo Hospitalario Universitario de Santiago de Compostela, Santiago, Spain, [21]Hospital Universitario Príncipe de Asturias, Madrid, Spain, [22]Hospital Universitario Juan Ramón Jimenez, Huelva, Spain, [23]Faculty of Medicine, University of Novi Sad, Novi Sad, Serbia, [24]Institute of Cardiovascular Diseases Vojvodina, Sremska Kamenica, Serbia, [25]Department for Statistical Analysis, University Heidelberg, Mannheim, Germany, [26]Department of Cardiology and Angiology, Bergmannsheil University Hospitals, Ruhr University of Bochum, Bochum, Germany

Background: Diabetes mellitus (DM) is one of the most frequent comorbidities in patients suffering from severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) with a higher rate of severe course of coronavirus disease (COVID-19). However, data about post-COVID-19 syndrome (PCS) in patients with DM are limited.

10.3389/fendo.2023.1167087

Methods: This multicenter, propensity score-matched study compared long-term follow-up data about cardiovascular, neuropsychiatric, respiratory, gastrointestinal, and other symptoms in 8,719 patients with DM to those without DM. The 1:1 propensity score matching (PSM) according to age and sex resulted in 1,548 matched pairs.

Results: Diabetics and nondiabetics had a mean age of 72.6 ± 12.7 years old. At follow-up, cardiovascular symptoms such as dyspnea and increased resting heart rate occurred less in patients with DM (13.2% vs. 16.4%; p = 0.01) than those without DM (2.8% vs. 5.6%; p = 0.05), respectively. The incidence of newly diagnosed arterial hypertension was slightly lower in DM patients as compared to non-DM patients (0.5% vs. 1.6%; p = 0.18). Abnormal spirometry was observed more in patients with DM than those without DM (18.8% vs. 13; p = 0.24). Paranoia was diagnosed more frequently in patients with DM than in non-DM patients at follow-up time (4% vs. 1.2%; p = 0.009). The incidence of newly diagnosed renal insufficiency was higher in patients suffering from DM as compared to patients without DM (4.8% vs. 2.6%; p = 0.09). The rate of readmission was comparable in patients with and without DM (19.7% vs. 18.3%; p = 0.61). The reinfection rate with COVID-19 was comparable in both groups (2.9% in diabetics vs. 2.3% in nondiabetics; p = 0.55). Long-term mortality was higher in DM patients than in non-DM patients (33.9% vs. 29.1%; p = 0.005).

Conclusions: The mortality rate was higher in patients with DM type II as compared to those without DM. Readmission and reinfection rates with COVID-19 were comparable in both groups. The incidence of cardiovascular symptoms was higher in patients without DM.

KEYWORDS

diabetes mellitus, post-COVID-19 syndrome, SARS-CoV-2, respiratory complications, reinfection, vaccination rate, long-term mortality

Introduction

Coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and is associated with significant morbidity and mortality (1).

Among other related diseases such as arterial hypertension and obesity, diabetes mellitus (DM) is identified as a risk factor for the severe course of COVID-19, developing sepsis, and mortality (2–4).

In patients suffering from COVID-19, SARS-CoV-2 binds the angiotensin-converting enzyme 2 (ACE2) receptor and uses it as a potential target for viral interventions (5). In diabetic mice, the expression of ACE2 is increased as compared to mice without DM. In addition, patients who suffered from insufficient glycemic control showed worse outcomes, such as more complications and higher mortality rates (6). New-onset DM and metabolic complications in patients suffering from manifested DM with high doses of insulin have been revealed in COVID-19 (7, 8). Furthermore, uncontrolled glycemic levels in DM patients cause organ injury and may be exacerbated in patients suffering from COVID-19 (9).

The international Health Outcome Predictive Evaluation for COVID-19 (HOPE COVID-19) Registry was initiated to investigate comorbidity and mortality of COVID-19 (10). In the Health Outcome Predictive Evaluation for COVID-19 II (HOPE-II COVID-19) Registry, we investigated readmission, reinfection, vaccination rate, cardiovascular, neuropsychiatric, respiratory, gastrointestinal, and other symptoms in hospitalized patients suffering from COVID-19 and concomitant DM type II. Complications related to COVID-19 and long-term mortality were also systematically analyzed.

Material and methods

Study design and participants

HOPE-II COVID-19 (NCT04334291) is an international project at 55 international centers. It is designed as a retrospective and prospective cohort registry to investigate post-COVID-19 syndrome without any conflict of interest. We included hospitalized patients with a confirmed diagnosis of COVID-19. There are no exclusion criteria, except for the patient's explicit refusal to participate. Initially, data on 8,828 hospitalized patients suffering from COVID-19 were

collected until 30th September 2021. In this study, we excluded 56 patients due to age <18 and 53 patients with DM type I. Data from 8,719 consecutive patients with COVID-19 regarding their concomitant DM type II status were analyzed.

Ethics approval

This study was executed in compliance with the Declaration of Helsinki regarding human subjects, and the study was approved by the center ethics committee of Hospital Clinico San Carlos (Internal Code: 21/128-E) and, when needed, in all involved centers.

DM type II

DM type II was known and diagnosed by medical physicians. Data were collected from the patient's medical records.

Post-COVID-19 syndrome

Patients suffering from post-COVID-19 syndrome describe new-onset symptoms following initial recovery from an acutely confirmed COVID-19 or ongoing from the initial illness. This condition occurs 3 months from the onset of COVID-19 with symptoms that last for at least 2 months and cannot be explained by an alternative diagnosis. Symptoms may also fluctuate or relapse over time (11).

Outcomes and follow-up

We described long-term mortality as a primary endpoint. Readmission, reinfection rate, respiratory complications, cardiovascular, neuropsychiatric, respiratory, gastrointestinal, and other symptoms as secondary endpoints were also evaluated. Follow-up for the overall population for mortality was 20 months (mean post-COVID-19; 2.6 ± 4.6).

Statistical analysis

Descriptive and comparative analyses were presented. Continuous variables were shown as mean ± standard deviation if the distribution was normal or median (min–max) if not. Categorical variables were presented as frequency rates and percentages. The Chi-square test was used for categorical variables for group comparisons. Quantitative variables were performed using the Mann–Whitney *U* test for nonparametric variables and the Student's *t*-test for parametric variables, as verified by the Kolmogorov–Smirnov test. We applied a propensity score (PS)-based matching method to control for confounding baseline variables due to the nonrandomized nature of the study and the different participating centers. In a multivariable logistic regression test, hazard ratio (HR) with 95% confidence intervals (95% CI) was calculated for the determination of risk factors for the endpoint. Predictors of mortality were identified by univariate analysis. Predictors with $p < 0.05$ were analyzed by logistic multivariable regression. The multivariable regression test was used to investigate predictors of mortality, adjusting for all significant variables: age; male as sex; obesity; comorbidities such as arterial hypertension, dyslipidemia, DM type II, renal insufficiency, heart disease, cerebrovascular disease, liver disease, and cancer disease; immunosuppression; home oxygen therapy; premedication; symptomatic; clinical parameters such as peripheral oxygen saturation (SpO_2) <92% and reduced blood pressure (systolic blood pressure <90 mmHg or diastolic blood pressure <60 mmHg); and laboratory parameters. *p*-value of <0.05 was recognized as statistically significant. Statistical analysis was performed with IBM SPSS Statistics version 27.

Results

Baseline characteristics and in-hospital complications

Data from 8,719 consecutive hospitalized patients ($n = 1,578$ with DM; $n = 7,141$ with non-DM) with confirmed COVID-19 were collected. The 1:1 propensity score matching (PSM) according to age and sex resulted in 1,548 matched pairs. The mean age of matched pairs was 72.6 ± 12.7 years old. Even more, the male sex was 63.5% in both groups. Diabetics suffered from more chronic conditions such as arterial hypertension (77.5% vs. 58.5%; $p < 0.0001$), renal insufficiency (13.6% vs. 8.1%; $p < 0.0001$), and liver disease (5.7% vs. 3.4%; $p = 0.002$). In-hospital complications were observed more in diabetics as compared to nondiabetics, for example, respiratory insufficiency (62.1% vs. 56.3%; $p = 0.001$), acute kidney injury (26.6% vs. 19.8%; $p < 0.0001$), and sepsis (15.4% vs. 12.8%, $p = 0.04$). Other baseline characteristics, immunosuppression, home oxygen therapy, premedication, symptomatic, clinical, and laboratory parameters, in-hospital complications, and intervention procedures during hospitalization are presented in Table 1.

Clinical outcomes at long-term follow-up

Mean follow-up (2.6 ± 4.6 months) data were available for 412 diabetics and 443 nondiabetics. The readmission rate due to any cause was similar in diabetics and nondiabetics, respectively (19.7% vs. 18.3%; $p = 0.61$). The reinfection rate with COVID-19 was also comparable in patients with DM than those without DM (2.9% vs. 2.3%; $p = 0.55$). Additionally, diabetics were vaccinated more than nondiabetics at follow-up with the same time to vaccination (11.9 ± 3.1 months in diabetics vs. 12.2 ± 2.9 months in nondiabetics) (57.3% vs. 51.7%; $p = 0.10$). At follow-up, cardiovascular symptoms such as dyspnea and an increase in resting heart rate after discharge occurred less frequently in patients suffering from DM (13.2% vs. 16.4%; $p = 0.01$) than those without DM (2.8% vs. 5.6%; $p = 0.05$), respectively. In addition, the mortality rate at the 20-month follow-up was significantly higher in DM than in non-DM patients (33.9% vs. 29.1%; $p = 0.005$). Cardiovascular, neuropsychiatric, respiratory, gastrointestinal, and other symptoms are presented in Table 2.

10.3389/fendo.2023.1167087

TABLE 1 Patients with diabetes mellitus type II as compared to patients without DM II, baseline characteristics, laboratory and radiographic findings, complications, and clinical outcomes.

Characteristic	Diabetics (*N* = 1,548)	Nondiabetics (*N* = 1,548)	*p*-value*
Age (mean ± SD (years))	72.6 ± 12.7	72.6 ± 12.7	1.00
Male as sex (no. (%))	983 (63.5)	983 (63.5)	1.00
Chronic conditions (no. (%))			
Arterial hypertension	1,200 (77.5)	906 (58.5)	**<0.0001**
Dyslipidemia	927 (59.9)	551 (35.6)	**<0.0001**
Obesity	486 (31.4)	221 (14.3)	**<0.0001**
Current smoking	86 (6.2)	83 (5.9)	0.73
Renal insufficiency[a]	211 (13.6)	126 (8.1)	**<0.0001**
Lung disease	362 (30.1)	329 (28)	0.26
Cardiac disease	538 (34.8)	414 (26.7)	**<0.0001**
Cerebrovascular disease	192 (12.4)	164 (10.6)	0.12
Connective tissue disease	48 (3.1)	43 (2.8)	0.60
Liver disease	88 (5.7)	52 (3.4)	**0.002**
Cancer disease	291 (18.8)	237 (15.3)	**0.009**
Immunosuppression[b]	134 (8.7)	112 (7.2)	0.14
Home oxygen therapy	74 (4.8)	71 (4.6)	0.80
Premedication (no. (%))			
ASA	453 (29.3)	263 (17)	**<0.0001**
Antiplatelet drug	119 (7.7)	75 (4.8)	**0.001**
Oral anticoagulation	251 (16.2)	220 (14.2)	0.12
Beta-blockers	420 (27.1)	287 (18.5)	**<0.0001**
ACEI/ARB	863 (55.8)	638 (41.2)	**<0.0001**
Symptomatic (no. (%))			
Asymptomatic	81 (5.2)	105 (6.8)	0.07
Dyspnea	961 (62.9)	911 (59.5)	**0.05**
Tachypnea > 22 breaths/min	485 (31.4)	455 (29.4)	0.24
Hemoptysis	26 (1.7)	32 (2.1)	0.42
Fatigue	727 (47)	718 (46.4)	0.75
Anosmia/hyposmia	55 (3.6)	67 (4.3)	0.27
Dysgeusia	66 (4.3)	73 (4.7)	0.54
Sore throat	117 (7.6)	159 (10.3)	**0.01**
Fever	1,102 (71.3)	1,150 (74.4)	0.06
Cough	950 (61.5)	944 (61.1)	0.83
Vomiting	107 (6.9)	95 (6.1)	0.38
Diarrhea	268 (17.3)	234 (15.1)	0.10
Erythromelalgia	369 (23.9)	443 (28.7)	**0.003**
Clinical parameters (no. (%))			
Peripheral oxygen saturation < 92%	690 (44.6)	604 (39.1)	**0.002**

(Continued)

10.3389/fendo.2023.1167087

TABLE 1 Continued

Characteristic	Diabetics (*N* = 1,548)	Nondiabetics (*N* = 1,548)	*p*-value*
Abnormal blood pressure[c]	139 (9.9)	116 (8.3)	0.13
GCS < 15	149 (11.8)	144 (11)	0.62
Laboratory parameters (no. (%) or median (min–max))			
Elevated D-dimer	953 (61.6)	903 (58.4)	0.07
Elevated procalcitonin	302 (19.5)	231 (14.9)	**0.0007**
Elevated CRP	1,382 (89.4)	1,343 (86.9)	**0.03**
Elevated TnI	206 (13.3)	165 (10.7)	**0.02**
Elevated transaminases[d]	505 (32.7)	579 (37.5)	**0.006**
Elevated ferritin	494 (32)	515 (33.3)	0.42
Elevated triglyceride	172 (11.1)	129 (8.3)	**0.009**
Elevated LDH	1,018 (65.9)	1,033 (66.8)	0.59
Creatinine (mg/dl)	1.02 (0.38–11.3)	0.96 (0.12–33.9)	**0.0005**
Leukocytes (10E9/L)	7,000 (550–90,004)	6,440 (440–88,400)	**<0.0001**
Lymphocytes (10E9/L)	960 (12–41,100)	930 (244–77,100)	0.25
Hemoglobin (g/dl)	13 (1–19.3)	14 (4–18)	**<0.0001**
Thrombocytes (10E9/L)	201,000 (13,000–716,000)	190,000 (10,000–980,000)	**<0.0001**
Natrium level (mmol/L)	137 (115–179)	138 (117–180)	**<0.0001**
In-hospital complication			
Respiratory insufficiency	958 (62.1)	867 (56.3)	**0.001**
Heart failure	191 (12.4)	128 (8.3)	**0.0002**
Acute kidney injury	411 (26.6)	305 (19.8)	**<0.0001**
Upper respiratory tract infection	257 (16.7)	240 (15.6)	0.41
Pneumonia	1,344 (89.4)	1,336 (88.4)	0.41
SIRS	389 (25.2)	355 (23)	0.16
Sepsis	238 (15.4)	197 (12.8)	**0.04**
Any relevant bleeding[e]	65 (4.2)	46 (3)	0.07
Embolic event	49 (3.2)	47 (3.1)	0.84
Oxygen therapy			
O_2 at the admission	1,238 (80.2)	1,157 (75.1)	**0.0007**
High-flow nasal cannula	347 (22.5)	336 (21.8)	0.65
Noninvasive mechanical ventilation	250 (16.2)	237 (15.4)	0.54
Invasive mechanical ventilation	163 (10.6)	123 (8)	**0.01**
Another medication or intervention procedures during the admission			
Prone position	196 (12.7)	169 (11)	0.14
ECMO	119 (7.7)	82 (5.3)	**0.007**
Use of glucocorticoids	546 (35.4)	526 (34.1)	0.47
Use of hydroxychloroquine	1,173 (76)	1,180 (76.6)	0.69
Use of antiviral drugs[f]	714 (46.2)	812 (53)	**0.0003**
Use of interferon	180 (11.7)	233 (15.1)	**0.005**

(Continued)

TABLE 1 Continued

Characteristic	Diabetics (*N* = 1,548)	Nondiabetics (*N* = 1,548)	*p*-value*
Use of tocilizumab	131 (8.5)	128 (8.3)	0.85
Use of antibiotics	1,181 (76.5)	1,113 (72.2)	**0.007**
ACEI/ARB[g]	476 (30.9)	354 (23)	**<0.0001**
Anticoagulation	856 (81.7)	791 (75.8)	**0.001**
Discharge			
ACEI/ARB	82 (30.8)	71 (24.7)	0.11
Antiplatelet drug	226 (14.7)	147 (9.6)	**<0.0001**
Anticoagulation	413 (26.8)	365 (23.7)	**0.05**

ASA, acetylsalicylic acid; ACEI/ARB, angiotensin-converting enzyme inhibitor/angiotensin-receptor blocker; CRP, C-reactive protein; GCS, Glasgow coma scale; ECMO, extracorporeal membrane oxygenation; SIRS, systemic inflammatory response syndrome; TnI, high-sensitivity troponin I (cardiac injury; troponin > 99th percentile upper reference limit).
[a]CrCL < 30.
[b]Immunosuppressive therapy for psoriatic arthritis, lung transplantation, kidney transplantation, or systemic lupus erythematosus; oncological diseases such as mamma-ca, prostate-ca, myelodysplastic syndrome, or gammopathy; glucocorticoid therapy caused by COPD; dialysis; HIV; or hepatitis.
[c]Systolic blood pressure < 90 mmHg or diastolic blood pressure < 60 mmHg.
[d]ALAT and ASAT.
[e]Rectorrhagia, hematuria, epistaxis, and popliteal aneurysm bleeding with relevant decreased hemoglobin > 2 mg/l.
[f]Lopinavir or/and ritonavir.
[g]Premedication with ACEI/ARB is not stopped.
*Statistical significance level is set at 0.05 and value of statistical significance is emphasized in bold.

TABLE 2 Follow-up in patients suffering from DM type II as compared to those without DM.

	Diabetics (*N* = 1,548)	Nondiabetics (*N* = 1,548)	*p*-value*
Follow-up (mean ± SD)			
Follow-up time (months (PCS))	2.6 ± 4.6	2.8 ± 4.9	0.77
Duration to recovery (months)	2.2 ± 4.6	2.4 ± 4.9	0.51
Duration to readmission (months)	2.5 ± 4.5	2.6 ± 4.6	0.95
Number of patients (*n*)	412	443	–
Readmission	81 (19.7)	81 (18.3)	0.61
Vaccination	236 (57.3)	229 (51.7)	0.10
Time to vaccination (months)	11.9 ± 3.1	12.2 ± 2.9	0.74
Reinfection with COVID-19	12 (2.9)	10 (2.3)	0.55
Clinical event after discharge	171 (43.1)	181 (42)	0.75
Cardiovascular symptoms			
Fatigue	114 (28.7)	125 (29)	0.93
Dyspnea	204 (13.2)	254 (16.4)	**0.01**
Dizziness	34 (8.6)	35 (8.1)	0.82
Chest pain	28 (7.1)	28 (6.5)	0.75
Acute coronary syndrome	3 (0.8)	4 (0.9)	1.00
Palpitation	24 (6.1)	37 (8.6)	0.16
Increase in resting heart rate	11 (2.8)	24 (5.6)	**0.05**
Syncope	2 (0.5)	8 (1.9)	0.11
Arrhythmias	27 (6.8)	22 (5.1)	0.30
Atrial fibrillation	21 (5.3)	26 (6)	0.65

(Continued)

10.3389/fendo.2023.1167087

TABLE 2 Continued

	Diabetics (*N* = 1,548)	Nondiabetics (*N* = 1,548)	*p*-value*
Perimyocarditis	1 (0.3)	2 (0.5)	1.00
Limb edema	13 (3.3)	18 (4.2)	0.50
New hypertension	2 (0.5)	7 (1.6)	0.18
New left ventricular dysfunction	5 (1.3)	7 (1.6)	0.66
Relevant bleeding	5 (1.3)	5 (1.2)	0.90
Neuropsychiatric symptoms			
Headache	11 (2.8)	21 (4.9)	0.12
Migraine	5 (1.3)	11 (2.6)	0.18
Ageusia	17 (4.3)	19 (4.4)	0.93
Anosmia	12 (3)	18 (4.2)	0.38
Attention disorder	16 (4)	25 (5.8)	0.24
Memory loss	31 (7.8)	34 (7.9)	0.97
Cognitive disorder	18 (4.5)	20 (4.6)	0.94
Anxiety	34 (8.6)	54 (12.5)	0.06
Depression	26 (6.6)	35 (8.1)	0.39
Tinnitus or hearing loss	9 (2.3)	14 (3.3)	0.39
Sleeping disorder	27 (6.8)	36 (8.4)	0.40
Mood disorder	22 (5.5)	31 (7.2)	0.33
Paranoia	16 (4)	5 (1.2)	**0.009**
Respiratory symptoms			
Cough	33 (8.3)	42 (9.7)	0.47
Reduce pulmonary diffusing capacity	28 (7.1)	44 (10.2)	0.11
Polypnea	15 (3.8)	19 (4.4)	0.65
Sleep apnea	13 (3.3)	9 (2.1)	0.29
Gastrointestinal symptoms			
Tongue involvement	1 (0.3)	7 (1.6)	0.07
Digestive disorder	20 (5)	17 (3.9)	0.45
Nausea/vomiting	10 (2.5)	8 (1.9)	0.51
Other symptoms			
Intermittent fever	8 (2)	10 (2.3)	0.76
Chills	6 (1.5)	8 (1.9)	0.70
Hair loss	20 (5)	18 (4.2)	0.55
Joint pain	19 (4.8)	25 (5.8)	0.52
Myalgia	26 (6.6)	32 (7.4)	0.62
Sweat	5 (1.3)	4 (0.9)	0.74
Weight loss	24 (6.1)	23 (5.3)	0.66
Cutaneous involvement	6 (1.5)	13 (3)	0.15
New diabetes	–	4 (0.9)	–
New renal insufficiency	19 (4.8)	11 (2.6)	0.09

(Continued)

TABLE 2 Continued

	Diabetics (*N* = 1,548)	Nondiabetics (*N* = 1,548)	*p*-value*
Pain	12 (3)	8 (1.9)	0.28
Red eyes	4 (1)	6 (1.4)	0.76
Flushing	4 (1)	2 (0.5)	0.43
Incident neoplasia	2 (0.5)	6 (1.4)	0.29
Management after discharge			
Home oxygen therapy	43 (10.8)	37 (8.6)	0.27
ASA	99 (24.9)	58 (13.5)	**<0.0001**
Antiplatelet drug	34 (8.6)	23 (5.3)	0.07
Anticoagulation	69 (17.4)	54 (12.5)	**0.05**
ACEI/ARB	140 (35.3)	113 (26.2)	**0.005**
Beta-blockers	75 (18.9)	69 (16)	0.27
Beta agonist inhalation therapy	34 (8.6)	46 (10.7)	0.31
Vitamin supplementation	72 (18.1)	80 (18.6)	0.88
Antidepressant	47 (11.8)	64 (14.9)	0.20
Statin	151 (38)	103 (23.9)	**<0.0001**
Diagnostic test after discharge			
Elevated di-dimer	137 (34.6)	151 (35.2)	0.86
Elevated CRP	167 (42.2)	183 (42.7)	0.89
Elevated procalcitonin	45 (11.4)	34 (7.9)	0.09
Elevated TnI	18 (4.6)	16 (3.7)	0.56
Elevated NT-proBNP	23 (5.8)	29 (6.8)	0.57
Elevated transaminases[a]	92 (23.2)	100 (23.3)	0.98
Abnormal spirometry	21 (18.8)	17 (13)	0.24
Any chest X-ray abnormality	99 (39.4)	103 (38.9)	0.99
Any CT abnormality	37 (35.6)	48 (35.3)	0.60
In-hospital mortality	492 (31.8)	426 (27.5)	**0.009**
Long-term mortality	524 (33.9)	451 (29.1)	**0.005**

PCS, post-COVID-19 syndrome; ASA, acetylsalicylic acid; ACEI/ARB, angiotensin-converting enzyme inhibitor/angiotensin-receptor blocker; CRP, C-reactive protein; TnI, high-sensitivity troponin I cardiac injury; troponin > 99th percentile upper reference limit. [a]ALAT and ASAT.
*Statistical significance level is set at 0.05 and value of statistical significance is emphasized in bold.
Summarized, - means not available.

PSM and predictors of mortality

The mortality rate at long-term follow-up was significantly higher in patients with DM than those without, in the overall cohort and in the matched cohort, respectively ($p < 0.0001$ and $p = 0.005$). The Kaplan–Meier curve with landmark analysis is displayed in Figure 1. In the multivariable analysis for mortality, age, and male sex were determined as predictors for mortality, respectively (HR: 2.34; $p < 0.0001$) (HRK 1.23; $p = 0.008$). Other predictors are performed in Table 3. Clinical outcomes before PSM are presented in the Supplementary Appendix.

Discussion

This study presents characteristics of PCS in patients suffering from DM as compared to those without DM. The main findings of this study are as follows: (1) readmission rate for any cause was similar in diabetics than nondiabetics at follow-up; (2) reinfection rate with COVID-19 was similar in both groups; (3) symptoms such as dyspnea and an increase of resting heart rate occurred less in diabetics as compared to nondiabetics; (4) The incidence of newly diagnosed arterial hypertension was less in diabetics than nondiabetics without statistical significance; (5) respiratory

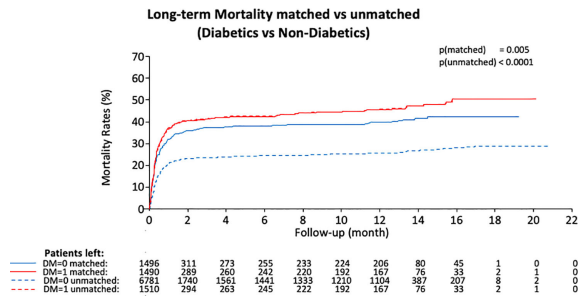


FIGURE 1
Kaplan–Meier curve for long-term mortality divided by diabetics vs. nondiabetics in the overall population and matched cohort. In both comparisons, a worse outcome in diabetics was detected.

complications were revealed in diabetics and nondiabetics; and (5) long-term mortality was higher in patients suffering from DM as compared to those without DM.

Recently, it has been reported that the progression of type II DM is associated with increased insulin resistance accompanied by chronic inflammation and endothelial and ß-cell dysfunction (12). On the other hand, the inflammatory response in infected patients with SARS-CoV-2 may worsen insulin resistance and endothelial dysfunction (13). The existence of both diseases may further enhance the inflammation and decrease interferon levels, neutrophil chemotaxis, and T lymphocyte-mediated immune response with impairment of cytokine production (14–16). That is associated with a severe course of COVID-19 in DM patients. Furthermore, ACE2 expression increases insulin resistance. This receptor and dipeptidyl peptidase 4 (DPP4), which may be a factor in the severity of COVID-19 infection, are present in several physiological processes and are modulated by hyperglycemia and pharmacological therapies that are common in DM patients (17). In addition, chronic hyperglycemia leads to chronic vascular and kidney disease. Other comorbidities, such as obesity and hypertension, are present in concurrent DM. These diabetes-related comorbidities may negatively impact outcomes in DM patients with COVID-19 (18, 19).

DM as a risk factor for post-COVID-19 syndrome

Our DM cohort had more comorbidities such as arterial hypertension, renal insufficiency, liver disease, and cardiac disease than patients without DM. Furthermore, respiratory insufficiency requiring oxygen therapy and invasive mechanical ventilation (MV) was observed more in diabetics as compared to nondiabetics. During hospitalizations, sepsis and acute kidney injury occurred more often in diabetics than nondiabetics. A prospective study showed that the persistence of symptoms was associated with the severity of the disease at the beginning and that the intensive care unit (ICU) admission was an independent risk factor for PCS (20). In addition, the need for MV was determined as a predictor for the development of PCS (21). However, it has been reported that 60% of low-risk patients for mortality with COVID-19 suffered from severe PCS (22). In patients with DM, optimizing hyperglycemia therapy improve metabolic function which may be beneficial for the long-term management of patients with PCS (23). In this study, PCS was slightly comparable despite the different comorbidities and in-hospital complications in both groups.

Cardiovascular symptoms

In our study, dyspnea and an increase in resting heart rate occurred more significantly in nondiabetics as compared to diabetics. Additionally, newly diagnosed arterial hypertension was also revealed slightly more in nondiabetics than diabetics. Regarding that, the persistence of cardiovascular symptoms was recently reported (24). In one of the studies from Wuhan, Huang et al. showed that patients infected with SARS-CoV-2 suffered from acute cardiac injury (25). Subclinical myocarditis with an increased risk of arrhythmias may play a role in PCS (26). Data about the comparison between diabetics and nondiabetics are limited.

Neuropsychiatric symptoms

This study presented neuropsychiatric symptoms generally more common in nondiabetics as compared to diabetics without statistical significance, for example, headache, sleeping disorder, and anxiety. However, paranoia was observed significantly more in diabetics than nondiabetics at a 3-month follow-up. Studies reported that headache and other neuropsychiatric symptoms occurred after 3 months in patients infected with SARS-CoV-2 (27, 28). Guedj et al. reported that more areas in the brain showed hypometabolism in patients with PCS as compared to healthy subjects (29). Controlled, randomized studies are needed to investigate the neuropsychiatric symptoms in patients with DM as compared to those without DM.

Respiratory symptoms

Renal insufficiency and cardiac disease were observed more in patients with DM than non-DM, while the rate of lung diseases was similar in matched pairs. At follow-up, our data presented a similar rate of sleep apnea in diabetics and nondiabetics. Furthermore, computer tomography (CT) and chest X-ray abnormalities were revealed in both groups, but dyspnea occurred significantly more in nondiabetics as compared to diabetics at follow-up. In one retrospective study with 77 days of follow-up, spirometry (9.3%) and chest radiology (19%) abnormalities were detected in 277 patients, of whom 51% had PCS (30). In 22 patients after COVID-19-associated acute respiratory distress syndrome (ARDS), signs of lung fibrosis were detected in 55% of patients at 3-month follow-up (31). In patients with critical COVID-19, 9.5% of patients needed home oxygen therapy after discharge at a 1-year follow-up (32). Furthermore, DM was identified as a risk factor for the requirement of oxygen therapy in patients suffering from COVID-19 (33). In our multivariable analysis, DM was not identified as a predictor for mortality.

10.3389/fendo.2023.1167087

TABLE 3 Predictors of mortality, multivariable analysis.

Variable	Univariable analysis		Multivariable analysis	
	HR	*p*-value	HR	*p*-value
Patient demographics				
Age ≥70	2.90	**<0.0001**	2.34	**<0.0001**
Male	1.19	**0.01**	1.23	**0.008**
Chronic conditions				
Dyslipidemia	1.12	0.07		
Diabetes mellitus	1.18	**0.01**		
Obesity	1.01	0.88		
Renal insufficiency	1.86	**<0.0001**	1.33	**0.003**
Cancer disease	1.46	**<0.0001**		
Immunosuppression	1.41	**0.0009**	1.40	**0.003**
Premedication				
ASA	1.38	**<0.0001**		
Oral anticoagulation	1.66	**<0.0001**		
Clinical parameters				
SpO_2 < 92%[a]	3.14	**<0.0001**	2.13	**<0.0001**
Abnormal blood pressure[b]	2.09	**<0.0001**	1.36	**0.002**
GCS < 15	2.67	**<0.0001**	1.50	**<0.0001**
Clinical presentation				
Dyspnea	1.48	**<0.0001**		
Tachypnea > 22 breaths/min	2.17	**<0.0001**	1.41	**<0.0001**
Dysgeusia	0.32	**<0.0001**	0.40	**0.001**
Sore throat	0.79	0.07		
Cough	0.77	**<0.0001**	0.84	**0.02**
Erythromelalgia	0.73	**<0.0001**		
Laboratory parameters				
Elevated procalcitonin	1.89	**<0.0001**	1.53	**<0.0001**
Elevated CRP	1.60	**<0.0001**		
Elevated LDH	1.50	**<0.0001**	1.20	**0.04**

HR, hazard ratio; ASA, acetylsalicylic acid; SpO_2, peripheral oxygen saturation; GCS, Glasgow coma scale.
[a]SpO_2 < 92% at admission.
[b]Systolic blood pressure < 90 mmHg or diastolic blood pressure < 60 mmHg.
Statistical significance level is set at 0.05 and value of statistical significance is emphasized in bold.

This study has some limitations. It has a retrospective character; not all laboratory tests were done on all patients. Furthermore, data on hemoglobin A1c (HbA1c), antihyperglycemic treatment including metformin and DPP-4 inhibitors, and statin therapy at baseline are missing. A strength of our study is the sample size of patients with COVID-19 and concomitant DM type II at 55 international centers. The results are therefore real-world evidence.

To summarize, PCS was observed in diabetics and nondiabetics. However, the mortality rate was higher in diabetics as compared to nondiabetics. DM was not determined as a risk factor for mortality at follow-up.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

This study was executed in compliance with the Declaration of Helsinki regarding in human subjects and the study was approved by the center Ethics Committee of Hospital Clinico San Carlos (Internal Code: 21/128-E) and, when needed, in all involved centers. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

MA, IG, IE-B, and IA made substantial contributions to the study's concept and design. All authors obtained ethical approval. Data were collected by MA, IG, MV-L, SR, RR, EA-R, AU, GF, VB-M, FS, MP, AM, JS-C, AG, FM, JL-P, EM, OC, CP, AM, LV, DC, AF-O, MA, and CW analyzed all the data. CW supported the descriptive statistics. IJNG and IA approved the statistical analysis. MA, IG, IE-B, and IA prepared the manuscript. All authors contributed to the article and approved the submission version.

Funding

Nonconditioned grant (Fundación Interhospitalaria para la Investigación cardiovascular, FIC, Madrid, Spain). This nonprofit institution had no role in the study design; collection, analysis, or interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

Acknowledgments

The authors thank all HOPE-II researchers.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2023.1167087/full#supplementary-material

References

1. Nunez-Gil IJ, Fernandez-Perez C, Estrada V, Becerra-Munoz VM, El-Battrawy I, Uribarri A, et al. Mortality risk assessment in spain and italy, insights of the hope covid-19 registry. *Intern Emerg Med* (2021) 16(4):957–66. doi: 10.1007/s11739-020-02543-5

2. El-Battrawy I, Nunez-Gil IJ, Abumayyaleh M, Estrada V, Manuel Becerra-Munoz V, Uribarri A, et al. Covid-19 and the impact of arterial hypertension-an analysis of the international hope covid-19 registry (italy-spain-germany). *Eur J Clin Invest* (2021) 51 (11):e13582. doi: 10.1111/eci.13582

3. Abumayyaleh M, Nunez-Gil IJ, El-Battrawy I, Estrada V, Becerra-Munoz VM, Uribarri A, et al. Sepsis of patients infected by sars-cov-2: real-world experience from the international hope-covid-19-registry and validation of hope sepsis score. *Front Med (Lausanne)* (2021) 8:728102. doi: 10.3389/fmed.2021.728102

4. Abumayyaleh M, Nunez Gil IJ, El-Battrawy I, Estrada V, Becerra-Munoz VM, Aparisi A, et al. Does there exist an obesity paradox in covid-19? insights of the international hope-covid-19-registry. *Obes Res Clin Pract* (2021) 15(3):275–80. doi: 10.1016/j.orcp.2021.02.008

5. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, et al. Sars-cov-2 cell entry depends on ace2 and tmprss2 and is blocked by a clinically proven protease inhibitor. *Cell* (2020) 181(2):271–280.e278. doi: 10.1016/j.cell.2020.02.052

6. Mader JK, Brix J, Aberer F, Vonbank A, Resl M, Pieber TR, et al. [hospital diabetes management (update 2019)]. *Wien Klin Wochenschr* (2019) 131(Suppl 1):200–11. doi: 10.1007/s00508-019-1447-z

7. Ren H, Yang Y, Wang F, Yan Y, Shi X, Dong K, et al. Association of the insulin resistance marker tyg index with the severity and mortality of covid-19. *Cardiovasc Diabetol* (2020) 19(1):58. doi: 10.1186/s12933-020-01035-2

8. Chee YJ, Ng SJH, Yeoh E. Diabetic ketoacidosis precipitated by covid-19 in a patient with newly diagnosed diabetes mellitus. *Diabetes Res Clin Pract* (2020) 164:108166. doi: 10.1016/j.diabres.2020.108166

9. Mrigpuri P, Sonal S, Spalgais S, Goel N, Menon B, Kumar R. Uncontrolled diabetes mellitus: a risk factor for post covid fibrosis. *Monaldi Arch Chest Dis* (2021) 91 (1). doi: 10.4081/monaldi.2021.1607

10. Nunez-Gil IJ, Estrada V, Fernandez-Perez C, Feltes G, Vedia O, Vergara-Uzcategui CE, et al. Health outcome predictive evaluation for covid 19 international registry (hope covid-19), rationale and design. *Contemp Clin Trials Commun* (2020) 20:100654. doi: 10.1016/j.conctc.2020.100654

11. Soriano JB, Murthy S, Marshall JC, Relan P, Diaz and WHOCCDWGoP-C-Condition JV. A clinical case definition of post-covid-19 condition by a delphi consensus. *Lancet Infect Dis* (2022) 22(4):e102–7. doi: 10.1016/S1473-3099(21) 00703-9

12. Schwarz PEH, Timpel P, Harst L, Greaves CJ, Ali MK, Lambert J, et al. Blood sugar regulation for cardiovascular health promotion and disease prevention: jacc health promotion series. *J Am Coll Cardiol* (2018) 72(15):1829–44. doi: 10.1016/j.jacc.2018.07.081

13. Bornstein SR, Rubino F, Ludwig B, Rietzsch H, Schwarz PEH, Rodionov RN, et al. Consequences of the covid-19 pandemic for patients with metabolic diseases. *Nat Metab* (2021) 3(3):289–92. doi: 10.1038/s42255-021-00358-y

14. Berbudi A, Rahmadika N, Tjahjadi and R Ruslami AI. Type 2 diabetes and its impact on the immune system. *Curr Diabetes Rev* (2020) 16(5):442–9.

15. Hodgson K, Morris J, Bridson T, Govan B, Rush and N Ketheesan C. Immunological mechanisms contributing to the double burden of diabetes and intracellular bacterial infections. *Immunology* (2015) 144(2):171–85. doi: 10.1111/imm.12394

16. Santos A, Magro DO, Evangelista-Poderoso and MJA Saad R. Diabetes, obesity, and insulin resistance in covid-19: molecular interrelationship and therapeutic implications. *Diabetol Metab Syndr* (2021) 13(1):23. doi: 10.1186/s13098-021-00639-2

17. Drucker DJ. Coronavirus infections and type 2 diabetes-shared pathways with therapeutic implications. *Endocr Rev* (2020) 41(3). doi: 10.1210/endrev/bnaa011

18. Yan Y, Yang Y, Wang F, Ren H, Zhang S, Shi X, et al. Clinical characteristics and outcomes of patients with severe covid-19 with diabetes. *BMJ Open Diabetes Res Care* (2020) 8(1). doi: 10.1136/bmjdrc-2020-001343

19. Zhu L, She ZG, Cheng X, Qin JJ, Zhang XJ, Cai J, et al. Association of blood glucose control and outcomes in patients with covid-19 and pre-existing type 2

diabetes. *Cell Metab* (2020) 31(6):1068–1077.e1063. doi: 10.1016/j.cmet.2020.04.021

20. Peghin M, Palese A, Venturini M, De Martino M, Gerussi V, Graziano E, et al. Post-covid-19 symptoms 6 months after acute infection among hospitalized and non-hospitalized patients. *Clin Microbiol Infect* (2021) 27(10):1507–13. doi: 10.1016/j.cmi.2021.05.033

21. Michelen M, Manoharan L, Elkheir N, Cheng V, Dagens A, Hastie C, et al. Characterising long covid: a living systematic review. *BMJ Glob Health* (2021) 6(9). doi: 10.1136/bmjgh-2021-005427

22. Dennis A, Wamil M, Alberts J, Oben J, Cuthbertson DJ, Wootton D, et al. Multiorgan impairment in low-risk individuals with post-covid-19 syndrome: a prospective, community-based study. *BMJ Open* (2021) 11(3):e048391. doi: 10.1136/bmjopen-2020-048391

23. Khunti K, Davies MJ, Kosiborod and MA Nauck MN. Long covid - metabolic risk factors and novel therapeutic management. *Nat Rev Endocrinol* (2021) 17(7):379–80. doi: 10.1038/s41574-021-00495-0

24. Carvalho-Schneider C, Laurent E, Lemaignen A, Beaufils E, Bourbao-Tournois C, Laribi S, et al. Follow-up of adults with noncritical covid-19 two months after symptom onset. *Clin Microbiol Infect* (2021) 27(2):258–63. doi: 10.1016/j.cmi.2020.09.052

25. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. *Lancet* (2020) 395(10223):497–506. doi: 10.1016/S0140-6736(20)30183-5

26. Maron BJ, Udelson JE, Bonow RO, Nishimura RA, Ackerman MJ, Estes NA3rd, et al. Eligibility and disqualification recommendations for competitive athletes with cardiovascular abnormalities: task force 3: hypertrophic cardiomyopathy, arrhythmogenic right ventricular cardiomyopathy and other cardiomyopathies, and myocarditis: a scientific statement from the american heart association and american college of cardiology. *Circulation* (2015) 132(22):e273–280. doi: 10.1161/CIR.0000000000000239

27. Goertz YMJ, Van Herck M, Delbressine JM, Vaes AW, Meys R, Machado FVC, et al. Persistent symptoms 3 months after a sars-cov-2 infection: the post-covid-19 syndrome? *ERJ Open Res* (2020) 6(4). doi: 10.1183/23120541.00542-2020

28. Wijeratne T, Crewther S. Post-covid 19 neurological syndrome (pcns); a novel syndrome with challenges for the global neurology community. *J Neurol Sci* (2020) 419:117179. doi: 10.1016/j.jns.2020.117179

29. Guedj E, Campion JY, Dudouet P, Kaphan E, Bregeon F, Tissot-Dupont H, et al. (18)f-fdg brain pet hypometabolism in patients with long covid. *Eur J Nucl Med Mol Imaging* (2021) 48(9):2823–33. doi: 10.1007/s00259-021-05215-4

30. Moreno-Perez O, Merino E, Leon-Ramirez JM, Andres M, Ramos JM, Arenas-Jimenez J, et al. Post-acute covid-19 syndrome. incidence and risk factors: a mediterranean cohort study. *J Infect* (2021) 82(3):378–83.

31. Truffaut L, Demey L, Bruyneel AV, Roman A, Alard S, De Vos N, et al. Post-discharge critical covid-19 lung function related to severity of radiologic lung involvement at admission. *Respir Res* (2021) 22(1):29. doi: 10.1186/s12931-021-01625-y

32. Gribensk A, Schneider A, Gallaher JR, Reid TS, Kindell DG, Charles AG, et al. Posthospitalization outcomes after extracorporeal membrane oxygenation (ecmo) for covid-19. *Surgery* (2022) 172(1):466–9. doi: 10.1016/j.surg.2022.01.044

33. Dinh A, Mercier JC, Jaulmes L, Artigou JY, Juilliere Y, Yordanov Y, et al. Safe discharge home with telemedicine of patients requiring nasal oxygen therapy after covid-19. *Front Med (Lausanne)* (2021) 8:703017.

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 21 June 2023
DOI 10.3389/fpubh.2023.1219271

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Edmond Puca,
Service of Infection Diseases University Hospital Center, Albania
Abdel-Hameed Al-Mistarehi,
Johns Hopkins Medicine, United States

*CORRESPONDENCE
Márcio Flávio Moura de Araújo
✉ marcio.moura@fiocruz.br

RECEIVED 08 May 2023
ACCEPTED 02 June 2023
PUBLISHED 21 June 2023

CITATION
de Araújo MFM, Monteiro FPM, de Araújo TM, Lira Neto JCG, Santos LFS, Rolim ILTP, Santos FS, Pascoal LM, Costa ACPDJ and Santos Neto M (2023) Individual and mutual effects of diabetes, hypertension, and obesity on acute respiratory distress syndrome mortality rates in clinical patients: a multicentre study.
Front. Public Health 11:1219271.
doi: 10.3389/fpubh.2023.1219271

Individual and mutual effects of diabetes, hypertension, and obesity on acute respiratory distress syndrome mortality rates in clinical patients: a multicentre study

Márcio Flávio Moura de Araújo[1]*, Flávia Paula Magalhães Monteiro[2], Thiago Moura de Araújo[2], José Cláudio Garcia Lira Neto[1], Lívia Fernanda Siqueira Santos[3], Isaura Letícia Tavares Palmeira Rolim[3], Floriacy Stabnow Santos[3], Livia Maia Pascoal[3], Ana Cristina Pereira de Jesus Costa[3] and Marcelino Santos Neto[3]

[1]Department of Family Health, Oswaldo Cruz Foundation, Eusébio, Brazil, [2]Health Science Institute, University for International Integration of the Afro Brazilian Lusophony, Redenção, Brazil, [3]Center for Social Science, Health and Technology, Federal University of Maranhão, Imperatriz, Brazil

Patients with comorbidities are more vulnerable to severe clinical cases of acute respiratory distress syndrome (ARDS) and COVID-19 require complex health care. To analyse the association between the individual and combined effects of diabetes, hypertension, and obesity on ARDS mortality rates among patients receiving clinical care. A multicentre study encompassing retrospective data analysis and conducted with 21,121 patients from 6,723 health services across Brazil, during the 2020–2022 time period. The sample group consisted of clinical patients of both sexes and different age groups who received clinical care and showed at least one comorbidity. The data collected were analysed using binary logistic regressions and the Chi-square test. The overall mortality rate was 38.7%, with a higher predominance among males ($p<0.001$), mixed-race individuals ($p<0.001$), and older adults ($p<0.001$). The main comorbidity variables associated with and leading to death from ARDS were arterial hypertension ($p<0.001$), diabetes mellitus ($p<0.001$), diabetes mellitus and arterial hypertension ($p<0.001$), cardiovascular diseases ($p<0.001$) and obesity ($p<0.001$). Both the patients who progressed to recovery (48.4%) and to death (20.5%) presented only one comorbidity (χ^2 (1,749)=8, $p<0.001$), respectively. The isolated comorbidities with the greatest impact on death outcomes were diabetes (95% CI 2.48–3.05, $p<0.001$), followed by obesity (95% CI 1.85–2.41, $p<0.001$) and hypertension (95% CI 1.05–1.22, $p<0.001$), even after adjusting for sex and number of simultaneous comorbidities. Diabetes and obesity, as isolated conditions, had a greater influence on the number of deaths of clinical patients with ARDS compared to those with mutual diagnosis of diabetes, hypertension and obesity.

KEYWORDS

COVID-19, comorbidity, diabetes, obesity, hypertension, epidemiology

1. Introduction

Since the beginning of the pandemic, chronic, non-communicable diseases (NCDs) have been associated with severe and lethal cases of COVID-19, given that NCDs are characterised by a set of pathologies with multiple causes and risk factors, long latency periods, and prolonged courses (1). Coronavirus infections have affected people of different health profiles. However, we have noticed a higher prevalence of severe clinical complications among those affected by pre-existing comorbidities. In this regard, diabetes mellitus, hypertension, and obesity are reported to be among the most lethal NCDs in the presence of viral infections in their most severe form, according to several publications in different scenarios around the world (2–4).

It has been unanimously confirmed by meta-analysis data that cardiovascular diseases, obesity, hypertension, and diabetes are the main comorbidities observed in COVID-19-infected patients (1, 5, 6). In addition, some studies have identified diabetes and hypertension as comorbidities associated with severe and fatal cases, respectively, of COVID-19 and acute respiratory distress syndrome (ARDS) in clinical patients (5, 6). To this end, another review identified that kidney disease is the most prevalent risk factor in cases resulting in patient death, while obesity, even though very prevalent, has shown no association with death in COVID-19 and ARDS cases (1).

As previously mentioned, in isolation, such comorbidities have potentially strong effects in the presence of ARDS, especially in the most severe stage of the disease. However, when these diseases are mutually present in a patient, the combined relationship with the virus, its pathophysiological effects, and other associated factors such as age and sex, is complex and, when consulting the literature, this link seems to not have yet been elucidated.

From this perspective, this study is justified by the need to address and understand the influence between the comorbidities mutually present and patient death outcomes, in order to establish safer and more effective guidelines in the management of hospitalised and/or outpatient ARDS patients. The study of these interactions also has the potential to elucidate mechanisms of hospital readmission for respiratory diseases in humans (7). Therefore, the main aim of this study was to analyse the influence of both the isolated and the combined presence of diabetes mellitus, systemic arterial hypertension, and obesity in ARDS patients on their death rates. We hypothesize that patients with ARDS receiving clinical care affected by these NCDs, either mutually or simultaneously, have a higher death chance in relation to other patients with different comorbidities.

2. Materials and methods

2.1. Ethical aspects

This investigation was authorized by the Human Research Ethics Committee of the Federal University of Maranhão (UFMA) under approval n° 4,227,396.

2.2. Research design

This is a multicentre study using a retrospective data analysis, conducted across 6,723 health services. The service providers are located across 19 healthcare area divisions in the state of Maranhão, including 2,049 primary health care units, 1,246 specialized outpatient clinics, 588 support units (hospitals, diagnosis centres, and treatment centres), and 404 health units. The state of Maranhão is located in the northeast region of Brazil and has approximately 7,075,181 inhabitants living across 217 municipalities. The data used in this study was from a database of patients and were collected between March 2020 and January 2022.

The research analysed the epidemiological surveillance clinical records of the Covid-19 Notification System of Maranhão, Brazil. The system records and tracks notifications of the influenza and the ARDS up to patient death.

The study comprised the period from March 2020 to January 2022, when 386,567 cases of ARDS were registered in the notification system in Maranhão, Brazil. Of these cases, 10,986 resulted in death.

2.3. Criteria for defining deaths related to SARS-CoV-2 infection

To define death by ARDS, we considered clinical, epidemiological and/or laboratory criteria.

The clinical criterion consisted of reported deaths of patients diagnosed with acute Covid 19 (clinical picture <7 days) or diagnosed with acute respiratory failure syndrome in a hospital setting. Epidemiological criteria consisted of death of patients who showed close or household contact with a laboratory-confirmed case for COVID-19 in the previous 7 days; death of patients who lived or worked in an area with a high risk of virus spread (nursing homes, homeless shelters, etc.), and who showed up to 7 days of clinical symptoms; or death of a health professional working in a hospital environment with up to 7 days of clinical symptoms. Finally, the laboratory criteria consisted of the death of patients with clinical symptoms in the last 7 days who had undergone the following laboratory tests: molecular biology (detectable positive result for SARS-CoV-2 by the real-time RT-PCR method), or antigen test (reagent for SARS-CoV-2 by the immunochromatography method).

2.4. Participants

In this investigation, general patients of both sexes, any age group, with a positive diagnosis for ARDS, and with some comorbidities were eligible. Based on these eligibility criteria, the final sample of this study was 21,121 patients.

2.5. Prediction variables

For this study, individual-related variables (sex, age group, race, type of health service, diagnostic criteria, and type of test), as well as pre-existing comorbidities were listed as prediction variables, namely: hypertension, diabetes, diabetes and hypertension, obesity, diabetes and obesity, cardiovascular diseases, respiratory diseases, neurological diseases, cancer, metabolic diseases, kidney diseases, smoking, gastric diseases, psychiatric disorders, HIV-AIDS, rheumatological diseases, autoimmune diseases, hepatitis, alcoholism, urological diseases, diabetes mellitus, leprosy, chemical dependency, haematological diseases, malformation, rare diseases and dermatological diseases.

10.3389/fpubh.2023.1219271

2.6. Outcome variables

The following variables were chosen as outcomes, namely:
- mortality rate in patients with comorbidities,
- mortality rate in patients with diabetes,
- mortality rate in patients with hypertension,
- mortality rate in patients with obesity,
- mortality rate in patients with diabetes and hypertension,
- mortality rate in patients with diabetes and obesity,
- mortality rate in patients with diabetes, hypertension, and obesity.

2.7. Data analysis

Uni- and bivariate exploratory analysis of the data were carried out with the aid of the free software JAMOVI version 1.6.

The specific mortality rate from the above selected causes was calculated as follows:

$$\frac{\begin{array}{l}Number\ of\ deaths\ of\ patients\ with\ a\ specific\ comorbity\\ between\ January\ 2020\ and\ March\ 2022\end{array}}{\begin{array}{l}Total\ number\ of\ deaths\ of\ patients\ with\ any\ comorbity\\ between\ January\ 2020\ and\ March\ 2022\end{array}} \times 100$$

The normality of the variables was examined via the Shapiro–Wilk test. We compared the number of commodities per group (deaths vs. recoveries) using the Mann–Whitney test. Through contingency tables, we observed differences in the frequency of comparisons between the predictor variables (comorbidities) and death outcomes ('yes' or 'no'), through a Chi-square test. In turn, for correlations, the Spearman correlation test was used.

In the associations where a $p < 0.001$ were observed, we performed a binary logistic regression analysis, respecting the assumption of the absence of multicollinearity of the variables and the adjustment indices of the models (variance inflation factor [VIF] < 5 and tolerance >0.1). To analyse the goodness of fit for the logistic regression model, we used the Hosmer-Lemeshow test, while to determine the degree of risk, odds ratios were calculated within a 95% confidence interval. Two-tailed Alpha (α) values below 0.05 were statistically significant.

3. Results

Regarding the profiling of the patients, it can be highlighted that the sample distribution in relation to sex was balanced, but as for the age group, we observed that the older adult group (> 60 years of age) was predominant (54.5%). A substantial portion of the patients had brown skin color, which, in a general Brazilian classification, corresponds to mixed-race individuals. Disease diagnostics and detections took place mostly in public health service providers via laboratory tests, in this case, rapid tests (Table 1).

The overall mortality rate was 38.7% in this study. Death cases were higher for men (22.5%) than for women (16.2%) [$\chi^2(425) = 1$, $p < 0.001$]. Regarding the group of patient cases who evolved to death, a directly proportional relationship was observed. With an age increase, there is also a statistically significant percentage increase in mortality from ARDS: 0–9 (0.1%), 10–19 (0.1%), 20–29 (0.4%), 30–39 (1.3%), 40–49 (2.6%), 50–59 (4.9%), 60–70 (10%), and > 70 years of age (19.3%) [χ^2 (3024) =8, $p < 0.001$]. Deaths were predominant in mixed raced (24.8%) and white (7.3%) individuals [χ^2 (1357) =5, $p < 0.001$].

TABLE 1 Profile of clinical patients with COVID-19 showing at least one comorbidity. Maranhão, Brazil, 2023.

Variable		*n*	%
Age group	0–9	115	0.5
	10–19	295	1.4
	20–29	709	3.4
	30–39	1,766	8.4
	40–49	2,938	14
	50–59	3,813	18
	60–70	4,838	23
	Above 70	6,605	31.5
Sex	Female	10,740	50.8
	Male	10,381	49.2
Race	Brazilian Yellow	2,958	14.9
	Brazilian White	3,409	17.2
	Indigenous	75	0.3
	Mixed race	11,721	59.3
	Black	1,598	8
Death	Yes	8,172	38.7
	No	12,943	61.3
Diagnostic criteria	Clinical examination	119	0.6
	Clinical examination and tomography	192	0.9
	Lab test	20,808	98.5
Laboratory	Public/National	18,821	90.7
	Private	1,915	9.3
Exam type	Serological	404	1.9
	Rapid test	13,813	68.1
	RT-PCR	6,605	30
Health service	Public/National	13,052	88.4
	Private	1,729	11.6

Regarding total deaths, the most significant comorbidities for mortality rates were arterial hypertension (63.7%), diabetes mellitus (39.5%), diabetes mellitus and arterial hypertension (26.2%), obesity (7.4%) and diabetes mellitus and obesity (1.8%). The main comorbidity variables associated with and resulting in death from ARDS were arterial hypertension ($p < 0.001$), diabetes mellitus ($p < 0.001$), diabetes mellitus and arterial hypertension ($p < 0.001$), cardiovascular diseases ($p < 0.001$) and obesity ($p < 0.001$) (Table 2).

In turn, for the correlation study, we identified that the comorbidities with the highest positive correlation with a death by ARDS outcome were diabetes mellitus (Spearman *rho* = 0.25 e $p < 0.001$), and the simultaneous presence of diabetes mellitus and arterial hypertension (Spearman *rho* = 0.22 e $p < 0.001$). We also observed that death outcomes presented a negative and statistically significant correlation (Spearman rho = −0.27 and $p < 0.001$) with the number of comorbidities per patient. Also, in the 75th percentile, it is possible to verify that the patients who died had 3 simultaneous

10.3389/fpubh.2023.1219271

TABLE 2 Distribution of COVID-19 mortalities among clinical patients with comorbidities. Maranhão, Brazil, 2023.

Comorbidities	Death		
	Yes *n* (%)	No *n* (%)	*p*-value[1]
Hypertension	5,208 (24.6)	7,428 (35.1)	**<0.001**
Diabetes	3,235 (15.3)	2,198 (10.4)	**<0.001**
Diabetes and hypertension	2,142 (10.1)	1,246 (5.9)	**<0.001**
Obesity	612 (2.9)	484 (2.3)	**<0.001**
Diabetes and obesity	149 (0.7)	28 (0.1)	**<0.001**
Diabetes, hypertension and obesity	110 (0.5)	18 (0.1)	**<0.001**
Cardiovascular diseases	1,309 (6.2)	2,565 (12.1)	**<0.001**
Respiratory diseases	543 (2.5)	1,479 (7)	**<0.001**
Neurological diseases	527 (2.5)	192 (0.9)	**<0.001**
Cancer	325 (1.5)	163 (0.7)	**<0.001**
Metabolic diseases	61 (0.2)	86 (0.4)	0.767
Kidney disease	477 (2.2)	312 (1.4)	**<0.001**
Smoking	399 (1.8)	96 (1.4)	**<0.001**
Gastric diseases	14 (0.0)	78 (0.3)	**<0.001**
Psychiatric disorders	52 (0.2)	143 (0.6)	**0.002**
HIV	55 (0.2)	41 (0.1)	**<0.001**
Rheumatological diseases	36 (0.1)	68 (0.3)	0.682
Autoimmune diseases	28 (0.1)	48 (0.2)	0.936
Hepatitis	70 (0.3)	46 (0.2)	**<0.001**
Ethanolism	61 (0.2)	12 (0.0)	**<0.001**
Urological diseases	3 (0.0)	20 (0.0)	0.041
Hansen's disease	21 (0.0)	19 (0.0)	0.199
Drug addiction	21 (0.0)	407 (1.9)	**<0.001**
Haematological disease	12 (0.0)	24 (0.1)	**<0.001**
Malformation	27 (0.1)	403 (1.9)	**<0.001**
Rare diseases	5 (0.0)	4 (0.0)	0.583
Dermatological diseases	–	10 (0.0)	**0.042**

[1]Chi-square test.
The bold values are statistically significant data.

comorbidities in relation to the group that did not evolve to death, whose comorbidity number was 1 ($p < 0.001$).

We also observed a statistically significant association between death outcome and the number of simultaneous comorbidities. In this case, both those who evolved to recovery (48.4%) and to death (20.5%) had only 1 comorbidity (χ^2 (1.749) =8, $p < 0.001$), respectively (Figure 1).

When comparing the percentiles for the variable number of comorbidities, we observed that in the 95th and 99th percentiles, the number of comorbidities was 4 and 5, respectively, in the group that evolved to death. In the group that evolved to recovery in the same position, the values found were 3 and 4, respectively (U-Mann Whitney Test, $p < 0.0001$).

The impact of comorbidities on the deaths of ARDS patients was tested in three binary logistic regression models. The most pronounced effect was observed in the comorbidity diabetes (95% CI 2.48–3.05, $p < 0.001$), followed by obesity (95% CI 1.85–2.41, $p < 0.001$) and hypertension (95% CI 1.05–1.22, $p < 0.001$) individually, even after adjustments for sex and number of concurrent comorbidities (Table 3).

4. Discussion

We found that the isolated presence of diabetes and obesity was more associated with the prevalence of ARDS deaths than the simultaneous presence of comorbidities. By the way, patients with COVID-19 who progress to death used to mostly have at least one comorbidity.

Data from other studies conducted in the Middle East found that the comorbidities most associated with death from COVID-19 were kidney injury, deep vein thrombosis, and tumors (8). Other researchers concluded that the triad of smoking, hypertension, and diabetes mellitus increases the mortality rate (9). In fact, both smoking and kidney injury are mentioned in association with the use of invasive mechanical ventilation, in the case of patients hospitalised for COVID-19 (10).

In this regard, it is important to highlight in this discussion the evidence of a meta-analysis on this topic, which implies that advanced age and the presence of two or more comorbidities are significantly impactful for the clinical evolution of ARDS cases in general hospitalised patients (11). Also, the conclusions of a recent published cohort study are interesting as they show that age and comorbidities can predict the outcome of ARDS, regardless of the severity of the patients' innate immune response (12). This means that, particularly when NCDs are individually or mutually present in patients, they become an additional burden on the vulnerability of older adults to the SARS-CoV-2. In our findings, we found a positive correlation between death outcomes and the number of comorbidities statistically significant, but a weak association. It was also observed that both in recovered patients and in cases of death, one comorbidity was predominant.

Reviewing the literature, we observed the relationship between the number of comorbidities and mortality from COVID-19 is not unanimously reported, as there are some data that show an exponential relationship, while others show an inverse relationship (13–15). However, the common agreement among the consulted studies is that patients affected by different NCDs and COVID-19 have a high and differentiated prevalence around the world that ranges from 33.8–56.6% of the investigated samples (2).

The inflammatory load related to the COVID-19 pathophysiology is interleukin 6 (IL-6), whose high levels cause cardiovascular damage and the deterioration of the clinical condition of patients, especially of those who already have pre-existing comorbidities, such as chronic NCDs. This is due to these patients having an inadequate immune response owing to other factors such as age, functional impairments, and lifestyle choices (16, 17).

In this sense, fatal cases of ARDS are closely related to the imbalance of the renin-angiotensin-aldosterone system and hyper-inflammation, which causes a dysregulated immune response and a subsequent activation and dysfunction of endothelial cells, which mutually can trigger a thrombotic event. Therefore, in patients with systemic arterial hypertension, for example, this cascade of dysfunctional events can be fatal (18).

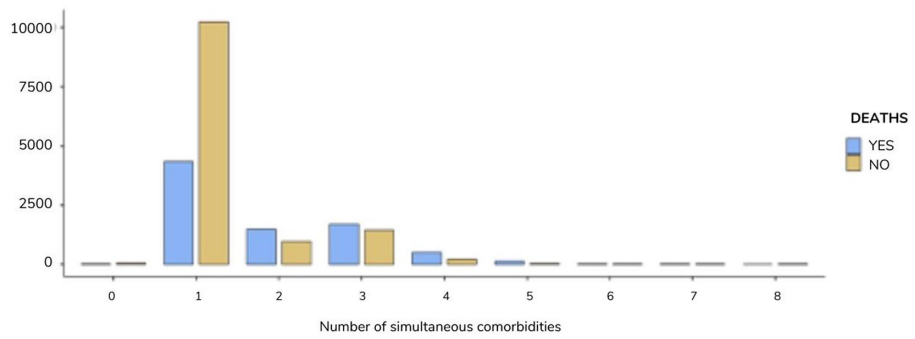


FIGURE 1
Distribution of the number of simultaneous comorbidities resulting in patient death by ADRS. Maranhão, Brazil, 2023.

TABLE 3 Impact of comorbidities (diabetes, hypertension, and obesity), isolated and/or combined, on the death of clinical patients by ADRS. Maranhão, Brazil, 2023.

Model	Predictor	Odds ratio	Confidence interval		*p*-value
			Lower	Upper	
1	Hypertension	1.30	1.21	1.39	**<0.001**
	Diabetes	3.03	2.73	3.36	**<0.001**
	Obesity	2.60	2.29	2.96	**<0.001**
	Diabetes and hypertension	1.14	1.06	1.30	**0.040**
	Diabetes and obesity	1.38	0.73	2.61	0.319
	Diabetes, hypertension and obesity	1.10	0.51	2.40	0.794
2	Hypertension	1.13	1.05	1.22	**<0.001**
	Diabetes	2.76	2.48	3.06	**<0.001**
	Obesity	2.11	1.85	2.41	**<0.001**
	Diabetes and hypertension	0.51	0.43	0.60	**<0.001**
3	Hypertension	1.13	1.05	1.22	**<0.001**
	Diabetes	2.75	2.48	3.05	**<0.001**
	Obesity	2.11	1.85	2.41	**<0.001**
	Diabetes and hypertension	0.51	0.43	0.60	**<0.001**

Model 1: no fits.
Model 2: fitted for amount of simultaneous comorbidities.
Model 3: fitted for amount of simultaneous comorbidities and sex.
The bold values are statistically significant data.

Through logistic regression analysis, we detected that mortality rates were higher in those diagnosed with isolated comorbidities of diabetes and obesity, respectively, and the odd ratios for death were more than double and almost doubled ($p < 0.001$) among them. This finding is specifically corroborated by an epidemiological observatory of ARDS related deaths in Brazil, with an interval of 34 weeks of follow-up, which pointed out diabetes and obesity, respectively, always being at the top of the list of factors associated with deaths (19).

In the realm of comorbidities, including hypertension, cardiovascular disease, and chronic obstructive pulmonary disease (COPD), among others, diabetes is considered as a crucial comorbidity for the survival of patients infected by the coronavirus (20). Previous published studies attest to the potent effects of diabetes, in relation to other chronic diseases, to cause deaths by COVID 19 in patients with comorbidities (13, 21). By specifically comparing diabetes and obesity, among the publications consulted, we identified that diabetes is a more potent risk factor for death from COVID 19 (14, 18, 22).

This disease causing a decrease in phagocytic activity, neutrophil chemotaxis, decreased T-cell function, and lower innate immunity in patients affected by Type 2 Diabetes. In addition, among these patients there are higher levels of angiotensin-converting enzyme-2 (ACE2), which serves as an entry receptor for the respiratory virus due to its high binding affinity expressed in pulmonary alveolar cells, cardiac cells, vascular endothelium, as well as other various sites (23).

In the case of obesity, it is well documented in the literature, through observational and review studies, that this condition alone is an intrinsic factor for respiratory failure and that it increases the risk of death in patients affected by COVID-19 by up to five times, as it aggravates previous chronic conditions and increases demand for mechanical ventilation (22–26). Overweight and obesity are characterised by an accumulation of abnormal or excessive fat that increases health risks and can trigger a series of other problems, such as cardiac and vascular conditions, diabetes, endocrinal complications, among others. These are all conditions associated with a high risk of mortality from COVID-19. In addition, the expression of ACE-2 receptors is higher in the visceral adipose tissue than in subcutaneous fat, allowing for a higher viral load and, consequently, more severe forms of the disease in patients with increased visceral fat (4). Inflammatory processes resulting from the infection and immune dysregulation in obese individuals potentialize this relationship (27).

Thus, the scenario presented through our findings poses the pertinent reflection that it is not necessarily the number of comorbidities, but instead, the influence that certain comorbidities can bring to ARDS patients that can lead to an increase in mortality rate.

To date, it is known that the COVID-19 virus binds to the ACE-2, decreasing the activity of this type of receptor and leading to more severe cases. Precisely in patients with diabetes mellitus and systemic arterial hypertension, this receptor is present in high levels. In addition, in the presence of diabetes mellitus, the SARS-CoV-2 virus

hijacks an endocrine pathway that plays a crucial role in regulating blood pressure, metabolism, and inflammation, which are aspects that accentuate cellular damage, hyperinflammation, and respiratory failure (28).

In obesity, this bidirectional relationship is no different, given the inflammatory load of the obese and their inefficient immune system in the face of COVID-19 infection. Scholars have stated that cytokine/adipokine levels and inflammatory markers, such as C-reactive proteins, are associated with a higher body mass index in COVID-19-positive patients, suggesting that the inflammatory background and immune dysregulation of obese patients may influence the immune response in this group (27).

Based on our final model, we think it is mandatory for healthcare professionals to consider the amount and type of concurrent comorbidities of patients with ARDS. Therefore, for the purpose of risk stratification, it is crucial for the health care team to understand the parameters that predispose patients with an ARDS to a more severe course of the disease, especially to adopt more appropriate health protocols. In addition, the recognition of these parameters and associated factors are important tools to characterize the typical behavior of the disease, as well as to guide decision-making in the context of public health policies and epidemiological surveillance.

It then becomes clear that the medical team giving care to general patients should adhere to glycemic management and care with nutrition and dietetics focus for people with NCDs and ARDS. This is due to these being triggers for important clinical changes that predict mortality, even when we consider the variable comorbidities. The performance of health workers teams in the management of the above diseases should be mandatory, given that there may be hospital readmission of patients already considered to have recovered from COVID-19 (7). In this sense, the nursing management should include a therapeutic plan focused both on the primary as well as secondary prevention of diabetes and overweight, since many patients may be people with undiagnosed diabetes mellitus or with poor nutritional management.

4.1. Limitations

Database-based studies present vulnerability points such as incomplete information or missing data. This was no different in our investigation. For example, we did not have access to the number of days until a death outcome was observed, or to an overview of the current stage of management of the patients' comorbidities aiming at a possible study of the interaction between the status of said comorbidities and a death outcome. Furthermore, other important data such as vaccination status for COVID-19 and types of medication in use were not part of our database and could provide us with a more complete picture of these patients.

Therefore, we only investigated complete data sheets to avoid the use of *missing data* and thus ensure the best fit of the regression model. In addition, this was a sample of considerable size, which allowed for an in-depth study of the influence of the simultaneous presence of important comorbidities. This investigation may help to generalize and extrapolate findings on mortality rates from ARDS in clinical patients with simultaneous comorbidities, which, to date, lacks in evidence in the current health literature.

5. Conclusion

The mortality rates of clinical patients with comorbidities, hospitalised due to ARDS with a simultaneous presence of diabetes, hypertension, and obesity, do not differ from that of those who presented either of these comorbidities individually, or a combination of hypertension and diabetes. On the contrary, a greater impact on death outcomes was observed in patients with the isolated presence of diabetes or obesity.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material. All deaths from COVID-19, obtained from the Maranhão COVID-19 Notification System (SNC-19 MA), linked to the Department of Monitoring and Health Evaluation of the Superintendence of Epidemiology and Disease Control, part of the State Health Department of Maranhão (SES/MA), were considered for the study. This system consolidates data from the notification forms for the flu syndrome and the form for severe acute respiratory syndrome due to COVID-19 (SES/MA, 2020). Data on the population estimate were collected from the IBGE website (https://sidra.ibge.gov.br/). Reference: SES/MA 2020: STATE SECRETARY OF HEALTH OF MARANHÃO (SES/MA). Technical Note No. 01/CIEVS/ /SECD/SAPAPVS/2020. Available at: https://www.saude.ma.gov.br/docs/nota-tecnica-no-01-cievs-secd-sapapvs-2020/.

Ethics statement

The studies involving human participants were reviewed and approved by Human Research Ethics Committee of the Federal University of Maranhão. The patients/participants provided their written informed consent to participate in this study.

Author contributions

MA: conceptualization, data curation, formal analysis, funding acquisition, methodology, and writing—original draft. FM and JL: conceptualization, validation, and writing—original draft. TA, LS, IR, FS, LP, AC, and MS: methodology and validation. All 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

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.

References

1. Ng WH, Tipih T, Makoah NA, Vermeulen JG, Goedhals D, Sempa JB, et al. Comorbidities in SARS-CoV-2 patients: a systematic review and Meta-analysis. *MBio*. (2021) 12:e03647–20. doi: 10.1128/mBio.03647-20

2. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the new York City area [published correction appears in JAMA. 2020 may 26;323(20):2098]. *JAMA*. (2020) 323:2052–9. doi: 10.1001/jama.2020.6775

3. Honardoost M, Janani L, Aghili R, Emami Z, Khamseh ME. The association between presence of comorbidities and COVID-19 severity: a systematic review and Meta-analysis. *Cerebrovasc Dis*. (2021) 50:132–40. doi: 10.1159/000513288

4. Vulturar DM, Crivii CB, Orăsan OH, Palade E, Buzoianu AD, Zehan IG, et al. Obesity impact on SARS-CoV-2 infection: pros and cons "obesity paradox"-a systematic review. *J Clin Med*. (2022) 11:3844. doi: 10.3390/jcm11133844

5. Wang Z, Deng H, Ou C, Liang J, Wang Y, Jiang M, et al. Clinical symptoms, comorbidities and complications in severe and non-severe patients with COVID-19: a systematic review and meta-analysis without cases duplication. *Medicine (Baltimore)*. (2020) 99:e23327. doi: 10.1097/MD.0000000000023327

6. Gold MS, Sehayek D, Gabrielli S, Zhang X, McCusker C, Ben-Shoshan M. COVID-19 and comorbidities: a systematic review and meta-analysis. *Postgrad Med*. (2020) 132:749–55. doi: 10.1080/00325481.2020.1786964

7. Akbari A, Fathabadi A, Razmi M, Zarifian A, Amiri M, Ghodsi A, et al. Characteristics, risk factors, and outcomes associated with readmission in COVID-19 patients: a systematic review and meta-analysis. *Am J Emerg Med*. (2022) 52:166–73. doi: 10.1016/j.ajem.2021.12.012

8. Karasneh RA, Khassawneh BY, Al-Azzam S, Al-Mistarehi A-H, Lattyak WJ, Aldiab M, et al. Risk factors associated with mortality in COVID-19 hospitalized patients: data from the Middle East. *Int J Clin Pract*. (2022) 2022:1–10. doi: 10.1155/2022/9617319

9. Al-Mistarehi AH, Al-Azzam S, Karasneh R, Al Sbihi A, Alomari S, Haj Hussein AA, et al. Risk factors for in-hospital mortality among adults hospitalized with COVID-19: a cross-sectional study from Jordan. *Chest*. (2022) 162:A659–60. doi: 10.1016/j.chest.2022.08.514

10. Kabbaha S, Al-Azzam S, Karasneh RA, Khassawneh BY, Al-Mistarehi AH, Lattyak WJ, et al. Predictors of invasive mechanical ventilation in hospitalized COVID-19 patients: a retrospective study from Jordan. *Expert Rev Respir Med*. (2022) 16:945–52. doi: 10.1080/17476348.2022.2108796

11. Cheng S, Zhao Y, Wang F, Chen Y, Kaminga AC, Xu H. Comorbidities' potential impacts on severe and non-severe patients with COVID-19: a systematic review and meta-analysis. *Medicine (Baltimore)*. (2021) 100:e24971. doi: 10.1097/MD.0000000000024971

12. Mohan AA, Olson LB, Naqvi IA, Morrison SA, Kraft BD, Chen L, et al. Age and comorbidities predict COVID-19 outcome, regardless of innate immune response severity: a single institutional cohort study. *Crit Care Explor*. (2022) 4:e0799. doi: 10.1097/CCE.0000000000000799

13. Toofan F, Hosseini SM, Alimohammadzadeh K, Jafari M, Bahadori M. Impact of comorbidities on mortality in hospitalized patients with COVID-19: an experience from Iran. *J Educ Health Promot*. (2021) 10:460. doi: 10.4103/jehp.jehp_1589_20

14. Singh P, Bhaskar Y, Verma P, Rana S, Goel P, Kumar S, et al. Impact of comorbidity on patients with COVID-19 in India: a nationwide analysis. *Front Public Health*. (2023) 10:1027312. doi: 10.3389/fpubh.2022.1027312

15. Roedl K, Jarczak D, Boenisch O, de Heer G, Burdelski C, Frings D, et al. Chronic critical illness in patients with COVID-19: characteristics and outcome of prolonged intensive care therapy. *J Clin Med*. (2022) 11:1049. doi: 10.3390/jcm11041049

16. Carvalho PR, Sirois P, Fernandes PD. The role of kallikrein-kinin and renin-angiotensin systems in COVID-19 infection. *Peptides*. (2021) 135:170428. doi: 10.1016/j.peptides.2020.170428

17. Ángeles Correa MG, Villarreal Ríos E, Galicia Rodríguez L, Vargas Daza ER, Frontana Vázquez G, Monrroy Amaro SJ, et al. Enfermedades crónicas degenerativas como factor de riesgo de letalidad por COVID-19 en México [Chronic degenerative conditions as risk factors for lethal COVID-19Doenças crônicas degenerativas como fator de risco para letalidade por COVID-19]. *Rev Panam Salud Publica*. (2022) 46:e40. doi: 10.26633/RPSP.2022.40

18. Peña JE, Rascón-Pacheco RA, Ascencio-Montiel IJ, González-Figueroa E, Fernández-Gárate JE, Medina-Gómez OS, et al. Hypertension, diabetes and obesity, major risk factors for death in patients with COVID-19 in Mexico. *Arch Med Res*. (2021) 52:443–9. doi: 10.1016/j.arcmed.2020.12.002

19. Siqueira LD, Muchon JD, Arruda JT, Paludo RL d R. Analysis of mortality from COVID-19 and obesity as a risk factor. *RSD*. (2022) 11:e10911123432. doi: 10.33448/rsd-v11i1.23432

20. Das SKRA, Birangal SR, Nikam AN, Pandey A, Mutalik S, Joseph A. Role of comorbidities like diabetes on severe acute respiratory syndrome coronavirus-2: a review. *Life Sci*. (2020) 258:118202. doi: 10.1016/j.lfs.2020.118202

21. Sindi AA, Tashkandi WA, Jastaniah MW, Bashanfar MA, Fakhri AF, Alsallum FS, et al. Impact of diabetes mellitus and co-morbidities on mortality in patients with COVID-19: a single-center retrospective study. *Saudi Med J*. (2023) 44:67–73. doi: 10.15537/smj.2023.44.1.20220462

22. Nagy É, Cseh V, Barcs I, Ludwig E. The impact of comorbidities and obesity on the severity and outcome of COVID-19 in hospitalized patients-a retrospective study in a Hungarian hospital. *Int J Environ Res Public Health*. (2023) 20:1372. doi: 10.3390/ijerph20021372

23. Pranata R, Henrina J, Raffaello WM, Lawrensia S, Huang I. Diabetes and COVID-19: the past, the present, and the future. *Metabolism*. (2021) 121:154814. doi: 10.1016/j.metabol.2021.154814

24. Rocha GV, Soares CEM, Filho LH d O, do AMVF, de Castro VE, Junior EA, et al. A influência da obesidade na mortalidade de adultos com COVID-19 / the influence of obesity on adult mortality with COVID-19. *Braz J Hea Rev*. (2021) 4:1405–18. doi: 10.34119/bjhrv4n1-119

25. Suresh S, Siddiqui M, Abu Ghanimeh M, Jou J, Simmer S, Mendiratta V, et al. Association of obesity with illness severity in hospitalized patients with COVID-19: a retrospective cohort study. *Obes Res Clin Pract*. (2021) 15:172–6. doi: 10.1016/j.orcp.2021.02.006

26. Ribeiro AC, Poli P, da SA USC. Increased risk of mortality from COVID-19 in people with obesity. *Rev Rene*. (2023) 24:e81453. doi: 10.15253/2175-6783.20232481453

27. Belchior-Bezerra M, Lima RS, Medeiros NI, Gomes JAS. COVID-19, obesity, and immune response 2 years after the pandemic: a timeline of scientific advances. *Obes Rev*. (2022) 23:e13496. doi: 10.1111/obr.13496

28. Bornstein SR, Rubino F, Khunti K, Mingrone G, Hopkins D, Birkenfeld AL, et al. Practical recommendations for the management of diabetes in patients with COVID-19. *Lancet Diabetes Endocrinol*. (2020) 8:546–50. doi: 10.1016/S2213-8587(20)30152-2

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 18 July 2023
DOI 10.3389/fpubh.2023.1150818

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Mohamad-Rodi Isa,
Universiti Teknologi MARA, Malaysia
Farzad Pourghazi,
Tehran University of Medical Sciences, Iran

*CORRESPONDENCE
Haewon Byeon
✉ bhwpuma@naver.com

RECEIVED 31 January 2023
ACCEPTED 10 April 2023
PUBLISHED 18 July 2023

CITATION
Byeon H (2023) Developing a nomogram for predicting depression in diabetic patients after COVID-19 using machine learning.
Front. Public Health 11:1150818.
doi: 10.3389/fpubh.2023.1150818

Developing a nomogram for predicting depression in diabetic patients after COVID-19 using machine learning

Haewon Byeon[1,2]*

[1]Department of Digital Anti-aging Healthcare (BK21), Graduate School of Inje University, Gimhae, Republic of Korea, [2]Department of Medical Big Data, College of AI Convergence, Inje University, Gimhae, Republic of Korea

Objective: This study identified major risk factors for depression in community diabetic patients using machine learning techniques and developed predictive models for predicting the high-risk group for depression in diabetic patients based on multiple risk factors.

Methods: This study analyzed 26,829 adults living in the community who were diagnosed with diabetes by a doctor. The prevalence of a depressive disorder was the dependent variable in this study. This study developed a model for predicting diabetic depression using multiple logistic regression, which corrected all confounding factors in order to identify the relationship (influence) of predictive factors for diabetic depression by entering the top nine variables with high importance, which were identified in CatBoost.

Results: The prevalence of depression was 22.4% (n=6,001). This study calculated the importance of factors related to depression in diabetic patients living in South Korean community using CatBoost to find that the top nine variables with high importance were gender, smoking status, changes in drinking before and after the COVID-19 pandemic, changes in smoking before and after the COVID-19 pandemic, subjective health, concern about economic loss due to the COVID-19 pandemic, changes in sleeping hours due to the COVID-19 pandemic, economic activity, and the number of people you can ask for help in a disaster situation such as COVID-19 infection.

Conclusion: It is necessary to identify the high-risk group for diabetes and depression at an early stage, while considering multiple risk factors, and to seek a personalized psychological support system at the primary medical level, which can improve their mental health.

KEYWORDS

depression, COVID-19 pandemic, CatBoost, machine learning, diabetic patients

1. Introduction

Depression is the most common psychiatric disease in diabetic patients (1). Lustman et al. (1) conducted a large-scale epidemiological study on type 1 and type 2 diabetic patients and reported that the prevalence of depression in diabetic patients was approximately 30%, which was twice of that in non-diabetics. It could be a result of the physical, mental, and economic

burden caused by diabetes and diabetic complications. However, the causal relationship between diabetes and depression has not yet been clearly understood, although approximately 30% of melancholic patients without diabetes are at risk of developing diabetes and depression is an independent risk factor for diabetes (2). Three mechanisms are suggested to explain why diabetes and depression are frequently accompanied: (1) stress due to increased intensity and repetition of diabetes treatment, (2) increased burden of other comorbidities and complications, and decreased quality of life due to the prolonged duration of diabetes, and (3) diabetes and depression share a common metabolic abnormality and are linked (3).

Depression has critically adverse effects on the prognosis of several chronic diseases (4). It especially leads to poor glycemic control by inducing diabetic patients to neglect self-care and reducing treatment compliance (4). Moreover, it eventually increases mortality by increasing the risk of microvascular complications and cardiovascular diseases (4). Since it has been reported that only 25% of diabetic patients are diagnosed with depression by medical personnel (5), early screening for depression is a very important issue for the patient's prognosis and diabetes management.

Depression and diabetes are two conditions that have been significantly impacted by the COVID-19 pandemic. Both conditions can be influenced by lifestyle factors such as diet, exercise, and social support, and the disruptions caused by the pandemic have made it more challenging for individuals to manage these conditions. Especially since depression is affected by complex interactions among various factors, such as lifestyle and social networks rather than a single factor (6), it is necessary to develop a predictive model that considers multiple risk factors simultaneously in order to efficiently predict groups vulnerable to depression. Nevertheless, only a few studies have investigated multiple risk factors for depression in diabetic patients.

Many recent previous studies (6–8) used a Bayesian nomogram as a way to identify a high risk of disease by considering multiple risk factors. A nomogram is a graph that visualizes a prediction function derived from a Bayesian model or a logit model in two dimensions so that healthcare workers can easily interpret the derived results, and it is widely used in the healthcare field, such as predicting the risk of cancer recurrence (9). In particular, since the logistic nomogram has the advantage of being able to predict the probability of occurrence due to multiple risk factors by adding up individual risk factors (6), it can be effective for predicting depression in community diabetic patients after the COVID-19 pandemic. Therefore, this study identified major risk factors for depression in diabetic patients within the community using machine learning techniques and developed predictive models to identify the high-risk group for depression in diabetic patients based on multiple risk factors.

2. Materials and methods

2.1. Data source

It is an epidemiological study using the 2020 Community Health Survey data as secondary data. The Community Health Survey is conducted under the supervision of the Korea Disease Control and Prevention Agency to produce health statistics necessary for establishing a regional healthcare plan and implementing health projects. Please see Byeon (10) for a more detailed explanation of the data collection method and others of the Community Health Survey. Briefly explaining, the 2020 survey targeted adults (≥19 years old) based on resident registration in cities, counties, and districts nationwide and sampled using the systematic sampling method by extracting sampling points assigned to each region from the sampling frame created by linking the resident registration population data and housing data, which were complete enumeration data, and identifying the number of households selected as the sampling points. The survey was conducted from August 16th to October 31st, 2020, and a trained researcher conducted a 1:1 interview with the survey subject using a laptop computer (Computer Assisted Personal Interviewing, CAPI) to collect data. CAPI minimizes human errors and ensures accuracy through automated data collection and analysis. Additionally, CAPI enables fast and efficient data collection. The process of creating surveys and collecting data is automated, resulting in time and cost savings. This study analyzed 26,829 adults living in the community who were diagnosed with diabetes by a doctor in the 2020 Community Health Survey.

2.2. Measurement and definition of variables

The prevalence of a depressive disorder was the dependent variable in this study. The Korean version of the Patient Health Questionnaire (PHQ-9) (11) was used to assess depressive disorder. PHQ-9 is a standardized depression screening test developed by Spitzer et al. (12) to diagnose mental health in primary health care centers. It is made up of nine items that correspond to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) diagnostic criteria for major depressive disorders. The PHQ-9 is a self-report test with high sensitivity and specificity (13). Furthermore, because it can simply check the severity of a depressive disorder using only nine items, it has the advantage of being highly likely to be applied to actual screening in epidemiological investigations as well as the medical field (13). The PHQ-9 asks a subject how frequently he or she has experienced anhedonia, depression, changes in sleep, fatigue, changes in appetite, guilt or worthlessness, decreased concentration, akathisia or feeling down, and suicidal thoughts in the previous 2 weeks. It is graded on a four-point scale: "never," "for a few days," "more than 1 week," and "almost every day." The total score ranges from 0 to 27, with a higher score indicating more severe depression. The threshold of depression was defined as 10 points (depression ≥ 10 points out of 27 points) based on the results of the previous studies (14, 15). Choi (16) reported that the sensitivity and specificity of PHQ-9 were 81.1 and 89.9%, respectively. Also, the reliability of the tool (Cronbach's α) was 0.89 in this study. Based on the findings of previous studies (14, 15), the depression threshold defined as 10 points (depression ≥10 points out of a possible 27 points).

The explanatory variables included changes in instant food consumption before and after the COVID-19 pandemic (increased, similar, or decreased; responses were categorized based on self-report), changes in delivery food consumption before and after the COVID-19 pandemic (increased, similar, or decreased), changes in drinking before and after the COVID-19 pandemic (increased, similar, or decreased), changes in smoking before and after the COVID-19 pandemic (increased, similar, or decreased), changes in the use of public transportation before and after the COVID-19

pandemic (increased, similar, or decreased), satisfaction with life after the COVID-19 pandemic (satisfied or dissatisfied), age (40–49 years, 50–59 years, or ≥ 60 years), gender (male or female), residing location (urban or rural), education level (elementary school graduation or below, middle school graduation, high school graduation, college graduation or above), mean monthly household income [<South Korean won (KRW) 1.00 million, KRW 1.00 million – 2.99 million, KRW 3.00 million – 4.99 million, or ≥KRW 5.00 million], smoking (current smoker, former smoker, or non-smoker), subjective health (good, moderate, or poor), fear of infection due to the COVID-19 pandemic (yes, moderate, or no), fear of death due to the COVID-19 pandemic (yes, moderate, or no), concern about reproach from people around you due to the expression of COVID-19 symptoms (e.g., coughing) (yes, moderate, or no), concern about infection of health-vulnerable people such as infants and older adults among family members due to the COVID-19 pandemic (yes, moderate, or no), concern about economic loss due to the COVID-19 pandemic (yes, moderate, or no), changes in the number of meetings with people around you due to the COVID-19 pandemic (increased, similar, or decreased), changes in sleeping hours due to the COVID-19 pandemic (increased, similar, or decreased), marital status (living with a spouse or not living with a spouse), time of first diagnosed with diabetes (<60 years old or ≥60 years old), current non-drug treatment for diabetes (e.g., exercise) (yes or no), current diabetes drug (e.g., oral hypoglycemic drug) treatment (yes or no), current insulin injection treatment (yes or no), number of HbA1c tests in the past year (1 or fewer or 2 or more), diabetic eye disease complication test (fundus examination) in the past year (yes or no), diabetic renal complication test (microalbuminuria test) (yes or no), economic activity (yes or no), awareness of own blood glucose level (yes or no), awareness of own blood pressure (yes or no), number of days of conducting moderate-intensity (e.g., yoga and cycling) physical activity at least 30 min per day in the past week (none, 1–2 days, or 3 days or more), number of days of walking at least 30 min per day in the past week (none, 1–2 days, or 3 days or more), the number of people you can ask for help in a disaster situation such as COVID-19 infection (none, 1 ~ 2, or 3 or more), and diagnosis with hypertension (yes or no).

2.3. Development of a predictive model: categorical boosting

Categorical boosting (CatBoost) is a boosting algorithm that was developed in 2017 (17). It is designed to handle categorical variables efficiently and minimize model overfitting by using an ordered boosting technique. With CatBoost, categorical variables can be used without the need to convert them into numbers. The algorithm also automatically applies a suitable encoding technique for categorical variables, such as one-hot encoding, target encoding, mean encoding, and response encoding (17). Additionally, CatBoost optimizes hyperparameters with an internal algorithm instead of using special hyperparameter optimization, making it easier to use compared to other algorithms that require hyperparameter tuning. This study set the regularization lambda, the number of trees, the limit depth of individual trees to 6, and the learning rate of CatBoost to 3, 100, 6, and 0.300, respectively. This study calculated the importance of variables based on the mean decrease in impurity and selected the top 9 variables with high importance.

2.4. Development and verification of logistic monogram

When the number of risks included in the nomogram increases, the number of cases needed to calculate the predicted probability also increases. This study developed a model for predicting diabetic depression using multiple logistic regression, which corrected all confounding factors in order to identify the relationship (influence) of predictive factors for diabetic depression by entering the top nine variables with high importance, which were identified in CatBoost. This study used an adjusted odds ratio (aOR) and 95% confidence interval (CI) to identify the independent relationship between predictors and diabetic depression.

The developed model for predicting depression in individuals with diabetes presented a graph by establishing a nomogram, which allows healthcare workers to easily interpret the probability of high-risk groups based on multiple risk factors. The nomogram developed in this study consisted of four lines. Firstly, the point line was placed at the top of the nomogram to derive scores corresponding to the categories of risk factors, and the point line of the Bayesian nomogram was between 0 and 100 points. Secondly, there were as many risk factor lines as the number of risk factors. Thirdly, the total point line was the sum of each individual risk factor and was located at the bottom of the nomogram. Finally, the probability line was placed at the bottom of the nomogram to derive the probability of depression in individuals with diabetes.

The prediction performance of the finally developed diabetic depression prediction nomogram was evaluated using the 10-fold cross-validation method. This study used F1-score, the area under the curve (AUC), general accuracy, precision, recall, and calibration plot as evaluation indices to confirm the predictive performance.

3. Results

3.1. General characteristics according to the depression prevalence in diabetic patients after the COVID-19 pandemic

Table 1 shows the characteristics of the subjects according to the depression prevalence in diabetic patients in South Korea. Among 26,829 diabetic patients, the prevalence of depression was 22.4% ($n = 6{,}001$). The results of chi-square test revealed that diabetic depression was significantly affected by changes in instant food consumption before and after the COVID-19 pandemic, changes in delivery food consumption before and after the COVID-19 pandemic, changes in drinking before and after the COVID-19 pandemic, changes in smoking before and after the COVID-19 pandemic, satisfaction with life after the COVID-19 pandemic, gender, residing location, education level, mean monthly household income, smoking, subjective health, fear of infection due to the COVID-19 pandemic, fear of death due to the COVID-19 pandemic, concern about reproach from people around you due to the expression of COVID-19 symptoms, concern about infection of health-vulnerable people due to the COVID-19 pandemic, concern about economic loss due to the COVID-19 pandemic, changes in sleeping hours due to the COVID-19 pandemic, marital status, time

TABLE 1 General characteristics according to depression prevalence in diabetic patients after the COVID-19 pandemic, *n* (%).

Variable	Depression		*p*
	Yes (*n* =6,001)	No (*n* =20,828)	
Age			0.262
40–49 years	471 (23.6)	1,527 (76.4)	
50–59 years	1,050 (22.7)	3,571 (77.3)	
≥60 years	4,448 (22.1)	15,656 (77.9)	
Gender			<0.001
Male	2,289 (17.3)	10,926 (82.7)	
Female	3,712 (27.3)	9,902 (72.7)	
Residing location			<0.001
Urban	3,151 (24.8)	9,565 (75.2)	
Rural	2,850 (20.2)	11,263 (79.8)	
Education level			<0.001
Elementary school graduation or below	2,722 (25.3)	8,020 (74.7)	
Middle school graduation	1,077 (22.8)	3,647 (77.2)	
High school graduation	1,427 (20.2)	5,721 (80.0)	
College graduation or above	765 (18.3)	3,415 (81.7)	
Mean monthly household income			<0.001
<KRW 1.00 million	1,923 (29.7)	4,555 (70.3)	
KRW 1.00 million – 2.99 million	1,938 (22.2)	6,811 (77.8)	
KRW 3.00 million – 4.99 million	760 (19.3)	3,181 (80.7)	
≥KRW 5.00 million	533 (17.3)	2,551 (82.7)	
Smoking			<0.001
Current smoker	961 (22.9)	3,238 (77.1)	
Former smoker	1,177 (17.5)	5,538 (82.5)	
Non-smoker	3,863 (24.3)	12,049 (75.7)	
Subjective health			<0.001
Good	735 (10.9)	6,009 (89.1)	
Moderate	2,124 (18.3)	9,482 (81.7)	
Poor	3,141 (37.1)	5,336 (62.9)	
Changes in instant food consumption before and after the COVID-19 pandemic			<0.001
Increased	332 (29.6)	788 (70.4)	
Similar	1,935 (21.0)	7,291 (79.0)	
Decreased	571 (24.4)	1,768 (75.6)	
Changes in delivery food consumption before and after the COVID-19 pandemic			<0.001
Increased	484 (25.9)	1,383 (74.1)	
Similar	1,188 (20.5)	4,596 (79.5)	
Decreased	419 (24.2)	1,309 (75.8)	
Changes in drinking before and after the COVID-19 pandemic			<0.001
Increased	148 (31.2)	327 (68.8)	
Similar	1,100 (19.0)	4,700 (81.0)	
Decreased	950 (20.4)	3,705 (79.6)	
Changes in smoking before and after the COVID-19 pandemic			<0.001
Increased	149 (35.6)	269 (64.4)	
Similar	804 (20.1)	3,194 (79.9)	
Decreased	307 (20.9)	1,161 (79.1)	
Changes in the use of public transportation before and after the COVID-19 pandemic			0.155
Increased	47 (30.7)	106 (69.3)	
Similar	1,011 (24.4)	3,126 (75.6)	
Decreased	2,484 (25.3)	7,336 (74.7)	
Satisfaction with life after the COVID-19 pandemic			<0.001
Dissatisfied	1,314 (27.2)	3,523 (72.8)	
Satisfied	4,630 (21.3)	17,103 (78.7)	
Fear of infection due to the COVID-19 pandemic			<0.001
Yes	4,653 (23.5)	15,166 (76.5)	
Moderate	810 (19.6)	3,323 (80.4)	
No	529 (18.5)	2,333 (81.5)	
Fear of death due to the COVID-19 pandemic			<0.001
Yes	3,660 (24.2)	11,459 (75.8)	
Moderate	1,051 (20.8)	3,994 (79.2)	
No	1,279 (19.3)	5,352 (80.7)	

(Continued)

TABLE 1 (Continued)

Variable	Depression		p
	Yes (*n* =6,001)	No (*n* =20,828)	
Concern about reproach from people around you due to the expression of COVID-19 symptoms			<0.001
Yes	4,814 (22.9)	16,229 (77.1)	
Moderate	617 (21.0)	2,318 (79.0)	
No	553 (19.7)	2,250 (80.3)	
Concern about infection of health-vulnerable people such as infants and older adults among family members due to the COVID-19 pandemic			<0.001
Yes	5,065 (23.0)	16,954 (77.0)	
Moderate	326 (18.7)	1,416 (81.3)	
No	218 (18.0)	993 (82.0)	
Concern about economic loss due to the COVID-19 pandemic			<0.001
Yes	5,083 (23.1)	16,940 (76.9)	
Moderate	483 (20.1)	1,918 (79.9)	
No	423 (17.7)	1,964 (82.3)	
Changes in the number of meetings with people around you due to the COVID-19 pandemic			0.578
Increased	23 (25.8)	66 (74.2)	
Similar	797 (21.4)	2,931 (78.6)	
Decreased	4,545 (21.7)	16,419 (78.3)	
Changes in sleeping hours due to the COVID-19 pandemic			<0.001
Increased	626 (26.3)	1,753 (73.7)	
Similar	4,598 (20.6)	17,754 (79.4)	
Decreased	776 (37.0)	1,320 (63.0)	
Marital status			<0.001
Living with a spouse	3,449 (19.2)	14,539 (80.8)	
Not living with a spouse	2,552 (28.9)	6,289 (71.1)	
Time of first diagnosis with diabetes			0.010
<60 years old	3,394 (22.9)	11,398 (77.1)	
≥60 years old	2,588 (21.6)	9,375 (78.4)	
Current non-drug treatment for diabetes (e.g., exercise)			0.241
Yes	2,160 (22.0)	7,671 (78.0)	
No	3,839 (22.6)	13,155(77.4)	
Current diabetes drug (e.g., oral hypoglycemic drug) treatment			0.239
Yes	5,468 (22.3)	19,080 (77.7)	
No	531 (23.4)	1,743 (76.6)	
Current insulin injection treatment			<0.001
Yes	596 (30.4)	1,367 (69.6)	
No	5,400 (21.7)	19,457 (78.3)	
Number of HbA1c tests in the past year			0.078
1 or less	2,768 (22.8)	9,357 (77.2)	
2 or more	3,195 (21.9)	11,376 (78.1)	
Diabetic eye disease complication test (fundus examination) in the past year			<0.001
Yes	2,614 (23.7)	8,425 (76.3)	
No	3,358 (21.4)	12,343 (78.6)	
Diabetic renal complication test (microalbuminuria test)			0.040
Yes	2,929 (22.9)	9,880 (77.1)	
No	3,017 (21.8)	10,812 (78.2)	
Economic activity			<0.001
Yes	2,159 (16.6)	10,844 (83.4)	
No	3,841 (27.8)	9,977 (72.2)	
Awareness of own blood glucose level			<0.001
Yes	4,444 (21.6)	16,140 (78.4)	
No	1,539 (24.8)	4,658 (75.2)	
Awareness of own blood pressure			<0.001
Yes	4,260 (21.4)	15,607 (78.6)	
No	1,723 (24.9)	5,186 (75.1)	
Number of days of conducting moderate-intensity (e.g., yoga and cycling) physical activity at least 30 min per day in the past week			<0.001
None	4,679 (23.4)	15,276 (76.6)	
1–2 days	381 (20.6)	1,465 (79.4)	

(Continued)

TABLE 1 (Continued)

Variable	Depression		p
	Yes (*n* =6,001)	No (*n* =20,828)	
3 days or more	937 (18.7)	4,068 (81.3)	
Number of days of walking at least 30 min per day in the past week			<0.001
None	1,928 (25.8)	5,553 (74.2)	
1–2 days	670 (24.5)	2,069 (75.5)	
3 days or more	3,402 (20.5)	13,198 (79.5)	
Number of people you can ask for help in a disaster situation such as COVID-19 infection			<0.001
None	1,724 (27.6)	4,517 (72.4)	
1–2 people	2,637 (23.0)	8,817 (77.0)	
3 people or more	1,633 (18.0)	7,453 (82.0)	
Diagnosis with hypertension			<0.001
Yes	3,910 (23.6)	12,648 (76.4)	
No	2,090 (20.4)	8,177 (79.6)	

of first diagnosed with diabetes, current insulin injection treatment, diabetic eye disease complication test in the past year, diabetic renal complication test, economic activity, awareness of own blood glucose level, awareness of own blood pressure, number of days of conducting moderate-intensity physical activity at least 30 min per day in the past week, number of days of walking at least 30 min per day in the past week, the number of people you can ask for help in a disaster situation such as COVID-19 infection, and diagnosis with hypertension ($p < 0.05$).

3.2. Predictors of depression in diabetic patients living in South Korean community

This study calculated the importance of factors related to depression in diabetic patients living in South Korean community using CatBoost to find that the top nine variables with high importance were gender, smoking status, changes in drinking before and after the COVID-19 pandemic, changes in smoking before and after the COVID-19 pandemic, subjective health, concern about economic loss due to the COVID-19 pandemic, changes in sleeping hours due to the COVID-19 pandemic, economic activity, and the number of people you can ask for help in a disaster situation such as COVID-19 infection.

Table 2 presents the results of logistic regression analysis for predicting depression in diabetic patients living in the South Korean community the top nine variables with high impact on model output in CatBoost. The analysis results of the adjusted model for predicting depression in South Korean diabetic patients showed that female (AOR = 1.78, 95% CI = 1.68, 1.89), current smoker (AOR = 1.39, 95% CI = 1.26, 1.53), concern about economic loss due to the COVID-19 pandemic (moderate: AOR = 1.19; yes: AOR = 1.39), changes in sleeping hours due to the COVID-19 pandemic (similar: AOR = 0.60; decreased: AOR = 1.38), changes in drinking before and after the COVID-19 pandemic (similar: AOR = 1.76; increased: AOR = 1.93), changes in smoking before and after the COVID-19 pandemic (similar: AOR = 2.09; increased: AOR = 2.20), subjective health (moderate: AOR = 2.62; poor: AOR = 4.81), and the number of people you can ask for help in a disaster situation such as COVID-19 infection (1–2 persons: AOR = 1.27; none: AOR = 1.74) were independent factors of depression in diabetic patients ($p < 0.05$).

3.3. Development and validation of depression predictive nomogram for diabetic patients living in the South Korean community

Figure 1 presents the depression predictive nomogram for diabetic patients living in the South Korean community. The nomograph (Figure 1) analyzed the high-risk group for depression in diabetic patients and predicted that female diabetic patients who had fewer sleeping hours after the COVID-19 pandemic, increased the frequency of smoking and drinking increased than before the pandemic, concerned about economic loss due to the COVID-19 pandemic, had no one to ask for help, and perceived subjective health as poor had an 88% predictive possibility of depression.

This study examined the predictive performance of the developed depression predictive nomogram for diabetic patients living in South Korea using calibration plot (Figure 2), AUC, and accuracy (Figure 3). This study compared the prediction probability and observation probability of the diabetic patient group with depression with those of the diabetic patient group without depression using calibration plot and chi-square test (Figure 2). The prediction probability and observation probability were not significantly different ($p < 0.05$). The results of 10-fold cross validation showed that AUC, general accuracy, precision, recall, and F1-score were 0.704, 0.780, 0.735, 0.780, and 0.712, respectively.

4. Discussion

This study identified the prevalence of depression among diabetic patients living in South Korean local communities using national survey data conducted after the COVID-19 pandemic and found that 22.4% of the subjects were diabetic patients with depression. The prevalence of depression among diabetic patients living in South Korean local communities was approximately twice the prevalence of depression among healthy people (12%) during the period (18). Although it cannot be directly compared with the results of this study, the meta-analysis of Anderson et al. (19), conducted before the COVID-19 pandemic, reported that depression in diabetic patients (28.5%) was 1.5 times higher than that in the general population (16.2%). Even though diabetic patients in local communities are at

TABLE 2 Predictors of depression in diabetic patients living in the South Korean community: AOR and 95% CI.

Variables	AOR	95%CI	*p*
Changes in drinking before and after the COVID-19 pandemic			
Increased	1.93	1.57, 2.37	<0.001
Similar	1.76	1.43, 2.17	<0.001
Decreased (ref)	1	1	
Changes in smoking before and after the COVID-19 pandemic			
Increased	2.20	1.77, 2.72	<0.001
Similar	2.09	1.65, 2.65	<0.001
Decreased (ref)	1	1	
Gender			
Male (ref)	1	1	
Female	1.78	1.68, 1.89	<0.001
Smoking			
Current smoker	1.39	1.26, 1.53	<0.001
Former smoker	0.92	0.85, 1.00	0.060
Non-smoker (ref)	1	1	
Subjective health			
Good (ref)	1	1	
Moderate	2.62	2.46, 2.80	<0.001
Poor	4.81	4.40, 5.25	<0.001
Concern about economic loss due to the COVID-19 pandemic			
Yes	1.39	1.24, 1.55	0.001
Moderate	1.19	1.07, 1.32	<0.001
No (ref)	1	1	
Changes in sleeping hours due to the COVID-19 pandemic			
Increased (ref)	1	1	
Similar	0.60	0.53, 0.69	<0.001
Decreased	1.38	1.25, 1.51	<0.001
Economic activity			
Yes (ref)	1	1	
No	1.93	1.82, 2.05	<0.001
Number of people you can ask for help in a disaster situation such as COVID-19 infection			
None	1.74	1.61, 1.88	0.001
1–2 people	1.27	1.18, 1.36	0.001
3 people or more (ref)	1	1	

high risk of depression, not enough active attention has been given to their emotional aspects. It has been reported that when depression accompanies diabetes, medical costs increase because glycemic control deteriorates, the incidence of chronic complications increases, and mortality rises (20–22). Furthermore, if a diabetic patient cannot properly perform health behaviors due to depression, it can adversely affect the long-term course of diabetes, such as the occurrence of chronic complications, as well as glycemic control (23). Consequently, in order to efficiently screen depression in diabetic patients at an early stage, studies need to identify the risk factors for depression.

The results of this study confirmed that gender, subjective health, increased health risk behaviors such as drinking and smoking, decreased sleeping hours, and the number of people whom you could seek help in a disaster situation such as COVID-19 infection were independent risk factors of depression. These results agreed with the results of previous studies (24–27). Female sex, marital status, childhood adversity, and social deprivation are general population risk factors for depression that also apply to people with diabetes (28).

Gender is known to be a major factor influencing diabetic patients. Adriaanse et al. (24) analyzed depressive symptoms in type 2 diabetic patients and reported that the prevalence of depression was significantly higher in women (15%) than in men (9.1%), which concurred with the results of this study. Moreover, a decrease in

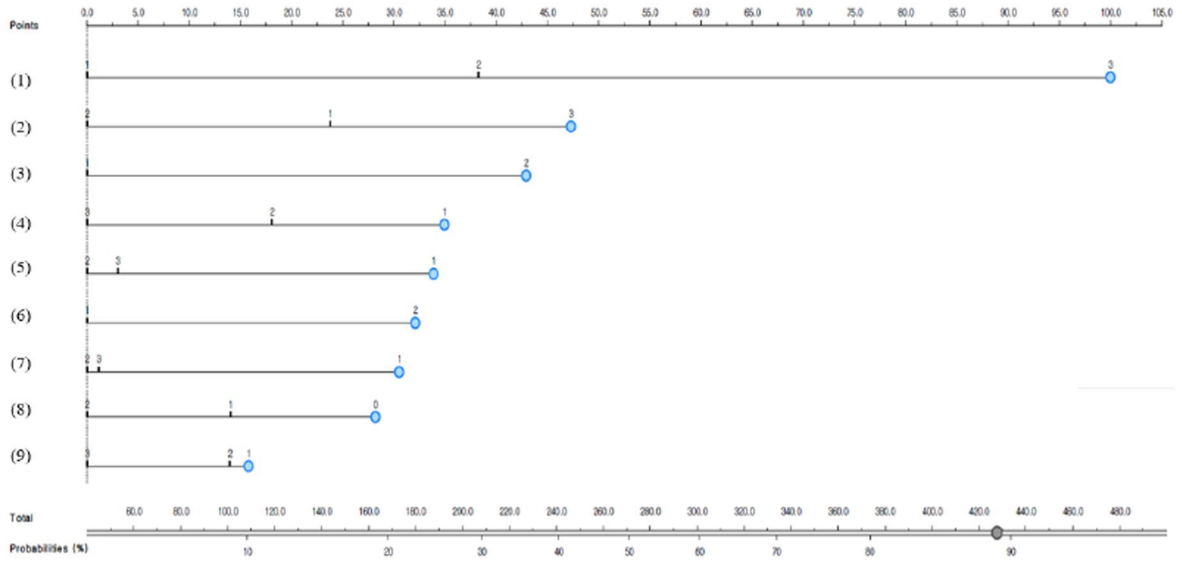


FIGURE 1
Depression predictive nomogram for South Korean diabetic patients; (1) Subjective health: 1, good; 2, moderate; or 3, poor; (2) changes in sleeping hours after the COVID-19 pandemic: 1, increased; 2, similar; or 3, decreased; (3) Gender: 1, male, or 2, female; (4) smoking status: 1, current smoker; 2, former smoker; or 3, non-smoker; (5) changes in drinking after the COVID-19 pandemic: 1, increased; 2, similar; 3, decreased; (6) economic activity: 1, yes; or 2, no; (7) changes in smoking after the COVID-19 pandemic: 1, increased; 2, similar; or 3, decreased; (8) people whom you can ask for help: 0, 0 people; 1, 1–2 people; or 2, 3 or more people; and (9) concern about economic loss due to the COVID-19 pandemic: 1, yes; 2, moderate; or 3, no.

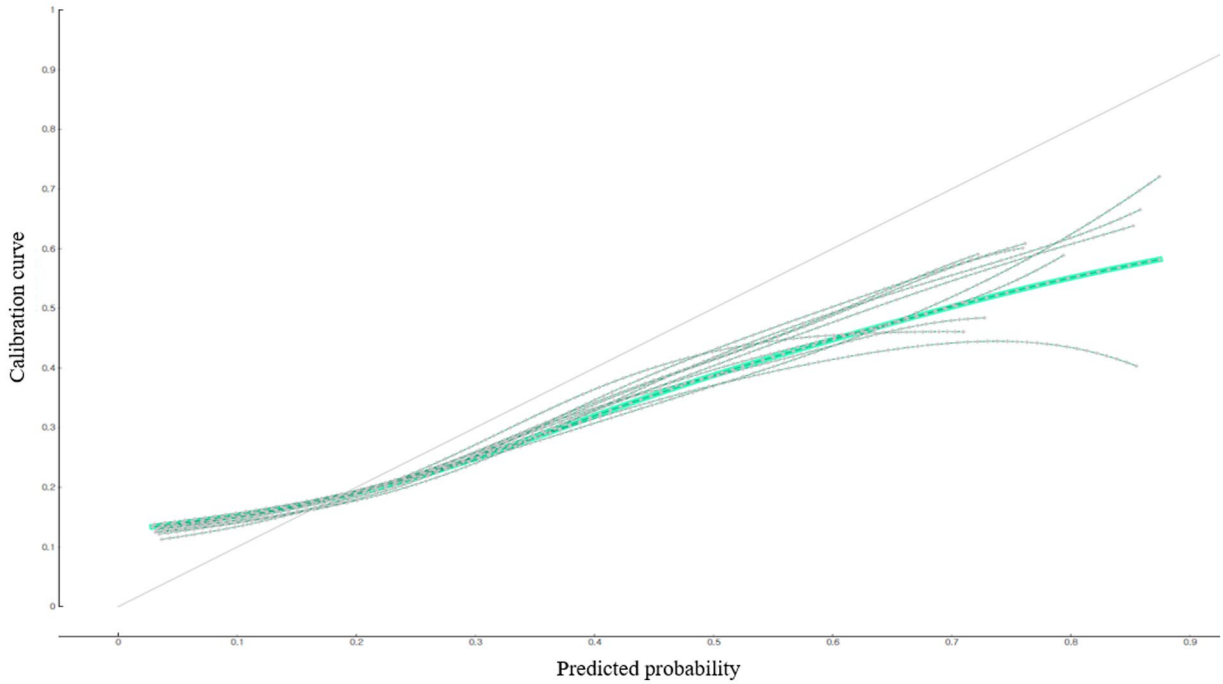


FIGURE 2
Predictive performance of the depression predictive nomogram for South Korean diabetic patients: calibration plot.

sleeping hours has been reported as a significant predictor of diabetic depression. Ghosh et al. (25) showed that a quarter of diabetic patients with depression experienced a decrease in sleeping hours. Particularly, the number of people subjects could ask for help in a disaster situation was a key factor related to diabetic depression. Social support is known to be another major risk factor for diabetic depression. Pibernik-Okanovic (26) found that diabetic patients who felt a lack of social support had a higher risk of developing depression, which supported the results of this study. Therefore, it is necessary to develop a psychological support program for diabetic patients in the

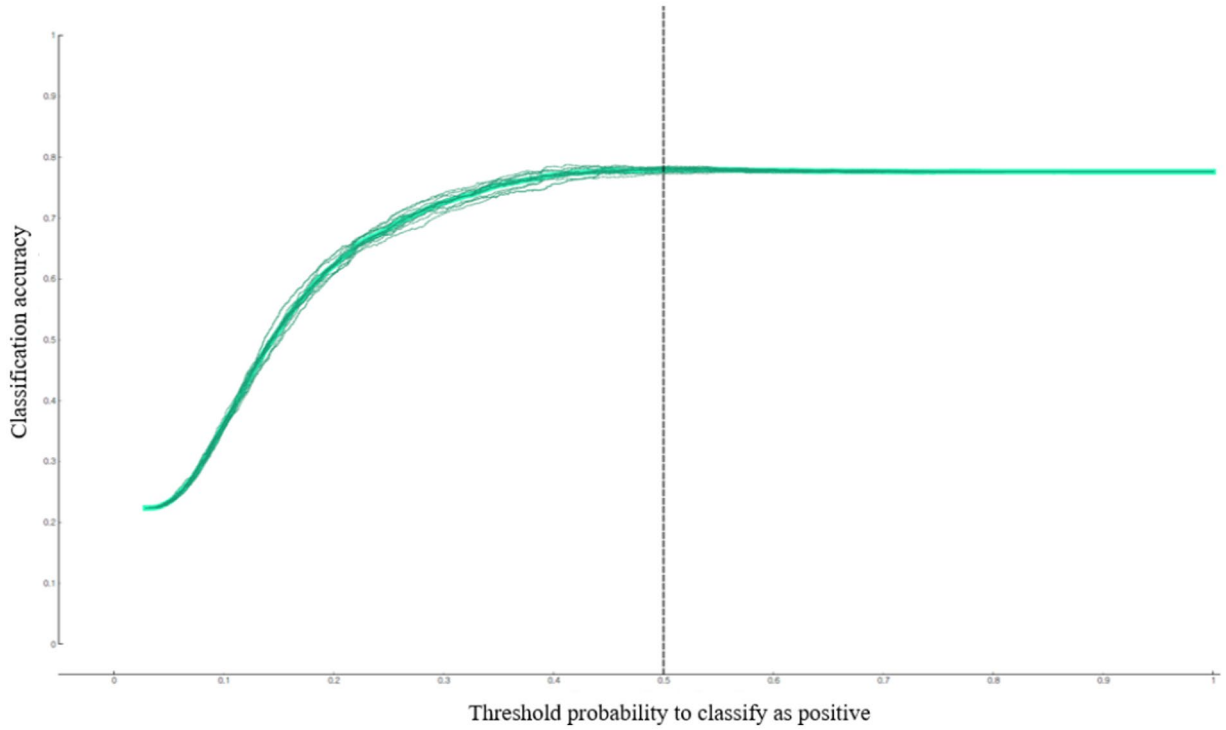


FIGURE 3
Predictive performance of the depression predictive nomogram for South Korean diabetic patients: accuracy.

community to increase their ability to cope with depression caused by social distancing in the era of COVID-19 and limited social contact and build a system that can continuously provide medical and social support for diabetic patients without sufficient social support to prevent depression in diabetic patients.

In this study, the number of diabetic complication tests in the past year, awareness of blood glucose levels, and diabetic treatment methods (e.g., oral hypoglycemic drug and insulin injection treatment) were not related to depression in diabetic patients, which did not agree with the results of previous studies (1, 29, 30). Numerous studies (1, 29) have proven the relationship between glycemic control and depression in diabetic patients. When glycemic control was poorer, depression symptoms were more severe (29). It is backed by the results that a higher level of glycated hemoglobin decreases the effects of antidepressants (1). It is speculated that the awareness of blood glucose levels or the number of diabetes complication tests based on a survey alone was not enough to directly identify the blood glucose management level of diabetic patients. Therefore, although previous studies (1, 29) reported that managing blood glucose is related to the depression of diabetic patients, the relationship was not significant in this study.

The mechanism underlying the high risk of depression in diabetic patients has not been clearly understood. There are some possible explanations: depression may induce insulin resistance by stimulating the secretion of insulin-antagonizing hormones (e.g., catecholamines, glucocorticoids, growth hormones, and glucagon) and inflammatory cytokines, or it can contribute to the development of diabetes by causing dysfunction of pancreatic beta cells. Moreover, diabetes may cause depression in association with inadequate glycemic control, the development of chronic complications due to diabetes, and a decline in socioeconomic status. However, since depression is caused by multiple factors rather than a single factor, future studies need to identify the relationship between blood glucose management and depression using clinical test data such as HbA1c level in addition to sociodemographic and psychological characteristics.

Another finding of this study was that the results of this study showed that "female diabetic patients who had fewer sleeping hours after the COVID-19 pandemic, increased the frequency of smoking and drinking increased than before the pandemic, concerned about economic loss due to the COVID-19 pandemic, had no one to ask for help, and perceived subjective health as poor had an 88% predictive possibility of depression," which was high. Since multiple risk factors for diabetic depression have not been clearly identified, it is needed to carry out future studies on multiple risk factors for diabetic depression based on large-scale cohort data. It is also necessary to continuously monitor depression in terms of primary care for diabetic patients with these multiple risk factors.

In the United States, the Centers for Disease Control and Prevention regularly conducts regular monitoring of comorbidities in diabetic patients and runs a chronic disease prevention program (31). On the other hand, South Korea lacks a systematic monitoring system for diabetic depression management, and previous studies mainly examined multi-center registry data for diabetic depression (32, 33). Particularly, in South Korea, mental health management and education such as depression for diabetic patients are mainly carried out in general hospitals (32). Considering the fact that general hospitals are playing a critical role in the emergency medical response system in a disaster situation such as the COVID-19 pandemic (34), it will be required to establish a systematic depression examination and monitoring system centered on primary care in the future for the sustainability mental health management of diabetic patients.

The strength of this study was that it identified a high-risk group for depression in diabetic patients using national survey data conducted after the COVID-19 pandemic and provided baseline data for preventing depression in diabetic patients. This study had several limitations. First, since this study analyzed secondary data by analyzing epidemiological data (survey data), clinical indicators such as insulin-antagonizing hormones and genes related to depression were not included. Second, the in-person survey may underestimate health risk behaviors such as smoking and drinking. Therefore, future studies need to reduce the possibility of recall bias including medical records in order to identify factors related to diabetic depression. Third, the Community Health Survey, the source data, did not survey the duration of diabetes. Future studies need to investigate the duration of diabetes and duration of diabetic complications additionally to develop a diabetes depression predictive model with higher predictive performance. Fourth, since this study was a cross-sectional study, even if risk factors for diabetic depression were identified in this study, it could not be interpreted as a causal relationship based on temporal precedence.

5. Conclusion

It is necessary to identify the high-risk groups for diabetes and depression at an early stage while considering multiple risk factors and provide a tailored psychological support system at the primary medical level to improve their mental health. Additionally, it is important to establish a system that can systematically monitor the high-risk groups for diabetes and depression, even in long-term disaster situations (e.g., pandemic), at the community level. Furthermore, additional longitudinal studies are needed to confirm the causal relationship between factors related to diabetic depression identified in this study.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://chs.kdca.go.kr/chs/index.do.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board (or Ethics Committee) of Korea Disease Control and Prevention Agency (protocol code: 117075 and date: 2021.07.01) and the study was conducted according to the guidelines of the Declaration of Helsinki. The patients/participants provided their written informed consent to participate in this study.

Author contributions

HB designed the manuscript, was involved in study data interpretation, preformed the statistical analysis, and assisted with writing the manuscript.

Funding

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07041091 and 2021S1A5A8062526).

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.

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.

References

1. Lustman PJ, Anderson RJ, Freedland KE, De Groot M, Carney RM, Clouse RE. Depression and poor glycemic control. A meta-analytic review of the literature. *Diabetes Care*. (2000) 23:934–42. doi: 10.2337/diacare.23.7.934

2. Knol MJ, Twisk JW, Beekman AT, Heine RJ, Snoek FJ, Pouwer F. Depression as a risk factor for the onset of type 2 diabetes mellitus. A meta-analysis. *Diabetologia*. (2006) 49:837–45. doi: 10.1007/s00125-006-0159-x

3. Fisher L, Chesla CA, Mullan JT, Skaff MM, Kanter RA. Contributors to depression in Latino and European-American patients with type 2 diabetes. *Diabetes Care*. (2001) 24:1751–7. doi: 10.2337/diacare.24.10.1751

4. Katon W, Fan M-Y, Unutzer J, Taylor J, Pincus H, Schoenbaum M. Depression and diabetes: a potentially lethal combination. *J Gen Intern Med*. (2008) 23:1571–5. doi: 10.1007/s11606-008-0731-9

5. Oh JY. Risk factors for depression in patients with diabetes mellitus. *Korean Clinic Diab*. (2008) 9:294–6. doi: 10.4093/kcd.2008.9.4.294

6. Byeon H. Developing a predictive model for depressive disorders using stacking ensemble and naive Bayesian nomogram: using samples representing South Korea. *Front Psych*. (2022) 12:12. doi: 10.3389/fpsyt.2021.773290

7. Toraih EA, Fawzy MS, Hussein MH, el-Labban MM, Ruiz EML, Attia AA, et al. MicroRNA-based risk score for predicting tumor progression following radioactive iodine ablation in well-differentiated thyroid cancer patients: a propensity-score matched analysis. *Cancers*. (2021) 13:4649. doi: 10.3390/cancers13184649

8. Karaismailoglu E, Karaismailoglu S. Two novel nomograms for predicting the risk of hospitalization or mortality due to COVID-19 by the naïve Bayesian classifier method. *J Med Virol*. (2021) 93:3194–201. doi: 10.1002/jmv.26890

9. Yi M, Meric-Bernstam F, Kuerer HM, Mittendorf EA, Bedrosian I, Lucci A, et al. Evaluation of a breast cancer nomogram for predicting risk of ipsilateral breast tumor recurrences in patients with ductal carcinoma in situ after local excision. *J Clin Oncol*. (2012) 30:600–7. doi: 10.1200/JCO.2011.36.4976

10. Byeon H. Influence of the COVID-19 pandemic on the subjective life satisfaction of South Korean adults: Bayesian nomogram approach. *Diagnostics*. (2022) 12:3. doi: 10.3390/diagnostics12030761

11. Park SJ, Choi HR, Choi JH, Kim KW, Hong JP. Reliability and validity of the Korean version of the patient health Questionnaire-9 (PHQ-9). *Anxiety Mood*. (2010) 6:119–24.

12. Spitzer RL, Kroenke K, Wiliams JBthe Patient Health Questionnaire Primary Care Study Group. Patient health questionnaire primary care study group. Valiation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *JAMA*. (1999) 282:1737–44. doi: 10.1001/jama.282.18.1737

13. Han CS, Jo SA, Kwak JH, Pae CU, Steffens D, Jo IH, et al. Validation of the patient health questionnaire-9 Korean version in the elderly population: the Ansan geriatric study. *Compr Psychiatry*. (2008) 49:218–23. doi: 10.1016/j.comppsych.2007.08.006

14. Lee H, Lee JA, Brar JS, Rush EB, Jolley CJ. Physical activity and depressive symptoms in older adults. *Geriatr Nurs*. (2014) 35:37–41. doi: 10.1016/j.gerinurse.2013.09.005

15. Werneck AO, Stubbs B, Szwarcwald CL, Silva DR. Independent relationships between different domains of physical activity and depressive symptoms among 60,202 Brazilian adults. *Gen Hosp Psychiatr*. (2020) 64:26–32. doi: 10.1016/j.genhosppsych.2020.01.007

16. Choi HS, Choi JH, Park KH, Joo KJ, Ga H, Koh HJ, et al. Standardization of the Korean version of patient health Questionnaire-9 as a screening instrument for major depressive disorder. *Korean J Fam Med*. (2007) 28:114–9.

17. Hancock JT, Khoshgoftaar TM. CatBoost for big data: an interdisciplinary review. *J Big Data*. (2020) 7:94–45. doi: 10.1186/s40537-020-00369-8

18. Korea Disease Control and Prevention Agency. *Community health survey 2020*. Osong: Korea Disease Control and Prevention Agency (2021).

19. Anderson RJ, Freedland KE, Clouse RE, Lustman PJ. The prevalence of comorbid depression in adults with diabetes: a meta-analysis. *Diabetes Care*. (2001) 24:1069–78. doi: 10.2337/diacare.24.6.1069

20. Roy T, Lloyd CE. Epidemiology of depression and diabetes: a systematic review. *J Affect Disord*. (2012) 142:S8–S21. doi: 10.1016/S0165-0327(12)70004-6

21. Lustman PJ, Griffith LS, Gavard JA, Clouse RE. Depression in adults with diabetes. *Diabetes Care*. (1992) 15:1631–9. doi: 10.2337/diacare.15.11.1631

22. Moulton CD, Pickup JC, Ismail K. The link between depression and diabetes: the search for shared mechanisms. *Lancet Diabetes Endocrinol*. (2015) 3:461–71. doi: 10.1016/S2213-8587(15)00134-5

23. Holt RI, De Groot M, Golden SH. Diabetes and depression. *Curr Diab Rep*. (2014) 14:1–9. doi: 10.1007/s11892-014-0491-3

24. Adriaanse MC, Dekker JM, Heine RJ, Snoek FJ, Beekman AJ, Stehouwer CD, et al. Symptoms of depression in people with impaired glucose metabolism or type 2 diabetes mellitus: the Hoorn study. *Diabet Med*. (2008) 25:843–9. doi: 10.1111/j.1464-5491.2008.02464.x

25. Ghosh A, Arora B, Gupta R, Anoop S, Misra A. Effects of nationwide lockdown during COVID-19 epidemic on lifestyle and other medical issues of patients with type 2 diabetes in North India. *Diabetes Metab Syndr*. (2020) 14:917–20. doi: 10.1016/j.dsx.2020.05.044

26. Pibernik-Okanovic M, Peros K, Szabo S, Begic D, Metelko Z. Depression in Croatian type 2 diabetic patients: prevalence and risk factors. A Croatian survey from the European depression in diabetes (EDID) research consortium. *Diabet Med*. (2005) 22:942–5. doi: 10.1111/j.1464-5491.2005.01528.x

27. Renn BN, Feliciano L, Segal DL. The bidirectional relationship of depression and diabetes: a systematic review. *Clin Psychol Rev*. (2011) 31:1239–46. doi: 10.1016/j.cpr.2011.08.001

28. Holt RI, de Groot M, Golden SH. Diabetes and depression. *Curr Diabetes Rep*. (2014) 14:491. doi: 10.1007/s11892-014-0491-3

29. Testa MA, Simonson DC. Health economic benefits and quality of life during improved glycemic control in patients with type 2 diabetes mellitus: a randomized, controlled double-blind trial. *JAMA*. (1998) 280:1490–6. doi: 10.1001/jama.280.17.1490

30. Egede LE, Ellis C. Diabetes and depression: global perspectives. *Diabetes Res Clin Pract*. (2010) 87:302–12. doi: 10.1016/j.diabres.2010.01.024

31. Albright AL, Gregg EW. Preventing type 2 diabetes in communities across the U.S.: the National Diabetes Prevention Program. *Am J Prev Med*. (2013) 44:S346–51. doi: 10.1016/j.amepre.2012.12.009

32. Park SW, Kim DJ, Min KW, Baik SH, Choi KM, Park IB, et al. Current status of diabetes management in Korea using National Health Insurance Database. *J Korean Diabetes Assoc*. (2007) 31:362–7. doi: 10.4093/jkda.2007.31.4.362

33. Lee YS. The current status of type 2 diabetes management at a university hospital. *Korean Diabetes J*. (2009) 33:241–50. doi: 10.4093/kdj.2009.33.3.241

34. Selvin E, Juraschek SP. Diabetes epidemiology in the COVID-19 pandemic. *Diabetes Care*. (2020) 43:1690–4. doi: 10.2337/dc20-1295

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 26 July 2023
DOI 10.3389/fendo.2023.1234256

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Gisha Mohan,
Centinela Hospital Medical Center, United States
Karthik Gangu,
University of Kansas Medical Center, United States
Sayak Roy,
King's College Hospital NHS Foundation Trust, United Kingdom

*CORRESPONDENCE
Asma Gorchane
dargorchene@gmail.com

RECEIVED 04 June 2023
ACCEPTED 30 June 2023
PUBLISHED 26 July 2023

CITATION
Gorchane A, Ach T, Sahli J, Abdelkrim AB, Mallouli M, Bellazreg F, Hachfi W, Chaieb MC and Ach K (2023) Uncovering the alarming rise of diabetic ketoacidosis during COVID-19 pandemic: a pioneer African study and review of literature. *Front. Endocrinol.* 14:1234256. doi: 10.3389/fendo.2023.1234256

Uncovering the alarming rise of diabetic ketoacidosis during COVID-19 pandemic: a pioneer African study and review of literature

Asma Gorchane [1,2]*, Taieb Ach [1,2,3], Jihene Sahli [2,4], Asma Ben Abdelkrim [1,2], Manel Mallouli [2,4], Foued Bellazreg [2,5], Wissem Hachfi [2,5], Molka Chadli Chaieb [1,2] and Koussay Ach [1,2]

[1]Department of Endocrinology, University Hospital of Farhat Hached, Sousse, Tunisia, [2]University of Sousse, Faculty of Medicine of Sousse, Sousse, Tunisia, [3]Laboratory of Exercice Physiology and Pathophysiology, Tunis, Tunisia, [4]Department of Community Medicine, University Hospital of Farhat Hached, Sousse, Tunisia, [5]Department of Infectious Diseases, University Hospital of Farhat Hached, Sousse, Tunisia

Introduction: Reports around the world indicate that COVID-19 pandemic may be contributing to an increase in the incidence of new onset diabetic ketoacidosis (DKA). This has yet to be studied in Africa. We aimed to compare the incidence trend of new onset DKA before and during the COVID-19 pandemic, with a focus on the type of diabetes mellitus (DM).Materials and methodsThis was a cross sectional analytical study, over a 4-year period, between March 2018 until February 2022 conducted in the referral center: diabetology department of university hospital Farhat Hached Sousse, Tunisia. The study population included patients hospitalized for new onset DKA divided in two groups: G1: before COVID-19 pandemic and G2: during COVID-19 pandemic. Patients younger than 14, new onset DM not presenting with DKA, other types of diabetes (monogenic, secondary or pancreatic diabetes) were not included. A statistical analysis of the monthly incidence trend was conducted using the Jointpoint software providing the average monthly percentage of change (AMPC).

Results: a total of 340 patients were included:137 registered before the pandemic and 203 during the pandemic, representing a 48.17% increase. The mean monthly incidence of new onset DKA during COVID-19 pandemic was statistically higher than that before COVID-19 pandemic (8.42 $\pm$ 4.87 vs 5.75 $\pm$ 4.29 DKA per month) (p=0.049). The temporal trend of DKA during the 4-year study showed a significant upward trend with a change in AMPC of +0.2% (p=0.037). The incidence of type 1 diabetes (T1D) and type 2 diabetes (T2D) increased by 50% and 44% respectively during COVID-19 pandemic. Anti-glutamic acid decarboxylase (anti-GAD) antibodies' titers significantly increased in G2 compared with G1 (median of 330 [Q1–Q3]=[58.5–1795]vs 92.5[Q1–Q3]=[22.5–1074] respectively)(p=0.021).

Discussion: The incidence trend of DKA showed an increase during the COVID-19 pandemic along with an increase of T1D and T2D implying that the pandemic may have been the underlying factor of this upward trend.

KEYWORDS

diabetic ketoacidosis, COVID-19, type 1 diabetes, type 2 diabetes, epidemiology

Introduction

The COVID-19 pandemic has had a profound impact on global health systems, overwhelming hospitals and healthcare workers with an unprecedented influx of patients (1). Africa was not spared from the rapid propagation of the disease adding to the burden of a precarious health system already plagued by endemic diseases (2). Reports from various regions of the world have suggested a possible association between COVID-19 and the development of new-onset diabetic ketoacidosis (DKA) (3).

DKA is a potentially life-threatening complication of diabetes mellitus (DM) affecting both type 1 (T1D) and type 2 diabetes (T2D) (4). Its incidence has been increasing over the last few decades, which has been attributed to a combination of factors, mainly the overall rising prevalence of DM (5). Furthermore, recent studies have suggested an additional increase in DKA during the COVID-19 pandemic, raising speculations about a potential involvement of COVID-19, whether it is directly or indirectly, in this upward trend (6–8).

On the one hand, the pandemic may have caused delays in diagnosis for fear of contracting the virus resulting in higher DKA cases (9). On the other hand, SARS-CoV-2 may have caused direct damage to pancreatic cells (10), triggered auto-immunity (11) or promoted insulin resistance (12). This underscores the relevance of investigating COVID-19's role in new onset DKA.

Our study sought to compare epidemiological aspects of new onset DKA including incidence of DKA specifically focusing on T1D and T2D, before and during COVID-19 pandemic within the context of an African country, aiming to provide valuable insights into the complex interplay between COVID-19 and new onset DM.

Materials and methods

Study design and setting

We conducted a cross-sectional descriptive and analytical study carried out in the Diabetology & Endocrinology department of Farhat Hached University Hospital of Sousse.

Inclusion criteria

It has included all the patients who had been hospitalized for new onset DKA over a period of 4 years between the year 2018 and 2022, that is before and during COVID-19 pandemic.

Abbreviations: COVID-19, coronavirus disease 2019; DKA, diabetic ketoacidosis; DM, diabetes mellitus; AMPC, average monthly change; T1D, type 1 diabetes; T2D, type 2 diabetes; anti-GAD, anti-glutamic acid decarboxylase; ELISA, Enzyme-Linked Immunosorbent Assay; ONMNE, Observatoire national des maladies nouvelles et émergentes; MPC, monthly percent change, MPC: monthly percent change; anti-IA2, anti-islet cell antigen; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; ACE2, angiotensin-converting enzyme, continuous glucose monitoring (CGM).

Since the second of March 2020 was the mark of patient zero in Tunisia declaring the start of the pandemic in Tunisia (13), the population was automatically divided in two groups:

- Group1 (G1): patients hospitalized before COVID-19 pandemic since the first of March of 2018 until first of March 2020.
- Group2 (G2): patients hospitalized during COVID-19 pandemic since second of March 2020 until 28th February 2022.

Non-inclusion criteria

Patients younger than 16 years old, patients with known DM, patients with new onset DM not presenting with DKA were not included in this study.

Exclusion criteria

All other types of diabetes (high presumption of monogenic diabetes, gestational, secondary or pancreatic diabetes) were excluded from this study.

We note that our study did not include genetic testing to specifically exclude monogenic diabetes. Instead, we excluded patients with a strong likelihood of monogenic diabetes based on clinical assessments and criteria (14).

Data collection

Data collection was conducted retrospectively by filling in a standardized information sheet collected by consulting medical records coded as "new onset DKA" for patients who were hospitalized two years before and during the COVID-19 pandemic.

Variables

Epidemiological data was compared between the two groups including incidence, the date of admission, month and season of discovery, duration of polyuria and polydipsia.

Islet antibodies and c-peptide were measured using the enzyme-linked immunosorbent assay (ELISA) in all patients at admission and their respective titers were compared between the two groups. In light of clinical and biological data such as islet antibodies and C-peptide, the new onset DM was eventually classified as T1D or T2D. The types of DM were compared between the two groups. Precipitating factors were compared between the two groups. It can include (15):

1-cardiovascular factors: myocardial infarction, stroke.

2-infections (urinary, pulmonary, COVID-19, otolaryngological, cutaneous, profound).

3-drugs that affect carbohydrate metabolism, such as corticosteroids.

4-psychological factors.

5-Excessive food intake.

The incidence of new onset DKA during COVID-19 pandemic was analyzed according to the pandemic waves which were specified according to National Observatory of New and Emerging Diseases (ONMNE) (16) as follows

1-First wave: from 17/08/2020 to 13/12/2020.

2-Second wave: from 14/12/2020 to 21/03/2021.

3-Third wave: 22/03/2021 to 16/05/2021.

4-Fourth wave: 17/05/2021 to 14/11/2021.

5-The fifth wave: 15/11/2021 to 28/02/2022.

The average monthly admission of DKA during COVID-19 according to the different waves were calculated and compared.

Time for DKA resolution expressed in hours, cumulative insulin dose as well as mean weight-based insulin at discharge were compared before and during COVID-19 pandemic.

Statistical analysis

The analysis of the incidence trend variations of DKA was performed using the JOINPOINT Version 116 4.5.0.1 software. Monthly data was used, and the software provided the monthly percentage of change (MPC and AMPC: Monthly Percent Change and Average Monthly Percent Change) with a 95% confidence interval. Data were analyzed using SPSS 26.0 software. Average monthly admission of DKA during the different pandemic waves were compared using ANOVA test. Quantitative variables were presented by means and standard deviation (SD) or median and quartiles [Q1–Q3] according to the normality of the distribution which was tested using Kolmogorov-Smirnov test. Our study respected all standards in ethics in research. The anonymity and data confidentiality of the patients' data were respected. We obtained approval from the Ethical Committee of University of Medicine of Sousse for our study, with the assigned number 4935/2023.

Results

A total of 340 patients were included. G1 counted 137 patients while G2 counted 203 patients.

The number of DKA cases witnessed an increase of 48.17% over a similar time interval.

The mean monthly incidence of DKA before COVID-19 pandemic (G1) was statistically different from that observed during COVID-19 pandemic (G2) with a mean of 5.75 ± 4.29 DKA per month in G1 vs 8.42 ± 4.87 DKA per month (p=0.049).

The study of the temporal trend of hospital cases of DKA between March 2018 and February 2022 showed a significant upward trend with a change in the average monthly percent change (AMPC) of +0.2%, with p=0.037 (Figure 1, Table 1).

The incidence of DKA according to pandemic waves during COVID-19 pandemic shows a decrease during the third wave with a re-increase during the fourth and fifth waves (Figure 2).

The average monthly admissions during each pandemic wave were as follows: 9.9 DKA/month for the first wave, 17 DKA/month for the second wave, 2.75 DKA/month for the third wave, 6 DKA/month for the fourth wave, and 12.5 DKA/month for the fifth wave. DKA admissions decreased during the third wave and increased during the fifth wave but the difference was not statistically significant with p=0.09.

The incidence of T1D increased by 50% during COVID-19 pandemic as 54 (38.42%) out of 137 patients were T1D in G1 vs 81 (40.30%) out of 201 patients in G2 while the incidence of T2D increased by 44% as 83 (60.58%) out of 137 patients were T2D in G1 vs 120 (59.70%) out of 201 patients in G2. However, the distribution of T1D and T2D frequencies was comparable between the two groups (p=0.871) (Figure 3).

Anti-glutamic acid decarboxylase (Anti-GAD) antibodies titers significantly increased during the pandemic period compared with the pre-pandemic period with a median value of 92.5 [Q1–Q3]=[22.5–1074] in G1 vs 330 [Q1–Q3]=[58.5–1795] in G2 (**p=0.021**) (Figure 4).

Anti-islet cell antigen (Anti-IA2) antibodies titers significantly increased as well during the pandemic period compared with the pre-pandemic period with a median value of 0 [Q1–Q3]=[0–104.75] in G1 vs 93 [Q1–Q3]=[0–3571] in G2 (p=**0.009**) (Figure 4).

The main precipitating factors of DKA before and during COVID-19 pandemic were comparable between the two groups: infections (29.10% in G1 vs 28.64% in G2), excessive food intake (2.24% in G1 vs 6.53% in G2), psychological stress (22.39% in G1 vs 31.16% in G2), no precipitating factor was found in 40.3% vs 31.66% in G2.

However, using subgroup analysis, the distribution of DKA precipitating factors in T2D differed significantly between the two groups (**p=0.003**) with a notable increase of stress (29.7% in G2 vs 17.1% in G1, AR=2) as well as excessive food intake notably hypertonic drinks (10.2% in G2 vs 3.7% in G1) at the expense of corticosteroid use as its accountability in precipitating DKA significantly decreased in G2 (2.5% in G2 vs 9.8% in G1, AR=-2.2) (Figure 5).

Even though the overall accountability of infectious causes did not differ between the two groups (36.4% in G2 vs 30.5% in G1, AR=0.9), infection sites had a significantly different distribution between the two groups ($p<10^{-3}$) with COVID-19 becoming the first infectious precipitating factor of DKA in G2 (AR=4.2).

Time for DKA resolution was considerably higher in G2 14 ± 6.2 hours in G1 vs 16.5 ± 1.4 hours in G2, p=0.022

We found no significant difference in cumulative insulin dose required for DKA resolution in G1 78 ± 6.5 UI vs 82.5 ± 77 UI in G2, p=0.142.

Insulin dose at discharge was also comparable between the two groups (0.40 [0.29-047] UI/Kg/day in G1 vs 0.41 [0.28-0.50] in G2, p=0.895)

Discussion

Since 2020, the COVID-19 pandemic has changed the world as we know it. It has led to changes in epidemiological and clinical

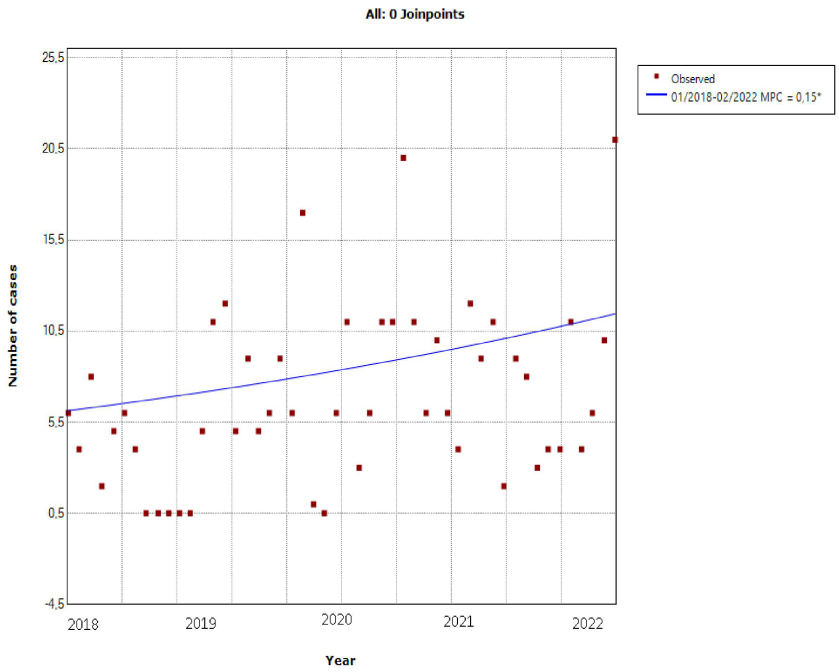


FIGURE 1
Temporal trend of hospital DKA between March 2018 and February 2022: it shows a significant upward trend with a change in the average monthly percent change (AMPC) of +0.2%, with p=0.037.

presentations of various comorbidities such as DM and has been incriminated in the increasing trend of DKA (17).

The study of the temporal trend of hospital cases of DKA between March 2018 and February 2022 showed a significant upward trend.

This finding has important implications (Table 2). First, the incidence of DKA admitted in our department has been increasing well before COVID-19 pandemic.

This has long been corroborated all over the world namely in a cohort study by Zhong et al. where they studied the trends in hospital admission for DKA in adults in England between 1998-2013 and found a significant rise of DKA among adults with T1D as well as T2D (22). This has been attributed to the increase of the overall incidence of T2D, increased prevalence of infection and ketosis-prone T2D in minority groups (22).

The second key finding is that the incidence of DKA showed a sustained increase during the COVID-19 pandemic, showing a notable rise of almost 50% compared to a similar period prior to the pandemic. Our study also identified a statistically significant difference in the mean monthly incidence of DKA between the two periods, implying that the pandemic may have been the underlying factor contributing to this upward trend.

This observation is consistent with the findings of other studies (9, 18–21), which have reported a rising trend of DKA cases during COVID-19 pandemic. Indeed, an international multicenter study based on data from 13 national diabetes registries by Birkebaek et al. found a significant increase in the proportion of presentations for DKA, with a rise of 39.4% in 2020 and 38.9% in 2021. This increase exceeded the predicted year-on-year rise in prevalence, which was predicted to be 32.5% for 2020 and 33.0% for 2021 (6).

TABLE 1 Monthly percent changes of new DKA cases.

Monthly Percent Change (MPC)							
Segment	Lower Endpoint	Upper Endpoint	MPC	Lower CI	Upper CI	Test Statistic (t)	Prob > |t|
1	01/2018	2022	0.2*	0.0	0.3	2.1	0.037
***Indicates that the monthly Percent Change (MPC) is significantly different from zero at the alpha =0.05 level.**							
Average Monthly Percent (AMPC)							
Range	Lower Endpoint	Upper Endpoint	AMPC	Lower CI	Upper CI	Test Statistic	P-Value~
2018-2022	01/2018	2022	0.2*	0.0	0.3	2.1	<0.1
***Indicates that the AMPC is significantly different from zero at the alpha=005 level.** **~ If the AMPC is within one segment, the t-distribution is used. Otherwise, the normal (z) distribution is used.**							

Many reasons could be behind this increase. The most apparent could be the delayed diagnosis of new onset DM due to the reluctance of individuals to seek medical attention for symptoms such as polyuria and polydipsia due to the fear of contracting the virus, leading to missed opportunities for earlier diagnosis. However, we found no significant difference in duration of polyuria and polydipsia preceding DKA in both study groups. This finding challenges the notion that delayed diagnosis has a significant influence on the increased incidence of DKA during the COVID-19 pandemic, arguably implying instead a direct effect of SARS-CoV-2 in the increased incidence of DKA during the pandemic.

Another possible reason is that SARS-CoV-2 may have caused direct damage to pancreatic beta cells, which could explain the observed increase in insulin-dependent diabetes and consequently in DKA cases, as hypothesized by Misra et al. (23).

The incidence of new onset DKA during COVID-19 pandemic according to pandemic waves shows a nadir during the third wave with a stronger re-increase during the fifth wave which saw the highest number of DKA case. The average monthly admission of DKA according to pandemic waves shows also a nadir during the third wave and a reincrease during the fifth wave without it being statistically significant (p=0.09).

Possible factors for the relative decline of DKA admitted in our department during the third wave include the relatively short duration of this wave, as defined by ONMNE, as well as strict social restrictions that were implemented during this wave. Moreover, during this particular wave, there was a simultaneous establishment of additional departments designated for the admission of COVID-19 patients, regardless of their DKA status. This occurrence potentially alleviated the burden on our department for a temporary period.

The pandemic variants of SARS-CoV-2 have been reported to change its viral characteristics. Mutations that alter the affinity of the virus to receptors may affect its ability to enter beta cells and cause cellular damage. For example, the Omicron variant is characterized by a higher ACE2 binding affinity (24). Therefore, it is plausible that the tropism for the pancreas differs according to SARS-CoV-2 variants (25). Although merely speculative, this can arguably explain the higher incidence of DKA during the fifth wave where Omicron variant was prevalent in Tunisia (26). However, data is insufficient to investigate whether the diabetogenicity of COVID-19 depends on its variants (25).

We found an increase of T1D incidence during COVID-19 pandemic which is in line with numerous studies around the world reported mainly among the pediatric population (9, 18–21, 27).

Although it was once believed that genetics played a significant role in T1D with over 50 genes identified, low concordance of T1D (<50%) in monozygotic twins implies that environmental factors, more specifically viruses, may be even more consequential than previously thought (28). Respiratory viruses have also been incriminated as shown by a large Norwegian cohort study published in June 2018 by Ruiz et al. who have investigated the risk of new onset DM subsequent to Influenza A (H1N1) pandemic in June 2009 and reported a twofold increased risk of new onset T1D (16).

The molecular mimicry hypothesis is the most appealing pathophysiological pattern mediating beta cells autoimmune injury as suggested by Andrade et al. in a study published in October 2022 where amino-acid sequences of human insulin and GAD65 as long as their epitopes were compared with the sequences of the SARS-CoV-2 proteins (S protein, Spike protein) (29). Epitope similarity between human insulin and SARS-CoV-2 and between GAD65 and SARS-CoV-2 ranged between 45 to 60%. This would

FIGURE 2
Incidence of DKA during COVID-19 according to pandemic waves: it shows a decrease during the third wave with a re-increase during the fourth and fifth waves.

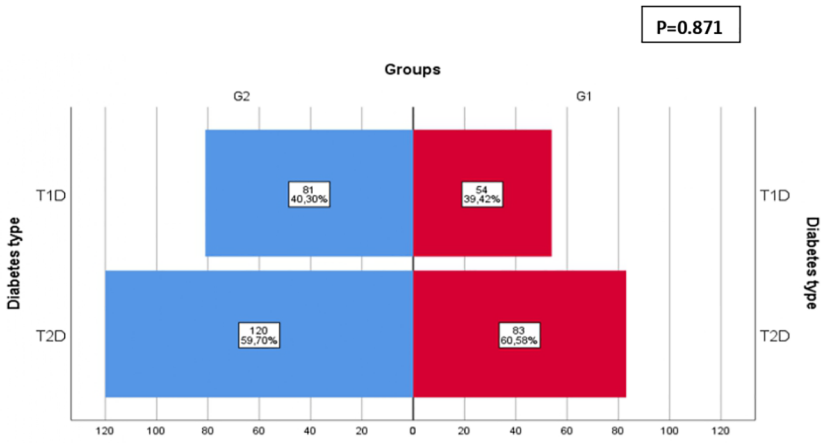


FIGURE 3
Distribution of diabetes type before (GI) and during COVID-19 (G2). The incidence of TID increased by 50% during COVID-19 pandemic as 54 (38.42%) out of 137 patients were TID in G1 vs 81 (40.30%) out of 201 patients in G2 while the incidence of T2D increased by 44% as 83 (60.58%) out of 137 patients were T2D in GI vs 120 (59.70%) out of 201 patients in G2. However, the distribution of TID and T2D frequencies was comparable between the two groups (p=0.871).

plausibly result in the development of an immune cross-reaction to self-antigens, thus triggering T1D (29). However, without clinical data from individuals with T1D or COVID-19, it is difficult to establish a direct causal relationship between SARS-CoV-2 and the triggering of T1D.

The other related mechanism is virus-induced beta cell injury causing the release of sequestered antigens which would eventually be expressed by antigen-presenting cells increasing the risk of autoantibodies generation (30).

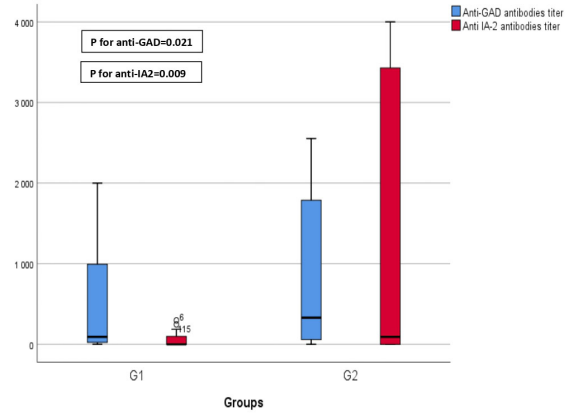


FIGURE 4
Comparison of antibodies' titers before (GI) and during COVID-19 (G2). This graphical representation provides a comparison of anti-glutamic acid decarboxylase (anti-GAD) and anti-islet cell antigen (anti-IA2) titers before and during the COVID-19 pandemic. The data reveals a statistically significant increase in anti-GAD titers during COVID-19 (GI) compared to the prepandemic period (p=0.021). Furthermore, it demonstrates a significant elevation in anti-IA2 titers during COVID-19 (G2) in contrast to the prepandemic period (p=0.009).

Although we found no significant difference between T1D frequencies before and during COVID-19 pandemic, autoantibodies such as Anti-GAD and Anti-IA2 titers interestingly increased during the pandemic period compared with the pre-pandemic period (p=**0.021**), (p=**0.009**) respectively. This would reasonably incriminate SARS-CoV-2 in triggering an auto-immune insulitis. Similarly, Wang et al. reported a marked increase in autoantibody reactivities in COVID-19 patients as compared to uninfected individuals reflected by a higher prevalence of autoantibodies against immunomodulatory proteins (including cytokines, chemokines, complement components and cell-surface proteins) (31). Even though this study also discusses the presence of tissue-associated autoantibodies in patients with COVID-19, targeting various organs and systems including vascular cells, coagulation factors, platelets, connective tissue, extracellular matrix components, anti-islet antibodies have not been investigated. Therefore, we cannot establish with certainty a direct link between COVID-19 and T1D.

We found an increase of new onset T2D cases by 44% during COVID-19 pandemic.

Insulin resistance is suspected to be directly induced by SARS-CoV-2 as speculated by a study by Montefusco et al. where it was found that compared to healthy controls, patients with COVID-19 had significantly higher levels of mean fasting insulin, proinsulin, and C-peptide, as well as higher values of the homeostasis model assessment of beta cell dysfunction (HOMA-B) and homeostasis model assessment of insulin resistance (HOMA-IR) which were correlated with inflammation markers indicating that COVID-19-related insulin resistance has an inflammatory basis suggesting that insulin resistance and beta cell dysfunction in COVID-19 may be triggered by a proinflammatory environment initiated by a cytokine storm (12).

The increase of accountability of stress in precipitating DKA can be explained by the levels of psychological distress associated with COVID-19 pandemic. Indeed, Xiong J et al. conducted a

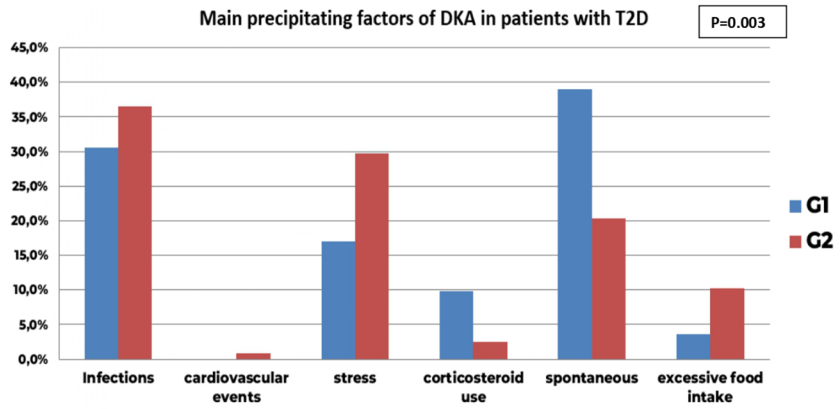


FIGURE 5
Comparison of main precipitating factors of DKA before (GI) and during COVID-19 pandemic in patients with new onset T2D (G2); the distribution of DKA precipitating factors in T2D differed significantly between the two groups (p=0.003) with a notable increase of stress (29.7% in G2 vs 17.1% in GI, AR-2) as well as excessive food intake notably hypertonic drinks (10.2% in G2 vs 3.7% in G1) at the expense of corticosteroid use as its accountability in precipitating DKA significantly decreased in G2 (2.5% in G2 vs 9.8% in GI, AR-2.2). The infectious precipitating factors remained comparable between the two groups (30.5% in G1 vs 36.4%, AR-0.9).

systematic review showing high levels of anxiety-related symptoms, depression, posttraumatic stress disorder, psychological distress in the general population in various countries during the COVID-19 pandemic. This can be explained by periods of lockdown where unemployment and marital problems peaked, not to mention frequently being exposed to worrisome news about COVID-19 (32).

The increased responsibility of excessive food intake as a trigger for DKA may be ascribed to modifications in eating behavior. According to a systematic review of longitudinal studies which compared eating habits before and during COVID-19 pandemic, an increased tendency towards eating snacks and a preference for sweets and ultra-processed food rather than fruits and vegetables was reported (33).

Among infectious precipitating factors, COVID-19 became the first infectious precipitating factor of DKA in G2 with a decrease of the accountability of pulmonary and other influenza like illness, the latter is likely due to the enforcement of COVID-19 public health protocols leading to a decline in the spread of common respiratory viruses (34).

DKA was managed classically by hydration, insulin infusion along with electrolyte supplementation. The time it took for DKA to resolve was considerably higher during COVID-19 pandemic compared to the prepandemic period (p=0.022). This supports the findings of Farzadfar et al. where longer time to DKA resolution was found in patients infected with COVID-19 (35).

This can be explained by various strategies implemented during the COVID-19 pandemic to effectively reduce the risk of healthcare workers being exposed to the virus while providing care to patients with COVID-19 (35). These strategies involved minimizing the use of venous insulin infusions for DKA whenever possible and decreasing the frequency of capillary glycemia checks for patients on subcutaneous insulin regimens (35). These measures underscored the urgent need for innovative technologies such as continuous glucose monitoring (CGM), especially in the context of a pandemic, as they would enable healthcare exposure to be minimized.

Regrettably, implementing such strategies is challenging in resource-limited settings, particularly in financially constrained regions like certain areas in Africa.

Our study is subject to several limitations that should be acknowledged. Firstly, it is a retrospective study, conducted at a single center, which may restrict the generalizability of our findings.

TABLE 2 Different studies comparing the incidence of new onset DKA before and during COVID-19 pandemic.

Authors	Country	Year	Patients	Increased incidence of DKA	Type of DM
Mastromauro et al. (18)	Italy	2022	172	Yes (55% vs 36%)	T1D
Jafari et al. (19)	USA	2022	175	Yes	T1D
Salmi et al. (9)	Finland	2021	315	Yes (6.25 → 20)	T1D
Vorgučin et al. (20)	Serbia	2022	231	Yes (36.12 → 42.42)	T1D
Khan et al. (21)	USA	2022	14630	Yes	T2D, T1D
Our study	Tunisia	2023	340	Yes	T1D, T2D

Additionally, the absence of computerized national registries for new onset DKA in our country limits our ability to access comprehensive data and needs manual collection from medical records.

Another important limitation is the challenge of accurately establishing the history of COVID-19 infection in our study population. While we documented cases of SARS-CoV-2 infection at the time of diagnosis, differentiating between a prior infection and a vaccinal status was not possible particularly as our study was conducted during a period when COVID-19 vaccinations became widely accessible.

Despite these limitations, our study provides valuable insights into the incidence and potential factors associated with DKA in our specific context of an African country. We believe that our findings contribute to the existing knowledge base and can serve as a foundation for future research efforts aimed at addressing these limitations and expanding our understanding of the relationship between COVID-19 and DKA.

Conclusion

The COVID-19 pandemic has had significant impacts on the epidemiology of DKA. While the increasing trend in DKA cases was observed even prior to the pandemic, the COVID-19 period has seen a significant rise in DKA incidence, which may be due to delayed diagnosis or arguably *via* an auto-immune triggering mechanism or an inflammatory milieu favorable for insulin resistance.

The potential long-term implications of these trends are concerning, particularly in terms of the burden on healthcare systems and the increased risk of complications associated with these conditions. As such, it is important for healthcare providers, particularly in Africa, to remain vigilant in monitoring and treating patients with DM, especially during times of crisis such as the COVID-19 pandemic.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Written and informed consent was obtained from the patients for publication of the submitted article. The anonymity and data confidentiality of the patients' data were respected. Approval from the Ethical Committee of University of Medicine of Sousse was obtained under the assigned number 4935/2023.

Author contributions

AG and TA drafted the manuscript. All authors have contributed significantly to this work, providing substantial intellectual input, and have thoroughly reviewed and approved the final manuscript.

Funding

The authors declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

Dr Hamdi Frikha played a role in validating the statistical analysis.

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. Richards F, Kodjamanova P, Chen X, Li N, Atanasov P, Bennetts L, et al. Economic burden of COVID-19: a systematic review. *ClinicoEconomics outcomes Res* (2022) 14:293–307. doi: 10.2147/CEOR.S338225

2. Tessema GA, Kinfu Y, Dachew BA, Tesema AG, Assefa Y, Alene KA, et al. The COVID-19 pandemic and healthcare systems in Africa: a scoping review of preparedness, impact and response. *BMJ Global Health* (2021) 6(12):71–9. doi: 10.1136/bmjgh-2021-007179

3. Gavkare AM, Nanaware N, Rayate AS, Mumbre S, Nagoba BS. COVID-19 associated diabetes mellitus: a review. *World J diabetes.* (2022) 13(9):729–37. doi: 10.4239/wjd.v13.i9.729

4. Maldonado MR, Chong ER, Oehl MA, Balasubramanyam A. Economic impact of diabetic ketoacidosis in a multiethnic indigent population: analysis of costs based on the precipitating cause. *Diabetes Care* (2003) 26(4):1265–9. doi: 10.2337/diacare.26.4.1265

5. Jensen ET, Stafford JM, Saydah S, D'Agostino RB, Dolan LM, Lawrence JM, et al. Increase in prevalence of diabetic ketoacidosis at diagnosis among youth with type 1 diabetes: the SEARCH for diabetes in youth study. *Diabetes Care* (2021) 44(7):1573–8. doi: 10.2337/dc20-0389

6. Birkebaek NH, Kamrath C, Grimsmann JM, Aakesson K, Cherubini V, Dovc K, et al. Impact of the COVID-19 pandemic on long-term trends in the prevalence of

diabetic ketoacidosis at diagnosis of paediatric type 1 diabetes: an international multicentre study based on data from 13 national diabetes registries. *Lancet Diabetes endocrinol* (2022) 10(11):786–94. doi: 10.1016/S2213-8587(22)00246-7

7. Beliard K, Ebekozien O, Demeterco-Berggren C, Alonso GT, Gallagher MP, Clements M, et al. Increased DKA at presentation among newly diagnosed type 1 diabetes patients with or without COVID-19: data from a multi-site surveillance registry. *J diabetes.* (2021) 13(3):270–2. doi: 10.1111/1753-0407.13141

8. Metwally AA, Mehta P, Johnson BS, Nagarjuna A, Snyder MP. COVID-19-Induced new-onset diabetes: trends and technologies. *Diabetes* (2021) 70(12):2733–44. doi: 10.2337/dbi21-0029

9. Salmi H, Heinonen S, Hästbacka J, Lääperi M, Rautiainen P, Miettinen PJ, et al. New-onset type 1 diabetes in Finnish children during the COVID-19 pandemic. *Arch Dis childhood.* (2022) 107(2):180–5. doi: 10.1136/archdischild-2020-321220

10. Yang JK, Lin SS, Ji XJ, Guo LM. Binding of SARS coronavirus to its receptor damages islets and causes acute diabetes. *Acta diabetologica.* (2010) 47(3):193–9. doi: 10.1007/s00592-009-0109-4

11. Gottesman BL, Yu J, Tanaka C, Longhurst CA, Kim JJ. Incidence of new-onset type 1 diabetes among US children during the COVID-19 global pandemic. *JAMA Pediatrics.* (2022) 176(4):414–5. doi: 10.1001/jamapediatrics.2021.5801

12. Montefusco L, Ben Nasr M, D'Addio F, Loretelli C, Rossi A, Pastore I, et al. Acute and long-term disruption of glycometabolic control after SARS-CoV-2 infection. *Nat Metab* (2021) 3(6):774–85. doi: 10.1038/s42255-021-00407-6

13. Organization WC-DGWH (2020). Available at: https://covid19.who.int/region/emro/country/tn.

14. Broome DT, Pantalone KM, Kashyap SR, Philipson LH. Approach to the patient with MODY-monogenic diabetes. *J Clin Endocrinol Metab* (2020) 106(1):237–50. doi: 10.1210/clinem/dgaa710

15. Kitabchi AE, Umpierrez GE, Miles JM, Fisher JN. Hyperglycemic crises in adult patients with diabetes. *Diabetes Care* (2009) 32(7):1335–43. doi: 10.2337/dc09-9032

16. Kayhan S, Hepsen S, Kalkisim HK, Sendur IN, Altay FA, Yalcindag A. The evaluation of pancreas β-cell autoantibodies in non-diabetic COVID-19 patients. *Arch Endocrinol Metab* (2022) 66(4):459–65. doi: 10.20945/2359-3997000000498

17. Banerjee M, Pal R, Dutta S. Risk of incident diabetes post-COVID-19: a systematic review and meta-analysis. *Primary Care diabetes.* (2022) 16(4):591–3. doi: 10.1016/j.pcd.2022.05.009

18. Mastromauro C, Blasetti A, Primavera M, Ceglie L, Mohn A, Chiarelli F, et al. Peculiar characteristics of new-onset type 1 diabetes during COVID-19 pandemic. *Ital J Pediatrics.* (2022) 48(1):26. doi: 10.1186/s13052-022-01223-8

19. Jafari K, Koves I, Rutman L, Brown JC. Impact of the COVID-19 pandemic on the severity of diabetic ketoacidosis presentations in a tertiary pediatric emergency department. *Pediatr Qual safety.* (2022) 7(2):e502. doi: 10.1097/pq9.0000000000000502

20. Vorgučin I, Savin M, Stanković Đ, Miljković D, Ilić T, Simić D, et al. Incidence of type 1 diabetes mellitus and characteristics of diabetic ketoacidosis in children and adolescents during the first two years of the COVID-19 pandemic in vojvodina. *Medicina* (2022) 58(8):1013. doi: 10.3390/medicina58081013

21. Khan F, Paladino L, Sinert R. The impact of COVID-19 on diabetic ketoacidosis patients. *Diabetes Metab syndrome.* (2022) 16(1):102389. doi: 10.1210/clinem/dgaa710

22. Zhong VW, Juhaeri J, Mayer-Davis EJ. Trends in hospital admission for diabetic ketoacidosis in adults with type 1 and type 2 diabetes in England, 1998-2013: a retrospective cohort study. *Diabetes Care* (2018) 41(9):1870–7. doi: 10.2337/dc17-1583

23. Misra S. Rise in diabetic ketoacidosis during the COVID-19 pandemic: several questions remain. *Lancet Diabetes endocrinol* (2022) 10(11):763–5. doi: 10.1016/S2213-8587(22)00272-8

24. Rangu R, Wander PL, Zraika S. Does diabetes risk after SARS-CoV-2 infection depend on the viral variant? *Diabetes Res Clin Pract* (2022) 191:110054. doi: 10.1016/j.diabres.2022.110054

25. Yonekawa A, Shimono N. Clinical significance of COVID-19 and diabetes: in the pandemic situation of SARS-CoV-2 variants including omicron (B.1.1.529). *Biology* (2022) 11(3):400–18. doi: 10.3390/biology11030400

26. ONMNE. *Bulletin de veille COVID-19 à la date du 12 décembre 2021* (2021). Available at: https://onmne.tn/?p=15695&fbclid=IwAR0w84Y2dEffX9Kn_ZJWl2zBveEPHDNWJ0XvjQU0cqUw57nSQpycETl8RsE.

27. Aggarwal S, Garcia-Telles N, Aggarwal G, Lavie C, Lippi G, Henry BM. Clinical features, laboratory characteristics, and outcomes of patients hospitalized with coronavirus disease 2019 (COVID-19): early report from the united states. *Diagnosis (Berlin Germany).* (2020) 7(2):91–6. doi: 10.1515/dx-2020-0046

28. Filippi CM, von Herrath MG. Viral trigger for type 1 diabetes: pros and cons. *Diabetes* (2008) 57(11):2863–71. doi: 10.2337/db07-1023

29. Andrade L, Bittencourt A, Oliveira L, Oliveira L, Oliveira G. Triggering type 1 diabetes post-covid: molecular mimicry? (2022). doi: 10.21203/rs.3.rs-2152791/v1

30. Caruso P, Longo M, Esposito K, Maiorino MI. Type 1 diabetes triggered by covid-19 pandemic: a potential outbreak? *Diabetes Res Clin Pract* (2020) 164:108219. doi: 10.1016/j.diabres.2020.108219

31. Wang EY, Mao T, Klein J, Dai Y, Huck JD, Jaycox JR, et al. Diverse functional autoantibodies in patients with COVID-19. *Nature* (2021) 595(7866):283–8. doi: 10.1038/s41586-021-03631-y

32. Xiong J, Lipsitz O, Nasri F, Lui LMW, Gill H, Phan L, et al. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. *J Affect Disord* (2020) 277:55–64. doi: 10.1016/j.jad.2020.08.001

33. González-Monroy C, Gómez-Gómez I, Olarte-Sánchez CM, Motrico E. Eating behaviour changes during the COVID-19 pandemic: a systematic review of longitudinal studies. *Int J Environ Res Public Health* (2021) 18(21):11130. doi: 10.3390/ijerph182111130

34. Doroshenko A, Lee N, MacDonald C, Zelyas N, Asadi L, Kanji JN. Decline of influenza and respiratory viruses with COVID-19 public health measures: Alberta, Canada. *Mayo Clinic Proc* (2021) 96(12):3042–52. doi: 10.1016/j.mayocp.2021.09.004

35. Farzadfar D, Gordon CA, Falsetta KP, Calder T, Tsegaye A, Kohn N, et al. Assessment of insulin infusion requirements in COVID-19-Infected patients with diabetic ketoacidosis. *Endocrine Pract* (2022) 28(8):787–94. doi: 10.1016/j.eprac.2022.05.006

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 14 August 2023
DOI 10.3389/fendo.2023.1237832

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Naina Mohamed Pakkir Maideen,
Dubai Health Authority, United Arab Emirates
Simonetta Bellone,
University of Eastern Piedmont, Italy

*CORRESPONDENCE
Chun-ping Zeng
zcp193@163.com
Wei-ping Sun
sunwp07@163.com

RECEIVED 10 June 2023
ACCEPTED 27 July 2023
PUBLISHED 14 August 2023

CITATION
Zhou X-y, Huang S-f, Lin J-x, Zhi H-n, Xiao L, Wang X-z, Guo K-h, Zhou L, Long T, You H-m, Lin M-r, Luo X-y, Sun W-p and Zeng C-p (2023) Clinical characteristics and acute complication of COVID-19 patients with diabetes: a multicenter, retrospective study in Southern China.
Front. Endocrinol. 14:1237832.
doi: 10.3389/fendo.2023.1237832

Clinical characteristics and acute complication of COVID-19 patients with diabetes: a multicenter, retrospective study in Southern China

Xiao-ying Zhou[1], Shao-feng Huang[1], Jun-xu Lin[2], Hai-ni Zhi[2], Lu Xiao[3], Xiang-zhu Wang[1], Kai-heng Guo[1], Lin Zhou[1], Tao Long[1], Hui-min You[1], Ming-run Lin[1], Xiang-ya Luo[1], Wei-ping Sun[3]* and Chun-ping Zeng[1]*

[1]Department of Endocrinology and Metabolism, The Fifth Affiliated Hospital of Guangzhou Medical University, Guangzhou, China, [2]Department of Endocrinology, The Affiliated Loudi Hospital, Hengyang Medical School, University of South China, Hengyang, Hunan, China, [3]Department of Endocrinology and Metabolism, Loudi Central Hospital, Loudi, China

Aims: This study aims to describe the clinical characteristics, laboratory data and complications of hospitalized COVID-19 patients with type 2 diabetes mellitus (T2DM) since epidemic prevention and control optimization was adjusted in December 2022 in China.

Methods: This retrospective multicenter study included 298 patients with confirmed type 2 diabetes mellitus with or without COVID-19. We collected data from the first wave of the pandemic in The Fifth Affiliated Hospital of Guangzhou Medical University, Loudi Central Hospital and The First People's Hospital of Xiangtan from December 1, 2022 to February 1, 2023. We extracted baseline data, clinical symptoms, acute complications, laboratory findings, treatment and outcome data of each patient from electronic medical records.

Results: For among 298 hospitalized patients with type 2 diabetes, 136 (45.6%) were COVID-19 uninfected, and 162 (54.4%) were COVID-19 infected. We found that the incidence of cough, fatigue, fever, muscle soreness, sore throat, shortness of breath, hyposmia, hypogeusia and polyphagia (all $p<0.01$) were significantly higher in the exposure group. They showed higher levels of ketone ($p=0.04$), creatinine ($p<0.01$), blood potassium ($p=0.01$) and more diabetic ketoacidosis ($p<0.01$). Patients with COVID-19 less use of metformin ($p<0.01$), thiazolidinediones ($p<0.01$) and SGLT2 ($p<0.01$) compared with patients without COVID-19.

Conclusion: COVID-19 patients with diabetes showed more severe respiratory and constitutional symptoms and an increased proportion of hyposmia and hypogeusia. Moreover, COVID-19 patients with diabetes have a higher incidence of acute complications, are more prone to worsening renal function, and are more cautious about the use of antidiabetic drugs.

KEYWORDS

COVID-19, diabetes, clinical characteristics, acute complications, renal insufficiency

1 Introduction

Since the optimization and adjustment of epidemic prevention and the orderly restoration of production and living in December 2022, COVID-19 has started the first wave of the pandemic in China. At present, SARS-CoV-2 Omicron BA.5.2 variants is progressively displacing other variants in southern CHINA, which displays a higher transmissibility than other Omicron subvariants (1). Since China no longer requires nucleic acid detection of SARS-CoV-2, the infection rate of the novel coronavirus is not available. The Centers for Disease Control and Prevention (CDC) in Sichuan (a province in southern CHINA) estimated the prevalence of COVID-19 approximately 65.5% as of December 26, 2022. According to the big data model of the National School of Development, as of January 11, 2023, the cumulative infection rate of COVID-19 in China is approximately 64%, and the cumulative number of infected people is approximately 900 million.

Several studies have reported that COVID-19 patients are more susceptible to type 2 diabetes, and the severity and mortality of COVID-19 in diabetes patients are higher than those in patients without diabetes (2, 3). In China, the reported prevalence of diabetes in patients with COVID-19 is similar to the national prevalence of T2DM, approximately 11% (4, 5). In New Delhi, India, the prevalence of diabetes among people with COVID-19 is 47%, which is far higher than the prevalence of T2DM in this region (6). Although the prevalence of diabetes among COVID-19 patients varies by region, studies have reported that the proportion of COVID-19 patients with diabetes is relatively high (7, 8). Moreover, COVID-19 patients with diabetes might be at increased risk of acute metabolic complications, especially an increase in diabetic ketoacidosis during the COVID-19 pandemic (9). The patients with diabetes also had abnormal blood glucose levels and increased the dose of insulin during hospitalization, which indicated their poor glycemic control (10). The increased prevalence, severity, and complications of type 2 diabetes in COVID-19 patients may be related to human angiotensin-converting enzyme 2 (ACE2) and transmembrane serine protease 2 receptors, which are expressed on pancreatic beta cells (11). SARS-CoV-2 infection has been shown to reduce insulin secretion levels and induce pancreatic β cell apoptosis (12).

We conducted a multicenter retrospective study in southern China, aiming to compare the basic information, laboratory examinations, clinical symptoms, acute complications, and medication in patients with diabetes with or without COVID-19, and ascertain the impact of COVID-19 on diabetes.

2 Materials and methods

This retrospective multicenter study included 298 patients with confirmed type 2 diabetes mellitus with or without COVID-19. We collected data from the first wave of the pandemic in The Fifth Affiliated Hospital of Guangzhou Medical University, Loudi Central Hospital and The First People's Hospital of Xiangtan from December 1, 2022 to February 1, 2023. This research was conducted with approval from the Fifth Affiliated of Guangzhou Medical University Research Ethics Committee. (GYWY-L2023-64).

Patients fulfilling the following criteria were included in this study: physician diagnosis of T2DM and aged over 18 years. Gravidas and patients with severe multi-organ dysfunction were excluded. We divided all patients into two groups depending on whether they were confirmed with COVID-19 after novel coronavirus nucleic acid testing or novel coronavirus antigen detection. All the enrolled patients were hospitalized due to poor blood glucose control or new-onset diabetes, among which the T2DM patients with COVID-19 had been clearly infected before hospitalization.

We extracted baseline data, clinical symptoms, acute complications, laboratory findings, treatment and outcome data of each patient from electronic medical records. All data collected were reviewed by the research team and double checked by experienced physicians. Patients with missing data or unknown medical records were excluded.

We used SPSS (version 26.0; IBM) statistical software to analyze and process the data. Quantitative data were expressed as x±s and differences between groups were compared using independent sample t-test if they were normally distributed. Quantitative data that did not conform to a normal distribution and ordinal data were expressed as medians or quartile ranges and differences between groups were compared using nonparametric tests. Qualitative data were described by frequency or percentage and differences between groups were compared using the χ^2 test. For all the statistical analyses, a p-value < 0.05 was considered statistically significant.

10.3389/fendo.2023.1237832

3 Results

A total of 298 hospitalized patients with type 2 diabetes were included in this retrospective study. Among these patients, 136 (45.6%) were COVID-19 uninfected and 162 (54.4%) were COVID-19 infected, respectively. Of all hospitalized patients, the median age was 64 years (IQR, 49-67) and the median duration of diabetes was 8.0 (IQR, 3.0-14.0). Compared with patients without COVID-19, patients with COVID-19 infected had older age (68.5 [IQR, 59.0-76.0]), longer diabetes duration (10.0 [IQR, 4.0-19.0]), and had no significant differences in either systolic blood pressure or diastolic blood pressure. The most common symptoms were cough (143 [48.0%]), fatigue (132 [44.3%]), fever (61[20.5%]), dizziness (77 [25.8%]), shortness of breath (61 [20.5%]), nausea (67 [22.5%]), polyuria (75[25.2%]) and polydipsia (132 [44.3]) at illness onset. Less common symptoms included muscle soreness, runny nose, sore throat, chest distress, diarrhea, hyposmia, hypogeusia, polyphagia and weight loss. We found that the incidence of cough (37[27.2%] vs 106[65.4%]; p<0.01), fatigue (37[27.3%] vs 95[58.6%]; p<0.01), fever (16[11.8%] vs 45[27.8%]; p<0.01), muscle soreness (3[2.2%] vs 22[13.6%]; p<0.01), sore throat (8[5.9%] vs 23 [14.2%]; p=0.02), shortness of breath (18[13.2%] vs 43[26.5%]; p<0.01), hyposmia (0[0%] vs 14[8.6%]; p<0.01), hypogeusia (1 [0.7%] vs 17[10.5%]; p<0.01) and polyphagia (1[0.7% vs 25 [15.4%]; p<0.01) were significantly higher in the exposure group than in the non- exposure group (Table 1).

The laboratory test results and the incidence of acute complications at admission are shown in Table 2. In all the patients, glycosylated hemoglobin and fasting plasma glucose were above the normal range, while the values of other laboratory indicators were within the normal range. Compared to non-exposure group, the exposure group showed higher levels of ketone (0.2 [IQR, 0.1-0.3] vs 0.2 [IQR, 0.1-0.9]; p=0.04), creatinine (69.3 [IQR, 53.7-89.3] vs 86.5 [IQR, 63.3-121.0]);

TABLE 1 Baseline characteristic and clinical signs and symptoms of patients infected with COVID-19.

			No.(%)		
		Total (n=298)	COVID-19 uninfected (n=136)	COVID-19 infected (n=162)	p value
Age, Median (IQR), years		64.0 (54.0-74.0)	60.0 (51.2-70.0)	68.5 (59.0-76.0)	0.000058
Diabetes duration, Median (IQR), years		8.0 (3.0-14.0)	7.0 (2-10)	10.0 (4.0-19.0)	0.003054
Systolic blood pressure, Median (IQR), mm Hg		133.0 (122.0-149.0)	133.0 (123.3-149.8)	133.5 (120.8-148.3)	0.556
Diastolic blood pressure, Median (IQR), mm Hg		78.0 (70.0-87.0)	77.5 (70.0-88.0)	79.0 (70.0-86.3)	0.724
Signs and symptoms	Cough	143 (48.0)	37 (27.2)	106 (65.4)	5.0998E-11
	Fatigue	132 (44.3)	37 (27.3)	95 (58.6)	5.5597E-8
	Fever	61 (20.5)	16 (11.8)	45 (27.8)	2.2877E-10
	Dizziness	77 (25.8)	44 (32.4)	33 (20.4)	0.696413
	Muscle soreness	25 (8.4)	3 (2.2)	22 (13.6)	0.000014
	Runny nose	4 (1.3)	1 (0.7)	3 (1.9)	0.150758
	Sore throat	31 (10.4)	8 (5.9)	23 (14.2)	0.019388
	Chest distress	53 (17.8)	22 (16.2)	31 (19.1)	0.506485
	Shortness of breath	61 (20.5)	18 (13.2)	43 (26.5)	0.004637
	Nausea	67 (22.5)	27 (19.9)	40 (24.7)	0.319801
	Diarrhea	25 (8.4)	6 (4.4)	19 (11.7)	0.023482
	Hyposmia	14 (4.7)	0 (0)	14 (8.6)	0.000455
	Hypogeusia	18 (6.0)	1 (0.7)	17 (10.5)	0.000438
	Polydipsia	132 (44.3)	58 (42.6)	74 (45.7)	0.600316
	Polyphagia	26 (8.7)	1 (0.7)	25 (15.4)	0.000297
	Polyuria	75 (25.2)	31 (22.8)	44 (40.3)	0.387774
	Weight loss	43 (14.4)	17 (12.5)	26 (16.0)	0.385904

TABLE 2 Comparison of laboratory parameters and complication between COVID-19 infected and uninfected diabetic patients.

		Median (IQR)		
	Normal range	COVID-19 uninfected (n=136)	COVID-19 infected (n=162)	p value
HbA1c, %	4.0-6.0	9.1 (7.5-11.2)	8.3 (7.4-11.0)	0.440
FPG, mmol/L	3.9-6.1	9.2 (6.7-11.2)	8.4 (6.1-12.5)	0.284
Ketone, mmol/L	0-0.3	0.2 (0.1-0.3)	0.2 (0.1-0.9)	0.035
WBC, $\times10^9$/L	3.5-9.5	6.9 (5.4-8.8)	6.4 (4.7-9.9)	0.274
Neut, $\times10^9$/L	1.8-6.3	4.11 (3.17-6.15)	4.6 (3.0-7.0)	0.492
Hemoglobin, g/L	115-150	131.0 (112.3-142.0)	125.0 (111.8-140.0)	0.297
Creatinine, umol/L	53-106	69.3 (53.7-89.3)	86.5 (63.3-121.0)	0.001
ALT, U/L	≤40	20.4 (13.4-31.6)	21.0 (14.1-31.3)	0.891
AST, U/L	≤41	25.4 (17.7-33.2)	29.6 (20.5-41.0)	0.559
Blood potassium, mmol/L	3.5-5.5	4.0 (3.7-4.3)	4.2 (3.7-4.6)	0.013
Blood sodium, mmol/L	135-145	138.2 (136.0-141.0)	137.4 (134.3-141.6)	0.455
Blood chlorine, mmol/L	96-168	103.1 (100.3-105.3)	101.8 (98.9-105.8)	0.576
Diabetic ketoacidosis, No. (%)	–	8 (5.9)	34 (21.0)	0.000194
Diabetic hyperosmolar coma, No. (%)	–	1 (0.7)	1 (0.6)	0.150758

HbA1c, Glycosylated Hemoglobin; FPG, Fasting Plasma Glucose; WBC, White cell count; Neut, Neutrophil count; ALT, Alanine Aminotransferase; AST, Aspartate Aminotransferase.

p<0.01), blood potassium (4.0[IQR, 3.7-4.3] vs 4.2[IQR, 3.7-4.6], p=0.01) and more diabetic ketoacidosis (8[5.9%] vs 34[21.0]; p<0.01). These laboratory data indicated that the COVID-19 patients with diabetes are more likely to develop diabetic ketoacidosis and are at greater risk of developing renal insufficiency than those without COVID-19 infection (Table 2).

During hospitalization, most patients were treated with insulin (192[64.1%]), followed by metformin (98[32.9%]), SGLT2 (71 [23.8%]), glucosidase inhibitors (48[15.4%]), DPP4 (41[13.8%]), thiazolidinediones (30[10.1%]), GLP1 (20[6.7%]), sulfnoylureas (7 [2.3%]) and glinides (3[1.0%]). There was no significant difference in the total daily dose of insulin used between patients with T2DM with and without COVID-19 infection. Compared with T2DM patients without COVID-19 infection, patients with COVID-19 were less likely to use metformin (88[64.7%] vs 10[6.2%]; p<0.01), thiazolidinediones (30[22.1%] vs 0[0%]; p<0.01) and SGLT2 (66 [48.5%] vs 5[3.2%]; p<0.01) (Table 3).

4 Discussion

This study was a retrospective study of 298 hospitalized T2DM patients with or without COVID-19 infection, and analyzed baseline data, clinical symptoms, acute complications, laboratory findings, treatment measures and outcome data. From the present results, the severity of T2DM patients with COVID-19 would be higher than those without COVID-19. This may be related to angiotensin-converting enzyme 2 (ACE2), whose expression is elevated in T2DM patients. SARS-CoV-2 binds to ACE2 receptors and can cause multiple organ damage. Moreover, high blood glucose activates inflammatory pathways and increases oxidative damage, impairing immune cell function (13, 14).

In many COVID-19 epidemic studies, COVID-19 with diabetes mostly occurred in the elderly (12, 15)..Recent studies have also indicated that the older the age of T2DM patients with COVID-19 is, the higher the incidence of severe clinical courses and increased mortality (16, 17). In the study, T2DM patients with COVID-19 had a longer duration of diabetes, indicating that this group of people was at higher risk from COVID-19 infection, which was consistent with the characteristics of such patients in other articles (18, 19). The cause was partly attributed to the immune system being impaired due to metabolic inflammation and the body's ability to deal with infections being reduced in patients with diabetes.

Diabetes is associated with hyperglycemia, and the common symptoms are polyuria, polydipsia, polyphagia and weight loss (20). Consistently, the data of this study showed that these symptoms were the most common in T2DM patients with or without COVID-19. Our study also found that cough was also common in T2DM patients with or without COVID-19, and we speculated that type 2 diabetes may be related to an increased prevalence of respiratory symptoms (21).

Hyperglycemia can trigger an inflammatory response, which leads to structural changes in lung tissue and impaired lung function (22). Such structural changes may also be associated with an increased risk of hospitalization for pneumonia in patients with diabetes (23). The occurrence of fatigue in T2DM patients is also related to the inflammatory response (24), and they showed high levels of inflammatory markers including IL-6, CRP, and neopterin, which plays a role in causing fatigue in T2DM

10.3389/fendo.2023.1237832

TABLE 3 Comparison of treatments between COVID-19 infected and uninfected diabetic patients.

		No. (%)		
	Total (n=298)	COVID-19 uninfected (n=136)	COVID-19 infected (n=162)	p value
Total daily dose of insulin, Median (IQR)	–	14.0 (0.0-28.0)	12.0 (0.0-25.3)	0.559
Insulin	192 (64.1)	89 (65.4)	102 (63.0)	0.657
Metformin	98 (32.9)	88 (64.7)	10 (6.2)	0.000218
Glucosidase inhibitors	48 (15.4)	39 (28.7)	7 (4.3)	0.514174
Thiazolidinediones	30 (10.1)	30 (22.1)	0 (0)	0.000027
Sulfonylureas	7 (2.3)	7 (5.1)	0 (1)	0.368240
Glinides	3 (1.0)	1 (0.7)	2 (1.2)	0.092488
DPP4	41 (13.8)	36 (26.5)	5 (3.1)	0.915298
SGLT2	71 (23.8)	66 (48.5)	5 (3.2)	0.000844
GLP1	20 (6.7)	17 (12.5)	3 (1.9)	0.472987

DPP4, Dipeptidyl peptidase-4; SGLT2, Sodium-glucose cotransporter2; GLP1, Glucagon-like peptide-1.

patients (25). Compared to T2DM patients without COVID-19, T2DM patients with COVID-19 had significantly more cough, fatigue, fever, muscle soreness, sore throat and shortness of breath. Studies have shown that patients with COVID-19 have significantly higher numbers of neutrophils (10) and elevated levels of inflammation-related biomarkers (26), which suggests that patients with diabetes are prone to develop an inflammatory storm that ultimately leads to worse symptoms of COVID-19. In addition, oxidative stress caused by persistent hyperglycemia was considered to be the main cause of lung injury in diabetes (27). Patients with diabetes tend to have lower forced vital capacity (FVC) and forced expiratory volume within one second, as well as lower diffusion capacity, which contributes to more severe COVID-19 symptoms (28). Dizziness and nausea were also common in T2DM patients with or without COVID-19, which could be partly explained by diabetic autonomic neuropathy and hypoglycemia caused by T2DM (29, 30). Compared with T2DM patients without COVID-19, hyposmia and hypogeusia were almost exclusively found in COVID-19 patients. Previous studies have shown that 41% and 38% of patients with COVID-19 have hyposmia and hypogeusia, respectively (31). The exact mechanism of hypogeusia and hyposmia in COVID-19 infected patients is not clear, and studies have noted that it may be related to the neuroinvasive potential of SARS-CoV-2 (32). Interestingly, the eating habits of T2DM patients with COVID-19 also changed - they became more polyophagous. This may be related to the COVID-19 lockdown and hyposmia. In a study from Italy, patients experienced a significant increase in appetite due to disruptions in daily work due to the COVID-19 lockdown and stress caused by reading news about COVID-19 from the media (33). Another study showed that patients were unable to perceive taste and flavor, resulting in no sensation of satiation and thus increased appetite (34).

Laboratory findings indicated that the level of creatinine significantly increased in patients with COVID-19, suggesting that kidney damage may have occurred. In a previous cohort study of 5,449 patients admitted to the hospital with COVID-19, 1,993(36.6%) patients developed AKI (35). However, the exact mechanism of COVID-19 on the kidney is unknown, and may be related to the direct damage of SARS-CoV-2 to renal tubules. SARS coronaviruses (including SARS-CoV-2) are detected in urine by PCR, which indicates that the virus interacts directly with or is exposed to renal tubules (36, 37). Furthermore, ACE as a viral receptor is only expressed in the proximal renal tubules, which is parallel to the damaged site of the kidney in patients with SARS-CoV infection (37, 38). Although the median values of creatinine in our study were still within the normal range, they were far from the extent of AKI. However, the difference between the exposure group and the non-exposure group was large, and the present study identified elevated creatinine as a significant predictor of all outcomes of interest (mortality, ICU admission and intubation), which needs to be considered (39).

We found higher ketones and higher rates of diabetic ketoacidosis in T2DM patients with COVID-19 infected, as has been reported in other research. Ketosis occurred in 6.4% of patients with COVID-19 and increased to 11.6% in patients with COVID-19 and diabetes, resulting in a higher mortality rate (33.3%) (40). In a CORONADO study, 11.1% of participants reported diabetes-related disorders at admission, including 40 cases of ketosis, 19 of which were ketoacidosis (2). ACE receptors are expressed in pancreatic tissue and β-cells, and SARS-CoV-2 has been found to bind to ACE2 receptors. Therefore, the metabolic disorders, including DKA may be caused by decreased insulin secretion due to severe insulin resistance and β-cell dysfunction (41, 42). For an unusually high number of patients with COVID-19 developing diabetic ketoacidosis and a hyperosmolar hyperglycemic state, a guideline has been released for the management of DKA (43).

In fact, T2DM patients with COVID-19 are most recommended to be treated with insulin (44). However, our results showed no significant difference between the exposure group and non-exposure group, which may be related to the fact that the proportion of hospitalized patients using insulin therapy was

already quite high. From the reported results, T2DM patients with COVID-19 were less likely to use metformin, sodium-glucose transporter-2 inhibitor (SGLT-2i) and thiazolidinediones than patients without COVID-19 infection. This may be associated with insulin therapy reduced expression of ACE2, while metformin, glucagon-like peptide-1 agonists and thiazolidinediones up-regulate ACE2 expression (45). Furthermore, practical recommendations indicated that discontinuing of metformin and SGLT-2i is recommended in patients with diabetes who have a severe course of COVID-19 (42). In randomized controlled trials, the risk of DKA after SGLT2 use was two times higher in patients with T2DM than in controls (46). In our results patients with DKA and increased creatinine were numerous in the exposure group and the use of metformin required close monitoring for acidosis and decreased renal function, so there was a decrease in metformin use. However, previous studies have shown that metformin and SGLT-2i are associated with reduced mortality in patients with COVID-19 and type 2 diabetes, possibly due to reduced release of inflammatory cytokines, so metformin and SGLT-2i can be used for asymptomatic and mild COVID-19 patients (42, 47). Thiazolidinediones have been found to reduce markers of inflammation in COVID-19 patients (48). However, as a second or even third line treatment, metformin is not explicitly recommended for T2DM patients with COVID-19, which may be the reason why it was not used in the exposure group in our study.

The strengths of this study are that it was a multicenter study with an adequate sample size and comprehensive clinical records. Furthermore, to the best of our knowledge, this is the first study to investigate the clinical characteristics and outcomes of hospitalized COVID-19 patients with diabetes in southern China since the COVID-19 policy adjustment. Our study also has certain limitations. First, it is difficult to assess risk factors for poor prognosis due to short-term outcome follow-up. Second, cases with mild symptoms who were treated at home were missed, so this study only represents patients with more severe COVID-19.

5 Conclusion

In conclusion, COVID-19 patients with diabetes showed more severe respiratory and constitutional symptoms and an increased proportion of hyposmia and hypogeusia. Moreover, T2DM patients with COVID-19 have a higher incidence of acute complications, are more prone to worsening renal function, and are more cautious about the use of antidiabetic drugs.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Fifth Affiliated of Guangzhou Medical University Research Ethics Committee (GYWY-L2023-64). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

C-pZ and W-pS performed the general development and design of the study and contributed to critical revisions. X-yZ is the first author who contributed to the design, statistical analysis and manuscript writing of the study. S-fH wrote sections of the manuscript. S-fH, J-xL, H-nZ, LX, X-zW, K-hG, LZ, TL, H-mY, M-rL, and X-yL contributed to data collection and helped perform the analysis with constructive discussion. All authors contributed to the article and approved the submitted version.

Acknowledgments

We thank all staff in department of endocrinology in The Fifth Affiliated Hospital of Guangzhou Medical University, Loudi Central Hospital and The First People's Hospital of Xiangtan.

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. Tian D, Nie W, Sun Y, Ye Q. The epidemiological features of the SARS-CoV-2 omicron subvariant BA.5 and its evasion of the neutralizing activity of vaccination and prior infection. *Vaccines (Basel)* (2022) 10(10):6. doi: 10.3390/vaccines10101699

2. Cariou B, Hadjadj S, Wargny M, Pichelin M, Al-Salameh A, Allix I, et al. Phenotypic characteristics and prognosis of inpatients with COVID-19 and diabetes: the CORONADO study. *Diabetologia* (2020) 63(8):1500–15. doi: 10.1007/s00125-020-05180-x

3. Shi Q, Zhang X, Jiang F, Zhang X, Hu N, Bimu C, et al. Clinical characteristics and risk factors for mortality of COVID-19 patients with diabetes in Wuhan, China: A two-center, retrospective study. *Diabetes Care* (2020) 43(7):1382–91. doi: 10.2337/dc20-0598

4. Singh AK, Khunti K. COVID-19 and diabetes. *Annu Rev Med* (2022) 73:129–47. doi: 10.1146/annurev-med-042220-011857

5. Wang L, Gao P, Zhang M, Huang Z, Zhang D, Deng Q, et al. Prevalence and ethnic pattern of diabetes and prediabetes in China in 2013. *JAMA* (2017) 317 (24):2515–23. doi: 10.1001/jama.2017.7596

6. Mithal A, Jevalikar G, Sharma R, Singh A, Farooqui KJ, Mahendru S, et al. High prevalence of diabetes and other comorbidities in hospitalized patients with COVID-19 in Delhi, India, and their association with outcomes. *Diabetes Metab Syndr* (2021) 15 (1):169–75. doi: 10.1016/j.dsx.2020.12.029

7. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *JAMA* (2020) 323(20):2052–9. doi: 10.1001/jama.2020.6775

8. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* (2020) 395(10223):497–506. doi: 10.1016/S0140-6736(20)30183-5

9. Kamrath C, Monkemoller K, Biester T, Rohrer TR, Warncke K, Hammersen J, et al. Ketoacidosis in children and adolescents with newly diagnosed type 1 diabetes during the COVID-19 pandemic in Germany. *JAMA* (2020) 324(8):801–4. doi: 10.1001/jama.2020.13445

10. Guo W, Li M, Dong Y, Zhou H, Zhang Z, Tian C, et al. Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes Metab Res Rev* (2020) 36(7): e3319. doi: 10.1002/dmrr.3319

11. Wu CT, Lidsky PV, Xiao Y, Lee IT, Cheng R, Nakayama T, et al. SARS-CoV-2 infects human pancreatic beta cells and elicits beta cell impairment. *Cell Metab* (2021) 33(8):1565–1576 e5. doi: 10.1016/j.cmet.2021.05.013

12. Muller JA, Gross R, Conzelmann C, Kruger J, Merle U, Steinhart J, et al. SARS-CoV-2 infects and replicates in cells of the human endocrine and exocrine pancreas. *Nat Metab* (2021) 3(2):149–65. doi: 10.1038/s42255-021-00347-1

13. Aluganti Narasimhulu C, Singla DK. Mechanisms of COVID-19 pathogenesis in diabetes. *Am J Physiol Heart Circ Physiol* (2022) 323(3):H403–20. doi: 10.1152/ajpheart.00204.2022

14. Pelle MC, Zaffina I, Provenzano M, Moirano G, Arturi F. COVID-19 and diabetes-Two giants colliding: From pathophysiology to management. *Front Endocrinol (Lausanne)* (2022) 13:974540. doi: 10.3389/fendo.2022.974540

15. Chen Y, Yang D, Cheng B, Chen J, Peng A, Yang C, et al. Clinical characteristics and outcomes of patients with diabetes and COVID-19 in association with glucose-lowering medication. *Diabetes Care* (2020) 43(7):1399–407. doi: 10.2337/dc20-0660

16. Yan Y, Yang Y, Wang F, Ren H, Zhang S, Shi X, et al. Clinical characteristics and outcomes of patients with severe covid-19 with diabetes. *BMJ Open Diabetes Res Care* (2020) 8(1). doi: 10.1136/bmjdrc-2020-001343

17. McGurnaghan SJ, Weir A, Bishop J, Kennedy S, Blackbourn LAK, McAllister DA, et al. Risks of and risk factors for COVID-19 disease in people with diabetes: a cohort study of the total population of Scotland. *Lancet Diabetes Endocrinol* (2021) 9 (2):82–93. doi: 10.1016/S2213-8587(20)30405-8

18. Baidya A, Singh SK, Bajaj S, Zargar AH, Singh P, Das S, et al. Diabetes and COVID-19: A review. *J ASEAN Fed Endocr Soc* (2020) 35(1):40–8. doi: 10.15605/jafes.035.01.06

19. Conway J, Gould A, Westley R, Raju SA, Oklopcic A, Broadbent A, et al. Characteristics of patients with diabetes hospitalised for COVID-19 infection-a brief case series report. *Diabetes Res Clin Pract* (2020) 169:108460. doi: 10.1016/j.diabres.2020.108460

20. Vieira R, Souto SB, Sanchez-Lopez E, Machado AL, Severino P, Jose S, et al. Sugar-lowering drugs for type 2 diabetes mellitus and metabolic syndrome-strategies for in Vivo administration: Part-II. *J Clin Med* (2019) 8(9):1. doi: 10.3390/jcm8091332

21. De Santi F, Zoppini G, Locatelli F, Finocchio E, Cappa V, Dauriz M, et al. Type 2 diabetes is associated with an increased prevalence of respiratory symptoms as compared to the general population. *BMC Pulm Med* (2017) 17(1):101. doi: 10.1186/s12890-017-0443-1

22. Glaser S, Kruger S, Merkel M, Bramlage P, Herth FJ. Chronic obstructive pulmonary disease and diabetes mellitus: a systematic review of the literature. *Respiration* (2015) 89(3):253–64. doi: 10.1159/000369863

23. Klein OL, Smith LJ, Tipping M, Peng J, Williams MV. Reduced diffusion lung capacity in patients with type 2 diabetes mellitus predicts hospitalization for pneumonia. *Diabetes Res Clin Pract* (2011) 92(1):e12–5. doi: 10.1016/j.diabres.2010.12.012

24. Bi Y, Zhang L, Li X, Kan Y, Li S, Zou Y, et al. Contributing factors of fatigue in patients with type 2 diabetes: A systematic review. *Psychoneuroendocrinology* (2021) 130:105280. doi: 10.1016/j.psyneuen.2021.105280

25. Stewart KJ. Role of exercise training on cardiovascular disease in persons who have type 2 diabetes and hypertension. *Cardiol Clin* (2004) 22(4):569–86. doi: 10.1016/j.ccl.2004.06.007

26. Wang Z, Du Z, Zhu F. Glycosylated hemoglobin is associated with systemic inflammation, hypercoagulability, and prognosis of COVID-19 patients. *Diabetes Res Clin Pract* (2020) 164:108214. doi: 10.1016/j.diabres.2020.108214

27. Zheng H, Wu J, Jin Z, Yan LJ. Potential biochemical mechanisms of lung injury in diabetes. *Aging Dis* (2017) 8(1):7–16. doi: 10.14336/AD.2016.0627

28. Klein OL, Krishnan JA, Glick S, Smith LJ. Systematic review of the association between lung function and Type 2 diabetes mellitus. *Diabetes Med* (2010) 27(9):977–87. doi: 10.1111/j.1464-5491.2010.03073.x

29. Freeman J. Management of hypoglycemia in older adults with type 2 diabetes. *Postgrad Med* (2019) 131(4):241–50. doi: 10.1080/00325481.2019.1578590

30. Vinik AI, Maser RE, Mitchell BD, Freeman R. Diabetic autonomic neuropathy. *Diabetes Care* (2003) 26(5):1553–79. doi: 10.2337/diacare.26.5.1553

31. Agyeman AA, Chin KL, Landersdorfer CB, Liew D, Ofori-Asenso R. Smell and taste dysfunction in patients with COVID-19: A systematic review and meta-analysis. *Mayo Clin Proc* (2020) 95(8):1621–31. doi: 10.1016/j.mayocp.2020.05.030

32. Li YC, Bai WZ, Hashikawa T. The neuroinvasive potential of SARS-CoV2 may play a role in the respiratory failure of COVID-19 patients. *J Med Virol* (2020) 92 (6):552–5. doi: 10.1002/jmv.25728

33. Di Renzo L, Gualtieri P, Pivari F, Soldati L, Attina A, Cinelli G, et al. Eating habits and lifestyle changes during COVID-19 lockdown: an Italian survey. *J Transl Med* (2020) 18(1):229. doi: 10.1186/s12967-020-02399-5

34. Hoier A, Chaaban N, Andersen BV. Possibilities for maintaining appetite in recovering COVID-19 patients. *Foods* (2021) 10(2):13. doi: 10.3390/foods10020464

35. Hirsch JS, Ng JH, Ross DW, Sharma P, Shah HH, Barnett RL, et al. Acute kidney injury in patients hospitalized with COVID-19. *Kidney Int* (2020) 98(1):209–18. doi: 10.1016/j.kint.2020.05.006

36. Chu KH, Tsang WK, Tang CS, Lam MF, Lai FM, To KF, et al. Acute renal impairment in coronavirus-associated severe acute respiratory syndrome. *Kidney Int* (2005) 67(2):698–705. doi: 10.1111/j.1523-1755.2005.67130.x

37. Soleimani M. Acute kidney injury in SARS-coV-2 infection: Direct effect of virus on kidney proximal tubule cells. *Int J Mol Sci* (2020) 21(9):2. doi: 10.3390/ijms21093275

38. Soler MJ, Wysocki J, Batlle D. ACE2 alterations in kidney disease. *Nephrol Dial Transplant* (2013) 28(11):2687–97. doi: 10.1093/ndt/gft320

39. Alguwaihes AM, Al-Sofiani ME, Megdad M, Albader SS, Alsari MH, Alelayan A, et al. Diabetes and Covid-19 among hospitalized patients in Saudi Arabia: a single-centre retrospective study. *Cardiovasc Diabetol* (2020) 19(1):205. doi: 10.1186/s12933-020-01184-4

40. Li J, Wang X, Chen J, Zuo X, Zhang H, Deng A. COVID-19 infection may cause ketosis and ketoacidosis. *Diabetes Obes Metab* (2020) 22(10):1935–41. doi: 10.1111/dom.14057

41. Hamming I, Timens W, Bulthuis ML, Lely AT, Navis G, van Goor H, et al. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J Pathol* (2004) 203(2):631–7. doi: 10.1002/path.1570

42. Bornstein SR, Rubino F, Khunti K, Mingrone G, Hopkins D, Birkenfeld AL, et al. Practical recommendations for the management of diabetes in patients with COVID-19. *Lancet Diabetes Endocrinol* (2020) 8(6):546–50. doi: 10.1016/S2213-8587(20)30152-2

43. Rayman G, Lumb A, Kennon B, Cottrell C, Nagi D, Page E, et al. Guidance on the management of Diabetic Ketoacidosis in the exceptional circumstances of the COVID-19 pandemic. *Diabetes Med* (2020) 37(7):1214–6. doi: 10.1111/dme.14328

44. Lim S, Bae JH, Kwon HS. COVID-19 and diabetes mellitus: from pathophysiology to clinical management. *Nat Rev Endocrinol* (2021) 17(1):11–30. doi: 10.1038/s41574-020-00435-4

45. Sinha S, Cheng K, Schaffer AA, Aldape K, Schiff E, Ruppin E, et al. In vitro and in vivo identification of clinically approved drugs that modify ACE2 expression. *Mol Syst Biol* (2020) 16(7):e9628. doi: 10.15252/msb.20209628

46. Scheen AJ. SGLT2 inhibition during the COVID-19 epidemic: Friend or foe? *Diabetes Metab* (2020) 46(5):343–4. doi: 10.1016/j.diabet.2020.06.003

47. Nguyen NN, Ho DS, Nguyen HS, Ho DKN, Li HY, Lin CY, et al. Preadmission use of antidiabetic medications and mortality among patients with COVID-19 having type 2 diabetes: A meta-analysis. *Metabolism* (2022) 131:155196. doi: 10.1016/j.metabol.2022.155196

48. Carboni E, Carta AR, Carboni E. Can pioglitazone be potentially useful therapeutically in treating patients with COVID-19? *Med Hypotheses* (2020) 140:109776. doi: 10.1016/j.mehy.2020.109776

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 06 September 2023
DOI 10.3389/fpubh.2023.1264099

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Narendra Kotwal,
Armed Forces Medical College, India
Rade Vukovic,
The Institute for Health Protection of Mother and Child Serbia, Serbia

*CORRESPONDENCE
Gordana Bukara-Radujkovic
✉ gocabr@inecco.com

RECEIVED 20 July 2023
ACCEPTED 18 August 2023
PUBLISHED 06 September 2023

CITATION
Bukara-Radujkovic G, Miljkovic V and Ljuboja O (2023) Evidence of increasing incidence of type 1 diabetes and ketoacidosis among children in the Republic of Srpska in period 2017–2022 with special focus on COVID-19 global pandemic years. *Front. Public Health* 11:1264099. doi: 10.3389/fpubh.2023.1264099

Evidence of increasing incidence of type 1 diabetes and ketoacidosis among children in the Republic of Srpska in period 2017–2022 with special focus on COVID-19 global pandemic years

Gordana Bukara-Radujkovic[1,2]*, Vesna Miljkovic[1] and Olivera Ljuboja[1,2]

[1]Pediatric Clinic, University Clinical Center of the Republic of Srpska, Banja Luka, Bosnia and Herzegovina, [2]Department of Pediatric, Faculty of Medicine, University of Banja Luka, Banja Luka, Bosnia and Herzegovina

Background and objectives: Primary focus of the research was to determine the incidence of type 1 diabetes mellitus in the period from 2017 to 2022, and whether COVID-19 had an impact on the increase in the number of newly diagnosed children with diabetes type 1 under the age of 15 in the Republic of Srpska (Bosnia and Herzegovina). In the period 2001–2016 the incidence of type 1 diabetes mellitus was 11/100,000, with an annual increasing rate of 14.2%.

Methods: Available data from pediatric endocrinology clinics, in the Republic of Srpska, on the number of newly diagnosed patients with diabetes mellitus in the period from January 1, 2017 until December 31, 2022 were used. A retrospective analysis was performed, and the capture-recapture method was used for the final assessment, and the obtained result corresponds to about 99% of the population.

Results: The total number of children in the group of 0–14 years of age diagnosed with type 1 diabetes mellitus in this period was 183, of which 96 (52.46%) were boys, and 87 (47.54%) were girls. The average age at which diabetes mellitus was diagnosed was 8.3 ± 3.9 years. Average incidence of diabetes in the period 2017–2022 was 19/100,000 (95% CI 13.1–25.0). The highest incidence was 28.7/100,000 in 2020, the first year of the global COVID-19 pandemic. Out of a total of 183 newly diagnosed cases in the period 2017–2022, 73 (39.9%) were diagnosed with ketoacidosis upon admission. The largest number of newly diagnosed children was recorded in the group of children aged 10–14 years.

Conclusion: In the last 6 years, there has been a significant increase in the incidence of type 1 diabetes mellitus in children under the age of 15. With an incidence of 19.4/100,000 in the Republic of Srpska, we entered the group of countries with high-risk for diabetes. Further steps must focus on the education of the entire society in order to recognize the symptoms of the disease in time and prevent the occurrence of ketoacidosis, which could significantly reduce the burden on health systems, especially in times of global pandemics, such as the COVID-19 pandemic.

KEYWORDS

type 1 diabetes mellitus, COVID-19, incidence, ketoacidosis, children

1. Introduction

Diabetes mellitus type 1 is a chronic disease defined by a complete lack of insulin caused by the autoimmune destruction of pancreatic beta cells, which leads to an increase in glycaemia. This requires lifelong insulin replacement and lifestyle changes with regular glycemic control (1). It is also one of the most common endocrine and metabolic diseases in childhood (2).

In 2021, the number of newly discovered cases of diabetes mellitus in children under the age of 19 was 355,900 (95% CI: 334,200–377,300), and estimation indicate that by 2050, this number could be higher for 100,000 children per year (3). Unfortunately, this unfavorable trend also reflects on our country. The previous study of type 1 diabetes mellitus incidence in the Republic of Srpska for the period from 2001 to 2016, showed that the incidence was 11/100,000 children, which put us in the group of countries with a medium risk for the disease (4). In that study, the largest number of patients with diabetes mellitus type 1 was in the group of children aged 10–14, unlike most countries where the largest number of patients were in the group of children under 5 years of age, which was explained by the "accelerator hypothesis" (5).

Continuing the previous research, we wanted to see if the growth trend continued in the coming period from 2017 to 2022. Compared to earlier research, this research is specific because it covered the entire period of the COVID-19 infection and it was interesting to see the possible impact of COVID-19 on the number of children with type 1 diabetes mellitus.

On March 11, 2020 the World Health Organization (WHO) declared a global pandemic caused by one strain of coronavirus—SARS CoV-2, calling it COVID-19. The first published studies suggested that children tolerated infection better than adults (6, 7). However, in an effort to prevent the spread and emergence of new cases of this infection in the Republic of Srpska, as in most countries of the world, restrictive measures (isolation, home learning, ban on extracurricular and sports activities) were adopted during 2020 and 2021, which included citizens of all age, especially children. The incidence of diabetes mellitus in children aged 0–14 years in the period of the previous 6 years is extremely interesting and significant data, because it tells us not only about the potential impact of COVID-19 on the occurrence of diabetes (8), but also about the impact of the restrictive measures. Compared to 2019, the estimated incidence in 2021 according to the International Diabetes Federation (IDF) has increased (9), which indicates that globally there was an increase in newly discovered cases of diabetes mellitus in the period of 2020, the first year of the global COVID-19 pandemic.

The potential increase in the number of newly discovered cases of diabetes mellitus during the first year of the global pandemic can be explained by the mechanism of interaction of the virus with the target cell or, more precisely, the receptor site to which the virus binds. Angiotensin-converting enzyme 2 (ACE2) receptor is the binding site of SARS-CoV-1 and-2 viruses and these receptors are strongly expressed in pancreatic cells, and initial findings hypothesized that the SARS-CoV-1 virus enters pancreatic cells via the ACE2 receptor and leads to the destruction of β-cells and, consequently, the appearance of a diabetes mellitus (10). Although most studies did not examine direct exposure to the SARS-CoV-2 virus and the consequent occurrence of diabetes, it is assumed that the infection itself leads to an increase in the number of newly diagnosed type 1 diabetes, accelerating the onset of autoantibodies that destroy β-cells (11). Children initially had milder symptoms of the virus infection, often with negative tests, which were not highly specific and internationally standardized, so there is a possibility that they were exposed to the virus itself, but unrecognized. Also, subsequent serological analyzes were not performed, which could confirm or reject this hypothesis.

Studies investigating the incidence of type 1 diabetes during the COVID-19 pandemic have shown mixed results. An analysis in London found an increase in newly diagnosed type 1 in the state of severe ketoacidosis with pronounced hypokalemia between March 23, 2020 and June 4, 2020, compared to the number of patients in the previous 5 years (11). In Canada during 2020, there was no significant increase in number of newly diagnosed patients comparing to 2019, but the number of ketoacidosis was significantly increased (12). In Germany, in the initial months of the COVID-19 infection (March, April, May 2020), the number of patients was lower compared to previous years, which was explained by restrictive measures and isolation that reduced exposure of infection of not just Sars-CoV-2 virus but, also, of the other common viruses in the pediatric population. However, the appearance of newly diagnosed patients with type 1 diabetes did not stop, which is explained by the fact that isolation led to the stress as a possible trigger for the onset of type 1 diabetes (13).

The most common complication of untreated diabetes mellitus in children is diabetic ketoacidosis, which can lead to death (1). Taking all of this into account, in this work we wanted to determine incidence of type 1 diabetes mellitus from 2017 to 2022, as well as the number of children with ketoacidosis. We will also determine whether there is a connection with the increased number of patients and the frequency of ketoacidosis in the period of 2020 and 2021, the period of COVID-19 global pandemic.

2. Materials and methods

The research covered the population of the Republic of Srpska from 0 to 14 years of age in the period 2017–2022. Republic of Srpska is an autonomous part of Bosnia and Herzegovina, which is located in the southeastern part of Europe, and according to data from the last census, it had 1,171,179 people. The last population census of Bosnia and Herzegovina was carried out in 2013, and since then the Republic Institute of Statistics of the Republic of Srpska estimates the number of people on the basis of demographic parameters every year and published the results in its yearbooks, and this data, as the only official one, were used in this research (14).

As there is no official register of patients with diabetes mellitus in the Republic of Srpska, yet, we used available data from pediatric endocrinology clinics in the country on the number of newly diagnosed patients with diabetes mellitus type 1 in the period from January 1, 2017 until December 31, 2022. Data were cross-referenced with official data from the Health Insurance Fund of the Republic of Srpska, which provides free insulin and blood sugar strips to patients up to 15 years of age. The capture-recapture method was used for the final assessment, and the obtained result corresponds to approximately 99% of the population. The criteria that the patients had to fulfill were: that they were diagnosed with type 1 diabetes in the period 2017–2022 by a doctor at one of the pediatric endocrinology clinics in the Republic of Srpska, that they were citizens of the Republic of Srpska

10.3389/fpubh.2023.1264099

at the time of diagnosis, and that they had between 0 and 14 years at the time of diagnosis. These criteria were taken from the EURODIAB study (15).

2.1. Statistical analysis

The incidence was calculated as the number of children diagnosed with type one diabetes mellitus per 100,000 children of the same age. The number of children of the same age was calculated based on the number that was on the census in 2013 and the ratio of the total number of people on the census to the estimated total number of people for each year according to the data from the Statistical Yearbook of the Republic Institute of Statistics of the Republic of Srpska. Numerical variables were examined based on measures of central distribution and variability, and categorical variables based on percentages and frequencies. Various tests were used in further statistical analysis, which will be presented in the results.

3. Results

One hundred and eighty three children in total were diagnosed with type 1 diabetes mellitus in period 2017–2022. 96 (52.46%) of them were boys, and 87 (47.54%) were girls. The average age at which diabetes mellitus was diagnosed was 8.3 ± 3.9 years.

As it can be seen from Table 1, in the period between 2017 and 2022, the incidence of diabetes mellitus in children aged 0–14 years was 19/100,000/yr., with a 95% confidence interval between 13.1 and 25.0 children (Poisson distribution). The highest incidence was established in 2020 at 28.7/100,000, and the lowest at 12.9/100,000 in 2017. After 2020 the incidence of type one diabetes is in a slight decline.

The Z-score for the incidence shows deviations from −2.41 to +4.18 in relation to the arithmetic mean, which clearly tells us that the incidence varied in this period. The year 2020 was the most interested one, because compared to 2019, the incidence increased by 28.13%, while in 2021, it decreased by 37% compared to 2020 (Figure 1).

TABLE 1 Incidence of type 1 diabetes mellitus per 100,000 children per year aged 0–14 in total and according to gender in the period from 2017 to 2022.

Year	Number of newly diagnosed children			Incidence per 100,000 children		
	All	Boys	Girls	All	Boys	Girls
2017	21	14	7	12.9	16.8	8.8
2018	25	13	12	15.5	15.7	15.3
2019	36	14	22	22.4	17.0	28.1
2020	46	22	24	28.7	26.8	30.8
2021	28	19	9	17.6	23.3	11.6
2022	27	14	13	17.0	17.2	16.8
Total	183	96	87	19.0	19.5	18.6

TABLE 2 Incidence of type 1 diabetes mellitus per 100,000 children per year by age groups in the period from 2017 to 2022.

Year	Number of diagnosed typed 1 diabetes by aged specific groups			Incidence per 100,000 children in aged specific groups		
	0–4	5–9	10–14	0–4	5–9	10–14
2017	4	6	11	7.6	11.3	19.5
2018	6	6	13	11.5	11.3	23.1
2019	8	12	16	15.3	22.7	28.6
2020	11	15	20	21.2	28.6	35.9
2021	6	9	13	11.7	17.3	23.5
2022	4	10	13	7.8	19.2	23.5
Total	39	58	86	12.5	18.4	26.1

3.1. Incidence by gender

Average incidence for boys in the period 2017–2022 was 19.5/100,000 of boys per year, while for girls it was 18.6/100,000. As we can see, in this period, boys have a higher incidence rate. However, if the standard arithmetic means of incidence for both groups are compared, it will be seen that both groups do not distinguish in average incidence rate (Levene's test $F = 3.376$, $p = 0.082$). Also, the age at which type 1 diabetes was diagnosed does not differ by gender specific group ($t = 0.288$, $p = 0.774$).

3.2. Incidence by age groups

We divided the sample in the three specific age groups, 0–4 years of age, then 5–9 and the last group of 10–14 years of age, and calculated the incidence for each group separately per 100,000 children. The results are given in Table 2. We can see that the lowest number of newly diagnosed type 1 diabetes is in the age group of 0–4, only 21% in relation to the total number of cases, therefore the incidence in this age group is also the lowest, 12.5/100,000 children per year. The largest number of cases is in the group of children aged 10–14, 86 (47% in relation to the total number of newly detected cases) with the highest incidence of 26.1/100,000 children per year.

Due to the small number of observed years ($n = 6$), we did not obtain statistically significant results in the incidence trend by age group.

3.3. Incidence depending on the state of admission before diagnosing the disease

In the past 6 years, a total of 73 cases of ketoacidosis or 39.9% of the total number of newly diagnosed type 1 diabetes mellitus were at the reception. Ketoacidosis, in relation to the total number of newly diagnosed cases of diabetes mellitus type 1, is the least present in the age group of 0–4 years of age (28.2%), while in the other two age groups it is present in approximately the same proportion. We can conclude that there is no statistically significant difference in the arithmetic means of groups of children who were admitted in a state of ketoacidosis or hyperglycemia in relation to age, because Levene's test shows $F = 1.26$ and $p = 0.723$. The largest number of ketoacidosis was detected in 2020, where the incidence is the highest. There was no statistically significant correlation between

gender and admission status, as Pearson's Chi-square test was 1.327, $p = 0.249$ (Table 3).

In further statistical analysis, we also took into account the distribution of ketoacidosis according to severity of it, based on blood pH and serum bicarbonate concentration (16). Mild ketoacidosis is defined as pH < 7.3 and serum bicarbonates HCO_3^- < 15 mEg/L, medium is defined as pH < 7.2 and serum bicarbonates HCO_3^- < 10 mEg/L, and severe as pH < 7.1 and serum bicarbonates HCO_3^- < 5 mEg/L. Table 4 shows the distribution of ketoacidosis by severity for every year of observed period.

Of the total number of ketoacidosis (73), which amounted to be the 39.9% of all cases of newly diagnosed diabetes mellitus type 1 in the period from 2017 to 2022 in the Republic of Srpska, the largest frequency was severe ketoacidosis (15.3% in total). The table shows that the percentage of ketoacidosis in total comparing to the number of all newly diagnosed type 1 diabetes, was higher in the period 2017–2019, or the pre-COVID-19 period. However, the most severe forms of ketoacidosis are significantly more prevalent in the COVID-19 period compared to the pre-COVID-19 period, while frequency of the mild forms of ketoacidosis have decreased in COVID-19 period compared to the total number of newly diagnosed type 1 diabetes.

Upon admission to hospitals in children diagnosed with ketoacidosis, the average value of HbA1c was 12.02%, while in children diagnosed only with hyperglycemia, HbA1c was 10.72%.

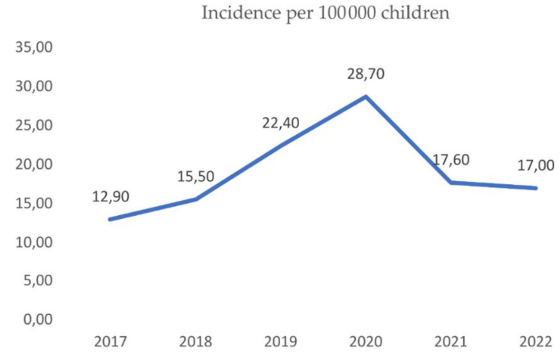


FIGURE 1
Incidence per 100,000 children in period of 2017–2022 in the Republic of Srpska, where we can see the linear trend till 2020, and decreasing in years after.

Also, the average BMI was lower in children diagnosed with ketoacidosis (16.64 kg/m^2), comparing to the children diagnosed only with hyperglycemia on admission (18.23 kg/m^2). We also analyzed if there was dependence of the severity of ketoacidosis by gender (Pearson Chi-Square 0.082, $p = 0.960$), as well as the year of admission (Pearson Chi-Square 5.794, $p = 0.670$), however, we did not obtain valid statistical significance between these variables.

3.4. Number of newly diagnosed cases per month of the years

In Figure 2, we see the distribution of newly diagnosed diabetes in children aged 0–14 according to the months of the year. It is noted that the largest number of newly diagnosed cases is in the winter months (December–March). In these periods, viral infections are commonly present, and it can be a trigger for the occurrence of diabetes in children (17), and this also coincides with the periods of the greatest restrictive measures during the COVID-19 pandemic in the Republic of Srpska.

4. Discussion

According to the results obtained in this study, we see that the incidence of diabetes mellitus type 1 in the period 2017–2022 increased compared to the period from 2001–2016, from 11/100,000 to 19.5/100,000 children aged 0–14 per year. The linear trend was observed in the period 2001–2016 ($r = 0.71$, $p \leq 0.002$) (4). Due to the small number of observed years ($n = 6$) in the period from 2017 to 2022, we did not obtain statistically significant results that would confirm a linear trend in this period ($r = 0.31$, $p \leq 0.54$). Combining the results from these two periods would probably yield statistically significant linear trends and this could be a topic for future research. The IDF estimates that in Bosnia and Herzegovina in 2021 the incidence rate is less than 10/100,000 children aged 0–14 (9), which does not correlate with our findings. This discrepancy is explained by the complicated state structure of Bosnia and Herzegovina, and the absence of a single national register of diabetes patients in the country, which would collect data in one place and facilitate planning and prevention of this disease (18). The annual growth of incidence is in line with study models that predict that by 2040, the number of people

TABLE 3 Number of ketoacidosis by age groups and years.

Year	Total number of newly diagnosed type 1 diabetes by specific age groups				Number of diagnosed ketoacidosis upon admission by specific age group (% of total number of newly diagnosed type 1 diabetes in that age group)			
	0–4	5–9	10–14	TOTAL	0–4	5–9	10–14	Total
2017	4	6	11	21	1 (25%)	3 (50%)	7 (63.6%)	11 (52.4%)
2018	6	6	13	25	3 (50%)	2 (33.3%)	8 (61.5%)	13 (52%)
2019	8	12	16	36	1 (12.5%)	2 (16.7%)	6 (37.5%)	9 (25%)
2020	11	15	20	46	4 (36.4%)	8 (53.3%)	6 (30%)	18 (39.1%)
2021	6	9	13	28	1 (16.7%)	5 (55.6%)	7 (53.8%)	13 (46.4%)
2022	4	10	13	27	1 (25.0%)	4 (40.0%)	4 (14.8%)	9 (33.3%)
Total	39	58	86	183	11 (28.2%)	24 (41.4%)	38 (44.2%)	73 (39.9%)

10.3389/fpubh.2023.1264099

TABLE 4 Number of ketoacidosis by severity per year (% of ketoacidosis in the total number of newly diagnosed cases per year).

Year	Mild	Medium	Severe	Total
2017	3 (14.3%)	3 (14.3%)	5 (23.8%)	11 (52.4%)
2018	6 (24.0%)	3 (12.0%)	4 (16.0%)	3 (52.0%)
2019	5 (13.8%)	2 (5.6%)	2 (5.6%)	9 (25.0%)
2020	5 (10.7%)	7 (15.2%)	6 (13.0%)	18 (39.1%)
2021	2 (7.1%)	5 (17.9%)	6 (21.4%)	13 (46.4%)
2022	2 (7.4%)	2 (7.4%)	5 (18.5%)	9 (33.33%)
Total	3 (12.6%)	22 (12.0%)	28 (15.3%)	73 (39.9%)

FIGURE 2
Number of newly diagnosed type 1 diabetes per month of diagnosis in period 2017–2022 in the Republic of Srpska.

suffering from diabetes mellitus type 1 will increase by 49% compared to 2020 (19). If we follow these trends in Bosnia and Herzegovina, there will be the increase not only in the incidence of diabetes but also in the number of ketoacidosis that will manifest when new cases are diagnosed. Currently, with an incidence of 19.5/100,000 children per year for the age group of 0–14 years in the period 2017–2022 we entered the group of high-risk countries for diabetes type 1 according to the WHO classification.

In comparison with available data from neighboring countries and regions, the incidence of type 1 diabetes mellitus in the Republic of Srpska is higher than in Serbia (14.3/100,000 for the period 2007–2017), Montenegro (19.2/100,000 for the period 2016–2020), Croatia (17.2/100,000 for the period 2004–2012), Vojvodina (11.9/100,000 for the period 2017–2021) (20–23). Our finding shows that the highest incidence of 28.7/100,000 was recorded in the first year of the global COVID-19 pandemic, i.e., in 2020, and compared to 2019, it increased by 28.13%. Growth trend like this, and even greater, can be also observed in the available data from the countries of the region. The recorded growth can be interpreted as the influence of external factors such as fear of the unknown, limited movement, reduction of contacts, increased intake of sweets in conditions of isolation, home learning. Other possible explanation is that the exposure to COVID-19 accelerate the onset of type 1 diabetes mellitus in children (11). The acceleration of the onset leads to more newly diagnosed patients in 2020 and fewer in 2021, and the fact that both of these years deviate from the linear trend seen in the years before 2020 suggests that COVID-19 can accelerate the onset of the type 1 diabetes. Also, this explains the decrease in 2021. Besides that, the decrease in incidence in 2021 can also be explained by the fact that the COVID-19 restrictive measures in this year were not so strict and the population got used to it. The children returned to their school duties, but still wore masks and thus protected themselves not only from COVID-19, but also from other common viruses that can be a trigger for a newly onset of type 1 diabetes (15). Another possible explanation of the decrease of newly diagnosed type one diabetes in 2021 is that in 2020 people started using more vitamin and mineral supplements and boosted their immune system. Effect of that supplement usage presented in 2021 in full potential. Connection between using the vitamins supplement during the COVID-19 pandemic and the onset of type one diabetes can be subject of some other research.

Out of the total number of patients, 39.9% were admitted with symptoms of ketoacidosis in this period as our results show. In America, new cases of ketoacidosis increased from 31 to 51% from 2008 to 2017, in Canada, the prevalence increased from 18.6 to 25.6% (12, 24). The percentage of children and adolescents presenting in a state of ketoacidosis at the time of diagnosis also varies between countries and depends on the average income. According to predictions, in poorer countries (in low-income countries) new patients with ketoacidosis should decrease from 80% in 1990 to 40% in 2050, while in highly developed countries the number of children should decrease to 20–30% in 2050 (3). Children may be misdiagnosed as newly diagnosed diabetes mellitus type 1 due to initial non-specific symptoms, which happened in as many as 24% of cases in America. In poorer countries, the percentage of wrong diagnosis is even higher. The highest percentage of newly diagnosed cases of ketoacidosis is expected in Africa and Asia, that is, countries with low incomes and poor countries (19).

Most of the published studies, which are still limited, record the increase in the incidence of type 1 diabetes and ketoacidosis during the pandemic period. We investigated the period before, during and after COVID-19 global pandemic and recorded a significant increase in the number of patients in the year of 2020 compared to previous years, and in 2021 we recorded a significant decrease in new cases. Also, a higher percentage of severe ketoacidosis was observed in the pandemic years compared to the previous period. Regular check-ups at physician were not possible due to the COVID-19 pandemic, so the parents probably overlooked the first symptoms of diabetes in children, and that could lead to increase number of severe ketoacidosis. In our country only one child had COVID-19 infection together with newly diagnosed type 1 diabetes. In that case COVID-19 was diagnosed on routine check-up after the patient admitted to hospital with the symptoms of severe ketoacidosis. In this patient the mild symptoms of COVID-19 infection did not require the steroids therapy. Since the serological assay for COVID-19 antibodies were not conducted on all newly diagnosed patient with type one diabetes, we do not have enough data that could connect the steroids usage as part of COVID-19 treatment and onset of type one diabetes. In America, only 4 children out of 187 patients (2.1%) had a COVID-19 infection at the time of illness. The results of this study are the most similar to ours, as the five-year retrospective incidence of type 1 diabetes before COVID-19 was shown, and it was showing the linear growth This growth continued during the COVID-19 pandemic, where the number of patients in a state of ketoacidosis also increased (25).

The results of the study in Canada did not show an increase in incidence of type 1 diabetes during the COVID-19 period, compared to pre-COVID-19 period. In Canada, a significant increase in the frequency of ketoacidosis among new patients was recorded in the pandemic compared to the pre-pandemic period (68.2% compared to 45.6% in the pre-pandemic period with $p < 0.001$) (13). The largest Polish National Diabetes Center did not record a statistically significant increase in 2020 compared to 2019, but a significant increase in diabetic ketoacidosis was recorded (35.2–47.53%, $p < 0.05$), which was accompanied by a significant increase in average HbA1c and pH values (26).

It will be interesting to observe the incidence in the next few years, in which we will get confirmation of whether COVID-19 accelerates the onset of type 1 diabetes. Our results show that this global pandemic has impacted the incidence of type 1 diabetes as well as the frequencies of ketoacidosis in our region. The exact mechanism of this influence has yet to be investigated.

5. Conclusion

The trend of increasing incidence of diabetes in the world is a well-known fact. However, what is worrying are prognostic data, according to which there will be even greater growth by 2050. In Republic of Srpska, as part of Bosnia and Herzegovina which is middle income country (27), it is expected that this increase should be higher than the European average, while the results of this study for the period 2017–2022 put us in the European average. Of particular concern is the fact that the average rate of increase in the incidence of type 1 diabetes in our region has significantly increased compared to previous observed periods in all age groups.

One of the ways of preventing diabetic ketoacidosis, which can be applied in all countries, is the education of the entire population on recognizing the symptoms of diabetes. This is especially important for any state of emergences such as the pandemic period, and therefore efforts should be made at the national level to prevent the occurrence of diabetic ketoacidosis. In case of any future pandemic, we should expect the higher number of newly diagnosed type 1 diabetes, and therefore the increased frequencies of ketoacidosis that could additionally burden the health systems, which are vulnerable in periods like this. So, findings of our research should warn us to have action plans in periods like this. Also, the prevention of the increase in the incidence of type 1 diabetes mellitus should be reflected in changing lifestyle habits, improving the quality of life through healthy eating and increased physical activity among young people.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee of University Clinical Center of the Republic of Srpska (protocol number: 01-19-82-2/23 from March 22, 2023). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

GB-R: Conceptualization, Supervision, Writing – original draft, Writing – review and editing, Investigation. VM: Formal analysis, Funding acquisition, Validation, Writing – review and editing. OL: Data curation, Investigation, Methodology, Writing – review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the study 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. Chiang JL, Maahs DM, Garvey KC, Hood KK, Laffel LM, Weinzimer SA, et al. Type 1 diabetes in children and adolescents: a position statement by the American Diabetes Association. *Diabetes Care*. (2018) 41:2026–44. doi: 10.2337/dci18-0023

2. Patterson C, Guariguata L, Dahlquist G, Soltész G, Ogle G, Silink M. Diabetes in the young - a global view and worldwide estimates of numbers of children with type 1 diabetes. *Diabetes Res Clin Pract*. (2014) 103:161–75. doi: 10.1016/j.diabres.2013.11.005

3. Ward ZJ, Yeh JM, Reddy CH, Gomber A, Ross C, Rittiphairoj T, et al. Estimating the total incidence of type 1 diabetes in children and adolescent aged 0-19 years from 1990 to 2050: a global simulation-based analysis. *Lancet Diabet Endocrinol*. (2022) 10:848–58. doi: 10.1016/S2213-8587(22)00276-5

4. Radujkovic Bukara G, Miljkovic V, Ljuboja O, Lakic S. Evidence for the increase in the incidence of type 1 diabetes in children aged 0-14 years in the Republika Srpska, 2001-2016. *Central Eur J Paed*. (2018) 14:61–7. doi: 10.5457/p2005-114.200

5. Wilkin TJ. The accelerator hypothesis: weight gain as the missing link between type 1 and type 2 diabetes. *Diabetologia*. (2001) 44:914–22. doi: 10.1007/s001250100548

6. Ludvigsson JF. Systematic review of COVID-19 in children shows milder cases and a better prognosis then adults. *Acta Paediatr*. (2020) 109:1088–95. doi: 10.1111/apa.15270

7. Borelli M, Corcione A, Castellano F, Fiori Nastro F, Santamarioa F. Coronavirus diasese 2019 in children. *Front Peadiatr*. (2021) 9:668484. doi: 10.3389/fped.2021.668484

8. DiMeglio LA, Albanese-O'Neill A, Munoz CE, Maahs DM. COVID-19 and children with diabetes – updates, unknowns and next steps: first, do not extrapolation. *Diabetes Care*. (2020) 43:2631–4. doi: 10.2337/dci20-0044

9. Ogle GD, James S, Dabelea D, Pihoker C, Svennson J, Maniam J, et al. Global estimates of incidence of type 1 diabetes in children and adolescent: results from the international diabetes federation atlas, 10th edition. *Diabetes Res Clin Pract*. (2022) 183:109083. doi: 10.1016/j.diabres.2021.109083

10. Yang JK, Lin SS, Ji XJ, Guo LM. Binding of SARS coronavirus to its receptor damages islets and causes acute diabetes. *Acta Diabetol*. (2010) 47:193–9. doi: 10.1007/s00592-009-0109-4

11. Unsworth R, Wallace S, Oliver NS, Yeung S, Kshirsagar A, Naidu H, et al. New-onset type 1 diabetes in children during COVID-19: multicenter regional findings in the U.K. *Diabetes Care*. (2020) 43:e170–1. doi: 10.2337/dc20-1551

12. Shulman R, Cohen E, Stukel TA, Diong C, Guttmann A. Examination of trends in diabetes incidence among children during the COVID-19 pandemic in Ontario, Canada, from march 2020 to September 2021. *JAMA Netw Open*. (2022) 5:e2223394. doi: 10.1001/jamanetworkopen.2022.23394

13. Tittel SR, Rosenbauer J, Kamarath C, Ziegler J, Reschke F, Hammarsen J, et al. For the DPV initiative. Did the COVID-19 lockdown affect the incidence of pediatric type 1 diabetes in Germany? *Diabetes Care*. (2020) 43:e172–3. doi: 10.2337/dc20-1633

14. Republic of Srpska Institute of Statistic. *Statistical Yearbook 2017-2022*. (2023). Available at: https://www.rzs.rs.ba

15. Karvonnen M, Vik-Kajander M. On behalf of the diabetes Mondiale (DIAMOND) project group. Incidence of childhood type 1 diabetes worldwide. *Diabetes Care*. (2000) 23:1516–26. doi: 10.2337/diacare.23.10.1516

16. Tumini S, Baki S, Kosteria I, Di Giuseppe I, Levantini G. Incidence of type 1 diabetes and factors associated with presence and severity of ketoacidosis at onset in children. *Acta Biomed*. (2022) 93:e2022009. doi: 10.23750/abm.v93i1.11694

17. Lönnrot M, Lynch KF, Larsson HE, Lernmark A, Rewers M, et al. On behalf of the TEDDY study group respiratory infections are temporally associated with initiation of type 1 diabetes autoimmunity: the TEDDY study. *Diabetologia*. (2017) 60:1931–40. doi: 10.1007/s00125-017-4365-5

18. Bak JCG, Serne EH, Kramer MHH, Nieuwdrop M, Vergheugt CL. National diabetes registry: do they make difference? *Acta Diabetol*. (2020) 58:267–78. doi: 10.1007/s00592-020-01576-8

19. Gregory GA, Robinson TG, Linklater SE, Wang F, Calagiuri S, de Beaufort C, et al. Global incidence, prevalence and mortality of type 1 diabetes in 2021 with projection to 2040: a modelling study. *Lancet*. (2022) 10:741–60. doi: 10.1016/S2213-8587(22)00218-2

20. Vukovic R, Jesic MD, Vorgucin I, Stankovic S, Folic N, Milenkovic T, et al. First report on the nationwide incidence of type 1 diabetes and ketoacidosis at onset in children in Serbia: a multicenter study. *Eur J Pediatr*. (2018) 177:1155–62. doi: 10.1007/s00431-018-3172-4

21. Raicevic M, Samardzic M, Soldatovic I, Curovic Popovic N, Vukovic R. Trends in nationwide incidence of pediatric type 1 diabetes in Montenegro during the last 30 years. *Front Endocrinol.* (2022) 13:991533. doi: 10.3389/fendo.2022.991533

22. Rojnic Putarek N, Ille J, Spehar Uroic A, Skrabic V, Stipancic G, Krnic N, et al. Incidence of type 1 diabetes mellitus in 0 to 14-yr-old children in Croatia-2004 to 2012 study. *Pediatr Diabetes*. (2015) 16:448–53. doi: 10.1111/pedi.12197

23. Vorgucin I, Savin M, Stankovic D, Miljkovic D, Ilic T, Simic MV, et al. Incidence of type 1 diabetes mellitus and characteristics of diabetic ketoacidosis in children and adolescent during the first two years of the COVID-19 pandemic in Vojvodina. *Medicina*. (2022) 58:2–9. doi: 10.3390/medicina58081013

24. Ebekozien OA, Noor N, Gallagher MP, Alonso GT. Type 1 diabetes and COVID-19: preliminary findings from a multicenter surveillance study in the U.S. *Diabetes Care*. (2020) 43:e83–5. doi: 10.2337/dc20-1088

25. Gottesman BL, Mas JY, Tanaka C, Longhurst CA, Kim JJ. Incidence of new-onset type 1 diabetes among US children during the COVID-19 global pandemic. *JAMA Pediatric*. (2022) 176:414–5. doi: 10.1001/jamapediatrics.2021.5801

26. Rusak E, Seget S, Macherski M, Furgal N, Dys P, Jarosz-Chobot P. Has the COVID-19 pandemic affected the prevalence of diabetic ketoacidosis in polish children with newly diagnosed type 1 diabetes? An example of the largest polish pediatric diabetes center (upper Silesia-Katowice, Poland). *Healthcare*. (2022) 10:348. doi: 10.3390/healthcare10020348

27. The World Bank. Overview of Bosnia and Herzegovina. (2023), Available at: https://data.worldbank.org/country/bosnia-and-herzegovina

frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 03 October 2023
DOI 10.3389/fendo.2023.1248676

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Mohamed Rafiullah,
King Saud University, Saudi Arabia
Cosmin Mihai Vesa,
University of Oradea, Romania

*CORRESPONDENCE
Lei Peng
penglei@njmu.edu.cn
Wei Gu
guwei154@163.com

RECEIVED 27 June 2023
ACCEPTED 15 September 2023
PUBLISHED 03 October 2023

CITATION
Li Y, Peng L and Gu W (2023) The published trend of studies on COVID-19 and diabetes: bibliometric analysis. *Front. Endocrinol.* 14:1248676. doi: 10.3389/fendo.2023.1248676

The published trend of studies on COVID-19 and diabetes: bibliometric analysis

Yuanyuan Li[1], Lei Peng[2*] and Wei Gu[1*]

[1]Department of Endocrinology, Children's Hospital of Nanjing Medical University, Nanjing, China, [2]Department of Gastroenterology, First Affiliated Hospital of Nanjing Medical University, Nanjing, China

Background: Since the COVID-19 pandemic outbreak, diabetes mellitus (DM) has been at the core of the confirmed risk factors for fatal or critical care unit-treated COVID-19 and COVID-19 related complications. Although relevant studies on DM have developed rapidly during the COVID-19 pandemic. However, the aforementioned research results have not been systematically quantified by means of bibliometric analysis.

Purpose: The purpose of this study is to provide a comprehensive analysis of the current status and trends of publications related to DM research during the COVID19 epidemic.

Methods: A bibliometric analysis was performed using the Web of Science database. In this study, we used citespace, R software and R-Bibliometrix to analyze keywords, most-cited authors, most-cited countries, most-cited global documents, and co-occurrence and co-citation networks.

Results: A total of 1688 publications was included in this study. Investigators from the United States contributed the most publications. The United States, China and Europe have the most collaboration with the other countries/regions. A total of 3355 institutions made contributions to this study. Of the top 10 institutions with the most publications, N8 Research Partnership showed the most centrality. Among the top 10 journals, Diabetes Research and Clinical Practice published the most articles. Among authors included, Khunti Kamlesh is rated first with 27 papers and has the highest centrality. The most frequently co-cited article is entitled "Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study". The most popular keywords included diabetes, mortality, diabetes, outcome, occurrences, risk, and type 1 diabetes.

Conclusion: This bibliometric study provides an overall picture of DM research and research trends during the COVID-19 pandemic and provides a basis for researchers to develop their next research strategies.

KEYWORDS

diabetes, COVID-19, bibliometric analysis, scientific collaboration, research trends

Introduction

The SARS-CoV-2 virus is acknowledged to be the causative agent for the acute respiratory infectious disease which has been named as coronavirus disease 2019 (COVID-19) by The World Health Organization Since 2019 (1, 2). In the three years to 24 June 2023, over 640 million clinically diagnosed infections have been reported on a global scale with 6.6 million coronavirus deaths, according to World Health Organization survey (3). To individuals, infection of COVID-19 wreaks havoc on health, and it also contributes an enormous burden to national health delivery system. We have witnessed high incidence of susceptibility crowd of COVID-19. Nonetheless, the prognosis of the elderly and patients with chronic diseases, such as cardiovascular or respiratory dysfunction, diabetes and cancer, seems to be significantly worse comparing with others with COVID-19 infection (4, 5).

By conducting researches of different scales in people of different racial origins, diabetes, as one of the most common chronic diseases worldwide, has been proven to be one independent risk factor associated with critical infection of COVID-19. Due to the damage of immune response, diabetic patients are susceptible to diverse types of infection, and may be at heightened risk for severe illness and more death (6). In a large-scale observational study conducted by Sweden, the proportion of critical cases with T2D was reported to be more than that in non-T2D patients after adjusting for age, gender socio-demographic factors, drug treatment and multiple comorbidities (7). Although the pathophysiological mechanisms between COVID-19 and diabetes are still being further explored, studies have confirmed the existence of a bidirectional interaction between the two disease states, with the relevant pathways mainly involving stress-induced pathways.

Bibliometric analysis is an interdisciplinary approach to conduct quantitative literature research. Based on published literature and references, researchers use this statistical analysis tool to establish connections between published literature and research hotspots and trends in certain academic fields, thus providing a quantitative investigation of the trends of a research topic (8). Compared with traditional reviews, bibliometric analysis has shown greater advantages in objectively presenting the internal conceptual structure and potential associations of a large body of literature. Despite the extensive research on DM conducted by scholars during the COVID-19 pandemic, there is still a lack of quantitative analysis to show the current status of the DM research literature related to COVID-19 to have a complete understanding of the relationship between the two. We aim to predict the publication trends in this research area by analyzing the countries, institutions, partnerships, co-cited papers and keywords of the published DM-related articles during the COVID pandemic based on the current research results.

Method

Data sources and search strategies

Web of Science Core Collection (WOS) database is commonly adopted to be used in bibliometric analysis, which provides comprehensive and multidisciplinary information statistically analysis (9). We comprehensively searched the included literature in the WOS database by publication time. All the relevant articles related to Diabetes and COVID-19 from 2019 to 2023 were searched (on June 24, 2023) by use of MeSH words. In this study, the following retrieval strategy was used: TI=(diabetes) AND (TI=(COVID 19) OR TI=(2019 novel coronavirus) OR TI=(coronavirus19) OR TI=(coronavirus disease 2019) OR TI=(2019-novel CoV) OR TI=(2019 ncov) OR TI=(COVID 2019) OR TI=(coronavirus 2019) OR TI=(nCoV-2019) OR TI=(ncovid19) OR TI=(2019-ncov) OR TI=(COVID-19) OR TI=(Severe acute respiratory syndrome coronavirus 2) OR TI=(SARS-CoV-2)).Literature restricted to the language and article type were further excluded. The detailed exclusion criteria were listed as follows: (1) meeting abstract, letter, editorial material, early access, corrections that were published as articles (2) the article was not written in English. Along the line, two researchers performed the literature search separately. Figure 1 showed the research flow chart.

Statistics and analysis

CiteSpace (6.1.R3) was used to analyze the included literature with the strongest citation bursts. By using CiteSpace (6.1.R3), co-citation analysis on countries, regions or institutions, co-citation analysis, biplot overlay of journals, and timeline view were performed. VOSviewer (1.6.18) was employed to provide visualization of publicly available data via analysis of bibliographic coupling, co-referencing or co-authorship affiliations. We carried out trend plots by analyzing the keywords' frequency using R software and the Bibliometrix package.The Bibliometrix package also was used to show the evolution of keyword topics over time and to make visualization of the features of the published issues.

Results

General characteristics of publications

As shown in Figure 1, a review of papers published from 2019 to 2023 was conducted, and 2749 publications were available by the search terms. Then we screened out 77 publications in languages other than English. Simultaneously, 979 publications (including conference abstracts, letters, editorial materials, early access, and corrections) were filtered out, resulting in a total of 1693 publications afterwards. After loading the data into CiteSpace, five publications in which there were formatting errors or duplicates were filtered out, resulting in a total of 1688 publications for inclusion in this study. The included publications had a total of 28921 citations, with an average of 17.13 citations per paper, and an H-index of 73.

Countries/regions

Publications included were from 112 countries or regions. Investigators from the United States contributed the most

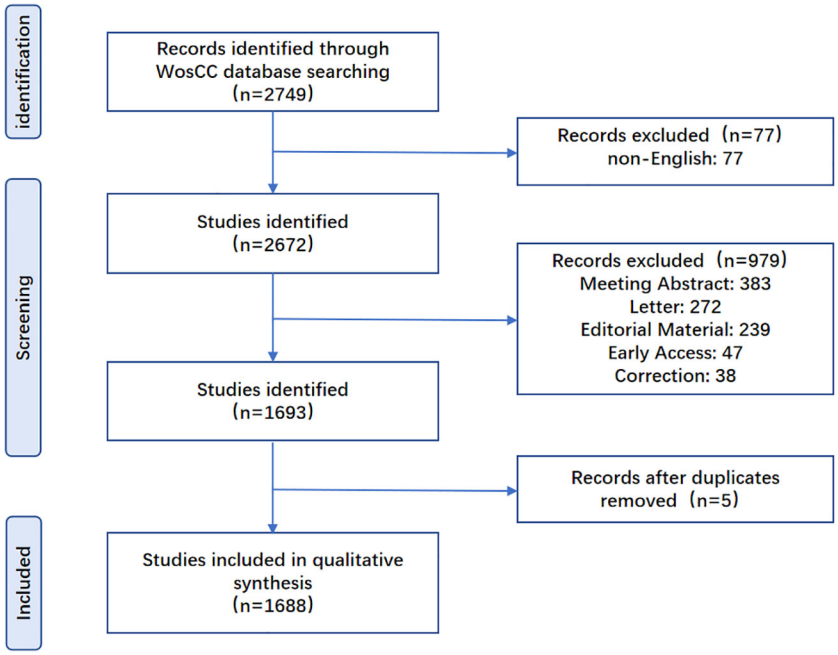


FIGURE 1
Flow chart of the data identification and screening results.

publications (n=341, 20.20% of the total; 5387 citations, mean 15.80 citations per paper), followed by China (n=172, 10.19%; 5423 citations, mean 31.53 citations per paper) and India (n=164, 9.72%; 4402 citations, mean 26.84 citations per paper) (Figure 2A). A total of 354 links and 112 nodes are depicted in Figure 2B to show the collaborative network between countries/regions. Each node symbolizes a country/region with a size proportional to the number of publications. The links between nodes represent the extent of collaboration between countries/regions. Of the top 10 countries/regions publishing the most papers, the Australia(0.24) displayed with the highest centrality, with USA and Italy(0.06) being the next highest, as shown in Table 1. Figures 2C, D display international collaboration between nations/regions. Of the top 10 published countries/regions, China, the United States and European countries have the most collaboration with the other countries/regions. Figure 2E shows that the tendency for global collaboration has become evident in targeting COVID-19.

Institutions

A total of 3355 institutions worldwide made contributions to these 1688 publications. CiteSpace generated a graphical visualization of the network of institutional collaborations, as shown in Figure 3. Those top 5 institutions with the most papers (Table 2) were the University of London, UDICE French Research University, Egyptian Knowledge Bank (EKB), Huazhong University of Science Technology, Institut National De La Sante Et De La Recherche Medicale (INSERM). The betweenness centrality (BC) value is a metric to assess the nodes' importance in a collaborative network. Of the top 10 institutions with the most publications, N8 Research Partnership showed the highest centrality (0.03), which was followed by the University of London, Universite Paris Cite and University of Oxford (0.02) as the most collaboration-oriented university.

Authors and co-citation authors

Figure 4A delineates the collaborative network among authors who authored more than 3 papers, with 334 authors conducting research on diabetes and COVID-19 from 2019 to 2023. In Table 3, Khunti Kamlesh is rated first with 27 papers. He has the most published papers in the field. and has the highest centrality (0.04). Through analysis of authors' co-citation networks, those with more than 20 citations were defined as key researchers (Figures 4B, C): Connections indicate collaboration between authors, and the size of the circles represents the amount of citations. Total link strength (TLS) indicates the impact of an author's publication on other contributing authors. The highest number of co-citations was recorded for Yang JK (n=410), followed by Guan WJ (n=327). The top 3 authors having the highest TLS were Yang JK, Guan WJ and Zhou F (Table 3).

Journals

From 2019 to 2023, 1,688 research articles related to COVID-19 and diabetes were published in 512 journals, 74 of which contained at least 5 articles. The 10 journals with the most published articles are listed in Table 4. Diabetes Research and Clinical Practice published the most articles (n=101), followed by Diabetes Metabolic Syndrome Clinical Research and Frontiers in Endocrinology (n=77 and 56,

10.3389/fendo.2023.1248676

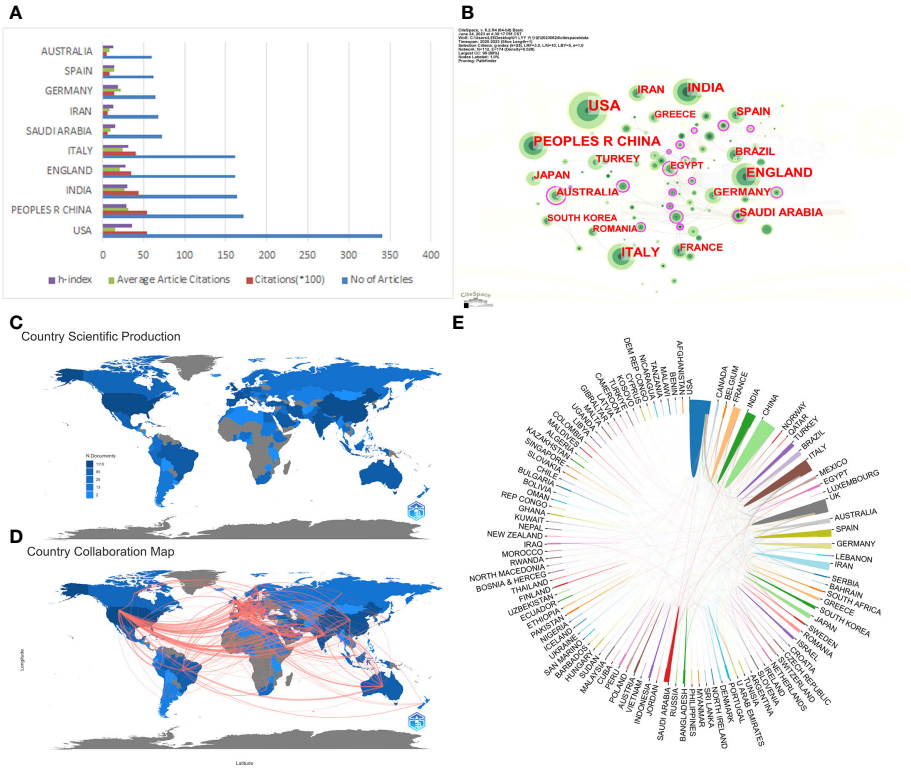


FIGURE 2
Cooperation network between countries. **(A)** The total publication number, total citations, average citation per paper, and H-index of the 10 most productive countries/regions; **(B)** The country collaboration network generated by Citespace; **(C, D)**: The country collaboration plotted on the world map; **(E)** Collaborative research between countries.

respectively). Additionally, the annual incidence of these 10 journals was generated by R-Bibliometrix to get a more specific picture of the trends in the number of publications of these journals across years (Figure 5A). A network visualization of the journal co-citation analysis was produced by VOS viewer, as shown in Figures 5B, C. Only journals that were visually cited at least 20 times were listed. Among the 461 journals that met the criteria, the top 3 journals that were frequently co-citation were DIABETES CARE (3435 times), DIABETES RESEARCH AND CLINICAL PRACTICE (1980 times), and NEW ENGLAND JOURNAL OF MEDICINE (1573 times) (Table 4).

TABLE 1 Top 10 countries/regions publishing the most papers in the field.

Rank	Country	No of Articles	Citations	Average Article Citations	H-index	Frequence	MCP_Ratio	Centrality
1	USA	341	5387	15.80	36	20.20%	0.177	0.06
2	PEOPLES R CHINA	172	5423	31.53	29	10.19%	0.151	0
3	INDIA	164	4402	26.84	30	9.72%	0.159	0
4	ENGLAND	162	3435	21.20	28	9.60%	0.355	0.03
5	ITALY	162	4015	24.78	31	9.60%	0.183	0.06
6	SAUDI ARABIA	72	654	9.08	15	4.27%	0.346	0.03
7	IRAN	68	559	8.22	13	4.03%	0.23	0
8	GERMANY	64	1450	22.66	19	3.79%	0.268	0
9	SPAIN	62	874	14.10	14	3.67%	0.114	0.03
10	AUSTRALIA	60	505	8.42	13	3.55%	0.257	0.24

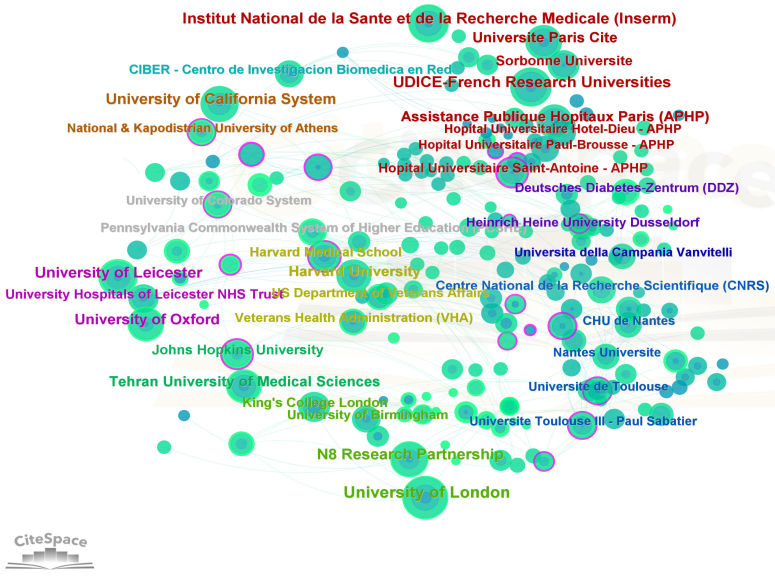


FIGURE 3
Cooperation network between major institutions.

Dual-map overlays

The literature of cited journals makes up the reference knowledge base, and the field of study of a highly cited journal represents an active interest or emerging field. We mapped and outlined the literature co-citation relationships for journal research areas using CiteSpace, with a graph of citing journals on the left and a graph of cited journals on the right. Remarkably, the colored paths shown in Figure 6 represent citation relationships in fields of highly active research. Published articles are concentrated in journals in the area of medicine, medical, and clinical, while most of the cited articles are published in journals in the area of molecular, biology, genetic, and health, nursing, and medicine.

Analysis of co-cited reference

In Figures 7A, B, a map of the 340 co-cited references that were cited more than 20 times is depicted. A cited reference is each represented by a node. The amount of co-cited references is characterized by the size of the node. Cross-references are indicated by the links between the nodes. The wider the connection, the higher the frequency of co-citations is indicated. The five most frequently co-cited references are listed in Table 5. The most frequently co-cited article is the article by Fei Zhou (2020) published in The Lancet (10), entitled "Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study". Then followed by Guo WN

TABLE 2 Top 10 institutions with the most papers in the field.

Rank	Institution	No of Articles	Citations	Average Article Citations	H-index	Frequence	Centrality
1	UNIVERSITY OF LONDON	44	651	14.80	14.8	2.61%	0.02
2	UDICE FRENCH RESEARCH UNIVERSITIES	37	1431	38.68	17	2.19%	0
3	EGYPTIAN KNOWLEDGE BANK EKB	35	426	12.17	12	2.07%	0
4	HUAZHONG UNIVERSITY OF SCIENCE TECHNOLOGY	34	3188	93.76	19	2.01%	0
5	INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE INSERM	34	1200	35.29	14	2.01%	0
6	UNIVERSITY OF LEICESTER	33	1579	47.85	14	1.96%	0
7	UNIVERSITY OF CALIFORNIA SYSTEM	32	549	17.16	11	1.90%	0.01
8	N8 RESEARCH PARTNERSHIP	31	299	9.65	11	1.84%	0.03
9	UNIVERSITE PARIS CITE	28	1305	46.61	15	1.66%	0.02
10	UNIVERSITY OF OXFORD	28	379	13.54	10	1.66%	0.02

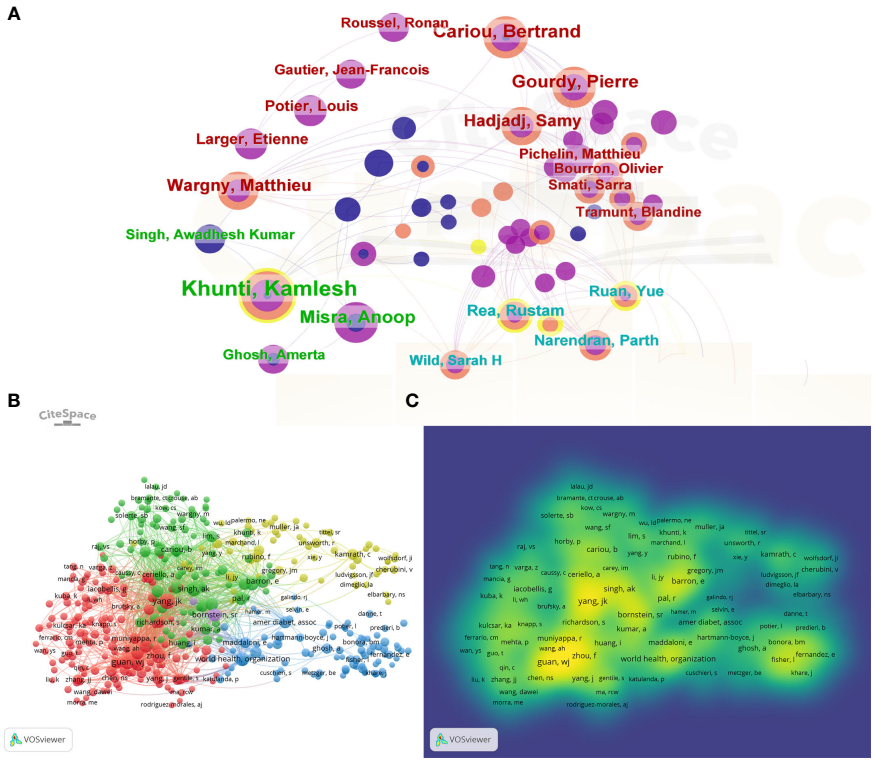


FIGURE 4
Map of collaboration networks of co-author analysis **(A)** Network visualization map of authors; **(B, C)** Overlay visualization map of authors. Network map showing authors' collaborations.

(2020), entitled "Diabetes is a risk factor for the progression and prognosis of COVID-19" was published in Diabetes Metab Res Rev (11); Yang JK(2010), entitled "Binding of SARS coronavirus to its receptor damages islets and causes acute diabetes" was published in acta diabetol (12); Barron E(2020), entitled "Associations of type 1 and type 2 diabetes with COVID-19-related mortality in England: a whole-population study" was published in lancet diabetes endo and Wu ZY (2020), entitled "Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention" was published in JAMA (5, 13).

Keywords

Table 6 shows the frequency of common occurrences of the top 10 key words. The most frequent word was diabetes mellitus (330 occurrences), followed by mortality (249 occurrences), mellitus (195 occurrences), outcome (183 occurrences), risk (168 occurrences), and type 1 diabetes (144 occurrences). Keywords with high centrality indicate hot spots in the field, with values between 0 and 1. In terms of centrality, the top 10 keywords were England, risk factors, coronavirus, pneumonia, mortality, coronavirus disease 2019, onset, impact, mellitus and receptor. The keyword cluster plots for COVID-19 and diabetes are shown in Figure 8A. The analysis of keywords was based on log-likelihood test cluster analysis of keyword co-occurrence analysis. There were 10 clusters obtained in this study. Details are shown in Figure 8B. the Q value (cluster module value) was 0.7892, indicating a significant clustering structure. In addition, the S-value, the mean profile value, was 0.9418, which indicated that the cluster members were highly homogeneous. The first 10 keyword clusters were chosen for analysis and are presented in Table 7. they are "covid-19 pandemic", "angiotensin-converting enzyme 2", "diabetic ketoacidosis", " wuhan", "diabetes distress", "type 1 diabetes", "diabetes mellitus", " inflammation", " risk factors", "gestational diabetes". The values for each cluster profile are greater than 0.5, suggesting a high degree of homogeneity and consistency in the clusters. Figure 8C shows the top 50 terms with the most intense outbreaks in the field. Figure 8D describes the trend of hot topics of the literature on COVID-19 and diabetes during 2020-2023.

Burstiness of keywords

CiteSpace was used for keyword burst detection (Figure 9). Keyword burstiness allows for representing new academic trends, foreshadowing future frontier research avenues, and highlighting potential topicalities in a discipline. Burstiness detection is shown as the red section of the blue timeline, indicating the onset year, finish year, and duration of the burst. A blue line is shown for the time

10.3389/fendo.2023.1248676

TABLE 3 Leading authors in the field.

Rank	Author	No of Articles	Citations	Average Article Citations	H-index	Frequence	Centrality
1	Khunti K	27	1540	57.04	12	1.60%	0.04
2	Cariou B	16	855	53.44	9	0.95%	0.01
3	Misra A	16	1554	97.13	12	0.95%	0.01
4	Gourdy P	15	854	56.93	9	0.89%	0.01
5	Wargny M	15	854	56.93	9	0.89%	0.01
6	Hadjadj S	14	845	60.36	8	0.83%	0.01
7	Holl RW	13	103	7.92	6	0.77%	0
8	Schiaffini R	12	288	24.00	7	0.71%	0
9	Schaan BD	11	158	14.36	4	0.65%	0
10	Yang Y	11	410	37.27	5	0.65%	0
Co-citation authors in the field							
Rank	Author	Co-citations	Total link strength				
1	Yang Jk	410	10181				
2	Guan Wj	327	6915				
3	Zhou F	303	6557				
4	Singh Ak	295	6522				
5	Guo Wn	250	5613				
6	Bornstein Sr	244	5613				
7	Pal R	243	4975				
8	Fadini Gp	188	4858				
9	Sardu C	181	4725				
10	Cariou B	202	4664				

line. We were particularly attracted to terms that were of research relevance in the top 25 keywords with the greatest outbreak intensity. These terms are representative of research trends for both the COVID-19 and diabetes fields (Figure 8). From 2020 to 2023, the highest outbreak intensity was observed for pneumonia (11), followed by receptors (6.56) and coronavirus (6.22). It is noteworthy that the burst of resistance and outbreak still continues.

Discussion

Since the COVID-19 pandemic, diabetes has been identified as a significant risk factor for increased mortality from severe COVID. To investigate the relationship between these two disease states, researchers have conducted a large number of studies and published numerous articles. This is the first time that a bibliometric review of all publications related to DM and COVID-19 has been conducted. The most published journals and the most cited articles were identified. The collaboration relationships across countries are depicted, and significant subjects in the field of research are discussed. Analysis of the leading journals and the most cited articles worldwide assisted in identifying potentially influential articles in the area. The networks of collaboration, trending keywords and thematic trends can be used as a reference for future research. In this study, a collection of 1688 articles was collected from the Web of Science Core Collection database. The United States published 341 papers related to diabetes and COVID-19 in the past three years, making it the country with the most research in related fields, followed by China (172 papers) and India (164 papers). Although the US was higher than China in terms of total number of papers, China ranked higher than the US in terms of total and average citation rates. We note that the 172 papers published by Chinese researchers in high impact factor journals include a large number of highly cited articles, and two of them have citation rates above 500 (14), suggesting the strong academic influence of China in the relevant research areas.

According to the latest data from the International Diabetes Federation, there are about 463 million people with diabetes worldwide, and the number is expected to reach 550 million in 2030, with China ranking first in the world in the number of people with diabetes (15, 16). The prevention and treatment of diabetes has become one of the common public health problems faced by the

10.3389/fendo.2023.1248676

TABLE 4 Top 10 journals with the most published and co-cited articles in the field.

Rank	Journal	No of Articles	Citations	Average Article Citations	h-index	Frequence
1	*DIABETES RESEARCH AND CLINICAL PRACTICE*	101	2399	23.75	27	5.98%
2	*DIABETES METABOLIC SYNDROME CLINICAL RESEARCH REVIEWS*	77	3858	50.10	27	4.56%
3	*FRONTIERS IN ENDOCRINOLOGY*	56	641	11.45	15	3.32%
4	*DIABETES CARE*	40	2095	52.38	23	2.37%
5	*PRIMARY CARE DIABETES*	37	309	8.35	10	2.19%
6	*ACTA DIABETOLOGICA*	34	526	15.47	14	2.01%
7	*INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH*	29	108	3.72	6	1.72%
8	*CUREUS JOURNAL OF MEDICAL SCIENCE*	27	33	1.22	4	1.60%
9	*DIABETIC MEDICINE*	21	339	16.14	10	1.24%
10	*JOURNAL OF DIABETES AND METABOLIC DISORDERS*	18	104	5.78	6	1.07%
Co-citation journals in the field						
Rank	Journal	Co-citations	Total link strength			
1	*DIABETES CARE*	3435	143531			
2	*DIABETES RES CLIN PR*	1980	85886			
3	*NEW ENGL J MED*	1573	76800			
4	*LANCET*	1461	70439			
5	*DIABETES METAB SYND*	1789	68390			
6	*JAMA-J AM MED ASSOC*	1308	62039			
7	*LANCET DIABETES ENDO*	1250	54733			
8	*DIABETOLOGIA*	908	47521			
9	*PLOS ONE*	672	37856			
10	*NATURE*	534	36196			

world (16, 17) By analyzing the data collected by the World Health Organization (WHO), we found that consistent with the results of the bibliometric analysis of those other COVID-19 studies not related to DM, the United States had the most studies on DM and COVID-related aspects, followed by China. Through bibliometric statistical analysis, we found that the majority of articles in relevant research publications on DM and COVID were also from the United States and China (18, 19). The analysis of research collaboration networks helps researchers to scientifically assess the next step in the scientific collaboration process and to select

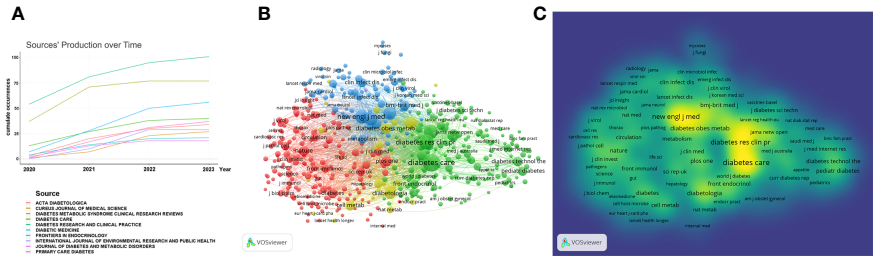


FIGURE 5
Map of collaboration networks of journal analysis. **(A)** The annual and cumulative numbers of research articles of journals; **(B, C)** The network visualization diagram of journal co-citation analysis of generated by VOSviewer. Overlay visualization map of journals.

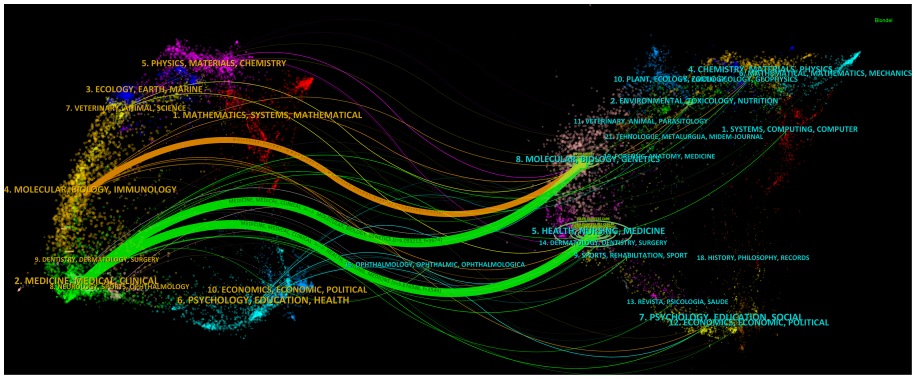


FIGURE 6
Dual map overlay of journals that contributed to publications on the use of artificial intelligence.

key collaborators. The analysis of this study found that among countries, China, the United States, and Europe collaborate most closely with other countries/regions, which may be related to their better economic status and higher research expenditures. In contrast, other developing countries remain the weak link in the current global collaboration due to their poorer economic level and insufficient research investment and infrastructure. This information may be crucial for researchers to select countries for further study exchanges or research collaborations.

At the research institution level, University of London, UDICE French Research University, Egyptian Knowledge Bank (EKB), Huazhong University of Science Technology, Institut National De La Sante Et De La Recherche Medicale(INSERM) are the five institutions with the highest number of COVID-related publications among all institutions. In England, there is a close interaction and cooperation between institutions, and N8 Research Partnership shows a central position. In terms of the number of published articles, the top ten research institutions are mainly universities of various countries, which may be related to the fact that universities invest more time, energy and resources in scientific research than other institutions and their relative emphasis on talent development. As far as the authors are concerned, Khunti Kamlesh has published a total of 27 papers and has made significant contributions to the field related to COVID and DM. Khunti Kamlesh's study on COVID-19 and DM primarily suggests that diabetes is an independent factor for in-hospital mortality in COVID-19 (5).

It was found that increased mortality associated with COVID-19 was associated with distant complications of diabetes, mainly with cardiovascular and renal complications. In addition, glycemic control and BMI were independent risk factors for elevated COVID-related mortality (20). Considering the continued epidemiological trend of COVID-19 and the continued increase in the prevalence of diabetic patients in the future, these findings are crucial for researchers to choose and adjust the future direction of their studies. By analyzing the co-cited articles of the authors, we found that Yang JK showed the greatest number of co-citations (n=410), followed by Guan WJ (n=327). Yang's study demonstrated that diabetes and environmental hyperglycemia were independent predictors of morbidity and mortality in SARS patients. The prognosis of SARS patients can be improved by metabolic control (21). A study by Zhou published in The Lancet showed risk factors associated with in-hospital mortality in adults with COVID-19, which was confirmed by the laboratories of Jinyintan Hospital and Wuhan Pulmonary Hospital (10). The analysis of literature co-citation rates and frequency of keyword occurrences can help to understand the main research directions, research hotspots and their evolution in related fields (22).

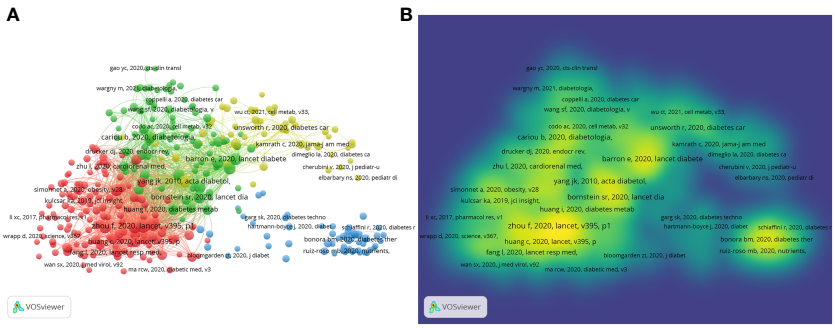


FIGURE 7
Map of co-cited references. **(A)** The network visualization diagram of references co-citation analysis; **(B)** References co-citation density chart.

TABLE 5 Top 10 most frequently co-cited references.

Rank	Author	Year	Journal	Citations	Total link strength	Cited reference
1	Zhou F	2020	*LANCET*	297	4968	Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study
2	Guo Wn	2020	*DIABETES-METAB RES*	250	4430	Diabetes is a risk factor for the progression and prognosis of COVID-19
3	Yang Jk	2010	*ACTA DIABETOL*	204	4008	Binding of SARS coronavirus to its receptor damages islets and causes acute diabetes
4	Barron E	2020	*LANCET DIABETES ENDO*	203	3204	Associations of type 1 and type 2 diabetes with COVID-19-related mortality in England: a whole-population study
5	Wu Zy	2020	*JAMA-J AM MED ASSOC*	201	3601	Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention
6	Bornstein Sr	2020	*LANCET DIABETES ENDO*	198	3438	Practical recommendations for the management of diabetes in patients with COVID-19
7	Yang Jk	2006	*DIABETIC MED*	197	4026	Plasma glucose levels and diabetes are independent predictors for mortality and morbidity in patients with SARS
8	Zhu Lh	2020	*CELL METAB*	196	3262	Association of Blood Glucose Control and Outcomes in Patients with COVID-19 and Pre-existing Type 2 Diabetes

Journal publication analysis can be useful in providing information to help researchers selecting appropriate journals for submitting their articles. In this research, the top 10 journals with 440 published articles related to DM and COVID-19 were found to be subspecialty journals. This could be due to the need for clinicians in this field to have subspecialty training. It is also exciting to note that subspecialty journals are not the only journals that publish relevant articles, but some general journals like Frontiers in Endocrinology also publish similar articles. Through in-depth analysis of keywords, we found that the primary keywords of high frequency in the present research focus on "COVID-19 and diabetes" were diabetes mellitus, mortality, mellitus, outcome, risk, and type 1 diabetes. The report shows that "increased COVID-19-related mortality associated with diabetes" is a hot topic for scholars these years. The high centrality of keywords such as "angiotensin-converting enzyme 2" and "inflammation" suggests that research on diabetes and COVID has partially shifted from epidemiology to pathogenesis studies. At this stage, there is also a significant increase

TABLE 6 Frequency and centrality of the top 10 key words.

Ranked by frequency				
Rank	Frequency	Centrality	Year	Key word
1	330	0	2020	diabetes mellitus
2	249	0.2	2020	mortality
3	195	0	2020	mellitus
4	183	0.22	2020	outcome
5	168	0.22	2020	risk
6	144	0.07	2020	type 1 diabetes
7	131	0	2020	type 2 diabetes
8	124	0.36	2020	coronavirus
9	118	0.07	2020	glycemic control
10	110	0.45	2020	impact

(Continued)

10.3389/fendo.2023.1248676

TABLE 6 Continued

Ranked by centrality				
Rank	Frequency	Centrality	Year	Key word
1	3	0.92	2023	england
2	86	0.76	2020	risk factors
3	124	0.72	2020	coronavirus
4	66	0.69	2020	pneumonia
5	249	0.68	2020	mortality
6	57	0.6	2020	coronavirus disease 2019
7	4	0.59	2023	onset
8	110	0.57	2020	impact
9	195	0.55	2020	mellitus
10	56	0.55	2020	receptor

in concern for adolescents and type 1 diabetes. Basic research and clinical trials related to the impact of the COVID-19 epidemic on adolescents and young adults (23), the effect on mood, leading to anxiety and depression (24), are actively being conducted along with the ongoing epidemic of COVID. Several investigations showed that non-enzymatic glycation of ACE2 receptors might be the pathogenic reason of the deteriorating outcome of COVID-19 disease in diabetic conditions (25, 26). Through analysis of keyword clusters, we identified that the hotspots of "COVID-19 and diabetes" research focus on type 1 diabetes and pathogenesis mechanism (angiotensin-converting enzyme 2, inflammation, cell). Recent studies have found increased morbidity and mortality in type 1 diabetes mellitus (T1DM) during the COVID-19 epidemic (27). The mechanism may be related to the excessive release of pro-inflammatory cytokines in the severe COVID-19 state. Notably, COVID-19 was demonstrated to lead to a severe imbalance in glucose homeostasis. Glucotoxicity can then synergize with inflammatory cytokine storms to promote oxidative stress, stimulate immune dysregulation, impair endothelial cell function and lead to a range of metabolic complications such as increased risk of thromboembolism and multi-organ damage, causing increased eventual patient mortality (28). It is a vicious circle. In

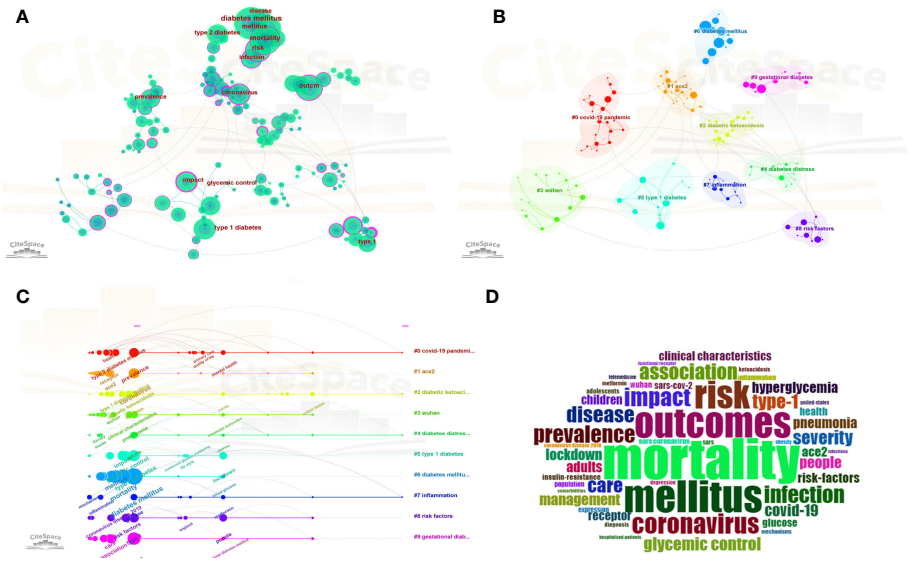


FIGURE 8
Map of Keyword clusters analysis. **(A)** Visualization of keyword co-occurrence; **(B)** map of Keyword clusters analysis; **(C)** timeline map of keyword; **(D)** word Cloud map regarding the keywords' frequency of occurrence generated from R studio Documents menu of Biblioshiny package.

10.3389/fendo.2023.1248676

TABLE 7 Top 10 keyword clusters in the field.

Cluster ID	Size	Silhouette	Cluster name
0	21	0.94	covid-19 pandemic
1	21	0.919	ace2
2	16	0.985	diabetic ketoacidosis
3	16	0.856	wuhan
4	12	0.98	diabetes distress
5	12	0.963	type 1 diabetes
6	11	1	diabetes mellitus
7	11	0.892	inflammation
8	10	0.894	risk factors
9	10	1	gestational diabetes

addition, COVID-19 can bind to the angiotensin-converting enzyme 2 (ACE2) receptor in pancreatic β-cells thereby leading to pancreatic β-cell destruction, which in turn promotes the development of diabetes (29). Also, some natural products such as quercetin, curcumin or other hypoglycemic agents would participant in COVID-19 and diabetes research (30–32). Nevertheless, long-term follow-up studies are still needed to assess the impact of COVID on the incidence, type, and complications of diabetes.

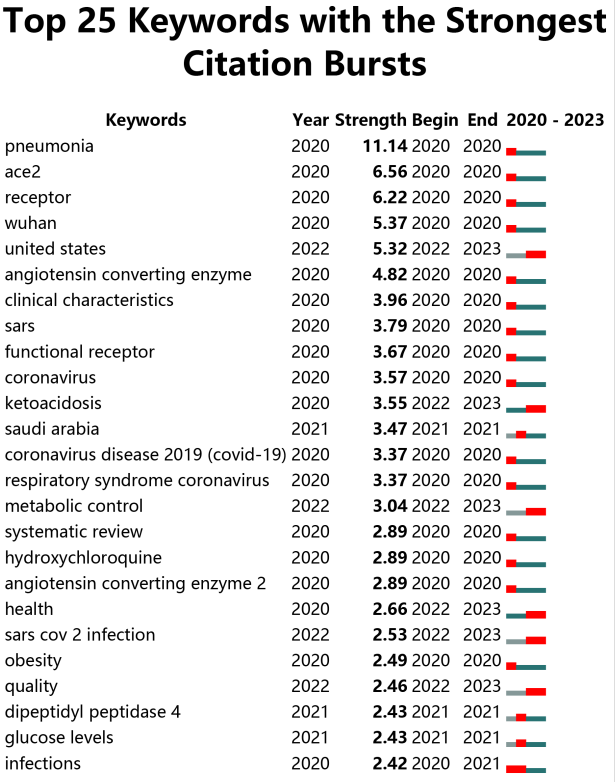


FIGURE 9
Visualization of Keyword Bursts.

Conclusion

This article provides the first presentation of a bibliometric evaluation of the publications on diabetes and COVID-19. This study also has some limitations, mainly in database selection, literature omissions due to time point limitations and citation analysis bias due to self-referencing. Despite these limitations, this bibliometric study still provides an overall picture of DM research and research trends during the COVID-19 pandemic and provides a basis for researchers to develop their next research strategies.

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 authors.

Author contributions

YL designed the study. YL and LP acquired the data and performed statistical analyses. YL and LP drafted the manuscript. YL and WG reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by National Nature Science Foundation of China (No. 82100595) and Natural Science Foundation of Jiangsu Province(No. BK20210958).

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. Akbari RAS, Mohammadi A, Koohi S. A new profiling approach for DNA sequences based on the nucleotides' physicochemical features for accurate analysis of SARS-CoV-2 genomes. *BMC Genomics* (2023) 24(1):266. doi: 10.1186/s12864-023-09373-7

2. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* (2020) 5(4):536–44. doi: 10.1038/s41564-020-0695-z

3. Zhu Y, Liu Y, Jiang H. Geriatric health care during the COVID-19 pandemic: managing the health crisis. *Clin Interv Aging* (2022) 17:1365–78. doi: 10.2147/CIA.S376519

4. Maddaloni E, D'Onofrio L, Siena A, Luordi C, Mignogna C, Amendolara R, et al. Impact of cardiovascular disease on clinical outcomes in hospitalized patients with Covid-19: a systematic review and meta-analysis. *Intern Emerg Med* (2021) 16(7):1975–85. doi: 10.1007/s11739-021-02804-x

5. Barron E, Bakhai C, Kar P, Weaver A, Bradley D, Ismail H, et al. Associations of type 1 and type 2 diabetes with COVID-19-related mortality in England: a whole-population study. *Lancet Diabetes Endocrinol* (2020) 8(10):813–22. doi: 10.1016/S2213-8587(20)30272-2

6. Rao KSS, Kaptoge S, Thompson A, Di Angelantonio E, Gao P, Sarwar N, et al. Diabetes mellitus, fasting glucose, and risk of cause-specific death. *N Engl J Med* (2011) 364(9):829–41. doi: 10.1056/NEJMoa1008862

7. Rawshani A, Kjolhede EA, Rawshani A, Sattar N, Eeg-Olofsson K, Adiels M, et al. Severe COVID-19 in people with type 1 and type 2 diabetes in Sweden: A nationwide retrospective cohort study. *Lancet Reg Health Eur* (2021) 4:100105. doi: 10.1016/j.lanepe.2021.100105

8. Ellegaard O, Wallin JA. The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics* (2015) 105(3):1809–31. doi: 10.1007/s11192-015-1645-z

9. Mehmood K, Mushtaq S, Bao Y, Saifullah, Bibi S, Yaseen M, et al. The impact of COVID-19 pandemic on air pollution: a global research framework, challenges, and future perspectives. *Environ Sci pollut Res Int* (2022) 29(35):52618–34. doi: 10.1007/s11356-022-19484-5

10. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* (2020) 395(10229):1054–62. doi: 10.1016/S0140-6736(20)30566-3

11. Guo W, Li M, Dong Y, Zhou H, Zhang Z, Tian C, et al. Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes Metab Res Rev* (2020) 36(7):e3319. doi: 10.1002/dmrr.3319

12. Yang JK, Lin SS, Ji XJ, Guo LM. Binding of SARS coronavirus to its receptor damages islets and causes acute diabetes. *Acta Diabetol* (2010) 47(3):193–9. doi: 10.1007/s00592-009-0109-4

13. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the chinese center for disease control and prevention. *JAMA* (2020) 323(13):1239–42. doi: 10.1001/jama.2020.2648

14. Zhu L, She ZG, Cheng X, Qin JJ, Zhang XJ, Cai J, et al. Association of blood glucose control and outcomes in patients with COVID-19 and pre-existing type 2 diabetes. *Cell Metab* (2020) 31(6):1068–1077.e3. doi: 10.1016/j.cmet.2020.04.021

15. Li MZ, Su LG, Liang BY, Tan JJ, Chen Q, Long JX, et al. Trends in prevalence, awareness, treatment, and control of diabetes mellitus in mainland China from 1979 to 2012. *Int J Endocrinol* (2013) 2013:753150. doi: 10.1155/2013/753150

16. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract* (2022) 183:109119. doi: 10.1016/j.diabres.2021.109119

17. Yan Y, Wu T, Zhang M, Li C, Liu Q, Li F. Prevalence, awareness and control of type 2 diabetes mellitus and risk factors in Chinese elderly population. *BMC Public Health* (2022) 22(1):1382. doi: 10.1186/s12889-022-13759-9

18. Zhao J, Zhu J, Huang C, Zhu X, Zhu Z, Wu Q, et al. Uncovering the information immunology journals transmitted for COVID-19: A bibliometric and visualization analysis. *Front Immunol* (2022) 13:1035151. doi: 10.3389/fimmu.2022.1035151

19. Ying J, Tan G, Zhang MW. Intellectual disability and COVID-19: A bibliometric review. *Front Psychiatry* (2022) 13:1052929. doi: 10.3389/fpsyt.2022.1052929

20. Holman N, Knighton P, Kar P, O'Keefe J, Curley M, Weaver A, et al. Risk factors for COVID-19-related mortality in people with type 1 and type 2 diabetes in England: a population-based cohort study. *Lancet Diabetes Endocrinol* (2020) 8(10):823–33. doi: 10.1016/S2213-8587(20)30271-0

21. Yang JK, Feng Y, Yuan MY, Yuan SY, Fu HJ, Wu BY, et al. Plasma glucose levels and diabetes are independent predictors for mortality and morbidity in patients with SARS. *Diabetes Med* (2006) 23(6):623–8. doi: 10.1111/j.1464-5491.2006.01861.x

22. Wang X, Li D, Huang X, Luo Q, Li X, Zhang X, et al. A bibliometric analysis and visualization of photothermal therapy on cancer. *Transl Cancer Res* (2021) 10(3):1204–15. doi: 10.21037/tcr-20-2961

23. Kamrath C, Rosenbauer J, Eckert AJ, Siedler K, Bartelt H, Klose D, et al. Incidence of type 1 diabetes in children and adolescents during the COVID-19 pandemic in Germany: results from the DPV registry. *Diabetes Care* (2022) 45(8):1762–71. doi: 10.2337/dc21-0969

24. Garcia-Lara RA, Gomez-Urquiza JL, Membrive-Jimenez MJ, Velando-Soriano A, Granados-Bolivar ME, Romero-Bejar JL, et al. Anxiety, Distress and Stress among Patients with Diabetes during COVID-19 Pandemic: A Systematic Review and Meta-Analysis. *J Pers Med* (2022) 12(9). doi: 10.3390/jpm12091412

25. Bassani D, Ragazzi E, Lapolla A, Sartore G, Moro S. Omicron variant of SARS-coV-2 virus: in silico evaluation of the possible impact on people affected by diabetes mellitus. *Front Endocrinol (Lausanne)* (2022) 13:847993. doi: 10.3389/fendo.2022.847993

26. Duarte T, Silva MM, Michelotti P, Barbosa N, Feltes BC, Dorn M, et al. The Drosophila melanogaster ACE2 ortholog genes are differently expressed in obesity/diabetes and aging models: Implications for COVID-19 pathology. *Biochim Biophys Acta Mol Basis Dis* (2022) 1868(12):166551. doi: 10.1016/j.bbadis.2022.166551

27. Aly HH, Fouda EM, Kotby AA, Magdy SM, Rezk AR, Nasef M. COVID-19-related multisystem inflammatory syndrome in children presenting with new-onset type 1 diabetes in severe ketoacidosis: A case series. *Diabetes Care* (2022) 45(4):983–9. doi: 10.2337/dc21-1094

28. Nouri-Keshtkar M, Taghizadeh S, Farhadi A, Ezaddoustdar A, Vesali S, Hosseini R, et al. Potential impact of diabetes and obesity on alveolar type 2 (AT2)-lipofibroblast (LIF) interactions after COVID-19 infection. *Front Cell Dev Biol* (2021) 9:676150. doi: 10.3389/fcell.2021.676150

29. Kountouri A, Korakas E, Ikonomidis I, Raptis A, Tentolouris N, Dimitriadis G, et al. Type 1 diabetes mellitus in the SARS-coV-2 pandemic: oxidative stress as a major pathophysiological mechanism linked to adverse clinical outcomes. *Antioxidants (Basel)* (2021) 10(5). doi: 10.3390/antiox10050752

30. Olah NKPS. The study of antioxidant capacity in extracts from vegetal sources with hypoglycaemic action. *Rev Chimie* (2019) 1(70):102–6. doi: 10.37358/RC.19.1.6860

31. Pawar A, Russo M, Rani I, Goswami K, Russo GL, Pal A. A critical evaluation of risk to reward ratio of quercetin supplementation for COVID-19 and associated comorbid conditions. *Phytother Res* (2022) 36(6):2394–415. doi: 10.1002/ptr.7461

32. Hassanizadeh S, Shojaei M, Bagherniya M, Orekhov AN, Sahebkar A. Effect of nano-curcumin on various diseases: A comprehensive review of clinical trials. *Biofactors* (2023) 49(3):512–33. doi: 10.1002/biof.1932

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 18 October 2023
DOI 10.3389/fpubh.2023.1297909

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Zeeshan Kaleem,
COMSATS University Islamabad,
Wah Campus, Pakistan
Fazal Wahab,
Northeastern University, China

*CORRESPONDENCE
Farman Ali
✉ farmankanju@gmail.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 20 September 2023
ACCEPTED 02 October 2023
PUBLISHED 18 October 2023

CITATION
Shoaib M, Sayed N, Shah B, Hussain T, AlZubi AA, AlZubi SA and Ali F (2023) Exploring transfer learning in chest radiographic images within the interplay between COVID-19 and diabetes.
Front. Public Health 11:1297909.
doi: 10.3389/fpubh.2023.1297909

Exploring transfer learning in chest radiographic images within the interplay between COVID-19 and diabetes

Muhammad Shoaib[1†], Nasir Sayed[2], Babar Shah[3], Tariq Hussain[4], Ahmad Ali AlZubi[5], Sufian Ahmad AlZubi[6] and Farman Ali[7*†]

[1]Department of Computer Science, CECOS University of IT and Emerging Sciences, Peshawar, Pakistan, [2]Department of Computer Science, Islamia College Peshawar, Peshawar, Pakistan, [3]College of Technological Innovation, Zayed University, Dubai, United Arab Emirates, [4]School of Computer Science and Technology, Zhejiang Gongshang University, Hangzhou, China, [5]Department of Computer Science, Community College, King Saud University, Riyadh, Saudi Arabia, [6]Faculty of Medicine, Jordan University of Science and Technology, Irbid, Jordan, [7]Department of Computer Science and Engineering, School of Convergence, College of Computing and Informatics, Sungkyunkwan University, Seoul, Republic of Korea

The intricate relationship between COVID-19 and diabetes has garnered increasing attention within the medical community. Emerging evidence suggests that individuals with diabetes may experience heightened vulnerability to COVID-19 and, in some cases, develop diabetes as a post-complication following the viral infection. Additionally, it has been observed that patients taking cough medicine containing steroids may face an elevated risk of developing diabetes, further underscoring the complex interplay between these health factors. Based on previous research, we implemented deep-learning models to diagnose the infection *via* chest x-ray images in coronavirus patients. Three Thousand (3000) x-rays of the chest are collected through freely available resources. A council-certified radiologist discovered images demonstrating the presence of COVID-19 disease. Inception-v3, ShuffleNet, Inception-ResNet-v2, and NASNet-Large, four standard convoluted neural networks, were trained by applying transfer learning on 2,440 chest x-rays from the dataset for examining COVID-19 disease in the pulmonary radiographic images examined. The results depicted a sensitivity rate of 98 % (98%) and a specificity rate of almost nightly percent (90%) while testing those models with the remaining 2080 images. In addition to the ratios of model sensitivity and specificity, in the receptor operating characteristics (ROC) graph, we have visually shown the precision vs. recall curve, the confusion metrics of each classification model, and a detailed quantitative analysis for COVID-19 detection. An automatic approach is also implemented to reconstruct the thermal maps and overlay them on the lung areas that might be affected by COVID-19. The same was proven true when interpreted by our accredited radiologist. Although the findings are encouraging, more research on a broader range of COVID-19 images must be carried out to achieve higher accuracy values. The data collection, concept implementations (in MATLAB 2021a), and assessments are accessible to the testing group.

KEYWORDS

COVID-19 disease, diabetes, transfer learning, disease detection, diagnosis using deep learning

1. Introduction

The intersection of COVID-19 and diabetes represents a multifaceted area of concern in contemporary healthcare. Diabetes, a chronic metabolic disorder characterized by high blood sugar levels, has emerged as a significant risk factor for severe COVID-19 outcomes (1). Emerging research has illuminated a complex relationship, revealing that individuals with diabetes are more susceptible to severe COVID-19 complications and adverse consequences, such as hospitalization and mortality. This heightened vulnerability is thought to be linked to the dysregulation of the immune system and impaired inflammatory response often associated with diabetes. The COVID-19 pandemic has raised concerns about the potential development of new-onset diabetes in individuals infected with the virus. Several studies have reported cases of acute or transient diabetes occurring in COVID-19 patients with no prior history of the condition (2). While the mechanisms behind this phenomenon remain under investigation, it is believed that the virus may directly impact pancreatic function or trigger an autoimmune response, resulting in temporary or long-term diabetes.

Beyond the realm of COVID-19, another facet of the diabetes narrative emerges in the context of cough medicines containing steroids (3). Steroids are known to influence blood sugar levels, and patients who require these medications to manage respiratory conditions such as asthma or chronic obstructive pulmonary disease (COPD) can face an increased risk of developing steroid-induced diabetes (4). Physicians must exercise caution and closely monitor patients with pre-existing diabetes or those at risk of developing the condition when prescribing such medications (5). However, the positive RT-PCR rate for the sample of nose swab samples is expected to be between 30 % and 60 % (30–60%) (6), resulting in undiagnosed patients that can infect a considerable amount of those people who are young and healthy (7). The daily use of the X-ray imaging method for diagnosing pneumonia is fast and straightforward. COVID-19 may be diagnosed with elevated Sensitivity using chest CT scans (8, 9). The images of chest X-ray images reveal sensory cues linked to the coronavirus (10). Multipolar involvement and opacities in the peripheral airspace are seen in chest imaging studies. Frosted glass (57 percent) and mixed mitigation (29 percent) are the most often mentioned opacities (11). A frosted glass pattern can be seen in areas bordering the pulmonary vessels at the start of COVID-19, which is challenging to determine visually (12). COVID-19 has also been linked to airspace opacities that are uneven or diffusely asymmetric (13). Expert radiologists are the only ones that can interpret these apparent anomalies. Automatic methods to detect these subtle anomalies may facilitate the diagnostic process and increase the early detection rate considerably, given many suspicious individuals and the small number of qualified radiologists.

The COVID-19 outbreak, generally regarded as the third coronavirus outbreak, affected over 209 countries, one of which was Pakistan. The COVID-19 epidemic, which first broke out in China, severely impacted the countries that border Pakistan, including China. China was also the country where the epidemic began. Italy had the highest mortality rate of COVID-19 in the western region, while Iran had the second-highest mortality rate in the northern part (14). Italy was also the country with the highest incidence of COVID-19. The COVID-19 virus was identified in Pakistan's first patient on February 26, 2020, by the Ministry of Health under the administration of the Pakistani government. The patient's location was determined to be in Karachi, which is the largest city and provincial capital of Sindh. On the same day, a second confirmed case was found in Islamabad, which is the location of the Federal Ministry of Health of Pakistan (15, 16). Within fifteen days, the total number of confirmed cases in the province of Sindh reached twenty (17) out of a total of 471 suspected cases. This was followed by the region of Gilgit Baltistan, which had the second-highest number of confirmed disease cases. All of the people whose cases have been verified have a history of having recently traveled from London, Tehran, or Syria. These reports are currently rising rapidly, which paints an even more dire picture of the situation than was previously presented.

The relationship between COVID-19 and diabetes is complex and multifaceted. People with diabetes are at increased risk of developing severe COVID-19, and COVID-19 can also worsen diabetes management. This is due to a variety of mechanisms, including increased ACE2 expression, insulin resistance, chronic inflammation, and cytokine storms. COVID-19 can also trigger new-onset diabetes in some people, and pregnant women with diabetes are at even higher risk of developing severe COVID-19. People with diabetes who have had COVID-19 may be more likely to experience long-term effects of the virus. It is important for people with diabetes to take steps to protect themselves from COVID-19 and to manage their diabetes carefully.

Artificial intelligence (AI) and deep learning solutions can be very effective in addressing these issues (18). Detailed reports documenting solutions for automated identification of coronavirus from chest X-ray images are not accessible at this time due to a shortage of public images of COVID-19 patients. A limited collection of data on images was recently obtained. This enables the researchers to create a machine-learning model that can diagnose COVID-19 *via* X-ray images of the chest (19). All of these photos were taken from research papers that reported on COVID-19 X-ray and C-Cmometric picture results. We re-labeled these X-ray images with a trained radiologist's aid, keeping just the simple sign of the coronavirus. Our radiologist defines these labeled X-ray images. Figure 1 shows three samples of images with their labeled regions. Then, as negative samples for COVID-19 identification, we used a subsection of medical images from the ChexPert dataset (20). The consolidated dataset (called COVID-Xray-3k) contains approximately 3,000 thoracic X-ray pictures, split into 2,100 training and 900 research samples.

In order to develop a reliable deep learning based COVID-19 detection model, the size of the dataset plays a significant role, and it has a direct impact on model generalization. For augmentation of the dataset, various image processing techniques were applied, including sharpening, blurring, contrast adjustment, intensity modification in the red, green, and blue channels, shearing effects, and rotation. The augmentation process enlarges the dataset size; the model receives a lot of COVID-19 image data to learn and recognize a broader spectrum of patterns and variations in chest X-ray images. Furthermore, data augmentation also contributes to clinical relevance. In medical imaging, patient diversity and variations in image quality are prevalent. Augmenting the dataset with various transformations helps the model better account for these real-world complexities. For instance, rotation and shearing effects mimic potential variations in patient positioning during imaging procedures, while adjustments in image intensity account for differences in equipment settings and patient characteristics.

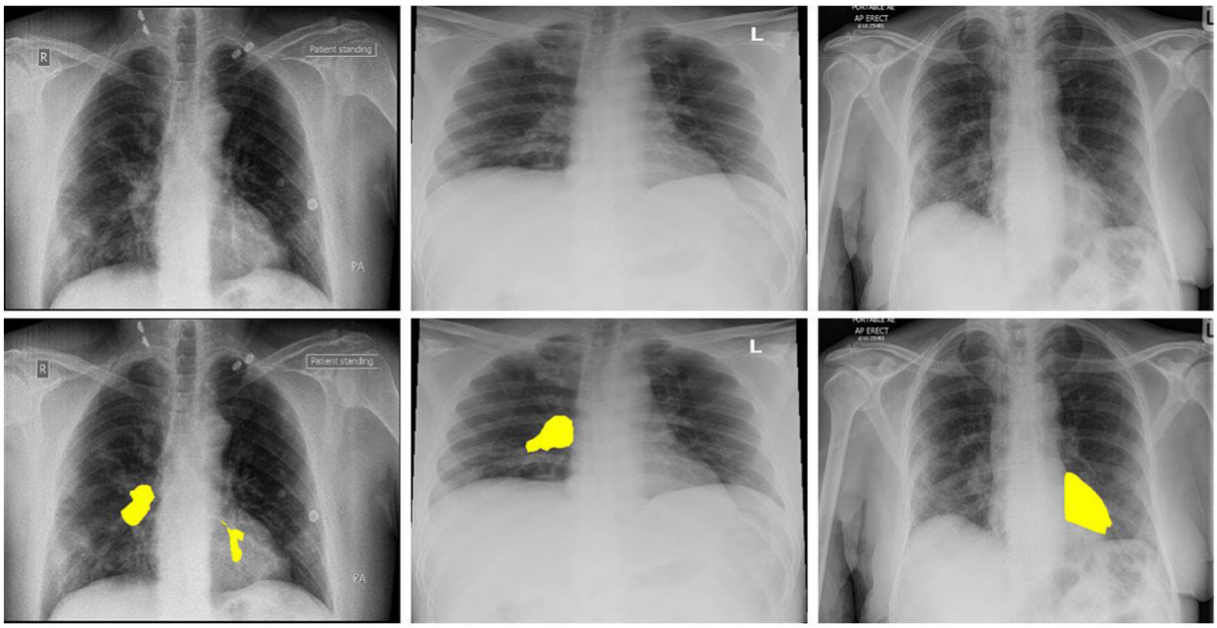


FIGURE 1
The above images are the 3 COVID-19 imageries samples and equivalent marked infected areas by our radiologist.

COVID-19 was predicted from thoracic X-ray images using a machine learning framework. We went in-depth on an end-to-end learning system that explicitly forecasts the raw images of COVID-19 diseases without the need to extract characteristics, in contrast to traditional approaches to the classification of images, also called medical image classification that adopts a two-step process (extraction of artisanal features – recognition). In recent years, studies (17, 21–23) have shown that in-depth learning-based models, i.e., Convolutive Neural Networks (CNN), surpass the traditional AI techniques in the domain of computer vision and medical image processing. Thus, these models are being applied to analyze problems ranging from classification, segmentation, and facial recognition to achieving high resolution and enhancing the images.

We use the COVID-Xray-3k dataset to create four standard convoluted networks that have shown promise in many tasks over the last few years and study their success in COVID-19 detection. The training steps could not be done from scratch for these networks since there are just a few widely accessible X-ray photos for the COVID-19 range. To resolve the issue of COVID-19 images absence, in this study, two techniques were used: We used the increased data to produce a modified version of COVID-19 pictures (such as spinning, a minor rotation, and inserting a small number of distortions) for increasing the images in the dataset by a factor of five. We optimize the former layer of a variant of the models on ImageNet rather than driving them from scratch. In this, the model can be built up with fewer tagged samples. These samples can be separated from each class in this manner. The two techniques described above aided in forming these networks using the accessible images and achieved good results on the test range of 30 0 0 images. We also quantify the trust interval of performance measurements since, in the COVID-19 class, the number of samples is small. The curves of receiver operating characteristics (ROC) and the region under or below the curve (AUC) of the proposed classification models are provided to summarize their output. Below are the article's significant contributions:

- To diagnose COVID-19 from pulmonary radiographs in the form of images, we prepared a data set of three thousand images with binary tags. For the testing group, this data collection should be used as a tool. A board-certified radiologist marks the pictures in the COVID-19 class. Only those images that were used for research purposes got clear and visible signs or marks.
- Using this dataset, we qualified four successful deep learning models and tested their output on a test collection of three thousand X-ray images. The top model that performed had a sensitivity rate of ninety-eight (98%) percent and a precision rate of ninety-two (92%) percent.
- We presented an experimental study based on the systematics of these models. This experimental study was a performance comparison between several CNN models where the performance evaluation is performed using the accuracy, F1-score, and the curve of ROC and AUC. The expected probability distribution for three classes is performed using the pie chart. Using a specific visualization method, we generated thermal maps of the most probable areas infected by COVID-19.
- This study leverages state-of-the-art CNN transfer learning models to design a sophisticated system capable of achieving heightened accuracy in the detection of two distinct categories: COVID-19 without comorbidity and COVID-19 with diabetes. Additionally, the system excels in precisely localizing the affected regions within X-ray images, providing valuable insights for medical diagnosis.

The objective of this study is to develop a deep-learning model for COVID-19 patient prediction. We are also working to identify clinical data characteristics that may influence the COVID-19 outcome

prediction. With the number of coronavirus-positive cases increasing daily, testing is impossible due to time and cost constraints. In recent years, machine learning in the medical field has become extremely reliable. The currently available models are developed on a relatively modest dataset, and the vast majority of the researchers have made use of a dataset that was not annotated by subject matter specialists (radiologists). The majority of the work that has been done in the field of machine learning has been accomplished through the use of hand-crafted methods and traditional approaches. The traditional methods have several performance flaws. To save human lives, a reliable and effective COVID detection system is required.

In the notice, although the results of this work are promising considering the volume of data tagged, they are only tentative, and a more definitive conclusion would take more studies in a broader dataset of COVID-19-labeled X-ray pictures. This study should be deemed as a starting point for potential research and comparisons.

The following is the outline for the remainder of this paper. Section 2 summarizes the prepared COVID-Xray-3k dataset. The proposed general structure has been explained in section 3. Experimental studies and parallels with previous work are presented in section 4. Lastly, the essay is closed in the 5th section.

2. The Xray-3k COVID dataset

The thoracic X-ray image from two datasets was combined to generate Covid X-ray 3,000 dataset images comprising 2,100 images for training and 900 images for testing purposes. The newly issued Covid-Chestxray-Dataset, which includes collecting X-ray images of articles published on the topic of coronavirus, was compiled by Cohen et al. https://github.com/ieee8023/ covid-chest-ray-dataset (2020). The dataset uses a mixed combination of CT scans with the images of chest X-rays. The dimension of CT images is 512x512x28 with a bit depth of 16 bits, and the file format is volumetric DICOM; similarly, the X-ray image size is 1024x1024x1 with a bit depth of 12 and 16 bits DICOM images. The images generated until May 3, 2020, contained two hundred and fifty X-ray images of corona-infected patients, with two hundred and three images corresponding to anteroposterior views. This data collection is continually modified according to the description. It also includes information about each patient, such as gender and age. Collecting images from both the CT scans and Xray diverse sources is a strategy employed in our study to enhance the comprehensiveness and robustness of our COVID-19 detection model. While domain adaptation and shifts pose challenges, our rigorous approach to data preprocessing, feature extraction, and model calibration is designed to mitigate these effects. By addressing these challenges head-on, we strengthen our model's reliability and real-world applicability, ultimately advancing the field of medical image analysis for COVID-19 diagnosis. This dataset provided us with all of our COVID-19 images. According to our accredited radiologist's recommendation, only anteroposterior X-ray samples are held to forecast COVID-19, as the previous samples were not considered appropriate for that reason. A qualified radiologist analyzed the anteroposterior images, and those lacking even the tiniest X-ray symbol of coronavirus were omitted from the data collection. 19 of the 203 COVID-19 indoor-outdoor X-ray images were discarded, leaving 184 for our radiologist to examine (which depicted clear indications of COVID-19). As a result, we would include a more accurately labeled data collection for the world. Among these images, 100 images per class are used for the testing (to achieve the highest value of confidence interval), while the remaining images are used as the training set. As previously mentioned, the data improvement is added to the learning kit to escalate the number of COVID-19 samples to 420.

Both patient X-ray images are transmitted only on one of the training courses, as we have ensured. Our radiologist highlighted the areas of clear Covid-19 signs due to the low number of images with no coronavirus collected on the dataset (20). This dataset includes 0.22 million images and three hundred and sixteen (224,316) chest X-ray images of sixty-five thousand two hundred and forty (65,240) patients. It is marked with the indication of 14 subcategories (non-finding, edema, pneumonia, etc.). We used only images from a single subcategory for non-COVID samples from the learning package, which consisted of seven hundred (700) pictures from the non-research class and one hundred (100) image from every other thirteen (13) subclasses, totaling two hundred (200) non-COVID images. We picked 1,700 images from the unsearched division.

We picked approximately a hundred (100) images from each of the other thirteen (13) subclasses in different sub-files for non-COVID samples from the research dataset, totaling 30,000 images. Table 1 shows the exact amount of X-ray images from each class used for preparation and research. Figure 2 displays 16 photos from the COVID-Xray-3k dataset, comprising four Coronavirus images (1st row), four regular ChexPert images (2nd row), and eight images of one of the 13 ChexPert images (3rd and 4th row).

It should be remembered that the resolution of the photos in this data collection varies significantly. Low-resolution COVID-19 images (less than 400 × 400 pixels) and high-resolution COVID-19 images (over 1900 × 1,400 pixels). This is a plus for models who can reach a reasonable precision level on this data collection, considering the variable image resolution and imaging technique. Although gathering all the photos in a highly controlled system, we desired to get ultra-sharp images with very high-resolution images; it is not always possible. As machine learning advances, more focus is put on the models and frames that will perform. On low-quality, small-scale tagged data sets, it performs reasonably well. Furthermore, the original vendor collects COVID-19 class images from various sites, showing dynamic variations (and even from ChexPert). However, the whole dataset is optimized to the same distribution in the testing phase to make the model less vulnerable to this.

Pursuing higher accuracy in COVID-19 diagnosis through deep learning models is challenging, and it necessitates an ongoing effort to access diverse and extensive datasets. To achieve this, researchers can explore several avenues. Public medical databases, such as the National Institutes of Health (NIH) Chest X-ray Dataset and the COVID-19 Image Data Collection, offer open-access repositories of

TABLE 1 Each category has no. of images in the Xray-3k COVID dataset.

Dataset Split	Non-COVID images	COVID-19 images	COVID-19 + Dibetic images
No. of training sets	700	700	700
No. of test sets	300	300	300

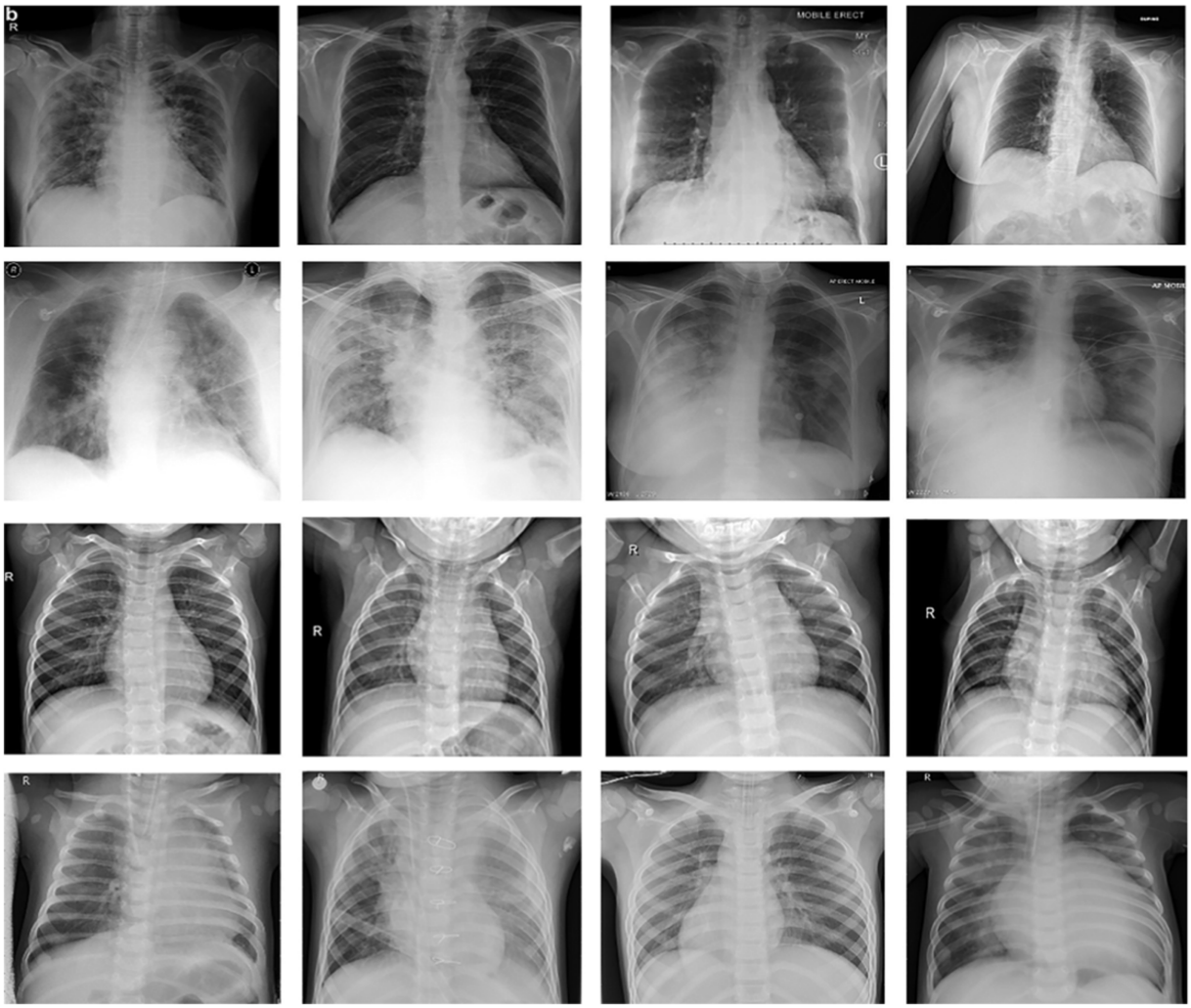


FIGURE 2
Sample of images from COVID-Xray-3k Dataset. First row corresponds to images with COVID-19. The second row corresponds to four sample images diagnosed with no COVID-19 infection from ChexPert, belonging to the no-finding category. The third and fourth row corresponds to images with eight samples belonging to all other subdomains in ChexPert.

radiographic images that can significantly augment existing datasets. Collaboration with medical institutions and hospitals can provide access to real-world patient data, capturing different COVID-19 manifestations and stages.

3. The proposed framework

Transfer learning was used to modify four deep neural networks and pre-trained images of the COVID-Xray-3k Dataset to solve the small data sizes. The choice of selection of the state-of-the-art transfer learning models in our study for COVID-19 detection using the x-ray images was based on their diverse architectural characteristics and well-established performance in image analysis tasks. These selected models are well known for their robustness, efficiency, and ability to transfer knowledge from large-scale datasets. This deliberate model selection aimed to comprehensively evaluate their suitability for COVID-19 detection and contribute to the advancement of medical image analysis.

3.1. Method of transfer learning

In this method, a model that has been educated on one task is reassigned to a similar task and is expected to respond to the new task. For, consider using an ImageNet model used to classify images (which includes billions of labeled images/pictures) to kickstart learning that will also be task-specific. This is used to detect COVID-19 on minor data collection. Transfer learning is most useful for those projects that require only a little effort to build models from the scrape, such as medical-based image recognition for evolving chronic diseases.

This is true, particularly for deep neural network-based models, which have many parameters to learn. In transfer learning, the setting of the model has better initial values, which needs a few minor improvements to make them more structured for the new mission. For each task, the pre-trained model is used in one of two ways. The first method is viewed as a model that extracts the characteristics, i.e., an extractor. In the second method, the model is trained to classify a classifier.

Another method involves purifying the whole network, or a subset of it, for the current mission. We simplify the end layer of complicated neural networks because the number of samples in the COVID-19 segment is relatively less. Consequently, the weights and biases of pre-trained CNN models are used to be a starting point for the proposed study, which are revised throughout the learning process. We use previously trained models as a characteristic extractor. ResNet-18 (24), ResNet-50 (25), Inception-ResNet-v2 (26), and NASNet-Large are four standard pre-formed models that we evaluate (27). The following segment gives a brief description of the models' design and their implementation to recognize coronavirus.

3.2. Inception-v3 and ShuffleNet based COVID-19 detection

The pre-designed Inception-v3 model, formed on the ImageNet dataset, is one of the models implemented in our research. Inception-v3 is one of the most common CNN architectures, and it won the 2015 ImageNet contest. It offers a more effortless gradient flow for more effective training. The implementation of an identity shortcut link that misses/skips one or more than one layer is at the heart of Inception-v3. This will enable the network to have a clear route to the network's first layers, rendering gradient changes far simpler for these layers. Supplementary Figure S1 depicts the Inception-v3 model's general theory scheme and its application to COVID-19 identification. The Inception-v2 design is similar to Inception-v3 but with a number of layers than the Inception-v3. The structure design of ShuffleNet CNN features learning and classification can be seen in Supplementary Figure S2. Supplementary Figures S4–S7 illustrates the probabilities estimated by the various CNN models when applied to the testing samples. This graphical representation provides valuable insights into the model's confidence scores and its decision-making process.

3.3. The inception-ResNet-v2 for COVID-19 detection

The Inception-ResNet-v2 is a small CNN model that obtains accuracy up to the AlexNet level (28) with 50 times more minor settings. Using these techniques, the biographers compressed Inception-ResNet-v2 to a smaller amount, i.e., smaller than 0.5 MB, making it prevalent for applications requiring lightweight models. They substitute one layer 1×1 that "tightens" the data entering the vertical dimension, followed by the sign of two parallel convoluted layers 1×1 and 3×3 that "extend" the data's depth again. Inception-ResNet-v2 services three effective strategies: replacing 3×3 filters with 1×1 filters, growing the number of input channels to 3×3 filters, and subsampling late in the network to ensure massive activation maps for convolution layers.

3.4. COVID-19 detection using NASNet-large

Another architecture introduced by (29) is the Neural Architectural Seach Convolutional Network (NASNet-Large), which won the ImageNet 2017 competition. Each layer in NASNet-Large receives additional entries from all preceding layers and transmits its function cards to all succeeding layers. Each layer gets all of the previous layers' accumulated information. The network can be thinner and more lightweight because every layer receives maps for every layer. Supplementary Figure S3 depicts the architecture of the NASNet-Large example.

3.5. Model training

The cross-entropy loss function, whose goal is to decrease the change between expected probability scores and field truth probabilities, is used to train all models.

$$L_{CE} = -\sum_{i=1}^{N} p_i \log q_i \quad (1)$$

Where *pi* denotes ground truth, whereas *qi* denotes predicted probabilities for every image. A stochastic gradient descent algorithm can then be used to minimize this loss function (and its variations). We tried to improve the loss feature by including regularization, but the resulting model did not improve.

4. Results

4.1. Hyper-parameters model

Each model has been trained with 100 Epochs. The loss function is optimized with the use of an ADAM optimizer having a learning rate of 0.0001. This optimizer has a size of 20. Since these models are typically created with a detailed image resolution, all the images are under 224*224 before being submitted to the neural network. All the experimental tasks are performed using the MATLAB deep learning framework. The confusion matrices for the four classification models, each tasked with classifying three distinct classes, are presented in Figures 3–6. These visual representations provide a comprehensive view of the models' performance in categorizing instances into "Normal," "Covid-19," and "Covid-19 + Diabetic" classes.

The Supplementary Table S1 displays the hyperparameters used, their corresponding values or methods, and the optimal

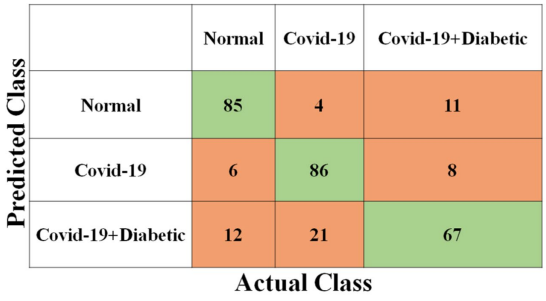


FIGURE 3
Shows the proposed Inception-v3 model's confusion matrix.

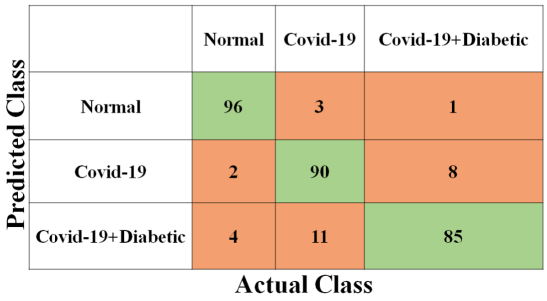


FIGURE 4
Shows the proposed ShuffleNet model's confusion matrix.

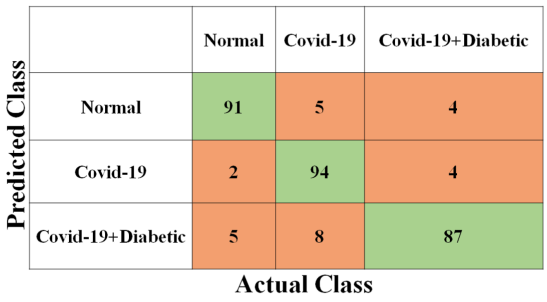


FIGURE 5
Shows the proposed Inception-ResNet-v2 CNN model confusion matrix.

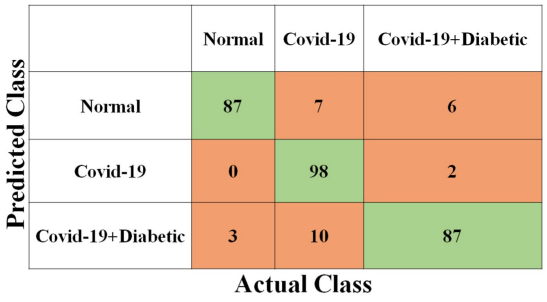


FIGURE 6
Shows the proposed NASNet-Large CNN model confusion matrix.

selections during the training of four transfer learning CNN models. The Inception-v3 CNN achieved an average accuracy of 79.62%, the Figure 3 displays correctly predicted samples in green, while incorrectly predicted samples are shown in red. Similarly, Figures 4–6 display the confusion matrices obtained when validating the test set with ShuffleNet, Inception-ResNet-v2, and NASNet-Large, each surpassing accuracy rates of 90.33, 90.67, and 90.67%, respectively. These remarkable accuracies underscore the effectiveness of the chosen optimal hyperparameters.

4.2. Evaluation metrics

Different metrics, including classification precision, Sensitivity, specificity, accuracy, and F1 ranking, may be used to evaluate classification models' success. Due to the unbalanced nature of the current test dataset (80 coronavirus infectious images vs. 2000 non-coronavirus infectious images), sensitivity and specificity are two critical indicators to report model performance:

$$\text{Sensitivity} = \frac{\text{The number of images correctly predicts COVID19}}{\text{The total COVID19 images}} \quad (2)$$

$$\text{Specificity} = \frac{\text{The total number of images correctly predicted as NonCOVID}}{\text{The total number of NonCOVID images}} \quad (3)$$

4.3. The predicted scores of models

We are based on four standard convoluted networks, as previously stated. All these models generate a probability score for every X-ray image. It also increases the probability factor of the disease being identified as COVID-19. We may develop a binary mark to indicate whether the image is COVID-19 or not. We can get this by the comparison of the binary Mars with a cut-off threshold. A perfect model can detect/predict the chance for every COVID-19 sample, which is found to be close to 1. Like this, an ideal model can predict the possibility of every non-COVID sample being close to 0. Tables 2–5 Present the Sensitivity and Specificity Achieved by Four CNN Models for the Detection of COVID-19 with Diabetes. Table 2 presents the sensitivity and specificity achieved by the Inception-v3 model across various threshold values. Meanwhile, Tables 3–5 provide sensitivity and specificity values for the ShuffleNet, Inception-ResNet-v2, and NASNet-Large models, respectively.

Supplementary Figures S4–S7 display the model's distributions of expected likelihood scores for the test set photos, respectively. We include the probability distribution of the expected three categories: COVID-19, Normal, and other diseases. Our study's non-COVID grouping consists of both standard cases and other forms of diseases. As can be said, non-COVID X-ray images of different types of infections have significantly better ratings than non-COVID examples without other types of diseases. The infected images of COVID-19 may have somewhat higher odds than non-COVID images, which is promising. We can see that Inception-ResNet-v2 is better at work than the other models. Table 6 provides a comprehensive overview of the class-specific performance metrics, and the average performance of four state-of-the-art CNN models used for chest radiography detection. The models were evaluated across three distinct classes: "Normal," "Covid-19," and "Covid-19 + Diabetic." The metrics examined include Accuracy, Precision, Recall, and F1-score, offering valuable insights into the models' capabilities for each class. Additionally, the table presents an "Average" row, summarizing the

10.3389/fpubh.2023.1297909

TABLE 2 The results of the Inception-v3 model in the form of sensitivity and specificity rates.

Sensitivity	Specificity	Threshold
100%	73.4%	0.19
99%	92.7%	0.18
96%	94.4%	0.22
93%	96.8%	0.23
87%	98.0%	0.31

TABLE 3 The results of the ShuffleNet model in the form of sensitivity and specificity rates.

Sensitivity	Specificity	Threshold
100%	79.2%	0.17
97%	90.2%	0.24
95%	95.3%	0.21
92%	98.5%	0.29
87%	98.4%	0.36

TABLE 4 The results of the Inception-ResNet-v2 model in the form of sensitivity and specificity rates.

Threshold	Sensitivity	Specificity
0.32	98%	91.2%
0.19	99%	90.4%
0.4	97%	95.8%
0.39	93%	98.2%
0.8	88%	99.7%

TABLE 5 The results of the NASNet-Large model in the form of sensitivity and specificity rates.

Threshold	Sensitivity	Specificity
0.16	97%	77.3%
0.22	94%	89.8%
0.29	92%	96.4%
0.38	81%	99.8%

TABLE 6 Class-specific Performance metrics and average performance of state-of-the-art CNN models for chest radiography detection.

Model	Class	Accuracy	Precision	Recall	F1-score
Inception-v3	Normal	85.85%	88.57%	89.47%	89.01%
	Covid-19	86.00%	86.00%	93.33%	89.33%
	Covid-19 + Diabetic	67.00%	67.00%	72.00%	69.33%
	Average	79.62%	80.52%	84.97%	82.72%
ShuffleNet	Normal	96.00%	96.00%	96.00%	96.00%
	Covid-19	90.00%	90.00%	92.00%	91.00%
	Covid-19 + Diabetic	85.00%	85.00%	87.00%	86.00%
	Average	90.33%	90.33%	91.67%	90.67%
Inception-ResNet-v2	Normal	91.00%	93.57%	95.56%	94.51%
	Covid-19	94.00%	94.00%	96.00%	95.00%
	Covid-19 + Diabetic	87.00%	87.00%	89.00%	88.00%
	Average	90.67%	91.52%	93.52%	92.52%
NASNet-Large	Normal	87.00%	90.55%	91.89%	91.21%
	Covid-19	98.00%	98.00%	98.00%	98.00%
	Covid-19 + Diabetic	87.00%	87.00%	89.00%	88.00%
	Average	90.67%	91.85%	92.96%	92.41%

collective performance of these models across all classes. These metrics serve as a vital reference point for evaluating the models' effectiveness in detecting and distinguishing between different chest radiography categories.

4.4. The sensitivity and specificity of four different models

Every model generates a probability score that indicates the likelihood of the image, i.e., the idea being COVID-19. These scores are then compared to a criterion to determine whether or not the picture is COVID-19. The value of the Sensitivity of all models and the importance of the specificity of all models were calculated using predicted labels. Tables 2–5 demonstrate sensitivity rates and specificity rates for various levels utilizing the four models. It can be shown that both of these models provide positive outcomes, with the strongest one achieving a sensitivity of 95% (95%) and specificity of 91% (91.06%). Inception-ResNet-v2 and Inception-v3 outperform the other models by a small margin.

The Inception-ResNet-v2 has the high sensitivity (98%) and specificity (91.2%) rates demonstrated by our top-performing model, which holds substantial clinical significance. These performances reflect the model capability of accurately detecting COVID-19 cases

and reducing the inaccurate diagnosis. The higher accuracy of the model assists in early disease diagnosis and treatment plans, making it a vital tool for radiologists and pulmonologists. Moreover, the model flexibility for different clinical scenarios is necessary, as indicated by various threshold options, to enhance its practical use in real-world applications, where balancing sensitivity and specificity is crucial for effective COVID-19 diagnosis.

4.5. The reliability of the model with a few cases of COVID-19

It should be mentioned whether the sensitivity and specificity rates shown earlier can be or cannot be accurate because there was a minimal amount of accurately annotated COVID-19 X-ray images by the experts who are available to date besides the fact that the COVID-19 dataset consists of several hundred X-ray samples. More studies on more test samples are required to get a more accurate estimate. To see the potential range of these values in every class, we will measure the confidence interval at 95% of sensitivity and specificity rates recorded here. The accuracy rate trust interval can be determined as follows:

$$r = \sqrt[z]{\frac{accuracy\left(1 - accuracy\right)}{N}} \tag{4}$$

Where z is the confidence interval's degree of significance, accuracy is the approximate accuracy (in our case, sensitivity rates and specificity rates), and N is the total number of samples. In this case, we used a 95 percent trust interval, which corresponds to a z-value of 1.96. Since a responsive model is critical for the COVID-19 diagnosis, we select a cut-off threshold for each model that fits a sensitivity rate of 98 % (98%) and can also evaluate their specificity values. Supplementary Table S2 shows how these four models performed throughout the test range. Since we have around three thousand samples for this class, the confidence interval for specificity values is minimal (around 1%). In contrast, the sensitivity rate has a somewhat higher confidence interval (about 2.7%) due to the smaller number of samples. The performance comparison is presented in Table 7, incorporating the latest advancements from state-of-the-art research.

4.6. The operating characteristics curve (ROC)

Since cut-off limits vary, it is challenging to equate various models. We ought to test all potential threshold values to see how these models compare overall. The precision-recall curve is one way to do this. Recall or Sensitivity is the Ratio of true positives to total (actual) positives in the data. Recall and Sensitivity are one and the same. Whereas the accuracy is calculated using the accurately detected +ve images and the total number of +ve images in the test set using the ROC curve. Figure 7 depicts the curve created using the precision and recall values of the proposed CNN models. The ROC curve is plotted by taking the precision values on the y-axis and recall values on the x-axis of the 2D line plot. Supplementary Figure S8 shows the ROC curves of these four models. Both versions work equally according to AUC.

TABLE 7 Comparison of the proposed model with existing state-of-the-art methods.

Model	Accuracy	F-Measure
CovidxNet-CT (30)	85%	86.06%
Optimized Resnet 101 (31)	95%	93.32%
UNet+ ResNet (32)	94%	92.3%
EfficientNet+SCO (32)	85%	87.66
Proposed Model (33)	96%	96.9%

It should be noted that the AUC might not be a suitable predictor of model success for very unbalanced test sets (because it can be very high) and that examining the medium accuracy curve and precision and recall may be a safer option in this case. For the sake of completeness, we have included all curves here. The confusion matrices of the two highest-performing CNN models, Inception-v3 and Inception-ResNet-v2, on a test set of 2080 Xrays can be observed in Figures 3, 5. These matrices provide an exact count of suitable samples, i.e., samples that are positive for COVID-19 and samples that are negative for COVID-19.

4.7. Hardware resources and simulation environment

The allocation of robust computational resources listed in Supplementary Table S3 was pivotal in successfully developing and training our deep learning models for COVID-19 diagnosis from chest X-ray images. Utilizing high-performance hardware components, including the Intel Core i7-12700K CPU and NVIDIA GeForce RTX 3080 Ti GPU, allowed us to efficiently process vast volumes of data and perform complex matrix computations, thus expediting the training process. This strategic choice significantly reduced training times and enabled the exploration of intricate model architectures. Furthermore, the abundant 32GB of RAM and the extensive 1 TB or more SSD storage were instrumental in ensuring the seamless loading of data, preventing potential bottlenecks, and accommodating the storage needs for our extensive dataset and model checkpoints.

Complementing our powerful hardware setup, the adoption of essential image processing, statistics and machine learning, and deep learning toolboxes provided in the MATLAB 2021a are used for developing, fine-tuning, and rigorously evaluating our deep neural networks. The Windows 10 operating system further contributed to a stable and reliable research environment. This fusion of computational resources and software tools facilitated our pursuit of precise COVID-19 diagnosis and laid the foundation for transparent, accessible, and collaborative research.

4.8. The infected regions

Thermal maps are acquired using thermal imaging camera sensors, which play a unique role in COVID-19 diagnosis. These images record the change in body temperature, which can be very

useful in the study and diagnosis of patients suffering from fever or other respiratory distress related to COVID-19. These images are overlapped with chest X-ray images to provide the radiologists with a multidimensional view, which assists in the localization of the affected region in the lungs. The fusion of thermal images with radiographic data dramatically improves the detection of COVID-19; in the case of subtle radiographic findings, it still achieves higher diagnosis accuracy. Moreover, the thermal maps assist in the ongoing monitoring of patient progress, which offers an early insight into disease treatment plans or disease deterioration, thereby assisting healthcare providers in making timely and informed decisions. When we detected COVID-19, we used an essential technique to see possibly contaminated regions—(34) work to imagine deep learning outcomes complex networks influenced this technique. We begin at the image's top-left corner, blocking a rectangular area of MxN or a square area of dimension M rows and N columns within the X-ray sample each time to predict the occlusal image. Suppose the model wrongly classifies a picture of COVID-19 as a picture of non-COVID due to this region's occlusion. In that case, this location will be called a likely polluted region in thoracic X-ray pictures. But if an area's occlusion has little effect on the model's projection, we should conclude that the region is free from contamination.

We can also have a sad map of infected areas detecting coronavirus by repeating this process for different slippery N x N windows and moving them each time with an S phase. Figure 8 shows the regions detected in six examples of COVID-19 photos from our test sample. In the last section, possible COVID-19 disease areas are identified and annotated in yellow color by our experts, who are certified by the Council of Radiology and Council of Medical Sciences. Regions annotated by the radiologist and experts in COVID-19 disease are in good agreement with the thermal mass produced.

5. Conclusion

For the sake of detecting COVID-19 and COVID-19 affected who are also diabetic, a standard dataset of 3k X-ray images is created and confirmed with the COVID-19 labels from the board-certified radiologist. The dataset is available for researchers and can be used as a benchmark dataset for COVID-19 prediction using machine-learning models. We reported that four pre-trained deep neural network models (Inception-v3, ShuffleNet, Inception-ResNet-v2, and NASNet-Large) are used to detect COVID-19 using X-ray images by fine-tuning the model's parameters. We conducted a detailed experimental analysis on the COVID-Xray-3k dataset test set to assess these four models' Sensitivity, specificity, ROC, and AUC performance. These models had an average specificity rate of about 90% for a sensitivity rate of 98 percent. This is encouraging because it shows promise for using X-ray images to diagnose COVID-19. This research used a set of publicly available images that included about 1,000 Normal images, 1,000 COVID-19 images, and 1,000 X-ray images of patients suffering from COVID-19 and also diabetic. The work presented here represents one of the earliest attempts at Covid-19 chest X-ray analysis and dataset preparation, which resulted in a time-sensitive correlation when the two aspects were combined. However, because there are only a few publicly available COVID-19 images, more experiments on a more extensive set of clearly labeled COVID-19 images are needed to estimate the accuracy of these models more reliably.

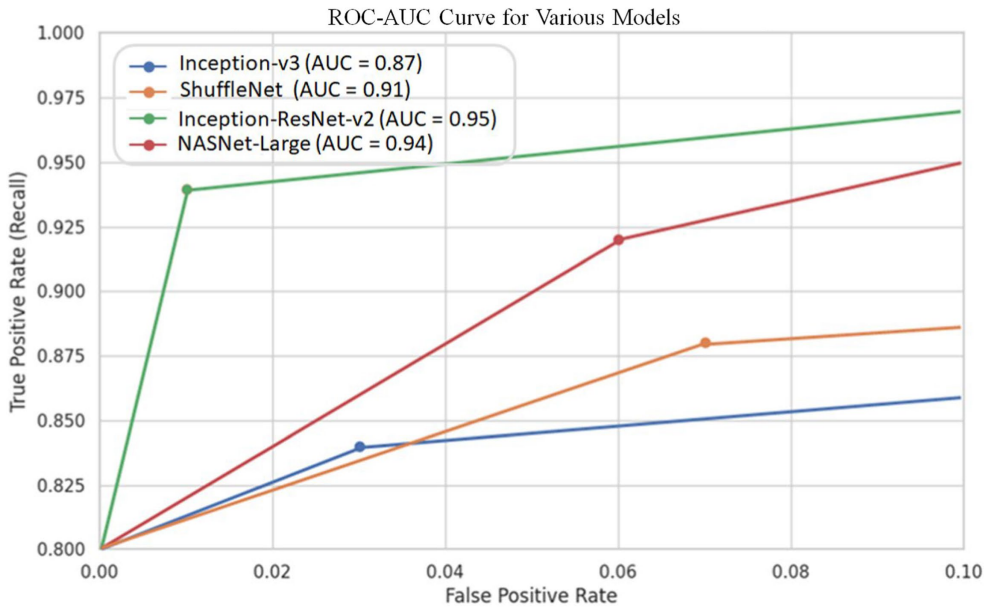


FIGURE 7
Shows the precision-recall curves of 4 CNN architectures for COVID-19 detection.

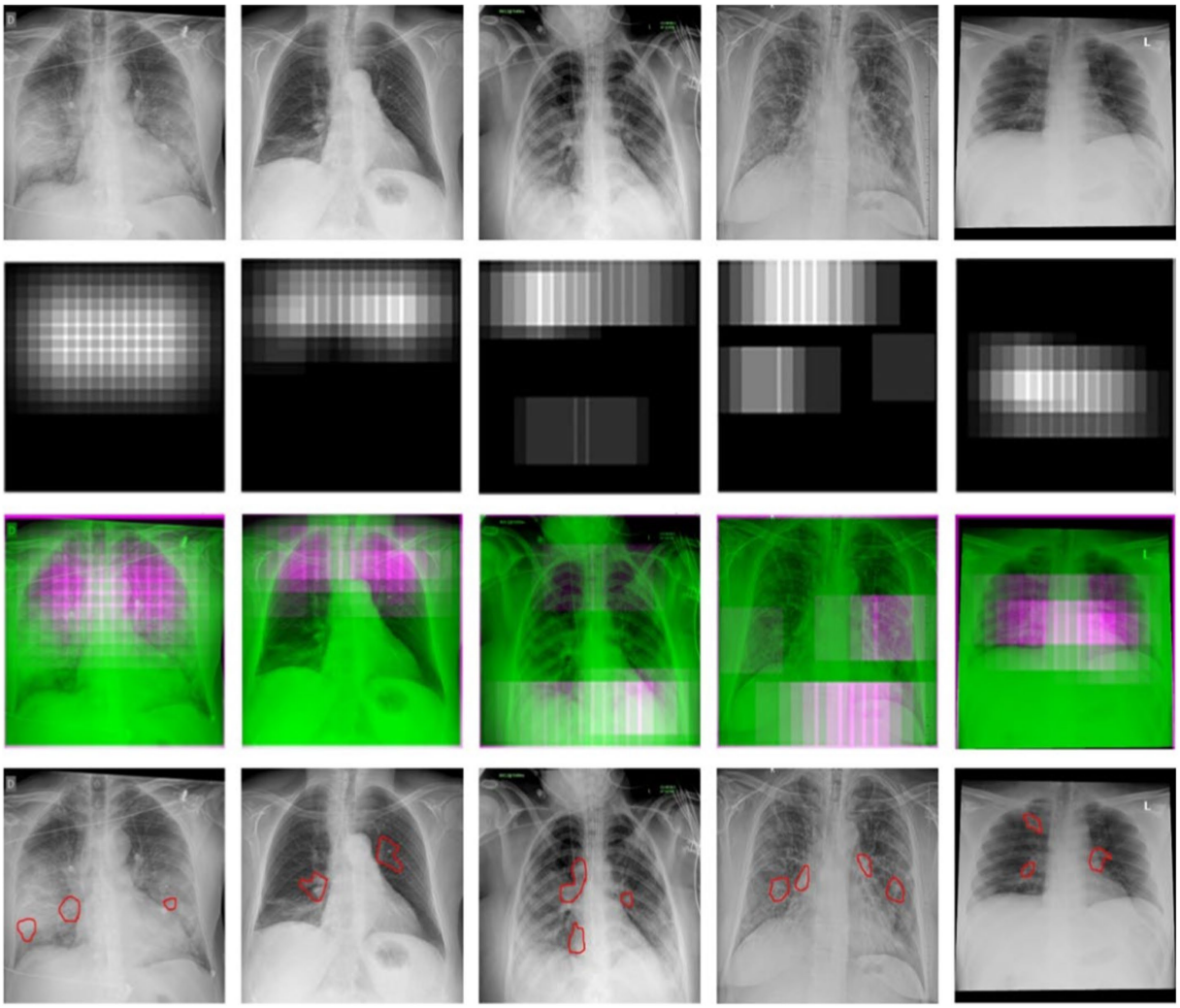


FIGURE 8
The detection of COVID-19-affected regions in the testing X-ray samples using the inception-v3 CNN model.

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.

Author contributions

MS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. NS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. BS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. TH: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. AA: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. SA: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. FA: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported

by the Researchers Supporting Project number (RSP2023R395), King Saud University, Riyadh, Saudi Arabia.

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2023.1297909/full#supplementary-material

References

1. Gerganova A, Assyov Y, Kamenov Z. Stress Hyperglycemia, diabetes mellitus and COVID-19 infection: risk factors, clinical outcomes and post-discharge implications. *Front Clin Diabetes Healthc*. (2022) 3:1–13. doi: 10.3389/fcdhc.2022.826006

2. Khunti K, Del PS, Mathieu C, Kahn SE, Gabbay RA, Buse JB. Covid-19, hyperglycemia, and new-onset diabetes. *Diabetes Care*. (2021) 44:2645–55. doi: 10.2337/dc21-1318

3. Bhandari S, Bhargava S, Samdhani S, Singh SN, Sharma BB, Agarwal S, et al. COVID-19, diabetes and steroids: the demonic trident for Mucormycosis. *Indian J Otolaryngol Head Neck Surg*. (2022) 74:3469–72. doi: 10.1007/s12070-021-02883-4

4. Pelle MC, Zaffina I, Provenzano M, Moirano G, Arturi F. COVID-19 and diabetes—two giants colliding: from pathophysiology to management. *Front Endocrinol (Lausanne)*. (2022) 13:1–13. doi: 10.3389/fendo.2022.974540

5. Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA - J Am Med Assoc*. (2020) 323:1843–4. doi: 10.1001/jama.2020.3786

6. Acuña-Castillo C, Barrera-Avalos C, Bachelet VC, Milla LA, Inostroza-Molina A, Vidal M, et al. An ecological study on reinfection rates using a large dataset of RT-qPCR tests for SARS-CoV- in Santiago of Chile. *Front Public Heal*. (2023) 11:1191377. doi: 10.3389/fpubh.2023.1191377

7. Yang Y, Yang M, Shen C, Wang F, Yuan J, Li J, et al. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections. *medRxiv*. (2020). doi: 10.1101/2020.02.11.20021493

8. Aminian A, Safari S, Razeghian-Jahromi A, Ghorbani M, Delaney CP. COVID-19 outbreak and surgical practice: unexpected fatality in perioperative period. *Ann Surg*. (2020) 272:E27–9. doi: 10.1097/SLA.0000000000003925

9. Kanne JP, Little BP, Chung JH, Elicker BM, Ketai LH. Essentials for radiologists on COVID-19: an update—radiology scientific expert panel. *Radiology* (2020) 296:E113–E114. doi: 10.1148/radiol.2020200527

10. Kumar R, Arora R, Bansal V, Sahayasheela VJ, Buckchash H, et al. Classification of COVID-19 from chest x-ray images using deep features and correlation coefficient. *Multimedia Tools and Applications* (2022) 81:27631–55. doi: 10.1007/s11042-022-12500-3

11. Kong W, Prachi P. Chest imaging appearance of COVID-19 infection. *Radiol Cardiothorac Imaging*. (2020) 2:e200028. doi: 10.1148/ryct.2020200028

12. Hansell DM, Bankier AA, Mcloud TC, Mu NL, Remy J. Fleischner society: glossary of terms for thoracic imaging. *Radiology*. (2008) 246:697–722. doi: 10.1148/radiol.2462070712

13. Rodrigues JCL, Hare SS, Edey A, Devaraj A, Jacob J, Johnstone A, et al. An update on COVID-19 for the radiologist - a British society of thoracic imaging statement. *Clin Radiol*. (2020) 75:323–5. doi: 10.1016/j.crad.2020.03.003

14. Saqlain M, Munir MM, Ahmed A, Tahir AH, Kamran S. Is Pakistan prepared to tackle the coronavirus epidemic? *Drugs Ther Perspect*. (2020) 36:213–4. doi: 10.1007/s40267-020-00721-1

15. Ali I. Pakistan confirms first two cases of coronavirus, govt says 'no need to panic. *Daw. News* (2020), Paksitan. Available at: https://www.dawn.com/news/1536792t

16. Mirza D. Pakistan latest victim of coronavirus *Geo News* (2021), Pakistan. Available at: https://www.geo.tv/latest/274482-pakistan-confirms-first

17. Badrinarayanan V, Kendall A, Cipolla R, Member S. SegNet: A deep convolutional encoder-decoder architecture for scene segmentation. *IEE*. (2016) 8828:1–14. doi: 10.1109/TPAMI.2016.2644615

18. Cohen JP, Morrison P, Dao L. COVID-19 Image Data Collection: Prospective Predictions are the Future. *Journal of Machine Learning for Biomedical Imaging* (2020) 1:1–38. doi: 10.59275/j.melba.2020-48g7

19. Cohen JP, Morrison P, Dao L. COVID-19 image data collection. *arXiv*. (2020). doi: 10.48550/arXiv.2003.11597

20. Irvin J, Rajpurkar P, Ko M, Yu Y, Ciurea-ilcus S, Chute C, et al. et al, CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. *arXiv*. doi: 10.48550/arXiv.1901.07031

21. Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell*. (2017) 39:1137–49. doi: 10.1109/TPAMI.2016.2577031

22. Dong C, Loy CC, He K, Tang X. *Learning a deep convolutional network for image super-resolution*. (2014)184–199. Springer, Cham

23. Minaee S, Abdolrashidi A, Su H, Bennamoun M, Feb CV. Biometrics recognition using deep learning: A survey. *Artif Intell Rev*. (2023) 56:8647–95. doi: 10.1007/s10462-022-10237-x

24. Zielonka M, Piastowski A, Czyżewski A, Nadachowski P, Operlejn M, Kaczor K. Recognition of emotions in speech using convolutional neural networks on different datasets. *Electron*. (2022) 11:3831. doi: 10.3390/electronics11223831

25. Vasan D, Alazab M, Wassan S, Naeem H, Safaei B, Zheng Q. IMCFN: image-based malware classification using fine-tuned convolutional neural network architecture. *Comput Netw*. (2020) 171:107138. doi: 10.1016/j.comnet.2020.107138

26. Shah SM, Zhaoyun S, Zaman K, Hussain A, Shoaib M, Lili P. A driver gaze estimation method based on deep learning. *Sensors*. (2022) 22:1–22. doi: 10.3390/s22103959

27. Yang S, Jiang L, Cao Z, Wang L, Cao J, Feng R, et al. Deep learning for detecting corona virus disease 2019 (COVID-19) on high-resolution computed tomography: a pilot study. *Ann Transl Med*. (2020) 8:450–0. doi: 10.21037/atm.2020.03.132

28. Wang B. Identification of crop diseases and insect pests based on deep learning. *Sci Program*. (2022) 2022:9179998. doi: 10.1155/2022/9179998

29. Szymak P, Piskur P, Naus K. The effectiveness of using a pretrained deep learning neural networks for object classification in underwater video. *Remote Sens*. (2020) 12:1–19. doi: 10.3390/RS12183020

30. Ulutas H, Sahin ME, Karakus MO. Application of a novel deep learning technique using CT images for COVID-19 diagnosis on embedded systems. *Alex Eng J*. (2023) 74:345–58. doi: 10.1016/j.aej.2023.05.036

31. Haennah JHJ, Christopher CS, King GRG. Prediction of the COVID disease using lung CT images by deep learning algorithm: DETS-optimized Resnet 101 classifier. *Front Med (Lausanne)*. (2023) 10:1157000. doi: 10.3389/fmed.2023.1157000

32. Althaqafi T, al-Ghamdi ASALM, Ragab M. Artificial intelligence based COVID-19 detection and classification model on chest X-ray images. *Healthcare*. (2023) 11:1–16. doi: 10.3390/healthcare11091204

33. Topff L, Sánchez-García J, López-González R, Pastor AJ, Visser JJ, Huisman M, et al. A deep learning-based application for COVID-19 diagnosis on CT: the imaging COVID-19 AI initiative. *PLoS One*. (2023) 18:1–17. doi: 10.1371/journal.pone.0285121

34. Zeiler MD, Fergus R, Khandakar A, Rahman T, Khurshid U, Musharavati F, et al. Visualizing and understanding convolutional networks. In: *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, Proceedings Part I 13*. (2014), 818–833. Springer International Publishing. doi: 10.1007/978-3-319-10590-1_53

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 06 November 2023
DOI 10.3389/fpubh.2023.1308404

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Inam Ullah,
Gachon University, Republic of Korea
Muzammil Khan,
University of Swat, Pakistan

*CORRESPONDENCE
Farman Ali
✉ farmankanju@gmail.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 06 October 2023
ACCEPTED 18 October 2023
PUBLISHED 06 November 2023

CITATION
Ahmad I, Merla A, Ali F, Shah B, AlZubi AA and AlZubi MA (2023) A deep transfer learning approach for COVID-19 detection and exploring a sense of belonging with Diabetes. *Front. Public Health* 11:1308404. doi: 10.3389/fpubh.2023.1308404

A deep transfer learning approach for COVID-19 detection and exploring a sense of belonging with Diabetes

Ijaz Ahmad[1†], Arcangelo Merla[2], Farman Ali[3*†], Babar Shah[4], Ahmad Ali AlZubi[5] and Mallak Ahmad AlZubi[6]

[1]Digital Transition, Innovation and Health Service, Leonardo da Vinci Telematic University, Chieti, Italy, [2]Department of Engineering and Geology (INGEO) University "G. d'Annunzio" Chieti-Pescara, Pescara, Italy, [3]Department of Computer Science and Engineering, School of Convergence, College of Computing and Informatics, Sungkyunkwan University, Seoul, Republic of Korea, [4]College of Technological Innovation, Zayed University, Dubai, United Arab Emirates, [5]Department of Computer Science, Community College, King Saud University, Riyadh, Saudi Arabia, [6]Faculty of Medicine, Jordan University of Science and Technology, Irbid, Jordan

COVID-19 is an epidemic disease that results in death and significantly affects the older adult and those afflicted with chronic medical conditions. Diabetes medication and high blood glucose levels are significant predictors of COVID-19-related death or disease severity. Diabetic individuals, particularly those with preexisting comorbidities or geriatric patients, are at a higher risk of COVID-19 infection, including hospitalization, ICU admission, and death, than those without Diabetes. Everyone's lives have been significantly changed due to the COVID-19 outbreak. Identifying patients infected with COVID-19 in a timely manner is critical to overcoming this challenge. The Real-Time Polymerase Chain Reaction (RT-PCR) diagnostic assay is currently the gold standard for COVID-19 detection. However, RT-PCR is a time-consuming and costly technique requiring a lab kit that is difficult to get in crises and epidemics. This work suggests the CIDICXR-Net50 model, a ResNet-50-based Transfer Learning (TL) method for COVID-19 detection via Chest X-ray (CXR) image classification. The presented model is developed by substituting the final ResNet-50 classifier layer with a new classification head. The model is trained on 3,923 chest X-ray images comprising a substantial dataset of 1,360 viral pneumonia, 1,363 normal, and 1,200 COVID-19 CXR images. The proposed model's performance is evaluated in contrast to the results of six other innovative pre-trained models. The proposed CIDICXR-Net50 model attained 99.11% accuracy on the provided dataset while maintaining 99.15% precision and recall. This study also explores potential relationships between COVID-19 and Diabetes.

KEYWORDS

COVID-19, deep learning, diabetes mellitus, chest x-ray, transfer learning, convolutional neural network, long-Covid

1. Introduction

COVID-19 is a severe and deadly disease caused by a newly discovered coronavirus. In late 2019, a strange sickness outbreak afflicted several people in Wuhan, China (1, 2). The precise reason for this widespread sickness outbreak was unknown, and the symptoms seemed strange. It has been determined that the virus has a unique coronavirus strain that was never found in

humans before (1, 3). Coronaviruses can cause various respiratory diseases, ranging from moderate to severe. COVID-19 can be identified clinically by several symptoms related to the respiratory system, including pneumonia, cough, dyspnea, and fever. However, these signs are not exclusive to COVID-19 and can be seen in various pneumonia cases, which presents a challenge for medical professionals. Real-Time Polymerase Chain Reaction is one of the most accurate coronavirus testing techniques (RT-PCR), which has been authorized by the WHO (World Health Organization). In RT-PCR, the RNA sequence is converted into DNA through a process called reverse transcription, which is then amplified (4, 5). Global health, economy, and well-being were all affected by the rapid spread of the COVID-19 epidemic (6, 7). Detecting COVID-19 is critical for patient care and public health, as the pandemic can be avoided and effectively managed by isolating infected patients (8). It is critical to separate people infected with COVID-19; thus, early detection is a significant challenge for preventing the further spreading of the infection (9, 10).

Computer Imaging techniques in medical science can help control the spread of infection more effectively, treat infected individuals, and reduce the mortality rate (11–13). Imaging modalities such as computed tomography (CT) scans and chest X-rays (CXR) are essential for diagnosing pulmonary diseases (14, 15). In clinical practice, CXR and CT are frequently used to detect COVID-19. Even though CT has better sensitivity in COVID-19 detection, the CXR is a popular imaging modality because of its many advantages, including its low price, low level of radiation exposure, straightforward operation, and easy availability in hospitals (16, 17). Radiologists consistently face a clinical dilemma for COVID-19 detection during this pandemic (18). Rapid and precise COVID-19 identification is critical for avoiding and treating this pandemic disease through quarantine and medical treatment.

The lack of accessible testing kits makes it challenging to determine if the disease has spread as the number of reported cases rises. Deep Learning (DL)--based methodologies have progressed to the point where they can compete with the most advanced approaches in computer-aided diagnostics (19). In establishing computer-aided detection (CAD) systems, significant progress has been made using medical images and robust DL algorithms. These schemes are designed to automatically examine disease characteristics (20) using DL approaches to assist radiologists in making more accurate diagnostics. Recently, researchers applied DL algorithms to explore and assess CXR images for COVID-19 detection. COVID-19 detection and diagnosis methods using deep learning-based algorithms are precise and effective. The advantages of supervised DL algorithms in medical imaging tasks have been demonstrated in numerous applications (21). These DL algorithms need many data to create an accurate model. Unfortunately, access to such large volumes of labelled data is another significant problem for machine-learning approaches in the medical domain (22–24). A deep CNN model already pre-trained with variable layers can compensate for the lack of labelled data problems (24–26). Most initial layers in the pre-trained model are fixed to generic reserve aspects of natural images and train higher-level layers on medical images (27). This process of taking the pre-trained model of one problem and applying it to other related problems by retraining higher layers is called Transfer Learning (TL). Compared to standard DL algorithms, TL is simple, efficient, and has minimal training cost (27), thereby overcoming the problem of limited datasets.

Wang et al. described a Deep CNN-based system called Covid-Net for detecting COVID-19 instances in chest CXR images (28). This study helps doctors improve their transparency and screening when utilizing COVID-Net for computer-assisted screening by highlighting the significance of the main characteristics of COVID-19 cases. Apostolopoulos and Mpesiana suggested a TL-based technique for automatically recognizing COVID-19 using CXR images to assess how modern CNN designs classify data (29). The results suggest that extracting key COVID-19 disease characteristics by combining DL and X-ray imaging may be possible. Sethy et al. used CXR pictures to adapt the COVID-19 detection based on the extensive feature extraction and implementation of a Support Vector Machine (SVM) as a classifier (30). The study used thirteen different DCNN-based pre-trained models as feature extractors, providing each feature to the SVM classifier. ResNet50 model with SVM performs better than the other selected twelve classification models.

Hemdan et al. proposed a new COVIDX-Net system that uses seven DL pre-trained models to detect and analyze COVID-19 in two-dimensional CXR images (31). The findings of the proposed COVIDX-Net showed that the VGG19 and DenseNet201 models achieved the most significant performance scores among the other DL classifiers. Manokaran et al. employed a DenseNet201-based model for CXR image classification that was created by substituting a new network for the final classifier layer utilizing TL methods (32). Chakraborty et al. proposed a transfer learning approach based on VGG-19 pre-trained architecture to classify COVID-19, pneumonia, and healthy patients using CXR images (33). Ozturk et al. suggested a DarkCovidNet model for the CXR image classification (34). Binary and multi-class classifications are both supported by the model. An experienced radiologist carried out an analysis of how well the DarkCovidNet model worked. Jain et al. presented a two-stage process for classifying COVID-19 CXR images of persons with bacterial pneumonia, viral pneumonia, and healthy individuals (35).

Further analysis of the X-ray scans of viral pneumonia was performed to identify the presence of COVID-19. Their proposed DL model performs remarkably well in multi and binary classification phases. Vaid et al. created a model using the VGG19-based TL technique to improve its accuracy in detecting COVID-19 from CXR images (36). Pathak et al. proposed a ResNet50-based approach for building a COVID-19 CT image classification model (37). In terms of efficiency, their proposed classification model surpasses supervised learning approaches and achieves high accuracy. Karacan et al. proposed a binary and trinary classification system using CXR images (38). Their proposed model includes MobileNetV2, DenseNet121, InceptionResNetV2, and Xception. These models were integrated with ensemble learning methods to improve their proposed model's performance further.

Narin et al. employed a Deep TL technique using chest radiographs to identify COVID-19 (39). The research used a technique known as five-fold cross-validation for three distinct binary classifications. Five models that had been pre-trained were applied to the three different datasets. According to their findings, The ResNet50 model offers the highest level of accuracy compared to the other four techniques included in the study. Using a precise weighted averaging ensemble model, Bhardwaj and Kaur attempt to detect COVID-19 and other pulmonary complications (40). Data augmentation strategies were implemented while training the four CNN models, DenseNet121, Xception, Inceptionv3, and InceptionResNetv2. The experiment's

binary classification accuracy was 98.33%, whereas, in the case of multi-class classification, they attained 92.36% accuracy.

The limitation of the previously mentioned research is that most of the studies used relatively limited CXR images, and others used very few COVID-19 radiographs. Some studies proposed binary classification models that cannot differentiate between bacterial and viral pneumonia. In this research, we proposed CIDICXR-Net50, a TL-based framework that uses a pre-trained model, ResNet50 architecture (41), for CXR image classification. Adopting the TL method reduces the impact of the problem of a restricted training dataset while providing us with the benefits of a shortened processing time, enhanced performance, and consistent results. The CIDICXR-Net50 model is built by substituting a new classification head for the last classifier layer in the ResNet50 model. The model is tested and trained using the dataset of 3,923 images, including 1,363 regular, 1,200 COVID-19, and 1,360 viral pneumonia Chest X-ray images, representing a sizeable dataset. The performance of the proposed CIDICXR-Net50 is assessed and compared with six other cutting-edge pre-trained models, including DenseNet-121, VGG-16, ResNet-101, VGG-19, InceptionV3, and MobileNetV2. The suggested CIDICXR-Net50 model achieved an accuracy of 99.11% on the provided dataset, with a 99.15% precision and recall rate.

- The current study offers a potential for cost-effective and swift diagnosis of Coronavirus disease using chest X-rays.
- This research presents a novel Deep Transfer Learning framework called CIDICXR-Net50, designed to aid radiologists in detecting COVID-19 from X-ray images with a high accuracy of 99.11%.
- In this study, we conducted a comprehensive performance evaluation of various deep learning architectures, offering insights into their accuracy in classifying COVID-19 based on an extensive X-ray image dataset.
- The proposed work facilitates collaborative efforts among interdisciplinary researchers to advance artificial intelligence methodologies within Computer-Aided Diagnosis (CAD) systems. To uncover potential linkages between COVID-19 and Diabetes Mellitus, thereby enhancing diagnostic precision and patient care strategies.

1.1. Relationship between COVID-19 and Diabetes

Diabetic patients, especially those with preexisting comorbidities or those in older age groups, have an elevated risk of COVID-19 infection (42). The trajectory of COVID-19 tends to be more severe for individuals with Diabetes, and they exhibit a markedly higher mortality rate (43). Through multivariable logistic regression analysis, Ciardullo et al. found that DM was an independent factor correlating with a rise in in-hospital mortality due to COVID-19 (44). Initial findings from China, subsequently supported by studies in the United States and Europe, revealed that the prevalence of Diabetes in individuals hospitalized with COVID-19 was as elevated as 20% (45–47). Emerging evidence indicates that Diabetes could potentially contribute as a risk factor for the occurrence of Post-Acute Sequelae of SARS-CoV-2 infection. After recuperating from the acute stage of COVID-19, certain individuals persistently suffer from symptoms over an extended duration, commonly known as "long COVID" or (PASC). Diabetic patients dealing with PASC may have difficulty controlling their blood sugar levels. There is still much to learn about the connection between COVID-19 and Diabetes, and research is underway. In order to provide Diabetes patients with the best care and outcomes possible throughout the pandemic, it is crucial to comprehend this link.

Once we confirm COVID-19 detection from the CXR image, then we can explore the relationship with Diabetes using some open datasets of electronic health records such as the National COVID Cohort Collaborative's (N3C) repository, COVID-19 can disrupt glycemic control in people with Diabetes. Infection and the body's immune response to the virus can lead to fluctuations in blood sugar levels, making it challenging for diabetic individuals to manage their condition effectively. Understanding the links between Diabetes and COVID-19 requires epidemiological, clinical, and molecular studies. Conditions that already exist, including a weakened immune response, viral replication, and persistent inflammation, are common contributors. These co-occurring conditions have also been linked to an amplified COVID-19 response. An impaired immune system is linked to poorly managed Diabetes. Individuals with Diabetes are at increased risk for severe complications from infections because their impaired immune systems cannot fight off the disease effectively. SARS-CoV-2, the virus responsible for COVID-19, may benefit from elevated blood glucose levels, speeding the course of the disease.

2. ResNet-50

The ResNet-50 is a 50-layer Residual Neural Network (RNN) variant trained on images from the ImageNet database. The main reason for proposing the ResNet-50 model was to avoid the vanishing gradients problem while constructing a deep neural network. Different variants of the ResNet model are available with varying layers. However, the most common model is called ResNet-50, and it comprises 49 Convolutional layers and a Fully Connected layer. ResNet altered the structure of CNNs by introducing the residual learning technique to train deep neural networks. ResNet50 was nominated as the ImageNet Large Scale Visual Recognition (ILSVRC) Challenge winner in 2015. Figure 1 shows the design of the ResNet model.

ResNet is 20 and 8 times deeper than AlexNet11 and VGG42, respectively, offering more accuracy. The ResNet network with 50, 101, and 152 layers performs significantly better than the ResNet network with 18 and 34 layers. The design of the ResNet-50 network is made up of sequences of convolutional blocks that use average pooling. As the final classification layer, Softmax is employed. ResNet has established shortcut connections between different layers to enable communication between the different layers. As a result of the layers' independence from parameters and data, non-residual functions can be characterized by them after a gated shortcut has been closed. In ResNet, shortcuts are never closed, but residual information is saved for good. Even as the search depth increases, it has lower computational complexity than VGG.

2.1. CIDICXR-Net50

The proposed CIDICXR-Net50 model is the modified version of the ResNet-50. In the CIDIXR-Net50 model, the network consists of one 7×7

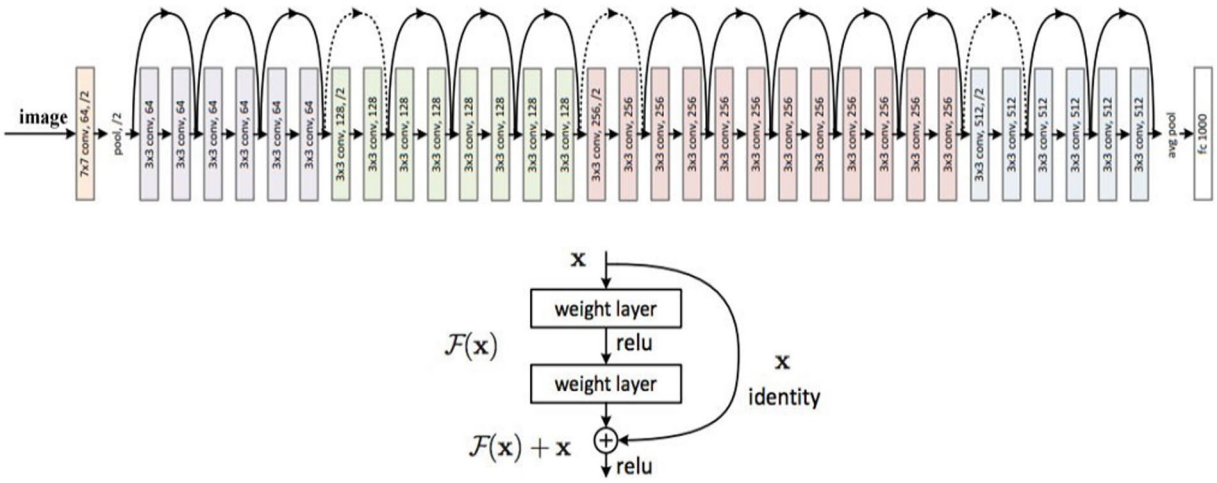


FIGURE 1
ResNet (Residual Network) architecture (41).

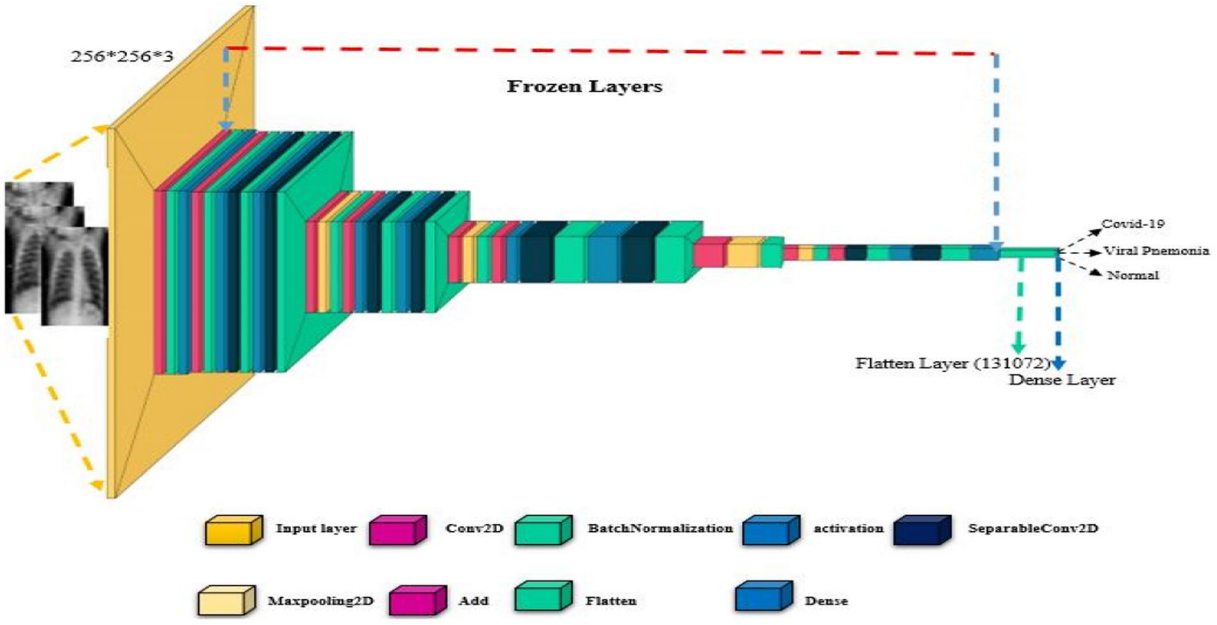


FIGURE 2
The architecture of the proposed CIDICXR-Net50.

convolutional layer, followed by three blocks of 1×1, 3×3, and 1×1 convolutional layers of size 56. Then, we have the same blocks of convolutional layers of sizes 28, 14, and 7. After that, instead of fully connected layers (as present in ResNet-50), we added flattened and dense layers. The flattened layer is used to transform the image into one dimension. The dense layer acts as a fully connected layer that uses Soft Max for multi-class classification (i.e., Normal, COVID-19, and Pneumonia). Figure 2 shows each layer of the original and proposed architectures sequentially. For both models, the first convolutional layer outputs a feature map of size 112 × 112 × 64 after applying 64 filters of size 7 × 7 × 3 over the input picture of size 224 × 224. The max-pooling layer processes the input feature map with a 3 × 3-pixel filter to create a 56 × 56 × 64 feature map. The initial 1×1 convolution layer is responsible for doing downsampling. However, in our proposed model CIDICXR-Net50, we use an input image of size 256 × 256 × 3 as it is the optimal image size. To effectively converge CNN models for training, hyper-parameters such as learning rate, optimization technique, dropout rate, and batch size must all be considered.

The proposed CIDICXR-Net50 model was trained using the Adam optimizer for 25 epochs, and the batch size was set to 32 with a learning rate 0.0001. A function called Early Stopping from the Keras library was also implemented. Model validation loss is tracked via this approach. If the model reaches its capacity and the validation loss continues to be constant, the best weights are kept, and the model is

terminated. Activities at various levels of training were carried out using the ModelCheckPoint and EarlyStop callback methods.

2.2. Experimentation preliminaries

This section discusses the experimental tools, frameworks, techniques, dataset, preprocessing steps, hyperparameters of the proposed model, and performance assessment parameters used to conduct the experimentation. The proposed CIDICXR-Net50 model is implemented in Python using the Keras framework and TensorFlow 2.8.0 in the backend. The experiments were conducted in a Google Colab (Collaboratory) environment with an NVIDIA GPU, 12GB RAM, and 2.3GHz Intel Xeon Processors.

2.3. Dataset

The proposed model was trained and evaluated using version 3 of the publicly available dataset 43 of CXR images. The dataset included 3,923 total CXR images, 1,363 Normal, 1,200 COVID-19, and 1,360 Viral Pneumonia. The images included in the dataset are two-dimensional and have three channels. Figure 3 illustrates a CXR of infected lungs with Viral Pneumonia, COVID-19, and normal lung Chest X-ray images. The dataset is divided into training, validation, and testing. Details of the dataset and subsets are shown in Table 1.

2.4. Data preprocessing

The dataset underwent preprocessing following the specifications of the suggested deep neural network model. Resizing and normalizing are the two essential procedures. Adjust the size of the CXR images to meet the specifications. The usual pre-trained models required fixed-size input images (such as 224 × 224, 227 × 227, 299 × 299), but the dataset contains images of varying sizes. As a result, all CXR images were resized to 256 × 256, and CXR images were normalized to [0,1] as an additional preprocessing step to meet the basic architecture's requirements.

2.5. Hyper-parameters of the CIDICXR-Net50 model

The proposed CIDICXR-Net50 model was trained for 25 epochs using a 0.0001 learning rate and 32 batch size. Adam (adaptive moment estimation) optimizer is used to develop the classification model. Adam Optimizer, proposed by Kingma and Ba (48), is robust against noisy gradients and flexible enough to be used with different neural network architectures and tasks. Adam combines Momentum and RMSprop's advantages to handle sparse gradients on noisy problems. The benefits of Adam Optimizer include Adaptive Learning Rates, Memory Efficiency, and Robust Variations. The Early Stopping function from the Keras library was implemented, monitoring the model's validation loss. When the model reaches saturation but the validation loss remains the same, the best weights are preserved, and the model is halted. In this research, we used ModelCheckPoint and EarlyStop callback mechanisms. During training, the ModelCheckPoint mechanism ensures that the model is preserved with minimum data loss. If the network enters a state of inactivity (no learning), the EarlyStop method will be employed to interrupt the training of the system. As the validation loss monitoring parameter, the patience value was set to 10 initially. Rather than immediately halting the training when the measure stops increasing, "patience"

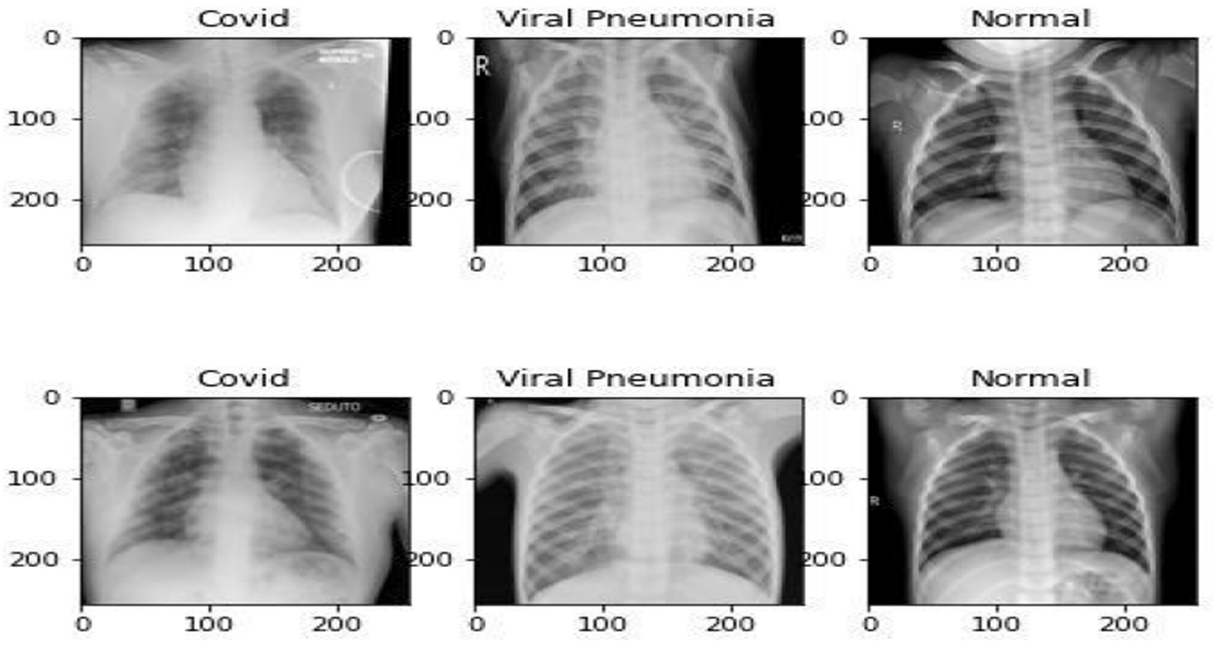


FIGURE 3
Preprocessed CXR images of Normal, Viral Pneumonia, and COVID-19.

TABLE 1 X-ray image dataset details.

Class	Training	Validation	Testing	Total CXR images
COVID-19	840	120	240	1,200
Normal	947	135	281	1,363
Viral pneumonia	956	135	269	1,360
Total CXR images	2,743	390	790	3,923

TABLE 2 Hyper-parameters of the proposed CIDICXR-Net50 model.

Hyper-parameters	Values
Optimizers	Adam
Learning rate	0.0001
EarlyStopping	patience = 10
Batch Size	32
Epochs	25
Callbacks	EarlyStop, ModelCheckPoint
Loss	Sparse Categorical Cross Entropy
Metrics	Accuracy

inserts a buffer. It is the number of epochs you will wait for the metric to improve again. The value of patience is set to 10, which means that if validation loss does not decrease for ten consecutive epochs, the training process will be stopped. The Accuracy rate was employed as a performance metric in this case. The hyper-parameters used to train the CIDICXR-Net50 model are given in Table 2.

2.6. Performance assessment metrics

For performance evaluation and comparison of the proposed CIDICXR-Net50 and other models, we used Precision, Recall, F-measure (F1-Score), and Accuracy. These metrics are generated using the confusion matrix values, i.e., True Positive, True Negative, False Negative, and False Positive. Equations (1–4) illustrate the performance metrics mentioned above.

$$\text{Accuracy} = \frac{TruePositive + TrueNegative}{TruePositive + FalsePositive + TrueNegative + FalseNegative} \quad (1)$$

$$\text{Precision} = \frac{TruePositive}{TruePositive + FalsePositive} \quad (2)$$

$$\text{Recall} = \frac{TruePositive}{TruePositive + FalseNegative} \quad (3)$$

$$\text{F1} - \text{Score} = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (4)$$

3. Results

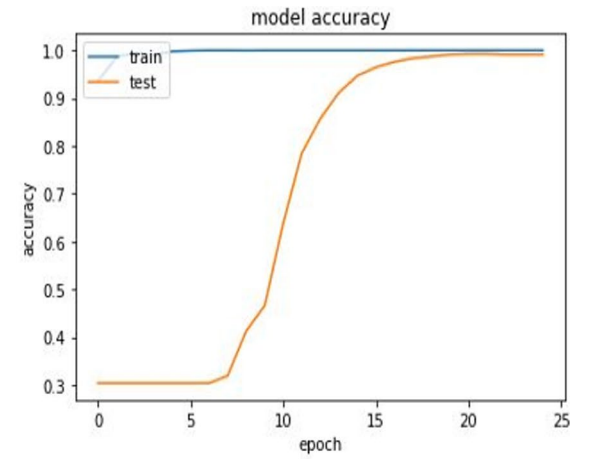
In this research, we proposed the CIDICXR-Net50 (COVID-19 Infection Detection In Chest X-Ray) model to detect COVID-19 using CXR radiographic images. The CIDICXR-Net50 model uses the base structure of the pre-trained ResNet-50 model using a transfer learning approach. The study demonstrates that deep learning can facilitate the diagnosis process as our proposed automated diagnostic tool, the CIDICXR-Net50 model, achieved an accuracy score of 99.11% overall. The suggested model was validated using 790 CXR images, which included 240 COVID-19, 281 normal CXR images, and 269 viral pneumonia. Figure 4 illustrates the proposed model's accuracy and loss.

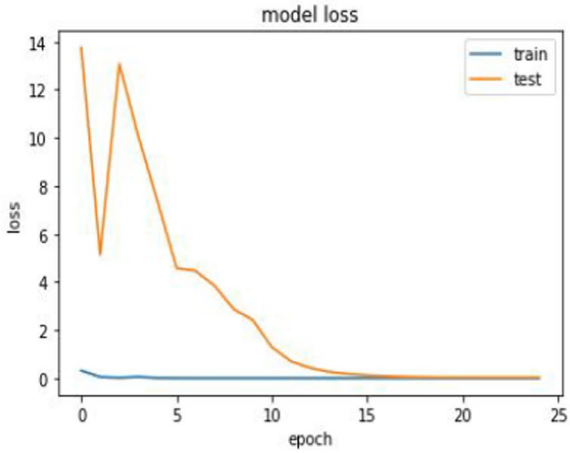
The trend graph shows the proposed model has no substantial overfitting or underfitting problems on the provided data. After 15 epochs, the performance curve for training and testing turns straight and progresses similarly. The proposed model accurately classified all 240 COVID-19 cases. Table 3 demonstrates each class's Precision, Recall, and F1-Score of the CIDICXR-Net50. The performance of the CIDICXR-Net50 model was also compared with other well-known Deep Learning models, including VGG-16, VGG19, DenseNet-121, InceptionV3, ResNet-101, and MobilNetV2. All the performance comparison experiments are conducted on the same dataset and its subsets (i.e., Training, Validation testing) with default parameters. The results show that the proposed CIDICXR-Net50 model classified 99.11% of CXR images accurately compared to other selected models on the given data. ResNet-101 and InceptionV3 are second and third best, with 98.99 and 98.61% accuracy. The performance of these classification algorithms in terms of recall, precision, f1-score, and accuracy is illustrated in Table 4. The confusion matrix for classification of COVID-19, normal and viral pneumonia using Different Deep Learning models are shown in Figure 5.

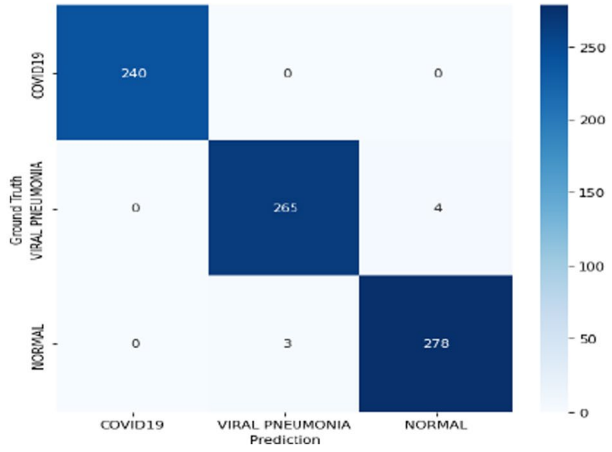
3.1. Discussion

This study aimed to construct a fully automated DL model called CIDICXR-Net50 to detect COVID-19 in chest X-ray images more accurately to classify COVID-19 CXR images from Viral Pneumonia and Normal CXR images. Previously, a hybrid technique was developed by Sethy et al. (30). Thirteen pre-trained DL models were used. An SVM classifier was trained using retrieved features from these models. ResNet-50 + SVM outperformed other classification models in ternary classification, with a sensitivity of 97.29% and an accuracy of 95.33%. The model was trained on 381 Chest X-ray images with an equal split across COVID-19, viral pneumonia, and normal. In contrast, the current CIDICXR-Net50 model accuracy and sensitivity are more significant by 3.78 and 1.82% on 11 times larger datasets.

Only 582 CT (Computed Tomography) scans have been used in the research. Multi-class classification accuracy of 99.11% and sensitivity of 99.15% are achieved by our proposed model, which was trained on a dataset five times larger (3923) than the one used by Pathak et al.DarkCovidNet automated model was suggested by Ozturk et al. (34) to detect COVID-19 in CXR. The model was developed for binary class (Normal and COVID) and multi-class classification (Normal, pneumonia, and COVID); it gained 87.02% for multi-class classification and 98.8% for binary classes. The CIDICXR-Net50







	precision	recall	f1-score	support
0	1.0000	1.0000	1.0000	240
1	0.9888	0.9851	0.9870	269
2	0.9858	0.9893	0.9876	281
accuracy			0.9911	790
macro avg	0.9915	0.9915	0.9915	790
weighted avg	0.9911	0.9911	0.9911	790

['0 = COVID19', '1 = VIRAL PNEUMONIA', '2 = NORMAL']

FIGURE 4
The overall result of the proposed CIDICXR-Net50 model.

TABLE 3 Precision, recall, and F1-Score of each class of the CIDICXR-Net50 model (proposed model).

CXR images class	F1-Score	Recall	Precision
Normal	0.99	0.99	0.99
COVID-19	1.00	1.00	1.00
Viral Pneumonia	0.99	0.99	0.99

model reported 99.11% accuracy and 99.15% sensitivity on a vast dataset (Three times larger). A more notable increase of 12.11 to 13.62% was also observed in accuracy and sensitivity for trinary classification. Chowdhury et al. (49) created a binary-class and multi-class classification framework for automatically recognizing COVID-19 using a pre-trained DenseNet-201-based transfer learning technique.

The study employed 3,487 CXR images, and the networks were trained using binary and trinary classification methods. The binary and trinary classification accuracy were 99.70 and 97.94%, respectively, while the proposed CIDICXR-Net50 model yields a 99.11% accuracy and a 99.15% sensitivity. The suggested model is trained and evaluated using a comparatively massive number of COVID-19 CXR images (1,200 versus 423). Apostolopoulos & Mpesiana (29) examined the five pre-trained models Inception, VGG19, Xception, InceptionResNetV2, and MobileNetV2 for detecting COVID-19 in Chest X-ray images. A total of 1,442 CXR images, including 224 verified cases of COVID-19, were used in the investigation, representing just 15.53% of the entire dataset. The primary goal of this research was to separate COVID-19 from normal lungs CXR and Bacterial/Viral Pneumonia CXR. The sensitivity and accuracy rates for MobileNetV2 were the highest, at 98.6 and 94.72%. In comparison, the proposed CIDICXR-Net50 has a sensitivity of 99.15% and an accuracy of 99.11% on a relatively large dataset. Wang et al.24 developed a COVID-Net model for multi-class classification using an open-source CXR images dataset COVIDx. Accuracy and sensitivity for the CIDICXR-Net50 were 5.81 and 8.15% higher than reported (93.3 and 91%). Hemdan et al. (31) proposed a binary classification model COVIDX-Net using seven different pre-trained frameworks, including InceptionV3, InceptionResNetV2, Xception, VGG19, DenseNet201, ResNetV2, and MobileNetV2. The researchers trained and tested their model on 50 CXR pictures from 25 COVID-19 instances.

Jain et al. (35) developed a two-step process for detecting COVID-19 in CXR. In phase I, their model uses ResNet-50 to

TABLE 4 Performance of selected classification models for Chest X-ray classification.

Models	Accuracy %	F1-Score %	Precision %	Recall %
CIDICXR-Net50 (Proposed)	99.11	99.15	99.15	99.15
ResNet-101	98.99	99.01	99.01	99.02
InceptionV3	98.61	98.67	98.68	98.66
VGG-19	98.48	98.52	98.52	98.53
DenseNet-121	98.48	98.53	98.53	98.53
VGG-16	98.35	98.37	98.43	98.32
MobileNetV2	83.04	82.90	87.59	84.05

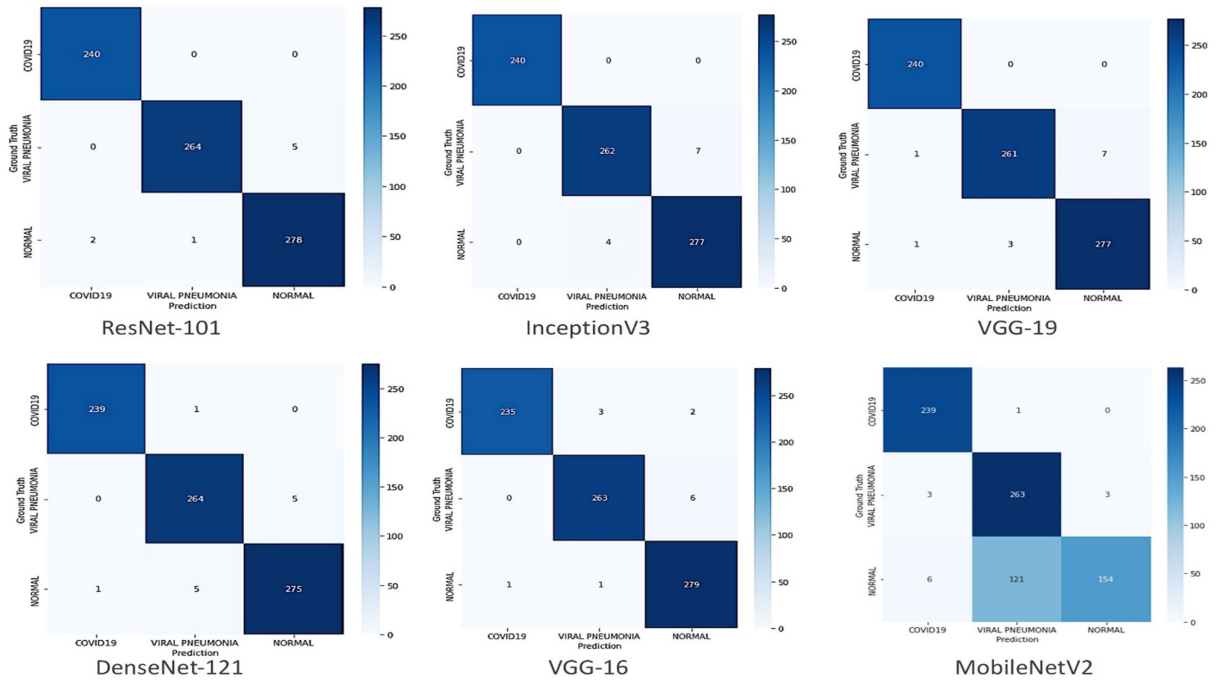


FIGURE 5
Confusion matrix for classification of COVID-19, normal and viral pneumonia using different deep learning models.

differentiate between bacterial and viral pneumonia in CXR images, including COVID-19. In multi-class classification, their model was 93.01% accurate. Phase II involved classifying COVID-19 CXR images from Viral Pneumonia using a pre-trained model based on ResNet-101. Experiments were performed on a dataset of 1,215 images that is publically accessible, and these experiments are supplemented further by data augmentation techniques. Their ResNet-101-based model achieved 97.22% accuracy. Manokaran et al. (32) suggested the DenseNet-201 base model, which detects COVID-19 CXR images with 94% accuracy. They used a dataset of 8,644 CXR images for experimentation, including 4,000 Normal, 4,000 Pneumonia, and only 644 COVID-19 cases. The results of their model outperformed other models by getting 92.19% accuracy. However, The key constraint was that their model was trained on a limited number of COVID-19 instances, and just 129 COVID-19 cases validated the conclusion.

In contrast, our proposed CIDICXR-Net50 model achieved 99.11% accuracy. Chakraborty et al. (33) proposed a VGG-19-based TL model to classify normal CXR, viral pneumonia CXR, and COVID-19 CXR. The accuracy of their model on the same dataset was 97.11%, whereas our proposed technique achieves 99.11% accuracy. The details of all previously mentioned research with their methods and accuracy are shown in Table 5.

4. Research limitations

Even though many medical imaging applications have achieved a good level of performance by utilizing deep learning models, many of these applications have failed clinical trials because of several issues, including a restricted training dataset, generalization, and overfitting. Training the CNN model on medical images instead of natural images (ImageNet) is recommended to obtain relevant medical characteristics. In this regard, a massive database of medical images is required to train the algorithm from scratch. Due to the recent disease outbreak and other factors, such as restrictions imposed by legal requirements that prohibit sharing patient CXR images, only a small amount of data is now available

TABLE 5 Comparative performance analysis of CIDICXR-NET50 against other leading techniques.

Model	Dataset	Methodology	Dataset size	Classification	Accuracy %
Sethy et al. (30)	CXR Images	TL ResNet50 + SVM	381	Multi-class	95.33
Pathak et al. (37)	CT Images	TL ResNet50	852	Binary-class	93.01
Ozturk et al. (34)	CXR Images	TL DarkCovidNet	1,127	Multi-class	87.02
Chowdhury et al. (49)	CXR Images	TL DenseNet-201	3,487	Binary-class	99.70
Chowdhury et al. (49)	CXR Images	TL DenseNet-201	3,487	Multi-class	97.94
Apostolopoulos & Mpesiana (29)	CXR Images	TL MobileNetV2	1,442	Multi-class	94.72
Wang et al. (28)	CXR Images	TL COVID-Net	13,975	Multi-class	93.3
Hemdan et al. (31)	CXR Images	TL COVIDX-Net	53	Binary-class	90
Jain et al. (Phase I) (35)	CXR Images	TL ResNet-50	1832	Multi-class	93
Jain et al. (Phase II) (35)	CXR Images	TL ResNet101	1832	Binary-class	97.78
Manokaran et al. (32)	CXR Images	TL DenseNet-201	8,644	Multi-class	92.19
Chakraborty et al. (33)	CXR Images	TL VGG-19	3,797	Multi-class	97.11
CIDICXR-Net50 (Proposed)	CXR Images	TL ResNet-50	3,923	Multi-class	99.11

in open sources, which is inadequate to train the model from scratch. However, This research delves into the interaction between COVID-19 and Diabetes. However, due to the absence of openly accessible and pertinent electronic health records (EHR), this study did not present its own compiled findings that led to clinical practice.

5. Conclusion and future work

This study proposed CIDICXR-Net50, a deep ResNet50 base model, using a sizeable balanced dataset of CXR images and a TL technique to classify images of viral pneumonia, COVID-19, and standard CXR images. This study further delved into the intricate connection between Diabetes mellitus and its association with COVID-19. It was underscored that diabetic patients exhibit a heightened vulnerability to contracting COVID-19 and are more likely to develop post-acute sequelae of COVID-19 (PASC).To determine how well the suggested model performs compared to six other pre-trained models, including VGG-16, VGG19, DenseNet-121, InceptionV3, ResNet-101, and MobilNetV2. The proposed model outperformed the other selected models' overall accuracy, efficiently separating patients diagnosed with COVID-19 from those diagnosed with normal or viral pneumonia. The results demonstrate that the proposed fully automated CIDICXR-Net50 model can detect COVID-19 infection with better accuracy. The CIDICXR-Net50 model proposed in this study can accurately detect COVID-19 from a dataset of ternary classes, another achievement of this research. The results of the experiments and assessments based on metrics show that the suggested model is suitable for use as a computer-aided diagnostics (CAD) system in hospitals and other medical facilities to diagnose COVID-19 disease in its early phases. This study supports the belief that deep learning algorithms have enormous potential for optimizing healthcare and improving diagnosis and treatment outcomes. The performance can be enhanced further in future work by increasing the dataset size. Collecting additional CXR images will increase the robustness and power of the proposed CIDICXR-Net50 model. To prevent overfitting issues and maximize generalizability, the developers of COVID-19 DL diagnostic models must train their models on vast and diverse datasets. Additionally, Because of the opaque nature of deep learning models, doctors may hesitate to rely on their results while making life-or-death decisions; therefore, Explainable Artificial Intelligence (XAI) techniques must be explored in the medical domain. Explanations are essential in the medical field, where every mistake might have severe consequences.

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.

Author contributions

IA: Conceptualization, Investigation, Methodology, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. AM: Conceptualization, Investigation, Methodology, Software, Supervision, Validation, Writing – original draft, Writing – review & editing. FA: Conceptualization, Investigation, Methodology, Software, Supervision, Writing – original draft, Writing – review & editing. BS: Conceptualization, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. AA: Conceptualization, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MA: Conceptualization, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Researchers Supporting Project number (RSP2023R395), King Saud University, Riyadh, Saudi Arabia.

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

1. Shereen MA, Khan S, Kazmi A, Bashir N, Siddique R. COVID-19 infection: origin, transmission, and characteristics of human coronaviruses. *J Adv Res*. (2020) 24:91–8. doi: 10.1016/j.jare.2020.03.005

2. Islam MZ, Islam MM, Asraf A. A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. *Informatics Med Unlocked*. (2020) 20:100412. doi: 10.1016/j.imu.2020.100412

3. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England J Med*. (2020) 382:1199–207. doi: 10.1056/NEJMOA2001316

4. Guaman-Bautista LP, Moreta-Urbano E, Oña-Arias CG, Torres-Arias M, Kyriakidis NC, Malcı K, et al. Tracking SARS-CoV-2: novel trends and diagnostic strategies. *Diagnostics*. (2021) 11. doi: 10.3390/DIAGNOSTICS11111981

5. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DKW, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eur Secur*. (2020) 25:1. doi: 10.2807/1560-7917.ES.2020.25.3.2000045

6. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan. *China Lancet*. (2020) 395:497–506. doi: 10.1016/S0140-6736(20)30183-5

7. Liao J, Fan S, Chen J, Wu J, Xu S, Guo Y, et al. Epidemiological and clinical characteristics of COVID-19 in adolescents and young adults. *Innov*. (2020) 1:100001. doi: 10.1016/J.XINN.2020.04.001

8. Qian ZS, Yang L, Xiang ZP, Bo LH, Liu F, Sheng ZR. Recommendations and guidance for providing pharmaceutical care services during COVID-19 pandemic: a China perspective. *Res Soc Adm Pharm*. (2021) 17:1819–24. doi: 10.1016/J.SAPHARM.2020.03.012

9. Vandenberg O, Martiny D, Rochas O, van Belkum A, Kozlakidis Z. Considerations for diagnostic COVID-19 tests. *Nat Rev Microbiol*. (2020) 19:171–83. doi: 10.1038/s41579-020-00461-z

10. Lee CYP, Lin RTP, Renia L, Ng LFP. Serological approaches for COVID-19: epidemiologic perspective on surveillance and control. *Front Immunol*. (2020) 11:879. doi: 10.3389/FIMMU.2020.00879/BIBTEX

11. Wang L, Zhang Y, Wang D, Tong X, Liu T, Zhang S, et al. Artificial intelligence for COVID-19: a systematic review. *Front Med*. (2021) 8:704256. doi: 10.3389/FMED.2021.704256/BIBTEX

12. Agrebi S, Larbi A. Use of artificial intelligence in infectious diseases. *Artif Intell Precis Heal*. (2020):415. doi: 10.1016/B978-0-12-817133-2.00018-5

13. Lei P, Huang Z, Liu G, Wang P, Song W, Mao J, et al. Clinical and computed tomographic (CT) images characteristics in the patients with COVID-19 infection: what should radiologists need to know? *J Xray Sci Technol*. (2020) 28:369–81. doi: 10.3233/XST-200670

14. Nishio M, Kobayashi D, Nishioka E, Matsuo H, Urase Y. Deep learning model for the automatic classification of COVID - 19 pneumonia, non - COVID - 19 pneumonia, and the healthy: a multi - center retrospective study. *Sci Rep*. (2022) 12:1–10. doi: 10.1038/s41598-022-11990-3

15. Jaiswal AK, Tiwari P, Kumar S, Gupta D, Khanna A, Rodrigues JJPC. Identifying pneumonia in chest X-rays: a deep learning approach. *Measurement*. (2019) 145:511–8. doi: 10.1016/J.MEASUREMENT.2019.05.076

16. Wang X, Deng X, Fu Q, Zhou Q, Feng J, Ma H, et al. A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT. *IEEE Trans Med Imaging*. (2020) 39:2615–25. doi: 10.1109/TMI.2020.2995965

17. Ko H, Chung H, Kang WS, Kim KW, Shin Y, Kang SJ, et al. COVID-19 pneumonia diagnosis using a simple 2D deep learning framework with a single chest CT image: model development and validation. *J Med Internet Res*. (2020) 22:e19569. doi: 10.2196/19569

18. Dai WC, Zhang HW, Yu J, Xu HJ, Chen H, Luo SP, et al. CT imaging and differential diagnosis of COVID-19. *Can Assoc Radiol J*. (2020) 71:195–200. doi: 10.1177/0846537120913033

19. Asiri N, Hussain M, Al Adel F, Alzaidi N. Deep learning based computer-aided diagnosis systems for diabetic retinopathy: a survey. *Artif Intell Med*. (2019) 99:101701. doi: 10.1016/J.ARTMED.2019.07.009

20. Gaál G, Maga B, Lukács A. Attention U-net based adversarial architectures for chest X-ray lung segmentation. *CEUR Workshop Proc*. (2020) 2692. Available at: https://arxiv.org/abs/2003.10304v1

21. Mohanasundaram R, Malhotra AS, Arun R, Periasamy PS. Deep learning and semi-supervised and transfer learning algorithms for medical imaging. *Deep Learn Parallel Comput Environ Bioeng Syst*. (2019):139–51. doi: 10.1016/B978-0-12-816718-2.00015-4

22. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. *Nat Med*. (2019) 25:24–9. doi: 10.1038/s41591-018-0316-z

23. Whang SE, Roh Y, Song H, Lee JG. Data collection and quality challenges in deep learning: a data-centric AI perspective. *VLDB J*. (2023) 32:791–813. doi: 10.1007/S00778-022-00775-9/FIGURES/22

24. Rana M, Bhushan M. Machine learning and deep learning approach for medical image analysis: diagnosis to detection. *Multimed Tools Appl*. (2023) 82:26731–69. doi: 10.1007/S11042-022-14305-W/TABLES/1

25. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Springer International Publishing*. (2021) 8:53. doi: 10.1186/s40537-021-00444-8

26. Atiya SU, Ramesh NVK, Reddy BNK. Classification of non-small cell lung cancers using deep convolutional neural networks. *Multimed Tools Appl*. (2023) 82:1–30. doi: 10.1007/S11042-023-16119-W/TABLES/6

27. Aggarwal P, Mishra NK, Fatimah B, Singh P, Gupta A, Joshi SD. COVID-19 image classification using deep learning: advances, challenges and opportunities. *Comput Biol Med*. (2022) 144:105350. doi: 10.1016/J.COMPBIOMED.2022.105350

28. Wang L, Lin ZQ, Wong A. COVID-net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Sci Rep*. (2020) 10:19549–12. doi: 10.1038/s41598-020-76550-z

29. Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Phys Eng Sci Med*. (2020) 43:635–40. doi: 10.1007/s13246-020-00865-4

30. Sethy PK, Behera SK, Ratha PK, Biswas P. Detection of coronavirus disease (COVID-19) based on deep features and support vector machine. *Int J Math Eng Manag Sci*. (2020) 5:643–51. doi: 10.33889/IJMEMS.2020.5.4.052

31. Hemdan EE-D, Shouman MA, Karar ME. *COVIDX-net: A framework of deep learning classifiers to diagnose COVID-19 in X-ray images*. (2020) doi: 10.48550/arXiv.2003.11055

32. Manokaran J, Zabihollahy F, Hamilton-Wright A, Ukwatta E. Detection of COVID-19 from chest x-ray images using transfer learning. *J Med Imaging*. (2021) 8:1–12. doi: 10.1117/1.jmi.8.s1.017503

33. Chakraborty S, Paul S, Hasan KMA. A transfer learning - based approach with deep CNN for COVID - 19 - and pneumonia - affected chest X - ray image classification. *SN Comput Sci*. (2022) 3:17. doi: 10.1007/s42979-021-00881-5

34. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra AU. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med*. (2020) 121:103792. doi: 10.1016/j.compbiomed.2020.103792

35. Jain G, Mittal D, Thakur D, Mittal MK. A deep learning approach to detect Covid-19 coronavirus with X-ray images. *Biocybern Biomed Eng*. (2020) 40:1391–405. doi: 10.1016/j.bbe.2020.08.008

36. Vaid S, Kalantar R, Bhandari M. Deep learning COVID-19 detection bias: accuracy through artificial intelligence. *Int Orthop*. (2020) 44:1539–42. doi: 10.1007/s00264-020-04609-7

37. Pathak Y, Shukla PK, Tiwari A, Stalin S, Singh S, Shukla PK. Deep transfer learning based classification model for COVID-19 disease. *IRBM*. (2020) 43:87–92. doi: 10.1016/j.irbm.2020.05.003

38. Karacan H, Eryilmaz F. Covid-19 detection from chest X-ray images and hybrid model recommendation with convolutional neural networks. *J Adv Res Nat Appl Sci*. (2021) 7:486–503. doi: 10.28979/jarnas.952700

39. Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Anal Applic*. (2021) 24:1207–20. doi: 10.1007/s10044-021-00984-y

40. Bhardwaj P, Kaur A. A novel and efficient deep learning approach for COVID-19 detection using X-ray imaging modality. *Int J Imaging Syst Technol*. (2021) 31:1775–91. doi: 10.1002/ima.22627

41. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. (2015):770–8. doi: 10.48550/arxiv.1512.03385

42. Sindi AA, Tashkandi WA, Jastaniah MW, Bashanfar MA, Fakhri AF, Alsallum FS, et al. Impact of diabetes mellitus and comorbidities on mortality in patients with COVID-19: a single-center retrospective study. *Saudi Med J*. (2023) 44:67–73. doi: 10.15537/SMJ.2023.44.1.20220462

43. Wargny M, Potier L, Gourdy P, Pichelin M, Amadou C, Dupuy O, et al. Predictors of hospital discharge and mortality in patients with diabetes and COVID-19: updated results from the nationwide CORONADO study. (2021). 64:778–94. doi: 10.1007/s00125-020-05351-w/Published

44. Ciardullo S, Zerbini F, Perra S, Muraca E, Cannistraci R, Lauriola M, et al. Impact of diabetes on COVID-19-related in-hospital mortality: a retrospective study from northern Italy. *J Endocrinol Investig*. (2021) 44:843–50. doi: 10.1007/s40618-020-01382-7

45. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med*. (2020) 382:1708–20. doi: 10.1056/NEJMOA2002032/SUPPL_FILE/NEJMOA2002032_DISCLOSURES.PDF

46. Grasselli G, Zangrillo A, Zanella A, Antonelli M, Cabrini L, Castelli A, et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy region, Italy. *JAMA*. (2020) 323:1574–81. doi: 10.1001/JAMA.2020.5394

47. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the new York City area. *JAMA*. (2020) 323:2052–9. doi: 10.1001/JAMA.2020.6775

48. Kingma DP, Ba JL. Adam: a method for stochastic optimization. *3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc*. (2015). doi: 10.48550/arXiv.1412.6980

49. Chowdhury MEH, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub ZB, et al. Can AI help in screening viral and COVID-19 pneumonia? IEEE. *Access*. (2020) 8:132665–76. doi: 10.1109/ACCESS.2020.3010287

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 12 December 2023
DOI 10.3389/fpubh.2023.1331517

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Sidra Abbas,
COMSATS University Islamabad, Pakistan
Gohar Ali,
Sur University College, Oman
Jagpreet Sidhu,
SVKM's Narsee Monjee Institute of Management Studies, India

*CORRESPONDENCE
Jaiteg Singh
✉ jaitegkhaira@gmail.com
Farman Ali
✉ farmankanju@gmail.com

RECEIVED 01 November 2023
ACCEPTED 21 November 2023
PUBLISHED 12 December 2023

CITATION
Thakur D, Gera T, Bhardwaj V, AlZubi AA, Ali F and Singh J (2023) An enhanced diabetes prediction amidst COVID-19 using ensemble models. *Front. Public Health* 11:1331517. doi: 10.3389/fpubh.2023.1331517

An enhanced diabetes prediction amidst COVID-19 using ensemble models

Deepak Thakur[1], Tanya Gera[1], Vivek Bhardwaj[2], Ahmad Ali AlZubi[3], Farman Ali[4]* and Jaiteg Singh[1]*

[1]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India, [2]Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur, Rajasthan, [3]Department of Computer Science, Community College, King Saud University, Riyadh, Saudi Arabia, [4]Department of Computer Science and Engineering, School of Convergence, College of Computing and Informatics, Sungkyunkwan University, Seoul, Republic of Korea

In the contemporary landscape of healthcare, the early and accurate prediction of diabetes has garnered paramount importance, especially in the wake of the COVID-19 pandemic where individuals with diabetes exhibit increased vulnerability. This research embarked on a mission to enhance diabetes prediction by employing state-of-the-art machine learning techniques. Initial evaluations highlighted the Support Vector Machines (SVM) classifier as a promising candidate with an accuracy of 76.62%. To further optimize predictions, the study delved into advanced feature engineering techniques, generating interaction and polynomial features that unearthed hidden patterns in the data. Subsequent correlation analyses, visualized through heatmaps, revealed significant correlations, especially with attributes like Glucose. By integrating the strengths of Decision Trees, Gradient Boosting, and SVM in an ensemble model, we achieved an accuracy of 93.2%, showcasing the potential of harmonizing diverse algorithms. This research offers a robust blueprint for diabetes prediction, holding profound implications for early diagnosis, personalized treatments, and preventive care in the context of global health challenges and with the goal of increasing life expectancy.

KEYWORDS

diabetes, COVID-19, ensemble models, classification, feature engineering, interaction, polynomial, correlation analysis

1 Introduction

The dawn of the twenty-first century has been illuminated by the transformative power of data-driven methodologies. This era, characterized by the amalgamation of technology and healthcare, has witnessed the birth and growth of innovative tools and techniques geared toward enhancing patient care, diagnosis, and management. These advancements have not only revolutionized medical treatments but have also given rise to predictive healthcare, an approach that leverages data to forecast medical outcomes, thereby enabling timely interventions.

As the world finds itself in the throes of the COVID-19 pandemic, a public health crisis of unparalleled magnitude, the significance of predictive healthcare is amplified manifold. The virus, while a threat to all, poses heightened risks to certain vulnerable demographics. Notably, individuals with pre-existing conditions, such as diabetes, have been identified as being particularly susceptible to severe manifestations of the virus (1, 2). This revelation underscores the criticality of early diabetes detection and management, both from a patient wellbeing perspective and from a broader public health standpoint (3).

Machine learning, with its deep-rooted capabilities in pattern recognition and data analytics (4), emerges as a beacon of hope in this scenario (5, 6). Its prowess in sifting through vast datasets, identifying hidden correlations, and predicting outcomes positions it as a formidable tool in the medical diagnostic toolkit. However, the multifaceted and often non-linear nature of medical data calls for an approach that goes beyond traditional algorithms. Feature engineering, a process that refines and transforms data attributes, presents itself as an indispensable ally in this quest. By generating interaction features, crafting polynomial attributes, and more, feature engineering seeks to enhance the richness of the dataset, making it more conducive to accurate predictions (7–9).

Yet, the path to optimal prediction is not solely paved by feature engineering. Ensemble models, which harmonize the strengths of multiple machine learning algorithms, offer a layer of sophistication and robustness. By synergizing diverse algorithms, ensemble models aspire to deliver predictions that are not only accurate but also consistent across varied scenarios.

With this contextual landscape as the backdrop, our research is anchored in a clear vision: to harness the combined might of machine learning, feature engineering, and ensemble models to redefine the standards of diabetes prediction, especially in the shadow of the COVID-19 pandemic (10).

The main objectives of this paper are:

1. To assess the performance of various classifiers, such as Support Vector Machines (SVM), Logistic Regression, Gradient Boosting and Random Forest, in predicting diabetes.
2. To delve into advanced feature engineering techniques, creating interaction features and generating polynomial attributes, aiming to capture hidden relationships in the data.
3. To gauge the efficacy of engineered features in relation to diabetes prediction outcomes, using visual tools like heatmaps.
4. To design and evaluate an ensemble model that integrates the strengths of diverse algorithms, targeting enhanced predictive accuracy.
5. To situate the findings within the broader context of the COVID-19 pandemic, examining the implications of accurate diabetes prediction in managing COVID-19 vulnerabilities.

Section 2 presents the Literature Review, where we critically examine previous studies that have utilized the Pima Indians Diabetes Database, delving into their methodologies, results, and conclusions. A key component of this section is the identification of the Research Gap. By pinpointing gaps or limitations in prior studies, we articulate the unique contributions our research aims to make. Section 3 outlines our Methodology. Within this section, we offer a detailed description of the dataset under Data Collection. We then discuss the various steps undertaken to refine the dataset in Data Preprocessing, from handling missing values to normalization and feature engineering. Section 4 showcases our Results. Here, we provide a statistical overview of each feature in the dataset through Descriptive Statistics. We then present the outcomes of our predictive models in Model Performance, illustrated through tables, charts, or graphs. The significance of different features in the prediction process is discussed in Feature Importance. This section, offers an interpretation of our results, drawing insights, and contextualizing our findings. Finally, Section 6 concludes the paper with a summary of our main findings, a discussion on the broader implications of our results, and suggestions for potential avenues for future research.

2 Literature review

A proposed e-diagnosis system leveraging machine learning (ML) algorithms was introduced for the Internet of Medical Things (IoMT) environment, specifically targeting type 2 diabetes diagnosis (11). Despite ML's promise, skepticism arises due to its opaque decision-making process, causing hesitancy in its adoption within some healthcare domains. The study employed three transparent ML models—Naïve Bayes, random forest, and J48 decision tree—using the Pima Indians diabetes dataset. Results indicated a preference for Naïve Bayes with select features, while random forests excelled with a richer feature set.

In the study (12), various methods were explored to determine the likelihood of diabetes mellitus. Four prominent classification approaches were initially assessed, namely Decision Tree, Neural Structures, Regression Analysis, and Probability-based Classification. Subsequently, aggregation strategies such as Bagging and Boosting were examined to enhance model stability. The Random Forest approach was also incorporated into the evaluations. Results indicated that the Random Forest method outperformed the others in disease risk determination. Based on these findings, an online tool was developed leveraging the Random Forest approach for diabetes risk categorization.

The primary objective of this research is to evaluate the efficacy of different algorithms in forecasting diabetes through data analysis techniques (13). This study assesses various computational classifiers, including the J48 Decision Tree, K-Nearest Neighbors, Random Forest, and Support Vector Machines, aiming to categorize individuals with diabetes mellitus. The algorithms were evaluated using data samples sourced from the UCI learning data archive. Their performance was analyzed on datasets both before and after data cleaning, and the results were benchmarked based on Accuracy, Sensitivity, and Specificity metrics.

The study in (14) introduces a technique to categorize patients with diabetes based on a set of features aligned with World Health Organization guidelines. By applying advanced data analysis algorithms to real-world data, a precision of 0.770 and a recall of 0.775 were achieved utilizing the HoeffdingTree method.

Historically, many clinical decision support systems, as documented in multiple studies, have been anchored in data mining techniques to predict diabetes onset and its progression. These traditional systems predominantly rely on singular classifier models or their uncomplicated combinations. However, a discernible trend in contemporary literature highlights the pivot toward ensemble classifiers. For instance, authors in (15) have delved into the efficacy of ensemble techniques, particularly emphasizing the adaboost and bagging methods, often utilizing decision trees like J48 (analogous to c4.5) as the foundational model. Moreover, specific research efforts have concentrated on the Canadian Primary Care Sentinel Surveillance network, aiming to classify individuals across various adult age groups based on diabetes

risk determinants. Such studies have collectively underscored the potential superiority of ensemble methods, especially adaboost, in enhancing prediction accuracy compared to conventional methods.

In the study (16), authors endeavors to meticulously review the infusion of machine learning and data mining paradigms in diabetes research. The focus areas being: (a) Prognostication and Diagnostic Processes, (b) Complications arising from Diabetes, (c) Interplay of Genetics and Environment, and (e) Healthcare Administration and Management. Notably, predictive and diagnostic applications have garnered heightened attention. The landscape of algorithms showcased a dominance of supervised learning techniques, constituting 85%, with the remaining 15% gravitating toward unsupervised methodologies, predominantly association rules. Among the gamut of algorithms, Support Vector Machines (SVM) emerged as the predominant choice.

In (17), the authors employed decision tree, random forest, and neural network algorithms to forecast diabetes mellitus using hospital examination datasets. Adopting a five-fold cross-validation and independent test experiments on a balanced sample of 68,994 records, the study addressed data imbalance through multiple random data extractions. Dimensionality reduction was achieved using principal component analysis (PCA) and minimum redundancy maximum relevance (mRMR). Notably, the random forest algorithm, utilizing the full attribute spectrum, showcased superior accuracy, registering at 0.8084.

In (18), the authors conducted an in-depth examination of complications and blood glucose prognosis in non-adherent T2D patients, sourcing data from inpatients at Sichuan Provincial People's Hospital between 2010 and 2015. Targeting T2D patients without recent monitoring or treatment adjustments, 18 predictive models were crafted using seven machine learning techniques, evaluated primarily through the area under the curve metric. Results revealed that out of 800 T2D patients, 165 qualified for the study, with 78.2% exhibiting poor glycemic control. Notable predictive performance was observed in areas like diabetic nephropathy (AUC = 0.902) and diabetic peripheral neuropathy (AUC = 0.859).

In (19), predictions of fasting plasma glucose levels were derived using a series of 100 bootstrap iterations, encompassing varied data subsets that mirrored biannual data influxes. Initial analyses, grounded in 6-month data snapshots, illuminated the primacy of the rudimentary regression model, recording the minimal RMSE at 0.838, trailed by RF, LightGBM, Glmnet, and subsequently XGBoost. As the data repository expanded, Glmnet showcased a noteworthy enhancement trajectory, peaking at an increment rate of 3.4%.

Utilizing Hadoop clusters, which are tailored for efficient processing and storage of vast datasets in a cloud setting, has been pivotal. Authors in (20), introduces a pioneering approach by integrating machine learning techniques within these Hadoop-based clusters, specifically for predicting diabetes. The outcomes underline the efficacy of these algorithms in yielding high-accuracy predictive systems for diabetes. For the assessment of the algorithm's functionality, the Pima Indians Diabetes Database from the National Institute of Diabetes and Digestive Diseases was employed.

In (21), researchers study delves into the comparative analysis of conventional classification techniques against neural network-driven machine learning approaches, specifically for a diabetes dataset. Furthermore, a plethora of performance metrics are assessed across multiple algorithms, such as K-nearest neighbor, Naive Bayes, extra trees, decision trees, radial basis function, and multilayer perceptron. The objective is to enhance the predictive accuracy for potential future diabetes cases in patients. From the findings, it becomes evident that the multilayer perceptron algorithm outperforms others, registering the peak accuracy, a minimal MSE at 0.19, and boasts the least instances of false positives and negatives, culminating in an impressive area under the curve of 86%.

The existing body of research predominantly operates in silos, either focusing on individual algorithms or generic feature engineering. There is limited exploration of harmonizing diverse algorithms in an ensemble model, especially in the context of diabetes prediction during global health crises. This presents an opportune avenue for innovation, highlighting the need for a comprehensive approach that seamlessly integrates state-of-the-art machine learning techniques with advanced feature engineering. Such an amalgamation not only promises enhanced predictive accuracy but also paves the way for more holistic patient care, encompassing early diagnosis, tailored treatments, and proactive preventive measures. Our research seeks to bridge this gap. We endeavor to amalgamate the strengths of proven algorithms, supplementing them with nuanced feature engineering techniques to craft a sophisticated model for diabetes prediction. Our focus remains steadfast on providing a solution that is not only academically rigorous but also clinically impactful, especially in the current global health landscape dominated by the challenges posed by COVID-19. Table 1 offers a concise representation of each study's focus and findings.

The word cloud depicted in Figure 1 generated from the literature survey provides a visual representation of the most frequently mentioned terms in the examined studies. Several observations can be drawn.

The most prominent terms, such as "machine learning," "diabetes," "algorithm," and "prediction," highlight the core focus of the literature, emphasizing the integration of advanced computational methods in diabetes diagnosis and prognosis (22, 23). Terms like "Hadoop," "cloud," and "IoMT (Internet of Medical Things)" indicate the contemporary shift toward integrating modern technological infrastructures with medical research, particularly in the realm of diabetes. The frequent appearance of words like "Random Forest," "Neural Network," "Decision Tree," and "Naive Bayes" underscores the popular machine learning algorithms employed in the studies (24, 25). Their prominence suggests their effectiveness or popularity in diabetes prediction tasks. The mention of "Pima Indians Diabetes Database" signifies its recurrent usage as a benchmark dataset for diabetes research, emphasizing its relevance and importance in the field (26–28). Words such as "accuracy," "AUC (Area Under the Curve)," and "MSE (Mean Squared Error)" highlight the key metrics used in the literature to evaluate the performance of predictive models. Their presence underscores the emphasis on quantitative assessment in the studies. The appearance of terms like "data imbalance,"

TABLE 1 Summary of work discussed in literature survey.

References	Key focus and techniques	Key findings and outcomes
Chang et al. (11)	E-diagnosis in IoMT using transparent ML models: Naïve Bayes, random forest, and J48.	Naïve Bayes preferred with certain features, but random forests excelled with a richer feature set.
Nai-Arun and Moungmai (12)	Predicting diabetes mellitus using various classification approaches and ensemble strategies.	Random Forest was the standout performer in risk determination.
Kandhasamy and Balamurali (13)	Evaluating different algorithms for diabetes forecasting using UCI data.	Multiple classifiers were assessed with performance metrics such as Accuracy, Sensitivity, and Specificity.
Mercaldo et al. (14)	Diabetes patient categorization using features aligned with WHO guidelines.	HoeffdingTree method achieved a precision of 0.770 and a recall of 0.775.
Perveen et al. (15)	Emphasis on the efficacy of ensemble techniques for diabetes onset prediction, focusing on adaboost and bagging.	Ensemble methods, particularly adaboost, were found to have potentially superior prediction accuracy.
Kavakiotis et al. (16)	Comprehensive review of ML and data mining in diabetes research.	Supervised learning dominated the landscape at 85%, with SVM emerging as the most popular algorithm.
Zou et al. (17)	Diabetes prediction using decision tree, random forest, and neural networks on hospital data.	Random forest showcased the highest accuracy of 0.8084 when leveraging the full set of attributes.
Fan et al. (18)	Examination of complications and blood glucose prognosis in non-adherent T2D patients.	Notable predictive performance in areas like diabetic nephropathy and diabetic peripheral neuropathy.
Kopitar et al. (19)	Predictions of fasting plasma glucose levels using multiple algorithms on biannual data influxes.	Initial analyses favored the simple regression model, but Glmnet showcased significant improvements as data increased.
Yuvaraj and Sripreethaa (20)	Diabetes prediction in Hadoop clusters leveraging ML.	Demonstrated the potential of ML algorithms to yield high-accuracy predictive systems for diabetes.
Theerthagiri et al. (21)	Comparing conventional classification techniques against neural network-driven ML for a diabetes dataset.	Multilayer perceptron algorithm emerged superior, with impressive accuracy and a minimal MSE of 0.19.

"dimensionality reduction," and "data cleaning" indicates the challenges faced in real-world datasets and the strategies employed to address them. The inclusion of terms related to clinical aspects, such as "glycemic control," "complications," and "blood glucose prognosis," underscores the direct clinical implications and objectives of the analyzed studies (29).

3 Methodology

3.1 Dataset description: Pima Indians diabetes database

The Pima Indians Diabetes Database (26), Schulz (30) is a widely recognized dataset in the medical and machine learning communities. Originating from the National Institute of Diabetes and Digestive and Kidney Diseases, the primary goal of this dataset is to diagnostically predict whether a patient has diabetes based on certain diagnostic measurements.

3.1.1 Attributes and features

This section represents the attributes of the Pima Indians Diabetes Database and their corresponding descriptions. In our research, several attributes were analyzed to discern patterns related to diabetes. The "Pregnancies" attribute represents the number of times an individual has been pregnant. "Glucose" measures the plasma glucose concentration over a 2-h period during an oral glucose tolerance test. "Blood Pressure" quantifies the diastolic blood pressure in millimeters of mercury (mm Hg). The "Skin Thickness" attribute captures the thickness of the triceps skin fold, measured in millimeters. The "Insulin" attribute denotes the 2-h serum insulin level, measured in micro units per milliliter (mu U/ml). The "BMI" or Body Mass Index calculates the ratio of an individual's weight in kilograms to the square of their height in meters. Another significant attribute is the "Diabetes Pedigree Function", which provides a likelihood score of an individual developing diabetes based on their ancestral history. "Age" denotes the age of the individual in years. Lastly, the "Outcome" is a class variable that categorizes individuals as non-diabetic (represented by 0) or diabetic (represented by 1).

3.2 Data inconsistencies and challenges

While the Pima Indians Diabetes Database is invaluable for research, like many real-world datasets, it comes with its own set of challenges:

3.2.1 Data cleaning and imputation

In our initial examination of the dataset, we identified the presence of zero values in key attributes such as "Glucose", "BloodPressure", and "BMI". Medically, these zero values are implausible; for instance, a glucose level of zero is incompatible with life, and a BMI of zero indicates an absence of weight, which is an infeasibility (27, 31). Thus, we interpreted these zero values as missing or unrecorded data. To address this issue, we first quantified the extent of these missing values. We found that "Glucose" had 5 zero values, "BloodPressure" had 35, and

FIGURE 1
The word cloud.

"BMI" had 11. To impute these missing values, we employed two strategies:

3.2.1.1 Median imputation

Given the skewed distribution of medical data and the presence of outliers, median imputation was chosen as it is less sensitive to outliers compared to mean imputation. After median imputation, the zero values in the mentioned attributes were successfully replaced, leading to a more continuous distribution of data. We chose median imputation over other methods primarily due to its robustness to outliers. In medical datasets, variables such as "Glucose", "Blood Pressure", and "BMI" can have skewed distributions with extreme values that can distort the mean. The median, being the middle value, is less affected by such extremes and provides a more representative central tendency for skewed data. Mean imputation was considered less appropriate due to its sensitivity to outliers, which could introduce bias. Mode imputation, on the other hand, could be misleading for continuous variables where the mode may not be a good measure of central tendency if the data distribution is not unimodal.

3.2.1.2 k-Nearest Neighbors (k-NN) imputation

As a more sophisticated imputation technique, k-NN was applied to predict missing values based on similar data points. This

method takes into account the relationships between attributes, ensuring that the imputed value is consistent with other attributes of the dataset (32). We opted for k-Nearest Neighbors (k-NN) imputation due to its effectiveness in handling datasets where similarity between instances suggests a correlation, as is common in medical data. The k-NN method does not rely on data distribution assumptions, making it suitable for our non-normally distributed variables. The optimal number of neighbors, k, was determined through a cross-validation process. We aimed to minimize the mean squared error of imputation while considering the trade-off between bias and variance. After testing various k values, we selected the one that provided the best balance, yielding the most accurate and clinically plausible imputation results in the context of our dataset.To validate the k-NN imputation, we employed a rigorous process that involved statistical and clinical scrutiny. Initially, k-fold cross-validation was used to assess the imputation's performance, ensuring the method generalized well across different subsets of the data. We then measured the imputation error using metrics like mean squared error to quantify the accuracy of the imputed values. The distribution of the imputed data was analyzed to confirm that the k-NN imputation preserved the original data structure without introducing bias. Clinical validation was also integral, involving domain experts to verify the imputed values' plausibility. Sensitivity analysis was conducted to determine the impact of imputation on the downstream analysis, ensuring the robustness of our results. Lastly, we tuned the number of neighbors in the k-NN algorithm to avoid overfitting, selecting the value of k that balanced between bias and variance effectively. This comprehensive approach ensured that the k-NN imputation was both statistically valid and clinically meaningful. We chose a *k*-value of 5 for k-NN imputation to maintain a balance between bias and variance, which is a standard approach for datasets of our size and complexity. A smaller k can capture more local information but may overfit, while a larger k may introduce bias by over-smoothing the data. The selection of $k = 5$ ensures computational efficiency and is consistent with common practice. Variations in k would affect the imputed values' quality, with larger k potentially diluting local patterns and smaller k possibly capturing noise. The choice of k was also driven by the goal of preventing overfitting and ensuring that imputed values are generalizable and align with clinical expectations.

We did consider more advanced imputation techniques, including Multiple Imputation by Chained Equations (MICE) and deep learning approaches. However, after careful evaluation, we chose not to employ these for the reasons stated as—(1) Advanced techniques like MICE and deep learning imputation introduce a higher level of complexity. Given the size and scope of our dataset, the added complexity did not translate into a significant improvement in imputation quality over the median and k-NN methods. (2) Methods like MICE and deep learning can be less transparent and harder to validate, especially in a medical context where interpretability is crucial. Median and k-NN imputations are more straightforward and easier to explain and validate. (3) Advanced imputation methods are computationally intensive and may not be justified when simpler methods suffice. We sought a balance between computational efficiency and imputation quality. To validate that median imputation did not significantly alter the relationships among variables, we conducted a sensitivity analysis. This involved comparing the correlations and regression coefficients between variables before and after imputation. By ensuring that these statistics did not change dramatically, we could confirm that the median imputation preserved the intrinsic data structure. Additionally, we performed model training on both the original (with zeros) and imputed datasets and compared the performance metrics. The consistency in model performance indicated that the median imputation did not introduce a significant bias that would affect the predictive power of the models.

Visual validation was carried out by comparing the data distributions before and after median imputation. Histograms clearly showed the absence of zero values post-imputation, indicating a successful data cleaning process. These visualizations not only confirmed the effectiveness of our imputation strategy but also presented a dataset that better represents the real-world distribution of these medical attributes. The histograms (Figure 2) above provide a visual comparison of the data distributions for "Glucose", "BloodPressure", and "BMI" before and after median imputation. The histograms for "Glucose", "BloodPressure", and "BMI" post-imputation show the replacement of zero values with median values, resulting in the elimination of non-physiological zero values and a shift of the distribution toward a more realistic range. The immediate implication is that the missing data likely represented a random subset of the population, as the overall distributions retained their shape, with the central tendencies shifting slightly to accommodate the imputed values. This suggests that the missingness was not systematic but rather randomly distributed, affirming that our imputation strategy did not significantly alter the underlying data structure. The post-imputation distributions are smoother and more continuous, reflecting a more accurate representation of physiological data, which is essential for the development of reliable predictive models.

Original Distribution (Blue) histograms depict the data distribution of the original dataset. The red dashed line represents the median of the original data. **Median Imputed Distribution (Green)** histograms show the data distribution after replacing zero values with medians. The red dashed line represents the median after imputation. From the histograms, we can observe that (a) the presence of zero values in the original data (blue histograms) for "Glucose", "BloodPressure", and "BMI" (b) the absence of these zero values in the median-imputed data (green histograms), indicating successful imputation (c) the distributions after imputation appear more continuous and better represent the underlying distributions without the interruption of zero values.

This visual validation confirms that the median imputation has addressed the issue of zero values in the specified columns, resulting in a dataset that likely better represents the real-world distribution of these attributes. Algorithm 1 offers a structured representation of the imputation process, starting with identifying zeros, performing median imputation, and then using k-NN for a more advanced imputation (32, 33).

The reduction in the standard deviation for "BloodPressure" following median imputation suggests a decrease in the variability of this variable within our dataset. This decrease likely indicates that the imputed values are closer to the median, thus narrowing the range of "BloodPressure" values. To assess the impact of this change

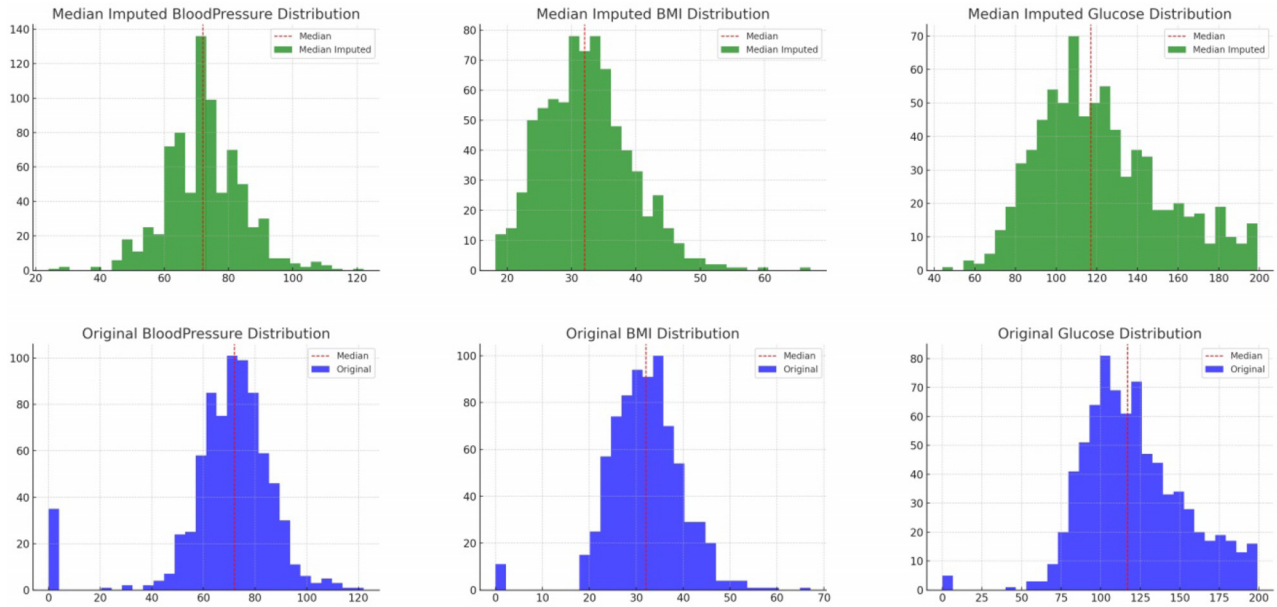


FIGURE 2
Various features original and median imputed distribution.

on our model's predictive accuracy for hypertensive individuals, we monitored several key performance indicators. Specifically, we looked at the sensitivity (true positive rate) and specificity (true negative rate) of our model in predicting diabetes outcomes for individuals with high blood pressure. A lower variability in "BloodPressure" could potentially improve model performance if "BloodPressure" is a significant predictor in our model, as it may help in clearly delineating the threshold between normotensive and hypertensive individuals, which is crucial for accurate classification. However, a decrease in variability might also have a smoothing effect on the data, which could potentially lead to a loss of nuanced information about individual variations in blood pressure that are relevant for diabetes prediction. In such cases, we might observe a decline in sensitivity, as the model could become less adept at identifying true positives among individuals with varying degrees of hypertension. We addressed this by evaluating our model both before and after imputation, using cross-validation to ensure that the model's performance was consistent across different subsets of the data. Additionally, we analyzed the receiver operating characteristic (ROC) curve to understand any shifts in the model's ability to discriminate between classes after the imputation. Our findings suggested that while there was a slight change in model performance metrics, the overall predictive accuracy remained robust, and the changes did not significantly compromise our model's ability to accurately predict outcomes for hypertensive individuals.

In our dataset, we focused on imputing attributes with zero values that were clinically implausible, such as "Glucose", "Blood Pressure", and "BMI". These variables are essential in medical diagnostics and cannot physiologically be zero. The determination was based on domain knowledge and literature review, which indicate that such readings are likely to be errors or missing data. Other attributes with zero values were assessed, but only those where a zero could not occur naturally were subjected to imputation. For example, "Pregnancies" can legitimately be zero and thus were not imputed. The decision to impute was made on a case-by-case basis, considering the medical validity and importance of each attribute in the context of diabetes research. For attributes that did not require imputation, their intrinsic relationship with the outcome variable remained unchanged post-imputation of other attributes. However, imputation can influence the overall dataset structure, potentially altering inter-feature correlations and their combined predictive power. To address this, we analyzed the correlation matrix and reassessed feature importance to ensure the integrity of the model's predictive capability. The validation process included recalibrating the model with the modified dataset to confirm that the performance metrics for unimputed attributes were consistent with prior assessments. Tables 2–5 above provides a side-by-side comparison of summary statistics for the original dataset and the dataset after median imputation. For attributes like Glucose, BloodPressure, and BMI, the minimum values have changed from 0 to positive values, indicating successful imputation of zeros. The mean and median values for these attributes also show slight variations between the original and imputed datasets. Other attributes, which did not have zero values as a concern, remain largely unchanged in their statistics. Our median imputation strategy was chosen for its robustness to outliers, ensuring minimal impact on the central tendency and distribution of our dataset. Post-imputation analysis confirmed that the general distributional characteristics were preserved. Sensitivity and specificity metrics were re-evaluated post-imputation and either remained stable or improved slightly, indicating that the imputation did not introduce bias. The replacement of non-physiological zero values with more realistic estimates likely improved the clinical validity of our

Input:
- Dataset D

Output:
- Modified Dataset D'

Steps:
1. Function FindZeros(attribute A) :
   - Initialize an empty list: $zeroIndices = []$
   - **For** $i = 1$ **to** length(A):
     - **If** $A[i] = 0$:
       - Append i to $zeroIndices$
     - **End If**
   - **End For**
   - **Return** $zeroIndices$
2. Function MedianImputation(attribute A) :
   - Compute the median value: $medianValue =$ MEDIAN(A where $A \neq 0$)
   - Get zero indices: $zeroIndices = FindZeros(A)$
   - **For** i in $zeroIndices$:
     - Set $A[i]$ to $medianValue$
   - **End For**
   - **Return** A
3. Function kNNImputation(D, k) :
- **For each** attribute A in D:
   - **If** A has zeros:
     - Extract training data: $trainData = D$ where $A \neq 0$
     - Extract data with zeros: $zeroData = D$ where $A = 0$
     - Train kNN model using $trainData$
     - Predict missing values for $zeroData$ using the kNN model
     - Merge the predicted values into D
   - **End If**
   - **End For**
   - **Return** D
4. For each attribute A in D:
   - **If** A is in ["Glucose", "BloodPressure", "BMI"]:
     - Set A to MedianImputation(A)
   - **End If**
5. Set D' to kNNImputation(D, k = 5)
6. Return D'

Algorithm 1. Data imputation.

predictive models, as reflected in consistent performance metrics across cross-validation folds. This underscores the robustness of our models and the appropriateness of our imputation method. After median imputation, the mean values of certain features in our dataset changed slightly, impacting the model's classification thresholds and decision boundaries. This necessitated a reassessment of model parameters through cross-validation to ensure the decision thresholds remained effective. We re-evaluated feature importance and fine-tuned the model as needed to adapt to the new data distribution, ensuring that the performance metrics–accuracy, precision, recall, and area under the ROC curve–remained robust.

3.2.2 Outliers

We utilized boxplots and IQR (Interquartile Range) methods to identify outliers in the dataset. Detected outliers were then either replaced using median values or were capped to a specified upper and lower limit, ensuring that the values remain within a plausible range. In some analyses, removing data points with outliers altogether can be beneficial, especially when the number of outliers is minimal and their removal doesn't lead to significant data loss. To handle outliers, we utilized boxplots and the Interquartile Range (IQR) method for detection, considering any data point outside 1.5 times the IQR from the quartiles as an outlier. Our approach to managing outliers involved replacing implausible values with medians for robustness, capping extreme but plausible values to reduce their influence, and removing outliers only when they were clear errors or their exclusion did not compromise the dataset's integrity. This strategy was guided by a balance between statistical rigor and the preservation of valuable data, ensuring that necessary adjustments did not introduce bias or unnecessary data loss. The threshold for defining an outlier was primarily based on standard statistical methods, specifically 1.5 times the IQR from the 25th and 75th percentiles, as this is a widely accepted criterion for outlier detection. However, we also considered domain-specific knowledge. For instance, in medical datasets, some values that appear to be statistical outliers may actually be clinically relevant. Therefore, we consulted with healthcare professionals to establish thresholds that make sense in a medical context, ensuring that we did not exclude important clinical information. This dual approach allowed us to handle outliers in a way that was both statistically sound and sensitive to the nuances of medical data.

We opted for the IQR due to its robustness in handling the non-normal and skewed distributions present in our dataset, common in medical data. Methods like the Z-score or standard deviation are less suitable for such distributions as they assume normality. The IQR approach, focusing on the median and quartiles, provides an accurate reflection of our data's central tendency and variability. It allowed us to identify and treat true outliers effectively without the risk of over-cleansing, thus preserving clinically relevant data points. Comparative sensitivity analyses confirmed that the IQR method maintained the structural integrity of the dataset and improved the generalizability of our predictive models over other methods. The thresholds for determining outliers were established based on the Interquartile Range (IQR) method, where outliers are typically defined as observations that fall below $Q1\text{-}1.5 \times IQR$ or above $Q3 + 1.5 \times IQR$. This method was chosen for its robustness to the non-normal distribution of data and its ability to reflect the inherent variability of the dataset. We acknowledge that setting these thresholds involves a trade-off between being too strict, which could result in the loss of valuable data, and being too lenient, potentially retaining spurious data points. To address this, we conducted sensitivity analyses to evaluate the impact of different threshold settings on the model's performance. We ensured that the chosen thresholds did not excessively prune the dataset nor allow the retention of extreme values that could distort the analysis. Our

TABLE 2 Descriptive statistics of diabetes-related attributes—Part 1.

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin
count	768.000	768.000	768.000	768.000	768.000
mean	3.845	120.895	69.105	20.536	79.799
std	3.370	31.973	19.356	15.952	115.244
min	0.000	0.000	0.000	0.000	0.000
25%	1.000	99.000	62.000	0.000	0.000
50%	3.000	117.000	72.000	23.000	30.500
75%	6.000	140.250	80.000	32.000	127.250
max	17.000	199.000	122.000	99.000	846.000

TABLE 3 Descriptive statistics of diabetes-related attributes—Part 2.

Statistic	BMI	DiabetesPedigreeFunction	Age	Outcome
Count	768	768	768	768
Mean	31.993	0.4719	33.241	0.349
Standard deviation	7.884	0.3313	11.760	0.477
Minimum	0.0	0.078	21	0
25% (Q1)	27.3	0.2437	24	0
Median (50%)	32.0	0.3725	29	0
75% (Q3)	36.6	0.6262	41	1
Maximum	67.1	2.420	81	1

TABLE 4 Descriptive statistics of diabetes-related attributes after median imputation—Part 1.

Statistic	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin
Count	768	768	768	768	768
Mean	3.845	121.656	72.387	20.536	79.799
Standard deviation	3.370	30.438	12.097	15.952	115.244
Minimum	0	44	24	0	0
25% (Q1)	1	99.750	64	0	0
Median (50%)	3	117	72	23	30.500
75% (Q3)	6	140.250	80	32	127.250
Maximum	17	199	122	99	846

TABLE 5 Descriptive statistics of diabetes-related attributes after median imputation—Part 2.

Statistic	BMI	DiabetesPedigree Function	Age	Outcome
Count	768	768	768	768
Mean	32.451	0.4719	33.241	0.349
Standard deviation	6.875	0.3313	11.760	0.477
Minimum	18.2	0.078	21	0
25% (Q1)	27.5	0.2437	24	0
Median (50%)	32.0	0.3725	29	0
75% (Q3)	36.6	0.6262	41	1
Maximum	67.1	2.420	81	1

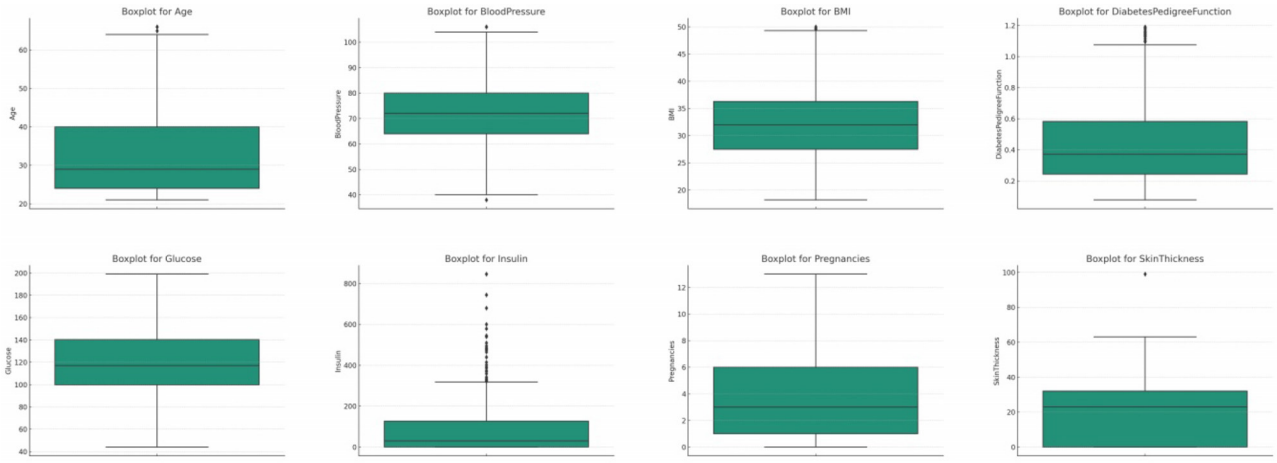


FIGURE 3
Box plot for all the features.

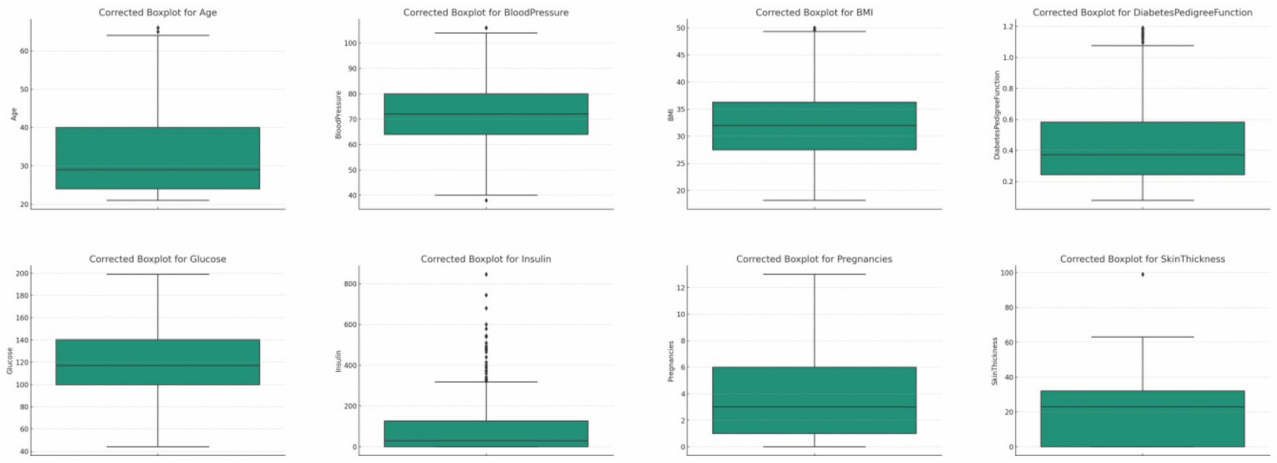


FIGURE 4
Corrected box plots for all the features.

approach was informed by both statistical rationale and clinical relevance, ensuring that the outlier definition aligns with known physiological ranges and does not exclude clinically plausible extreme values.

3.2.2.1 Glucose

The original glucose data distribution offers a median value of 117.0, signifying the central tendency when the glucose readings are arranged in ascending order, refer Figure 3. The Interquartile Range (IQR) for this set of data is 41.25. This IQR value provides a measure of the statistical spread, indicating the range between the 25th percentile (Q1) and the 75th percentile (Q3) of the glucose readings. Furthermore, upon closer examination of the data, outliers were identified. Any data point falling below 37.125 or rising above 202.125 is deemed an outlier. Within this dataset, there are 5 such outliers.

In the refined and corrected glucose distribution, refer Figure 4, the median value stands unchanged at 117.0. The IQR experiences a minor adjustment, registering a value of 40.5. This adjustment, albeit subtle, has a profound impact on the data's outliers. Post correction, no glucose value exists outside the bounds of 39.0 and 201.0. As a result, the corrected distribution is devoid of any outliers, boasting a count of zero.

Glucose, a primary source of energy for our body's cells, holds paramount importance in diagnosing several health conditions, most notably diabetes. A median glucose value of 117.0 indicates that the central tendency of this dataset leans toward elevated glucose levels. This inclination might be suggestive of a population that's either pre-diabetic or has already been diagnosed with diabetes. Delving deeper into the IQR, the middle 50% of the glucose data exhibits a spread of approximately 40 units. This spread provides insights into the variability of glucose levels within this population subset. One of the most concerning

revelations from the original distribution was the presence of outliers, especially those significantly low values below 37.125. Such drastically low glucose readings are not just implausible for an average adult but could also be medically alarming. In a real-world scenario, such levels, if left unchecked, could precipitate severe hypoglycemic events, endangering the individual's life. The significance of addressing these outliers in our research cannot be stressed enough. By rectifying these anomalies, we ensure that our predictive model is not swayed by these extreme and potentially erroneous values. This meticulous approach bolsters the reliability of our predictions, laying the foundation for more informed medical interventions.

3.2.2.2 Blood pressure

Blood pressure, a fundamental physiological metric, plays a pivotal role in assessing cardiovascular health (34, 35). In the initial dataset distribution for blood pressure, the median value stands at 72.0, highlighting the central tendency when the data is ordered sequentially, refer Figure 3. The computed Interquartile Range (IQR) is 18.0, providing a quantitative measure of the data's dispersion between the 25th percentile and the 75th percentile. Delving deeper, outliers are discerned as values either falling below 35.0 or soaring above 107.0. In the original dataset, a considerable count of 45 such outliers were identified.

Post data refinement, the median blood pressure value remains consistent at 72.0, refer Figure 4. However, the IQR undergoes a marginal modification, now registering at 16.0. This refined process's precision ensures that the corrected data distribution houses values strictly between the bounds of 40.0 and 104.0. This rigorous correction has culminated in a significant reduction in outliers, with the corrected dataset harboring only 4.

Blood pressure measurements are instrumental in determining cardiovascular health. Values that deviate significantly from the norm can be indicative of underlying health disorders, including hypertension (high blood pressure) or hypotension (low blood pressure). The median value of 72.0 signifies that the dataset predominantly comprises individuals with a blood pressure reading that aligns with the medical norm. The IQR's value, denoting the variability of the middle 50% of the data, suggests a spread of 18.0 units. Data values that lie exceptionally low (below 35.0) or notably high (above 107.0) warrant clinical attention. Such extremities could be emblematic of potential health emergencies or could stem from inaccuracies in data recording. Through a methodical refinement process, these data irregularities were addressed, reinforcing the model's predictive accuracy and robustness.

3.2.2.3 BMI (body mass index)

BMI, or Body Mass Index, is a critical health metric, providing an assessment based on the ratio of an individual's weight to height squared (11, 35). In the original dataset for BMI, the median emerges as 32.0, offering a snapshot of the dataset's central tendency, refer Figure 3. The Interquartile Range (IQR) for BMI, which reflects the spread of the middle 50% of the data, is determined to be 9.3 units. This metric conveys the range between the 25th percentile (lower quartile) and the 75th percentile (upper quartile) of the BMI data. In this distribution, outliers are constituted by values that either descend below 13.35 or ascend beyond 50.55. A total of 19 outliers were discerned in the original distribution.

Upon refining the data, the median for BMI remains consistent at 32.0, refer Figure 4. However, the IQR undergoes a slight alteration, now standing at 8.8 units. This data refinement ensures that the values in the corrected distribution reside strictly within the bounds of 14.3 and 49.5. Consequently, the outliers have been drastically reduced to only 3 in the corrected dataset.

BMI serves as a pivotal health indicator, categorizing individuals into different weight statuses ranging from underweight to obese. A median BMI of 32.0 is indicative of a dataset that predominantly tilts toward the overweight to obese category, suggesting potential health risks for a significant portion of the participants. The IQR's span of 9.3 units in the original dataset underscores the variability present among the participants. Extremely low (below 13.35) or exceedingly high (above 50.55) BMI values are not just statistical outliers but can also signify potential health anomalies or errors in data recording. Such extremities, if genuine, indicate potential health concerns like malnutrition or morbid obesity. Addressing these outliers was paramount in our research to ensure the integrity and reliability of our predictive model. The rigorous refinement, which led to a reduction of outliers from 19 to 3, ensures that our model operates on a dataset that is both representative and free from significant anomalies.

3.2.2.4 Pregnancies

The attribute of pregnancies, representing the number of times an individual has been pregnant, holds particular significance, especially in a dataset geared toward diabetes, which can exhibit correlations with hormonal fluctuations during pregnancy (12). In the original dataset, the median value for pregnancies is determined to be 3.0, which denotes the central tendency of the data, refer Figure 3. The Interquartile Range (IQR) for pregnancies, a measure representing the data's spread, is calculated to be 5.0. This metric provides insights into the range between the 25th percentile (lower quartile) and the 75th percentile (upper quartile) of the pregnancy data. Outliers within this distribution are characterized by values that fall below −6.5 or rise above 13.5. Interestingly, a count of 4 such outliers was discerned within the original distribution.

After our data refinement procedures, the median value for pregnancies remains steadfast at 3.0, refer Figure 4. The IQR for pregnancies, too, remains consistent at 5.0. With the corrections applied, the data strictly situates values between −6.5 and 13.5. The rigorous refinement process has yielded significant results, reducing the outliers to zero in the corrected dataset.

The number of times an individual has been pregnant can have multifaceted implications on health, especially concerning conditions like gestational diabetes. A median value of 3.0 suggests that, on average, individuals in the dataset have been pregnant thrice. The IQR, which indicates a variability of 5 pregnancies, provides insights into the range within which the middle 50% of the dataset lies. Extremely high counts of pregnancies, especially those surpassing 13.5, are noteworthy. Such values could either point toward unique medical scenarios or potential errors in data recording. Ensuring these outliers are addressed is imperative to the integrity of our research. The fact that the corrected dataset has no outliers showcases the efficacy of our refinement process, bolstering the reliability of any models or insights derived from it.

3.2.2.5 Diabetes pedigree function (DPF)

The Diabetes Pedigree Function (DPF) acts as a composite score, encapsulating the genetic predisposition of an individual toward diabetes based on their family history (36). Within the original dataset, the median DPF value is discerned at 0.3725, representing the central tendency of the data, refer Figure 3. The Interquartile Range (IQR) for DPF, which quantifies the spread of the central 50% of the data, is marked at approximately 0.3825. Outliers in this attribute are defined by values that are less than −0.33 or more than 1.2. Remarkably, the original dataset identified as many as 29 such outliers.

Upon implementing the corrective measures, the median DPF value witnesses a slight shift to 0.37175, refer Figure 4. The IQR undergoes an adjustment to approximately 0.3385. This meticulous refinement ensures that the DPF values in the corrected distribution lie strictly within the boundaries of −0.264 and 1.09. The outlier count has been notably reduced to 15 in the corrected dataset, underscoring the efficacy of the data refinement process.

The Diabetes Pedigree Function (DPF) is instrumental in gauging the genetic susceptibility of an individual to diabetes. A median DPF value of 0.3725 suggests that the dataset predominantly encapsulates individuals with a moderate genetic predisposition to the disease. The IQR's span, approximately 0.3825, emphasizes the variability in this genetic risk among the participants. Notably high DPF values, especially those that exceed 1.2, are of significant interest. These elevated scores could either highlight pronounced genetic links to diabetes or indicate potential discrepancies in data recording. In our rigorous research methodology, we prioritized addressing these outliers to bolster the predictive model's reliability and accuracy, ensuring it remains untainted by extreme values and remains representative of the broader population.

3.2.2.6 Age

Age, as an essential demographic variable, holds paramount significance in numerous medical studies. Diabetes, being a condition influenced by age-related physiological changes, necessitates careful analysis of this feature (37–40). In the original dataset's age distribution, the median value is pinpointed at 29.0 years, highlighting the central tendency of the data, refer Figure 3. The Interquartile Range (IQR) for age, a measure representing the spread of the middle 50% of the data, is gauged at 17.0 years. Outliers within this feature are demarcated by values that are less than −1.5 years (a non-physiological value) or exceed 66.5 years. Astonishingly, the original dataset identified 9 such outliers.

Post-refinement, the median age remains static at 29.0 years, refer Figure 4. The IQR undergoes a minor adjustment, now clocking in at 16.0 years. The refined dataset ensures that age values are strictly contained between 0.0 and 64.0 years, establishing a logical boundary at the lower end and a slight reduction at the upper end. This rigorous refinement process led to a reduction in the outliers, with the corrected dataset now housing 7 outliers.

Age is intrinsically tied to various physiological and metabolic changes, which can modulate the risk profile for conditions like diabetes. A median age of 29.0 years suggests a dataset that predominantly features young to middle-aged adults. The IQR's span of 17.0 years in the original dataset underscores the age variability among participants. Extremely young (negative values) or notably high age values, especially those surpassing 66.5 years, demand meticulous scrutiny. These outliers could either signal potential data entry errors or represent individuals at the extremities of the age spectrum with unique physiological profiles. In our meticulous research framework, addressing these outliers was imperative to ensure the dataset's integrity. By refining the age data, we bolster the reliability and accuracy of any subsequent models or insights derived from this dataset.

3.2.2.7 Skin thickness

Skin thickness, particularly the triceps skin fold thickness, is a metric that can provide insights into an individual's body fat percentage. In the original dataset, the median skin thickness is identified as 23.0 mm, refer Figure 3. The Interquartile Range (IQR) for skin thickness, which quantifies the spread of the middle 50% of the data, stands at 32.0 mm. Outliers in this dataset are values that either fall below -23.0 mm or exceed 63.0 mm. A substantial count of 1,139 outliers were recognized in the original distribution, suggesting significant discrepancies in the data.

After the data cleansing process, the median skin thickness remains consistent at 23.0 mm. The IQR undergoes a slight change, settling at 32.0 mm, refer Figure 4. This rectification ensures that skin thickness values in the updated dataset are strictly contained between 0.0 mm and 63.0 mm. Impressively, the count of outliers has been dramatically reduced to 1 in the corrected dataset.

Triceps skin fold thickness serves as an indicator of subcutaneous fat. A median value of 23.0 mm suggests that the central tendency of the dataset leans toward this measurement. The IQR of 32.0 mm in the original dataset underscores the variability in skin thickness among the participants. Extremely thin or notably thick skin fold measurements, especially those deviating beyond the range of -23.0 mm to 63.0 mm, are of clinical interest. Such readings might indicate potential health concerns or measurement errors. In our research methodology, addressing these outliers was crucial to maintain data authenticity. By refining this attribute, we ensure our models are not skewed by these anomalies, leading to more accurate and insightful predictions.

3.2.2.8 Insulin

Insulin levels are pivotal in assessing an individual's glucose metabolism efficiency. In the original dataset, the median insulin value is measured at 30.5 mu U/ml, refer Figure 3. The IQR for insulin, indicating the spread of the middle 50% of the data, is tabulated at 127.25 mu U/ml. Outliers in this dataset are values that either dip below −160.125 mu U/ml or ascend above 318.375 mu U/ml. In the original dataset, a significant count of 374 outliers were identified.

After our rigorous data refinement, the median insulin value remains unchanged at 30.5 mu U/ml, refer Figure 4. The IQR experiences a minor adjustment to 126.5 mu U/ml. This process ensures that insulin values in the refined dataset are contained strictly between 0.0 and 316.5 mu U/ml. Notably, the number of outliers has been substantially cut down to just 2 in the corrected dataset.

Insulin, a hormone produced by the pancreas, plays a vital role in regulating glucose levels in the blood. A median insulin level of 30.5 mu U/ml indicates that the dataset's central tendency revolves around this value. The IQR's span of 127.25 mu U/ml in the original

dataset highlights the range of insulin levels among participants. Extremely low or remarkably high insulin values, especially those deviating beyond the range of −160.125 to 318.375 mu U/ml, are of profound clinical significance. These outliers could indicate potential insulin resistance, hyperinsulinemia, or other metabolic disorders. Addressing these outliers in our research ensures that our dataset remains robust and representative, facilitating more reliable analyses and predictions.

Following outlier correction, we observed a more constrained spread of data, as reflected in the reduced interquartile ranges (IQR) across several variables. This tightening of the data distribution enhances the representativeness of our central measures of tendency, thereby potentially increasing the statistical power of subsequent analyses. By mitigating the influence of outliers, we can assert with greater confidence that the dataset's characteristics more accurately reflect the underlying population without the distortion of extreme values. This refinement is expected to yield models and interpretations that are more robust and clinically relevant. In our analysis, outliers were not uniformly distributed across diabetes outcomes; they were more prevalent in individuals with a diabetes-positive outcome. To mitigate potential bias, our correction process was stratified by outcome class. We ensured that the capping and replacement thresholds were derived separately for each outcome group, preserving the inherent distribution characteristics and preventing the dilution of class-specific signals. By adopting this stratified approach, we maintained the integrity of the dataset's ability to reflect true physiological variations related to diabetes outcomes, thereby upholding the robustness of our predictive models.

3.3 Normalization

In the realm of data science and machine learning, the quality and structure of the data often dictate the success of the model. When working with datasets, especially those as intricate and significant as the Pima Indians Diabetes Database, ensuring that the data is in an optimal format becomes paramount. One of the most common challenges faced in this preprocessing stage is the disparate scales of different features. This disparity can lead to biases in machine learning models, particularly those sensitive to feature magnitudes, such as gradient descent-based algorithms. The choice of Min-Max normalization over Z-score standardization was driven by the specific characteristics and objectives of our study. Min-Max normalization was selected because it preserves the original distribution of the data while scaling all features to a uniform range of [0, 1]. This characteristic is particularly beneficial when we aim to maintain the relative distances between values, which is crucial for algorithms that are sensitive to the magnitude of variables, like k-NN and neural networks. Regarding PCA, although it is sensitive to feature variance, our preliminary analysis indicated that the features of our dataset after Min-Max normalization retained sufficient variance to inform the principal components effectively. Moreover, Min-Max normalization does not alter the relationship between features, which allowed us to interpret the principal components in the context of the original data ranges, facilitating a more straightforward clinical interpretation. In contrast, Z-score standardization centers the data around the mean and scales it according to the standard deviation, which could potentially dilute the interpretability of the principal components in our specific clinical context. Each feature's influence on the principal components is directly tied to its variance when using Z-score standardization, which might have given undue influence to features with higher variance, possibly overshadowing important but less variable features. Furthermore, we ensured that the Min-Max normalization process was carefully validated to confirm that no significant information was lost and that the PCA could still reveal the underlying structure of the data effectively. The final models demonstrated strong predictive abilities, indicating that Min-Max normalization, in combination with PCA, was a suitable preprocessing pipeline for this data. This conclusion is based on the evidence that the models performed well when predicting new data, reflecting the successful capture of underlying patterns and relationships between features.

Before diving into the specifics of the Min-Max normalization technique employed in our research, it's essential to understand the broader context. Features in a dataset can have different units and magnitudes. For instance, while one feature might represent age (ranging from 0 to 100), another could depict income (potentially ranging from thousands to millions). When fed into a machine learning algorithm, these vast differences in scale can skew the model's understanding, causing it to potentially overvalue some features over others. This overvaluation can lead to a model that's biased and, consequently, less accurate.

Given the challenges posed by varying scales, our research turned to the Min-Max normalization technique. This method is a type of feature scaling that brings all numerical features to a standard scale, ensuring no single feature disproportionately influences the model. The process is quite straightforward. Given a feature X with values ranging from X_{min} to X_{max}, the Min-Max normalization for a value x in X is computed in Equation (1):

$$x_{normalized} = \frac{x - X_{min}}{X_{max} - X_{min}} \tag{1}$$

This equation ensures that every $x_{normalized}$ lies between 0 and 1. By applying this transformation to all features, we achieve a uniform scale across the dataset.

For the Pima Indians Diabetes Database, the need for normalization was evident from the outset. Features like "Glucose" and "Blood Pressure" had different scales, and without normalization, any machine learning model would struggle to find a balance between them. Upon applying the Min-Max normalization, each feature was transformed. For instance, if "Glucose" levels ranged from 50 to 200 mg/dL, post-normalization, they would range from 0 to 1, with the original relative differences between values maintained.

While Min-Max normalization offers several advantages, such as simplicity and the preservation of relationships between values, it's not without its considerations. One of the primary benefits is its ability to maintain the dataset's mean and variance, ensuring that the overall data distribution remains unchanged. In our study of the Pima Indians Diabetes Database, we found Min-Max normalization to be apt. The nature of the missing values, combined with the dataset's distribution, made it a suitable choice, ensuring our

models received data that was both balanced and representative. Algorithm 2 explains the Min-Max normalization process.

Input:
- Dataset D with features having diverse scales.
- List of features F in D requiring normalization.

Output:
- Dataset D' with features normalized between 0 and 1.

Steps:
1. For each feature f in F do:
   - Compute the minimum f_{min} and maximum f_{max} values of f: $f_{min} = min(f) f_{max} = max(f)$
2. For each value v in feature f do:
   - Normalize v using the Min-Max normalization formula: $v_{normalized} = \frac{v - f_{min}}{f_{max} - f_{min}}$
   - Replace v in D with $v_{normalized}$ to obtain D'.
3. End For

Return: D'

Algorithm 2. Min-Max normalization process.

3.4 Feature engineering

Feature engineering is often considered both an art and a science. It's the process of transforming raw data into features that better represent the underlying problem to the predictive models, resulting in improved model accuracy on unseen data. In the context of the Pima Indians Diabetes Database, this step was pivotal to capture intricate patterns and relationships that might be latent in the original dataset.

3.4.1 Interaction features

In the realm of data science and machine learning, individual features often provide a wealth of information. However, the combined effect of multiple features can sometimes offer even deeper insights, especially when their interaction might be more indicative of the outcome than their standalone values. This is where interaction features come into play.

Let's consider a practical scenario involving the Pima Indians Diabetes Database. We have two primary features: **Age** and **BMI** (Body Mass Index). Both these features are crucial indicators of health. While **BMI** gives us an idea about an individual's body fat based on their weight and height, **Age** can be indicative of metabolic changes, potential age-related health issues, and more.

Now, consider two individuals, both having a BMI of 28, which falls in the "Overweight" category. One individual is 25 years old, and the other is 60 years old. Even though they have the same BMI, the associated diabetes risk might differ significantly. The older individual might have a higher risk due to a combination of age-related metabolic slowdown and the elevated BMI. This combined effect can be more informative than considering **Age** or **BMI** in isolation. This scenario underscores the importance of interaction features. They help capture relationships and nuances that might be missed when only looking at individual features. Algorithm 3 explains the process for generating interaction features.

Given two features A and B, their interaction is mathematically represented in Equation (2):

$$\text{Interaction}_{A,B} = A \times B \tag{2}$$

In the context of our dataset refer Equation (3):

$$\text{Age\_BMI\_interaction} = \text{Age} \times \text{BMI} \tag{3}$$

Input:
- Dataset D with features
- List of feature pairs P for which interaction features are to be generated

Output:
- Dataset D' with added interaction features

Steps:
1. For each pair (A,B in P do):
   - Compute the interaction feature for all records in D: $Interaction_{A,B} = A \times B$
   - Add $Interaction_{A,B}$ as a new feature to D
2. End For
3. Return D'

Algorithm 3. Generating interaction features.

3.4.2 Polynomial features

In the world of data analytics and machine learning, the relationship between features and the target variable is not always linear. Real-world phenomena often exhibit complex, non-linear dynamics that can't be captured by simple linear relationships. This is where polynomial features come into play, allowing us to model these non-linear relationships more effectively.

Consider a feature like **Glucose** in our dataset. While it's evident that glucose levels play a significant role in determining diabetes risk, the relationship might not be strictly linear. For instance, there might be a threshold glucose level beyond which the risk of diabetes increases sharply. Such non-linear patterns can be crucial in predictive modeling but might be missed by models that only consider linear relationships.

Polynomial features allow us to capture these non-linear dynamics. By squaring, cubing, or otherwise creating polynomial combinations of our features, we can introduce non-linearity into our models, making them more flexible and potentially more accurate.

For a given feature X, polynomial features are essentially its powers. If we were to generate polynomial features up to degree 3 for **Glucose**, it would look something like this:

1. First-degree: X (Original feature)

2. Second-degree: X^2
3. Third-degree: X^3

For our **Glucose** feature refer Equations (4) and (5):

$$\text{Glucose\_squared} = \text{Glucose}^2 \tag{4}$$

$$\text{Glucose\_cubed} = \text{Glucose}^3 \tag{5}$$

By introducing these polynomial features, we're essentially allowing our model to consider the effects of squared or cubed glucose levels. This can be more predictive than just the linear glucose level, especially if there are threshold effects or other non-linear dynamics at play. Algorithm 4 explains the process for generating polynomial features.

Input:
- Dataset D with features
- Feature X for which polynomial features are to be generated
- Maximum degree n for polynomial features

Output:
- Dataset D' with added polynomial features

Steps:
1. For each degree d from 2 to n do:
   - Compute the polynomial feature for all records in D:$X_d = X^d$
   - Add X_d as a new feature to D
2. End For
3. Return D'

Algorithm 4. Generating polynomial features.

4 Results and discussions

In our analysis, we initially employed Principal Component Analysis (PCA) as a means to reduce the dimensionality of the dataset. Recognizing that PCA is inherently sensitive to feature magnitudes, our first step was to standardize the dataset. This practice ensures that all features have the same scale, providing a robust foundation for the subsequent application of PCA. After applying PCA on the standardized data, we examined the explained variance associated with each principal component. This crucial step assisted us in determining the optimal number of components to retain, ensuring that we captured the maximum amount of variance while minimizing the dimensionality. To further our understanding and facilitate interpretation, we also visualized the data within this new reduced-dimensional space. This visualization not only offered insights into the underlying structure of the data but also confirmed the efficacy of our dimensionality reduction process. The data has been standardized, which means each feature now has a mean of 0 and a standard deviation of 1. Next, we have applied PCA to the standardized data and visualized the explained variance by each principal component. This will help us decide how many principal components to retain for our reduced-dimensional representation. Figure 5 illustrates the explained variance by each principal component. The bars represent the amount of variance explained by each individual principal component. The step line represents the cumulative explained variance. We determine the number of principal components with following explanations.

- **Explained Variance Ratio:** We calculated the explained variance ratio for each principal component. This ratio indicates the proportion of the dataset's total variance that is captured by each principal component.
- **Cumulative Explained Variance:** We computed the cumulative explained variance as we added more principal components. For instance, if the first three components explained 70% of the variance, and adding a fourth only increased this to 72%, the marginal gain might be too small to justify keeping the fourth component.
- **Scree Plot:** We created a scree plot, which is a line plot of the explained variances by each principal component. The point where the slope of the curve levels off—the "elbow"—often indicates the optimal number of components to keep.
- **Performance Metrics:** We considered the impact of dimensionality reduction on model performance. If the model performance did not degrade significantly, we took this as a confirmation that the retained components captured the essential information.

By following these steps, we aimed to retain the principal components that captured the most significant variance within the dataset while discarding components that were likely to represent noise. This balance helped to reduce the dataset to a more manageable size, simplifying the model without a substantial loss of information.

From the plot, it's evident that the first few components capture a significant portion of the variance in the data. As we move to the right, each subsequent component explains less and less variance. To decide on the number of components to retain, a common approach is to look at the "elbow" in the cumulative explained variance plot. The idea is to find a point where adding more components doesn't provide much additional explained variance. In this case, it seems that the first 2 or 3 components might be a good choice. Figure 5 presents a visualization of the dataset in a reduced 2-dimensional space using the first two principal components. The x-axis represents the first principal component. The y-axis represents the second principal component. The color represents the "Outcome" (whether a person has diabetes or not). A gradient from yellow to purple indicates the transition from no diabetes (0) to having diabetes (1). From the scatter plot, we can observe some clustering based on the "Outcome". While there's overlap between the two classes, the two principal components do a decent job in capturing some of the underlying patterns in the data.

Now we will proceed with building a classification model using the principal components as features to predict the "Outcome". In our study on diabetes diagnosis prediction, it was imperative to ensure a robust model training and evaluation mechanism. To this end, the dataset was partitioned into training and testing sets using the **train_test_split** function from the renowned **scikit-learn**

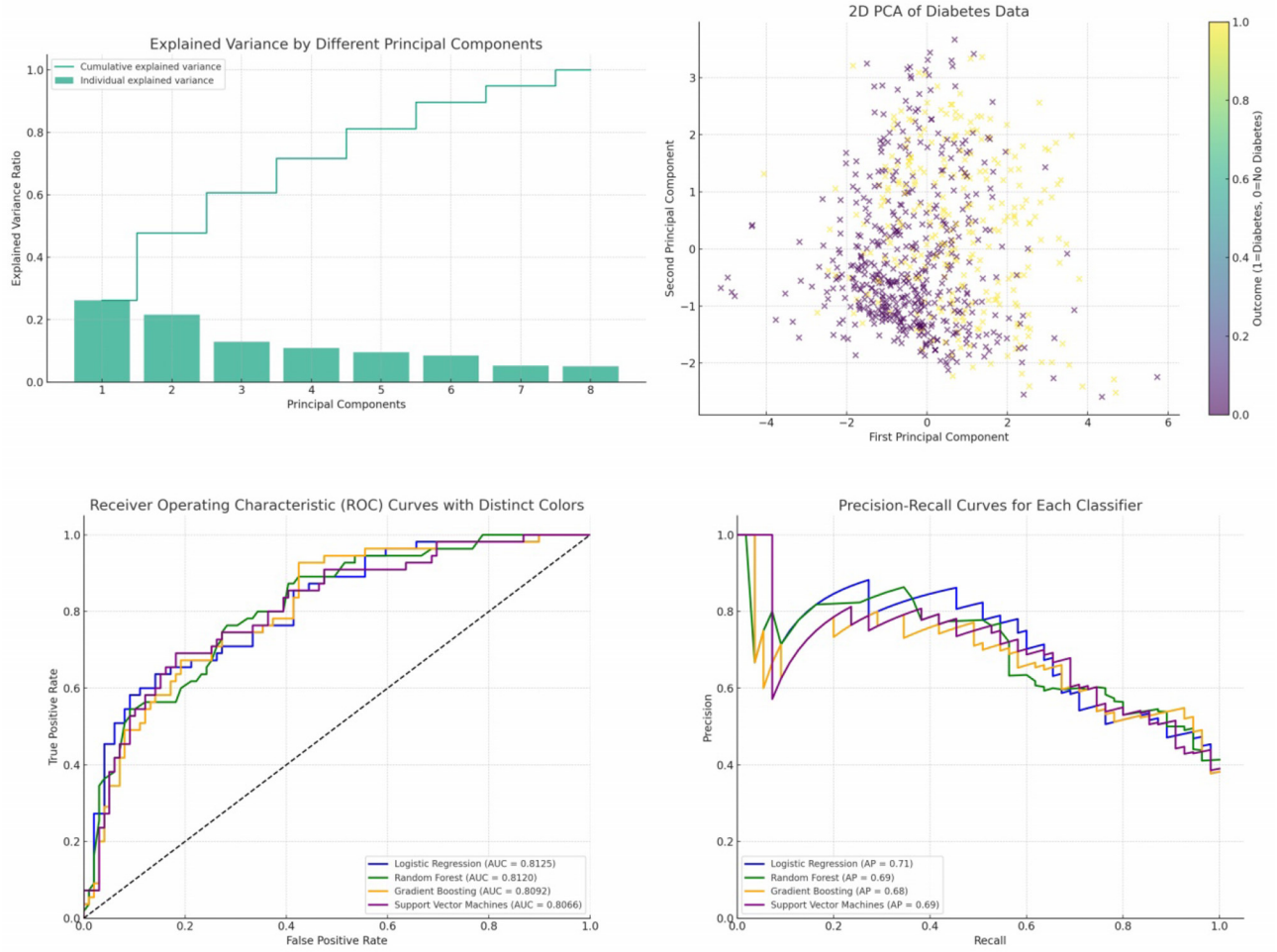


FIGURE 5
Cumulative and individual explained variance by different principal components, 2D PCA of diabetes data, ROC curves and PR curves of four different classifiers—Logistic Regression, Random Forest, Gradient Boosting, and Support Vector Machines.

library. The function was configured such that **X** represents the feature matrix encompassing all input variables, and **y** denotes the target variable, indicating the diabetes outcome. A division ratio was set with the parameter **test_size=0.2**, ensuring 20% of the dataset was reserved for testing, while the remaining 80% was used for training. To guarantee reproducibility in our experiments, a fixed seed (**random_state=42**) was used for the random number generator, ensuring that subsequent data splits would remain consistent. This data partitioning approach was instrumental in offering a comprehensive training regimen for our models while also providing an accurate evaluation framework. By training on a significant portion of the data, the models were exposed to diverse examples, enhancing their robustness. Meanwhile, the testing set, being distinct from the training data, offered insights into the models' real-world performance and generalization capabilities, a critical aspect in predictive medical analytics.

The Receiver Operating Characteristic (ROC) curve is a graphical representation that illustrates the diagnostic ability of a binary classifier as its discrimination threshold varies. The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) for various threshold values. The area under the ROC curve, termed as the Area Under Curve (AUC), provides a scalar value of the overall performance of the classifier, where a value of 1 indicates perfect classification and a value of 0.5 indicates that the classifier performs no better than random guessing.

4.1 Performance analysis of four different classifiers

In the provided Figure 5, the ROC curves of four different classifiers—Logistic Regression, Random Forest, Gradient Boosting, and Support Vector Machines—are depicted. Each curve represents the TPR vs. FPR for its respective classifier across different thresholds. Logistic Regression (LR) is represented by the blue curve. Random Forest (RF) is depicted by the green curve. Gradient Boosting (GB) is illustrated by the orange curve. Support Vector Machines (SVM) is shown by the purple curve.

The diagonal dashed line represents a classifier that predicts outcomes entirely by chance, without any learned insights from the data. An effective classifier's ROC curve will bow toward the top-left corner of the plot, indicating higher true positive rates for

lower false positive rates. The AUC values (provided in the legend) reveal the overall performance of each classifier. Higher AUC values indicate better classifier performance. The curves for the classifiers are above the diagonal line, suggesting that all four classifiers perform better than a random guess. Among the classifiers, Logistic Regression and Random Forest have very similar performance, with nearly identical AUC values. Gradient Boosting and Support Vector Machines have slightly lower AUC values, but they still indicate good classification performance.

Precision-Recall (PR) curves are a graphical representation that showcases the trade-off between precision and recall for different thresholds of a binary classifier, particularly useful when classes are imbalanced. Precision measures the accuracy of positive predictions, while recall (or sensitivity) measures the proportion of actual positives that were correctly identified. In the provided Figure 5, the PR curves of four different classifiers—Logistic Regression, Random Forest, Gradient Boosting, and Support Vector Machines—are presented. Each curve plots precision against recall for its respective classifier across different thresholds. Logistic Regression (LR) is represented by the blue curve. Random Forest (RF) is depicted by the green curve. Gradient Boosting (GB) is illustrated by the orange curve. Support Vector Machines (SVM) is shown by the purple curve.

The Average Precision (AP) values, provided in the legend, offer a summary measure of the PR curve, indicating the classifier's average precision value for all possible recall levels. All classifiers exhibit curves that are significantly above the baseline, indicating that they provide meaningful predictions beyond random guessing. The curves for Logistic Regression and Random Forest are closer to the top-right corner, suggesting that they might offer a better balance between precision and recall for certain threshold values compared to Gradient Boosting and Support Vector Machines. The AP values suggest that the classifiers have comparable performances, with Logistic Regression and Random Forest having slightly higher AP values than Gradient Boosting and Support Vector Machines. As expected, there's an evident trade-off between precision and recall. As recall increases, precision tends to decrease and vice versa. This is a typical characteristic of classifiers, and the optimal balance depends on the specific application and its requirements.

The confusion matrix for the Logistic Regression classifier shows a balanced prediction across both classes, refer Figure 6. The number of True Positives suggests that this model has a reasonable ability to correctly predict the positive class (patients with diabetes). The True Negatives indicate that the model also effectively identifies the negative class (patients without diabetes). However, the presence of False Positives and False Negatives means the model does make mistakes, especially in instances where patients without diabetes are incorrectly classified as having diabetes and vice versa.

Random Forest, an ensemble learning method, shows a similar trend in its confusion matrix, refer Figure 6. The model exhibits a robust performance in predicting both positive and negative classes. Nevertheless, there are instances where the model misclassifies, indicating areas for potential improvement, possibly through hyperparameter tuning or feature engineering.

Gradient Boosting, another ensemble technique, has its confusion matrix showcasing a different pattern, refer Figure 6. While the model has a commendable number of True Positives, there are noticeable False Negatives, suggesting that there are cases where patients with diabetes are incorrectly predicted as not having diabetes. This could be a cause for concern in a medical setting, as missing a positive diagnosis can have significant repercussions.

Support Vector Machines, a powerful linear classifier, displays a distinct pattern in its confusion matrix, refer Figure 6. The model seems to have a conservative approach, with a higher number of True Negatives. However, this also results in a considerable number of False Negatives, indicating that while the model is cautious about false alarms (FP), it might miss out on some actual positive cases (FN).

Table 6 resents a comparative evaluation of four distinct machine learning classifiers—Logistic Regression, Random Forest, Gradient Boosting, and Support Vector Machines—employed for diabetes prediction. Each classifier's performance is quantified using five pivotal metrics: Accuracy, Precision, Recall, F1-Score, and ROC-AUC.

In our evaluation of various classifiers for diabetes prediction, the Logistic Regression (LR) model exhibited an accuracy of 0.7468, implying it correctly predicts the diabetes outcome around 74.68% of the time, serving as a reflection of the model's overall correctness. Its precision of 0.6379 reveals that about 63.79% of the diabetes-positive predictions were accurate, showcasing the model's exactness. With a recall value of 0.6727, the LR model identified roughly 67.27% of all genuine diabetes-positive instances, indicating its capability to capture positive cases. An F1-Score of 0.6549, which represents the harmonic mean of precision and recall, infers a balanced trade-off between these two metrics. The ROC-AUC score for the LR model stands at 0.8125, highlighting its proficient ability to differentiate between positive and negative classes.

Moving on to the Random Forest (RF) model, it achieved an accuracy of 0.7208, suggesting it accurately predicts in approximately 72 out of every 100 instances. A precision of 0.6071 insinuates that nearly 60.71% of its positive predictions are correct. It boasts a recall of 0.6182, which might be perceived as moderate, capturing about 61.82% of actual positive cases. Its F1-Score of 0.6126 hints at a balanced model performance, with potential areas for improvement in both precision and recall. With an ROC-AUC of 0.8120, the RF model manifests a robust capacity to distinguish between the classes.

Interestingly, the Gradient Boosting (GB) model displayed metrics identical to the LR model. This parallelism is noteworthy, suggesting that, given this dataset and its configuration, both LR and GB offer similar performance dynamics.

Support Vector Machines (SVM) classifier registers the highest accuracy among the evaluated models at 0.7662, translating to nearly 76.62% correct predictions. It also leads in precision with a score of 0.7209, making its positive predictions considerably reliable. However, its recall is the least at 0.5636, pointing to potential misses in actual positive cases. The F1-Score of 0.6327 insinuates a tilt toward precision, possibly at the cost of recall. While its ROC-AUC score of 0.8066 is marginally lower than the others, it still represents a commendable capability in class separation.

The SVM classifier displays the highest accuracy, making it potentially the most reliable in general predictions. SVM prioritizes

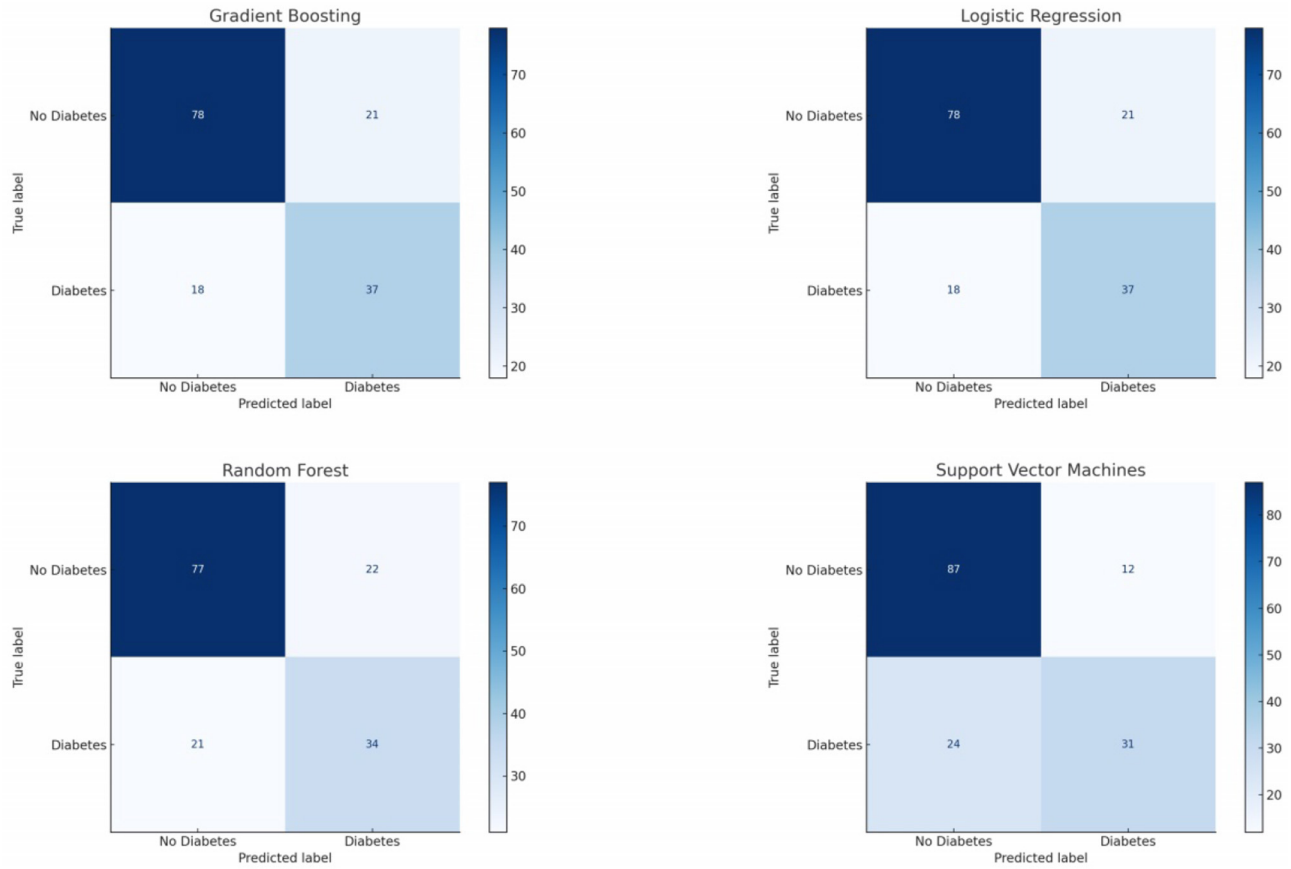


FIGURE 6
Confusion matrix for the classifiers used in this work.

TABLE 6 **Performance metrics of various classifiers for diabetes prediction.**

Classifier	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic regression	0.7468	0.6379	0.6727	0.6549	0.8125
Random forest	0.7208	0.6071	0.6182	0.6126	0.8120
Gradient boosting	0.7468	0.6379	0.6727	0.6549	0.8092
Support vector machines	0.7662	0.7209	0.5636	0.6327	0.8066

precision over recall, making its predictions more trustworthy but possibly missing out on some true positive cases. In contrast, Logistic Regression and Gradient Boosting offer a more balanced trade-off. All classifiers exhibit AUC scores above 0.8, suggesting that each of them has a strong capability to differentiate between positive and negative classes.

4.2 Feature engineering and correlation analysis

This section offers a comprehensive account of the feature engineering and correlation analysis conducted in the study. In the quest to enhance the predictive prowess of our model, we delved into advanced feature engineering techniques. These techniques aimed to unearth hidden relationships and patterns in the data that might be pivotal for accurate diabetes prediction.

4.2.1 Interaction features

One of the salient techniques employed was the creation of interaction features. These features represent interactions between pairs of existing attributes, capturing the combined effect of two variables on the outcome. A quintessential example from our dataset is the interaction between "Pregnancies" and "Age". The rationale behind such interactions is that the combined effect of two variables might be different from the sum of their individual effects.

4.2.2 Polynomial features

To unravel non-linear relationships inherent in the data, we ventured into polynomial feature generation. By squaring or cubing attributes, we aimed to encapsulate intricate patterns that linear terms might overlook. Notable instances from our dataset include squared terms for "Glucose" and "BMI".

Post this rigorous feature engineering, our dataset was enriched with both interaction and polynomial features, amplifying its information content.

Table 7 lists the original features alongside their correspondingsquared (polynomial) and interaction feature names. The table provides a comprehensive overview of the transformed features in the dataset, allowing for a clearer understanding of their nature and potential utility in modeling.

We generated a heatmap to visualize the correlation of all transformed features with the "Outcome" variable as depicted in Figure 7. This will help us see which features have the strongest relationship with the target. The heatmap visualizes the correlation of the top 10 and bottom 10 transformed features (based on their absolute correlation with the "Outcome" variable). In addressing multicollinearity within our highly correlated features, we set a correlation threshold at 0.85, above which we evaluated the need for feature removal through the Variance Inflation Factor (VIF), with a cut-off value of 10 indicating significant multicollinearity. Concurrently, we analyzed feature importance and assessed the impact on model performance to ensure that any exclusion would not compromise predictive accuracy. Clinically significant features were retained or adjusted based on domain knowledge, with a careful balance between model complexity and interpretability. When necessary, dimensionality reduction techniques like PCA were employed to condense correlated features into principal components, maintaining robustness without losing essential information.

The colors range from blue (negative correlation) to red (positive correlation). The strength and direction of the correlation between pairs of variables are represented by the color intensity and the annotated values. The diagonal line (from the top left to the bottom right) represents each feature's correlation with itself, which is always 1. The first row/column represents the correlation of each feature with the "Outcome" variable. The features at the top have the highest positive correlation with the outcome, while those at the bottom have the lowest (or highest negative).

Some features, like "Glucose", Glucose^2 Glucose^2, and "Glucose × Age", have a strong positive correlation with the "Outcome". This suggests that as these feature values increase, the likelihood of having diabetes (Outcome = 1) also increases. We can also observe the correlation between features. For example, "Glucose" and Glucose^{2}Glucose^2 are highly correlated, which is expected. Using this heatmap, we can prioritize features based on their correlation with the target variable.

TABLE 7 Transformed features, along with their classification as either "Polynomial" or "Interaction".

Transformed features	Feature type
Pregnancies^2	Polynomial
Glucose^2	Polynomial
BloodPressure^2	Polynomial
SkinThickness^2	Polynomial
Insulin^2	Polynomial
BMI^2	Polynomial
DiabetesPedigreeFunction^2	Polynomial
Age^2	Polynomial
Pregnancies × Glucose	Interaction
Pregnancies × BloodPressure	Interaction
Pregnancies × SkinThickness	Interaction
Pregnancies × Insulin	Interaction
Pregnancies × BMI	Interaction
Pregnancies × DiabetesPedigreeFunction	Interaction
Pregnancies × Age	Interaction
Glucose × BloodPressure	Interaction
Glucose × SkinThickness	Interaction
Glucose × Insulin	Interaction
Glucose × BMI	Interaction
Glucose × DiabetesPedigreeFunction	Interaction
Glucose × Age	Interaction
BloodPressure × SkinThickness	Interaction
BloodPressure × Insulin	Interaction
BloodPressure × BMI	Interaction
BloodPressure × DiabetesPedigreeFunction	Interaction
BloodPressure × Age	Interaction
SkinThickness × Insulin	Interaction
SkinThickness × BMI	Interaction
SkinThickness × DiabetesPedigreeFunction	Interaction
SkinThickness × Age	Interaction
Insulin × BMI	Interaction
Insulin × DiabetesPedigreeFunction	Interaction
Insulin × Age	Interaction
BMI × DiabetesPedigreeFunction	Interaction
BMI × Age	Interaction
DiabetesPedigreeFunction × Age	Interaction

4.3 Performance of ensemble model

Given the intricacies of predicting diabetes outcomes based on physiological measurements, we opted for a complex ensemble model, combining the strengths of various machine learning algorithms. This model integrates decision trees, gradient boosting, and support vector machines to harness their collective predictive power.

Recognizing the inherent relationships between physiological parameters, we employed polynomial and interaction feature

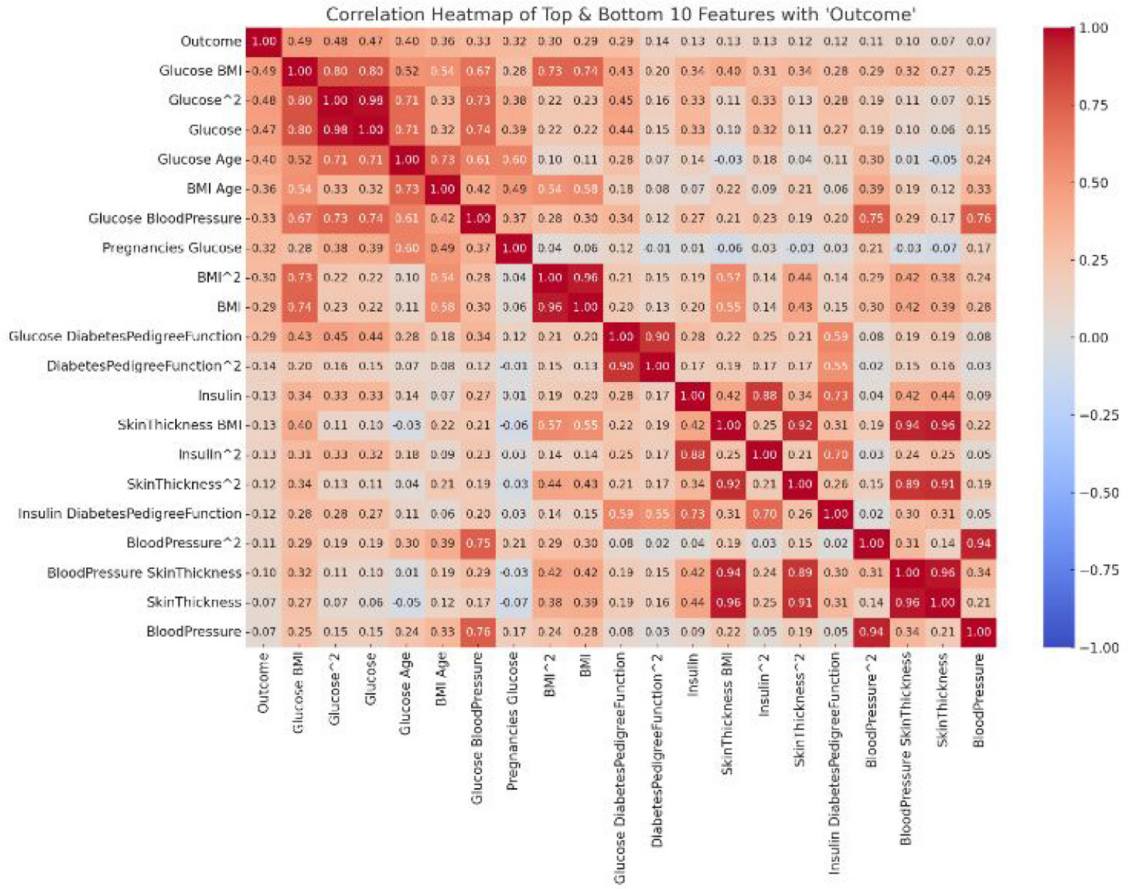


FIGURE 7
Correlation matrix.

engineering. This approach allowed the model to capture non-linear relationships and interactions that might be lost in simpler models. For instance, interactions between "Pregnancies" and "Age" or polynomial features like "Glucose^2" were introduced to better represent the underlying complexities of diabetes onset.

Due to the class imbalance evident in our dataset, we used a stratified sampling approach, ensuring each training batch had a representative mix of both diabetes outcomes. Additionally, the model was trained using a five-fold cross-validation strategy to ensure robustness and minimize overfitting. Although a higher number of folds could offer a marginally more stable performance estimate, we found that five-fold cross-validation provided a sufficient reduction in variance while maintaining low bias, without imposing excessive computational demands. This choice aligns with common practices in the literature, ensuring comparability across studies. Our analyses indicated that the variance in performance metrics was acceptably low across the folds, leading us to conclude that the benefits of additional folds would be minimal relative to the increased computational cost. The complex ensemble model, trained on the enhanced dataset, achieved an accuracy of over 93%.

This high accuracy, although promising, was scrutinized further using other metrics like precision, recall, and the F1-score. The model outperformed simpler classifiers and showed significant predictive power, especially when compared to models trained without the engineered features.

Figure 8 provides a holistic view of the model's performance. While accuracy gives an overall sense of correctness, precision and recall focus on the model's performance concerning each class. The F1-score is the harmonic mean of precision and recall, offering a balance between the two. The Area Under the Receiver Operating Characteristic curve (AUC-ROC) evaluates the model's ability to differentiate between the classes. The Matthews Correlation Coefficient (MCC) provides a balanced measure even when the classes are of very different sizes. Class 1 refers to individuals diagnosed with diabetes, and Class 0 refers to individuals without diabetes.

While accuracy provides a quick snapshot of overall correctness, it's vital to recognize its limitations, especially in datasets with class imbalances. Our dataset had a more significant number of non-diabetic individuals, which could bias the accuracy metric. Therefore, to truly appreciate the model's efficacy, we considered other metrics. Precision for Class 1 (diabetic individuals) stood at 91.0%, indicating that of all individuals the model labeled as diabetic, 91.0% were correctly classified. However, recall, which measures the model's ability to correctly identify all

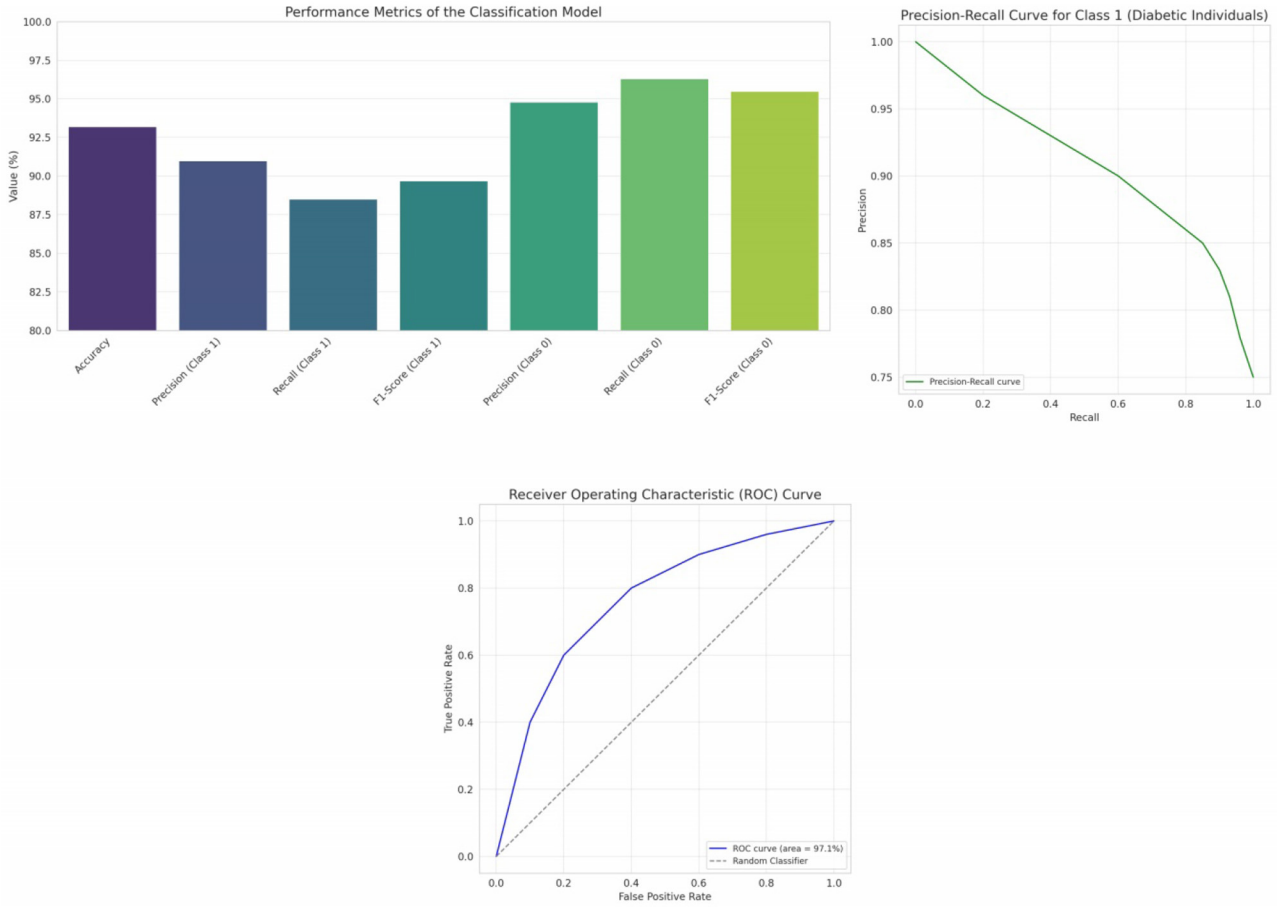


FIGURE 8
Visual representation of the model's performance across various metrics, trade-off between precision and recall at various thresholds and receiver operating characteristic (ROC) curve provides insights into the model's ability to discriminate between the classes.

diabetic individuals, was 88.5%. This difference underscores the classic precision-recall trade-off, refer Figure 8. For Class 0 (non-diabetic individuals), both precision (94.8%) and recall (96.3%) were high, affirming the model's prowess in correctly classifying non-diabetic individuals. Striking a balance between precision and recall, the F1-scores for Class 1 and Class 0 were 89.7% and 95.5% respectively. This harmonic mean provides a more holistic view of the model's performance, emphasizing its capability to maintain a balance between false positives and false negatives. The high AUC-ROC value, i.e., 97.1% signifies the model's strong ability to differentiate between diabetic and non-diabetic individuals, further reinforcing its diagnostic potential, refer Figure 8.

MCC, which takes values between −1 and 1, offers a balanced measure of binary classification, especially for imbalanced datasets. Our model's MCC of 0.87 indicates a strong correlation between the observed and predicted classifications, showcasing the model's reliability. Our research underscores the significance of feature engineering and complex ensemble modeling in enhancing diabetes prediction. In our work, we rigorously validated the impact of feature engineering on model performance by employing the paired *t*-test, a statistical method appropriate for comparing the means of two related groups. This test was particularly suited for our analysis as it allowed us to assess the significance of performance changes before and after the introduction of engineered features, using the same dataset. A *p*-value was computed from the t-statistic, with a threshold of 0.05 to determine statistical significance. Our analysis yielded a *p*-value well below this threshold, firmly establishing that the enhancements in performance metrics attributable to feature engineering were statistically significant and not merely a product of random variation.

5 Conclusion

In the contemporary healthcare landscape, accentuated by the pressing challenges of the COVID-19 pandemic, rapid and accurate diagnostics have never been more pivotal. One such critical area of focus is diabetes, a condition that has been identified as a significant vulnerability in the face of the virus. Our research, set against this global backdrop, embarked on a mission to enhance diabetes prediction using state-of-the-art machine learning techniques. Initially, we evaluated a gamut of classifiers to serve as our baseline. The SVM classifier emerged as the frontrunner in terms of accuracy, boasting a commendable

rate of 76.62%. While its precision was also the highest among peers, its recall hinted at potential misses, possibly overlooking some true positive cases. In contrast, both Logistic Regression and Gradient Boosting classifiers offered a more balanced performance dynamic, with metrics almost mirroring each other. Random Forest, while robust, showcased areas of potential enhancement, especially when juxtaposed against its peers. Collectively, these evaluations provided a foundational understanding, setting the stage for further enhancements. Our next endeavor led us to the realms of advanced feature engineering. By creating interaction features and generating polynomial attributes, we sought to capture hidden patterns and intricate relationships pivotal for prediction accuracy. This intensive process enriched our dataset, amplifying its informational depth and breadth. Subsequently, correlation analysis, depicted through heatmaps, shed light on the relationships between the engineered features and the outcome. It reaffirmed the significance of attributes like Glucose and highlighted the potential of newly generated features. Incorporating the insights from our initial evaluations and the subsequent feature engineering, we proposed an ensemble model that integrated the strengths of Decision Trees, Gradient Boosting, and Support Vector Machines. This model, with an accuracy of 93.2%, showcases the potential of harmonizing diverse algorithms.

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

DT: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing—original draft. TG: Conceptualization, Methodology, Validation, Writing—original draft. VB: Data curation, Formal analysis, Methodology, Resources, Writing—review & editing. AA: Conceptualization, Funding acquisition, Resources, Validation, Writing—review & editing. FA: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Writing—original draft. JS: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing—original draft, Writing—review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Researchers Supporting Project number (RSP2023R395), King Saud University, Riyadh, Saudi Arabia.

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

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.

References

1. Adamidi ES, Mitsis K, Nikita KS. Artificial intelligence in clinical care amidst COVID-19 pandemic: a systematic review. *Comput Struct Biotechnol J*. (2021) 19:2833–50. doi: 10.1016/j.csbj.2021.05.010

2. Krishnamoorthi R. A novel diabetes healthcare disease prediction framework using machine learning techniques. *J Healthc Eng*. (2022) 2022:9872970. doi: 10.1155/2022/1684017

3. Saleem K, Saleem M, Ahmad RZ, Javed AR, Alazab M, Gadekallu TR, et al. Situation-aware BDI reasoning to detect early symptoms of covid 19 using smartwatch. *IEEE Sens J*. (2022) 23:898–905. doi: 10.1109/JSEN.2022.3156819

4. Jalil Z, Abbasi A, Javed AR, Khan MB, Hasanat MHA, Malik KM, et al. COVID-19 related sentiment analysis using state-of-the-art machine learning and deep learning techniques. *Front Public Health*. (2022) 9:812735. doi: 10.3389/fpubh.2021.812735

5. Cahn A. Prediction of progression from pre-diabetes to diabetes: development and validation of a machine learning model. *Diabetes Metab Res Rev*. (2020) 36:3252. doi: 10.1002/dmrr.3252

6. Nadeem MW, Goh HG, Ponnusamy V, Andonovic I, Khan MA, Hussain M. A fusion-based machine learning approach for the prediction of the onset of diabetes. *Healthcare*. (2021) 9:1393. doi: 10.3390/healthcare9101393

7. Wu YT. Early prediction of gestational diabetes mellitus in the Chinese population via advanced machine learning. *J Clin Endocrinol Metab*. (2021) 106:1191–205. doi: 10.1210/clinem/dgaa899

8. Sonar P, Jayamalini K. Diabetes prediction using different machine learning approaches. In: *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*. Erode (2019). p. 367–71. doi: 10.1109/ICCMC.2019.8819841

9. Sharma R, Rani S, Gupta D. Stress detection using machine learning classifiers in internet of things environment. *J Comput Theor Nanosci*. (2019) 16:4214–9. doi: 10.1166/jctn.2019.8502

10. Shakeel T, Habib S, Boulila W, Koubaa A, Javed AR, Rizwan M, et al. A survey on COVID-19 impact in the healthcare domain: worldwide market implementation, applications, security and privacy issues, challenges and future prospects. *Complex Intell Syst*. (2023) 9:1027–58. doi: 10.1007/s40747-022-00767-w

11. Chang V, Bailey J, Xu QA, Sun Z. Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms. *Neural Comput Appl*. (2023) 35:16157–73. doi: 10.1007/s00521-022-07049-z

12. Nai-Arun N, Moungmai R. Comparison of classifiers for the risk of diabetes prediction. *Proc Comput Sci*. (2015) 69:132–42. doi: 10.1016/j.procs.2015.10.014

13. Kandhasamy JP, Balamurali S. Performance analysis of classifier models to predict diabetes mellitus. *Proc Comput Sci*. (2015) 47:45–51. doi: 10.1016/j.procs.2015.03.182

14. Mercaldo F, Nardone V, Santone A. Diabetes mellitus affected patients classification and diagnosis through machine learning techniques. *Proc Comput Sci*. (2017) 112:2519–28. doi: 10.1016/j.procs.2017.08.193

15. Perveen S, Shahbaz M, Guergachi A, Keshavjee K. Performance analysis of data mining classification techniques to predict diabetes. *Proc Comput Sci*. (2016) 82:115–21. doi: 10.1016/j.procs.2016.04.016

16. Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine learning and data mining methods in diabetes research. *Comput Struct Biotechnol J*. (2017) 15:104–16. doi: 10.1016/j.csbj.2016.12.005

17. Zou Q, Qu K, Luo Y, Yin D, Ju Y, Tang H. Predicting diabetes mellitus with machine learning techniques. *Front Genet*. (2018) 9:515. doi: 10.3389/fgene.2018.00515

18. Fan Y, Long E, Cai L, Cao Q, Wu X, Tong R. Machine learning approaches to predict risks of diabetic complications and poor glycemic control in nonadherent type 2 diabetes. *Front Pharmacol*. (2021) 12:665951. doi: 10.3389/fphar.2021.665951

19. Kopitar L, Kocbek P, Cilar L, Sheikh A, Stiglic G. Early detection of type 2 diabetes mellitus using machine learning-based prediction models. *Sci Rep*. (2020) 10:11981. doi: 10.1038/s41598-020-68771-z

20. Yuvaraj N, Sripreethaa KR. Diabetes prediction in healthcare systems using machine learning algorithms on Hadoop cluster. *Cluster Comput*. (2019) 22:1–9. doi: 10.1007/s10586-017-1532-x

21. Theerthagiri P, Ruby AU, Vidya J. Diagnosis and classification of the diabetes using machine learning algorithms. *SN Comput Sci*. (2022) 4:72. doi: 10.1007/s42979-022-01485-3

22. Saru S, Subashree S. Analysis and prediction of diabetes using machine learning. *Int J Emerg Technol Innovat Eng*. (2019) 5:1–9.

23. Palimkar P, Shaw RN, Ghosh A. Machine learning technique to prognosis diabetes disease: random forest classifier approach. In: Bianchini M, Piuri V, Das S, Shaw RN, editors. *Advanced Computing and Intelligent Technologies. Lecture Notes in Networks and Systems, Vol. 218*. Singapore: Springer (2022).

24. Olisah CC, Smith L, Smith M. Diabetes mellitus prediction and diagnosis from a data preprocessing and machine learning perspective. *Comput Methods Programs Biomed*. (2022) 220:106773. doi: 10.1016/j.cmpb.2022.106773

25. Ganie SM, Malik MB. An ensemble machine learning approach for predicting type-II diabetes mellitus based on lifestyle indicators. *Healthcare Anal*. (2022) 2:100092. doi: 10.1016/j.health.2022.100092

26. Mousa A, Mustafa W, Marqas RB, Mohammed SHM. A comparative study of diabetes detection using the Pima Indian diabetes database. *J Duhok Univ*. (2023) 26:277–88. doi: 10.26682/sjuod.2023.26.2.24

27. Sivasankari SS, Surendiran J, Yuvaraj N, Ramkumar M, Ravi CN, Vidhya RG. Classification of diabetes using multilayer perceptron. In: *2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*. Ballari (2022). p. 1–5. doi: 10.1109/ICDCECE53908.2022.9793085

28. Jaganathan SCB. Machine learning for smartphone-based early detection of diabetic disease in Pima Indians diabetes database. *J Algebr Stat*. (2022) 13:780–96.

29. Pujari P. Classification of Pima Indian diabetes dataset using support vector machine with polynomial kernel. In: *Deep Learning, Machine Learning and IoT in Biomedical and Health Informatics*. CRC Press (2022). p. 55–67.

30. Schulz LO. Effects of traditional and western environments on prevalence of type 2 diabetes in Pima Indians in Mexico and the US. *Diabetes Care*. (2006) 29:1866–71. doi: 10.2337/dc06-0138

31. Zou Y. Development and internal validation of machine learning algorithms for end-stage renal disease risk prediction model of people with type 2 diabetes mellitus and diabetic kidney disease. *Ren Fail*. (2022) 44:562–70. doi: 10.1080/0886022X.2022.2056053

32. Iparraguirre-Villanueva O, Espinola-Linares K, Flores Castañeda RO, Cabanillas-Carbonell M. Application of machine learning models for early detection and accurate classification of type 2 diabetes. *Diagnostics*. (2023) 13:2383. doi: 10.3390/diagnostics13142383

33. Dong Z. Prediction of 3-year risk of diabetic kidney disease using machine learning based on electronic medical records. *J Transl Med*. (2022) 20:1–10. doi: 10.1186/s12967-022-03339-1

34. Nemat H, Khadem H, Eissa MR, Elliott J, Benaissa M. Blood glucose level prediction: advanced deep-ensemble learning approach. *IEEE J Biomed Health Inform*. (2022) 26:2758–69. doi: 10.1109/JBHI.2022.3144870

35. Gupta H, Varshney H, Sharma TK, Pachauri N, Verma OP. Comparative performance analysis of quantum machine learning with deep learning for diabetes prediction. *Complex Intell Syst*. (2022) 8:3073–87. doi: 10.1007/s40747-021-00398-7

36. Flores-Dorantes MT, Díaz-López YE, Gutiérrez-Aguilar R. Environment and gene association with obesity and their impact on neurodegenerative and neurodevelopmental diseases. *Front Neurosci*. (2020) 14:863. doi: 10.3389/fnins.2020.00863

37. Vishwakarma A, Saini A, Guleria K, Sharma S. An early prognosis of lung cancer using machine intelligence. In: *2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1)*. Bangalore (2023). p. 1–6. doi: 10.1109/ICAIA57370.2023.10169432

38. Gupta S, Sharma HK, Kapoor M. Introduction to internet of medical things (IoMT) and its application in smart healthcare system. In: *Blockchain for Secure Healthcare Using Internet of Medical Things (IoMT)*. Cham: Springer (2022). p. 13–25.

39. Kumar S, Behl T, Sachdeva M, Sehgal A, Kumari S, Kumar A, et al. Implicating the effect of ketogenic diet as a preventive measure to obesity and diabetes mellitus. *Life Sci*. (2021) 264:118661. doi: 10.1016/j.lfs.2020.118661

40. Arora A, Behl T, Sehgal A, Singh S, Sharma N, Bhatia S, et al. Unravelling the involvement of gut microbiota in type 2 diabetes mellitus. *Life Sci*. (2021) 273:119311. doi: 10.1016/j.lfs.2021.119311

frontiers | Frontiers in Public Health

TYPE Mini Review
PUBLISHED 18 December 2023
DOI 10.3389/fpubh.2023.1213033

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
J. Carolyn Graff,
University of Tennessee Health Science Center (UTHSC), United States

*CORRESPONDENCE
Paweł Gać
✉ pawelgac@interia.pl;
✉ pawel.gac@umw.edu.pl

RECEIVED 27 April 2023
ACCEPTED 30 November 2023
PUBLISHED 18 December 2023

CITATION
Wolińska IA, Kraik K, Poręba R, Gać P and Poręba M (2023) Environmental factors of obesity before and after COVID-19 pandemic: a review.
Front. Public Health 11:1213033.
doi: 10.3389/fpubh.2023.1213033

Environmental factors of obesity before and after COVID-19 pandemic: a review

Irena Anna Wolińska[1], Krzysztof Kraik[2], Rafał Poręba[3], Paweł Gać[4]* and Małgorzata Poręba[5]

[1]Divison of Pathophysiology, Department of Physiology and Pathophysiology, Wroclaw Medical University, Wrocław, Poland, [2]Students' Scientific Association of Cardiovascular Diseases Prevention, Wroclaw Medical University, Wrocław, Poland, [3]Department of Internal Medicine, Occupational Diseases and Hypertension, Wroclaw Medical University, Wrocław, Poland, [4]Division of Environmental Health and Occupational Medicine, Department of Population Health, Wroclaw Medical University, Wrocław, Poland, [5]Department of Paralympic Sport, Wroclaw University of Health and Sport Sciences, Wrocław, Poland

In past decades the prevalence of overweight and obesity had grown rapidly. There are numerous factors contributing to this unfavorable change in people's health. This review article investigates the environmental factors which may play a role in the prevalence of overweight and obesity and additionally the novel factors which appeared after the beginning of the COVID-19 pandemic, which caused the increase in BMI during the lockdown period. Most of the studies reveal that the COVID-19 pandemic and lockdown contributed to the growth of BMI in numerous countries and, eventually the prevalence of overweight and obesity increased. Studies suggest that the physical activity was decreased while sleep time and screen time were increased and the amount of food consumed increased, additionally more processed food with long shelf life was consumed. The diverse environmental factors may have an impact on obesity and overweight development taking into account policy and local school policy issues, socioeconomic status, lifestyle including physical activity, diet habits, and amongst others, more trivial causes such as uninteresting neighborhoods, lack of sense of security outside the place of residence or a long distance from shops. Still, this is the object of debate if air pollution is an environmental risk factor influencing the unfavorable trends towards increasing body weight.

KEYWORDS

obesity, overweight, COVID-19 pandemic, environmental factors, eating habits, pollution

Introduction

Nowadays, overweight and obesity are serious healthcare problems in most countries. The prevalence of overweight and obesity has been increasing globally continuously for several decades (1) and it seems that this trend will not change soon. Overweight is usually recognized when the BMI of an adult person is in the range of 25.0–29.9 while obesity is recognized when BMI is equal to 30 or it is higher. It is worth noting that the localization of adipose tissue is often overlooked in statistics. Central obesity, also known as visceral obesity may be in such cases neglected and, because of that, metabolically obese normal-weight people are not included in statistics, so the real prevalence of obesity might be higher. Also in some Asian countries the norms of weight BMI should be lower than it is accepted in Western countries (1, 2). Normal

BMI in Asian populations is accepted as 18.0–22.9, the overweight range is 23–24.9 and obesity is when BMI is equal to or higher than 25 (2).

Prior to the COVID pandemic nearly 1 in 3 people worldwide was classified as overweight or obese (1). The number of people with too high body weight was rising rapidly. The prevalence of overweight and obesity doubled in the years 1980–2015. The rise in prevalence was most intense in the years 1992–2002. Obesity was more frequently affecting women and older people. In wealthy countries it affected mostly people with low socioeconomic status while in poor countries it affected mostly middle-aged people living in wealthy urban environments (3). Before the age of 45 women are less often obese than men but after that age women were more often obese than men. It might be linked to menopause. The main causes of obesity were identified as diet, lifestyle and socioeconomic status (4). The significant changes in mean adults' BMI in different regions of the world in the years from 1975 to 2016 according to WHO's Global Health Observatory (5) are presented in Figure 1.

Numerous factors may increase the risk of development of overweight and obesity including genetic, environmental, behavioral, biological, social and psychogenic ones (2). Among those, the most significant factors are physical activity, alcohol consumption and socioeconomic status. Moreover, the interplay of genes and environment further increases the risk of the development of overweight or obesity (6).

In late 2019 the SARS-CoV-2 virus responsible for the disease called COVID-19 was recognized in China. It spread rapidly to other countries and became a danger worldwide. As a result, on the day 11 March 2020, COVID-19 was declared a pandemic by WHO. To slow down the rate at which the virus spreads the governments of many countries declared a lockdown and encouraged people to stay at home and to keep a social distance. In most countries the lockdown had been expected to last a few weeks only, but then it was extended several times, which, eventually, enforced a change in people's habits. It affected various daily routines including eating behaviors and physical activity. In addition, remote education was introduced in public schools and remote work became more common. It is possible that this change of behaviors affected people's weight all over the world and, as a result, it could also contribute to the increase in the prevalence of overweight and obesity, which are conditions related to a higher risk of cardiovascular diseases, cancers and diabetes mellitus – the leading death causes worldwide (7, 8). Furthermore, both overweight and obesity also contribute to a more severe course of other diseases including COVID-19. Several studies confirmed that most of the patients admitted to Intensive Care Units in the pandemic era were overweight or obese and it had been found that both conditions increase the risk of respiratory failure in COVID-19 patients (9).

The goal of this study is to investigate the up-to-date knowledge on environmental risk factors of obesity and overweight, especially considering the influence of the COVID-19 pandemic.

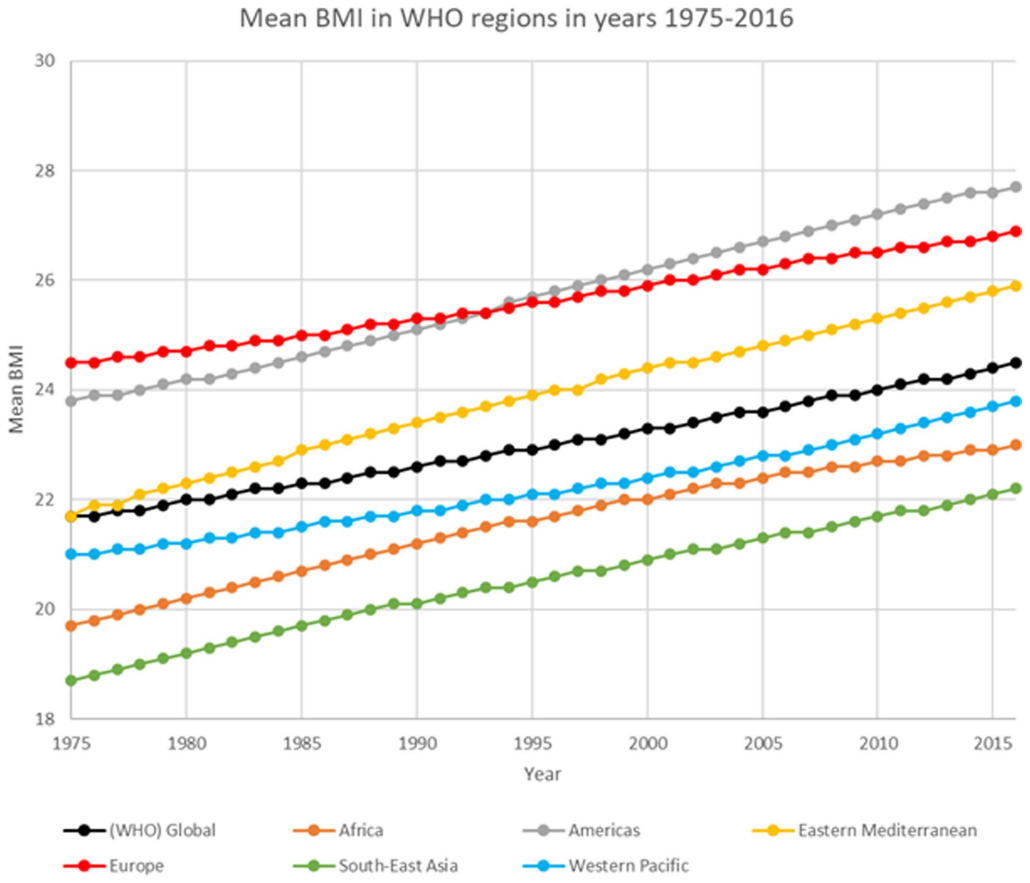


FIGURE 1
Mean BMI in WHO regions based on WHO's Global Health Observatory data (5).

10.3389/fpubh.2023.1213033

Methods

The non-systematic literature review was conducted using the following databases: PubMed, Cochrane Library, Embase and Google Scholar. Articles published between 1st January 2003 and 30th June 2023 were included. Different types of articles were included: systematic reviews, meta-analyses, reviews, clinical trials, randomized controlled trials, books and documents. A special focus was placed on systematic reviews and meta-analyses published since 2018 as these articles contain up-to-date information and they have the highest level of evidence.

In the search conducted in the databases we used a combination of groups of phrases to find publications related to the subject investigated by us. The first group of phrases included: "overweight," "obesity," "body weight," "weight gain" and "food consumption." The second group of phrases included: "environmental factors," "environment," "risk factors," "epidemiology," "pandemic," "COVID-19," "SARS-CoV-2," "lockdown," "coronavirus," "air pollution," "water pollution," "pollution," "pollutants," "smoking," "e-cigarettes," "work," "shift work," "night work," "circadian rhythm," "eating habits," "transport," "rural area," "urban area," "climate," "global warming," "daylight hours," "depression," and "stress." We used the conjunction "AND" in databases search boxes to connect both groups of phrases. We connected one phrase from the first group and one or more phrases from the second group in a single search. The duplicates were removed.

In the next step articles' titles and abstracts were screened to qualify them to full-text reading. The inclusion criteria were: (a) studies related to the investigated subject, (b) English or Polish language, (c) studies published in peer-reviewed journals, (d) human studies. The exclusion criteria were: (a) animal studies, (b) abstracts without full-text article, (c) conference proceedings. We made an exception to one study (10) investigating the effect of nanocolloids in drinking water on obesity in mice due to the lack of similar studies performed on the human population.

We obtained the full text of articles that initially met our criteria and during the full text read articles that not met all inclusion criteria or met any of the exclusion criteria in the full text were removed and finally 58 articles were included in this review. Types of articles and the number of articles of a given type included in this review are presented in Table 1.

The entire process of selection of articles was presented in Figure 2.

TABLE 1 Qualitative list of articles included in the current review.

Type of the article	Number of articles of a given type
Systematic review	8
Meta-analysis	11
Original research	24
Review	13
Editorial	1
Comment	1

Environmental risk factors of obesity and overweight

General features

In the last century, various changes including the industrialization of food production (2) have been introduced, which made the world a more obesogenic place. There are some types of overweight and obesity risk factors. The prevalence of overweight and obesity is influenced by age, sex, race, and socioeconomic status (11). The environment in which people live has many components that increase the risk of developing these conditions. The environmental risk factors include geography, food availability, work environment and transport-related factors. The prevalence of obesity is higher in some regions than in others. Even regions of the same country may have different rates of obesity (12). In the United States the rate of obesity is higher in rural areas than in urban areas (12, 13), which may be an astonishing fact. However, the influencing factors include access to healthy foods and, paradoxically, fewer opportunities to be physically active (13). Moreover, it may be caused by differences in education level and income of residents of these areas and by the local infrastructure (11). Food availability is determined by how easily people can get a certain type of food. When healthy food is not easily accessible it might contribute to the growth of obesity prevalence (14). Difficult accessibility may be related to both prices and distance to the store being the source of food (11). It has been found that decreasing the distance to the shop by opening a new one in a nearer area positively affects people's diet (15). Furthermore, when unhealthy food is easily accessible, for example in a nearby fast-food restaurant, it may also increase the risk of obesity development (16).

Advertising of fast food and calorie-rich food is another factor that increases caloric intake, and it especially affects children (17). Children are a vulnerable group that can be manipulated easily into buying certain products so many advertisements are aimed at them. Advertisements may shape their needs and preferences. It results in children buying products that have unfavorable effects on their health, eating more snacks and if that repeats, what is one of the aims of advertisements, they may carry harmful dietary habits to adult life or even develop overweight or obesity (18). Moreover, they may reduce the consumption of healthy food (19).

Work environment

The development of technology caused changes in the work environment. Physical labor is less common than it was in the past and work-related screen time increased. Simultaneously, people do not need to expend that much energy during work time and eventually it is associated with increasing body weight (20).

The popularization of shift work was another change that increased the risk of developing overweight and obesity, especially abdominal obesity (2, 21, 22). The most adverse effect was observed in people working permanently on night shifts. Mechanisms proposed to explain body weight gain are circadian rhythm disorders associated with the inability to adapt to working at night and sleeping during the day and sleep deprivation (21, 22). Other mechanisms included more opportunities to eat during night shifts, hormonal disturbances and fatigue, which promotes eating more and reduces physical activity (22). Bonham et al.'s study found that the energy intake in the groups of shift workers and day workers was similar so the weight gain may

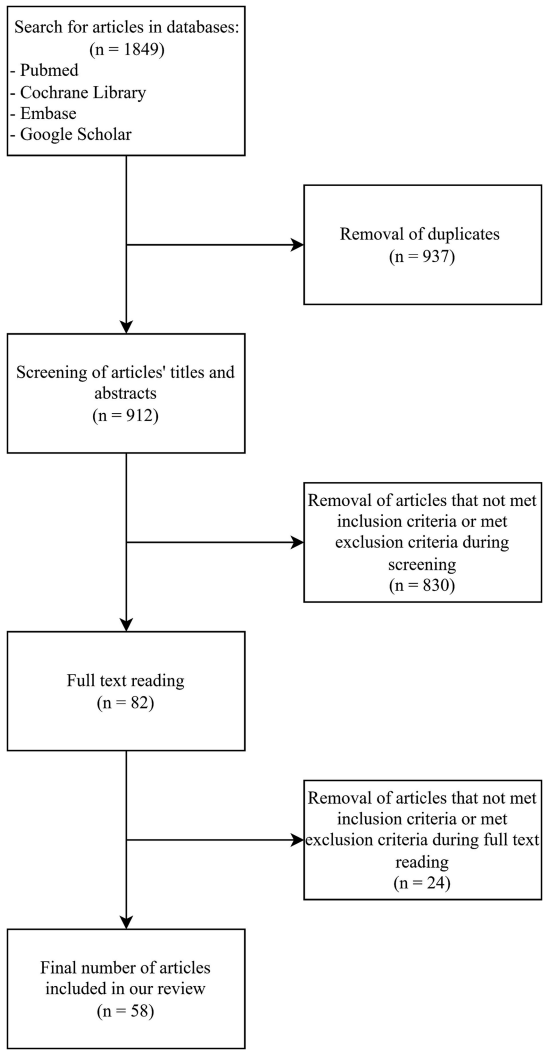


FIGURE 2
Summary of the article selection process.

be caused by meal timing, the type of consumed food and circadian rhythm disturbance (23).

Natural vs. built environment and transport

The means of transport also affect the prevalence of obesity (11). In areas in which people are more willing to walk the overweight and obesity rates are lower (24). People prefer to walk in areas with a good landscape, which have good pedestrian infrastructure including sidewalks and paths and which have parks and recreational facilities (25). However, people are more reluctant to be physically active in places that are dangerous because of high crime rates (26) and traffic-related risks (27). Because of that people living in well-kept locations with extensive pedestrian infrastructure are less likely to be overweight and obese while people living in areas that are neglected or have high crime rates or huge traffic are more likely to develop obesity. It is also known that environmental factors interact with the individual factors of a person (11). A study conducted in Nigeria found that people living in developing countries in Africa are affected by similar overweight and obesity risk factors to those in developed countries. The neighborhood which was inviting to go out was linked to lower overweight and obesity rates while poor and dangerous areas were associated with higher overweight prevalence. The presence of garbage, unpleasant smell, crime and long distance to shops were factors linked to being overweight. There were also other factors that were significant but only to males or only to females. Lack of good pedestrian infrastructure and low residential density increased the overweight rate in males. Meanwhile, heavy traffic and the lack of interesting surroundings in the neighborhood were associated with higher overweight prevalence in females. Other factors that might contribute to the occurrence of overweight in developing countries are bad transport infrastructure, lower income and the status of being married. It was estimated that environmental factors increased the risk of being overweight by 40 to 60% (27).

Leisure time

In other studies authors suggest that numerous other factors might be associated with the prevalence of obesity (28). The lack of recreational facilities may increase the chance of obesity in younger children by 68% (29). People who spend 3 h daily watching TV have two times greater prevalence of overweight than people who do not watch TV (28). These people are also more likely to be obese. Spending much time using smartphones and playing video games is even more likely to contribute to developing obesity because during these activities people often eat junk food which contains many obesogenic ingredients (28).

Smoking and eating habits

It has been found that smoking before and during pregnancy increases two times the risk of developing obesity during childhood (28). Moreover, gaining weight after smoking cessation is a very common phenomenon (30, 31). The cause of gaining weight after quitting smoking is excessive calorie intake, decreased resting metabolism rate, decreased physical activity and increased lipoprotein lipase activity (32, 33). However, smoking is harmful to the extent that the health damage from weight gain is less than the damage from continued smoking (32). Fortunately, there are interventions to prevent or reduce body weight gain, e.g., using bupropion (33), modifying diet or exercising (34). The role of e-cigarettes in terms of body weight is still unclear and requires more research on the human population. The conclusions of the current research are contradictory (35, 36). Some studies found that people using e-cigarettes had a higher prevalence of obesity than the normal-weight population. However, no significant causal link was found between e-cigarettes and obesity (36, 37).

Eating faster (38) and huge portions (39) of food are also factors that might contribute to higher calorie intake occurrence of obesity. Consumption of sweet beverages both with sugar and artificial sweeteners also increases the risk of body weight gain (28, 40). Poverty which is linked to low income and low education level also contributes to increasing obesity prevalence (41, 42). Social norms, prices of different types of food and fashion may both increase or decrease the rate of obesity occurrence (28). Families in which parents are overweight or obese have greater chances of having overweight children (43). This relationship is independent of genetic factors (28).

Climate, sun exposure, depression and stress

Changing climate and global warming also might be factors that increase the risk of obesity (44). There are reports that more energy is expended to digest colder food, and simultaneously it means, more calories are acquired by eating food at higher temperature than eating the same food at cold temperature. However, the potential effect of global warming on body weight is not large and is even less marked than the effect of owning a microwave (45).

The low number of daylight hours may also contribute to body weight gain by developing depression which increases the amount of food consumed by affected people (46). According to Luppino et al.'s meta-analysis depression increases the risk of developing obesity in both men and women due to hormonal changes (chronic activation of the hypothalamic–pituitary–adrenal axis), usage of antidepressants and lifestyle changes including the decreased amount of physical activity, switching to an unhealthy diet and eating an excessive amount of food when they feel bad (47). It is worth noting that obesity also may contribute to the development of depression. This reciprocal association has been found in many studies (47–49). However, according to Mannan et al.'s study, the risk of developing obesity due to depression is higher than the risk of developing depression due to obesity (48). Furthermore, Kanellopoulou et al.'s study found an association between depression and obesity in children (50).

Similarly, chronic stress also may contribute to excessive body weight gain. Nowadays, due to the constant rush and ambition, people are almost constantly exposed to chronic stress. Stress affects weight in many ways including overeating, eating calorie-rich food, decreasing the level of physical activity, decreasing the amount of sleep, disrupting intentional weight control, disrupting HPA axis, disrupting the reward center, changes in the gut microbiome and modifying the amount of synthesized regulatory peptides and hormones (neuropeptide Y, leptin, ghrelin). Furthermore, stigmatizing obese people increases the amount of stress they experience (51, 52). Moreover, some people may be more susceptible to stress due to individual factors such as the level of glucocorticosteroids and their sensitivity to glucocorticosteroids (53).

The impact of pollution on the prevalence of overweight and obesity

Because of the industrial development of the world the natural environment is becoming more and more degraded. The exploitation of the environment leads to climate change and the emission of pollutants, which decrease the quality of the air, water and soil. All these factors may cause adverse effects on human health.

Air pollution is one of the most important environmental problems related to human health. Air pollutants are responsible for health problems including cardiovascular system diseases, neoplastic diseases and respiratory system diseases. These three types of diseases are the leading cause of death worldwide. The most important air pollutants are carbon monoxide (CO), lead, nitrogen oxides (NOx), ground-level ozone (O3), particulate matter (PM), and sulfur oxides (SOx). The impact of air pollution depends on sex, age and which pollutant is present in the air. There are some hypothetical mechanisms in which air pollution contributes to weight gain. The pollutants may cause oxidative stress and inflammations which leads to metabolic disorders. These metabolic disorders may further contribute to the development of obesity. Pollutants also contribute to other diseases like asthma which make people less capable of physical activities and as a result abstain from exercise. People are also less likely to go outside and exercise when they know that the air is polluted and when they see that there is smog. It also decreases the amount of physical activity. It is to some extent controversial if air pollution contributes to obesity, as a similar number of studies support or deny this idea. However, it should be highlighted that fewer studies have shown that air pollution contributes to decreased risk of obesity (54). Studies investigating the impact of air pollution on body mass changes according to An et al. are presented in Figure 3.

It appears that children are more susceptible to the obesogenic impact of air pollution than adults (54). An association between PM2.5 exposure and increased adult BMI was found while no association was observed when it comes to PM10 and NO2 exposure (55). It is likely that skipping activity due to pollution is the main mechanism by which air pollution affects the obesity rate in adults (54). It is also found that air pollution may slow down the metabolism of young adults and increase the risk of obesity (56). It has been found that even prenatal exposure to pollutants in the air may affect obesity (55). Children who were affected prenatally by polycyclic aromatic hydrocarbons (PAHs), NO2, PM2.5 or benzopyrene were more likely to be obese in their childhood. Smoking is a source of PAHs so children of women who smoke during pregnancies are more likely to be overweight or obese. Traffic air pollution may affect the metabolism of newborns and lead to obesity. Children living close to places with large traffic like major streets had higher BMI than those who live further to these places (55, 56). Another possible mechanism is affecting the endocrine system by pollutants (55). It is hypothesized that high levels of NOx from traffic may cause inflammatory changes. There are studies that found no associations between low-level NOx exposure and overweight or obesity rates in children. Most of the research point that air pollution increases obesity in children (54–56). The pollutants that are most significant in childhood obesity are PM2.5, PM10, and NO2 (56). PM may also cause sleep disorders that contribute to weight gain.

Water pollution also may affect the prevalence of overweight and obesity. Both organic and inorganic water pollutants may create the nanocolloids which are contributing to obesity. It was found that exposure to nanocolloids increased the weight of mice. After the exposure there was a change in gut microbes, namely toward the status that is commonly present in obese individuals. These microbes generate then long-chain fatty acids. Also, the level of leptin increased, and the expression of adiponectin decreased. Nanocolloids are also responsible for disorders in blood lipid metabolism (10).

Overweight and obesity prevalence during the COVID-19 pandemic

Children and youths

In 2020 the hypothesis was made that lockdown might affect children's weight similarly to summer vacations because of the school closures (57). The hypothesis was based on the study which had been carried out prior to the pandemic. This study included a group of

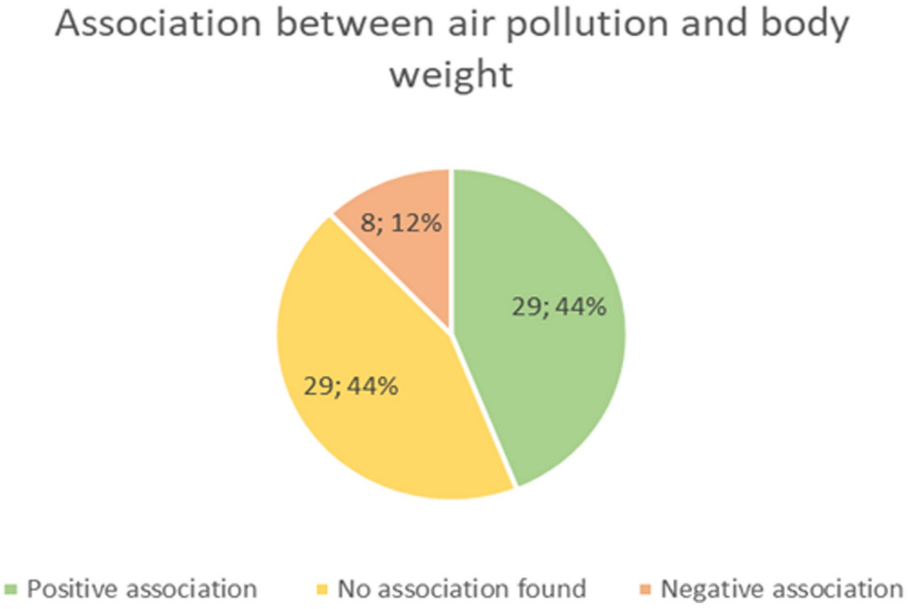


FIGURE 3
A number of studies which found positive, negative or no association between air pollution and body weight based on An et al. (54).

children whose BMI was monitored from kindergarten to second grade and it revealed that children's BMI increased faster during summer vacation compared to the school year and that prevalence of overweight and obesity increased only during summer vacations (58), Table 2.

That suggested that school attendance might reduce the impact of risk factors causing the growth of BMI. The hypothesis assumed that the pandemic would increase the screen time and consumption of snacks and shelf-stable food, which is usually highly processed and less healthy, and that social distancing will reduce physical activity in children, especially those who live in an urban environment (57).

Low physical activity and high screen time are likely to be risk factors for overweight and obesity in children (59). The low sample (41 participants) longitudinal study which was carried out in Italy based on telephone interviews with parents of obese children supports this hypothesis. The food consumption during the lockdown increased in this group: the consumption of unhealthy food (red meat, potato chips, and sweet beverages) increased significantly. The consumption of fruits also increased but its significance is not marked as clearly as the increased consumption of unhealthy food. Also, the number of meals consumed every day increased, especially in the group of males. There was also observed a change in the amount of time spent on different activities: sleeping time and screen time increased while the amount of time spent on doing sports decreased (60).

Another large sample (10,082 participants) retrospective study carried out in China based on a social media survey supports the statement that lockdown contributed to weight gain in youths. The study included youths between the age of 16 and 28 years. The average age of participants was 19.8 years. The data about BMI, the prevalence of overweight (defined in that study as BMI ≥23) and obesity (defined as BMI ≥27) and the lifestyle of youths before and during the lockdown were collected. The mean BMI increased from 21.8 before the lockdown to about 22.6 during the lockdown. The prevalence of both overweight and obesity increased. The screen time and sleeping time increased. Most of the participants kept a moderate level of physical activity. However, the rest of them rather decreased their physical activity due to the lockdown. Also, the amount of time spent on transport-related actions like walking and cycling decreased during the lockdown (61).

TABLE 2 A change of investigated parameters during school years and summer vacations based on von Hippel et al.'s study (58).

	Mean BMI	Prevalence of overweight	Prevalence of obesity
School years	↑	↓	↓
Summer vacations	↑↑↑	↑	↑

Furthermore, the meta-analysis encompassing 12 studies (including the two mentioned before) revealed that children's body weight and BMI have increased during the lockdown. Also, the prevalence of overweight and obesity increased in studied groups during the pandemic, especially in younger children aged from 5 to 9 years. The weight increase in the group of children affected by diabetes mellitus was not statistically significant. It was stated that the COVID-19 pandemic has worsened the epidemic of childhood obesity (62).

Adults

The lockdown and COVID-19 caused unfavorable changes in adults as well. Both the COVID-19 pandemic and all the rules introduced by different countries to prevent infection like lockdown or social distancing caused the change in people's diet and activity forms. In many cases the amount of physical activity decreased while sleep time and screen time increased. Additionally, the amount of food consumed increased. People were eating more processed food with long shelf life. Because of that it was more difficult for people to

control their body weight. These changes likely contributed to body weight gain (62–66).

The study conducted in the UK revealed that due to the lockdown adults encountered many barriers which were hindering them from maintaining the proper body weight. There were 2002 participants who completed the questionnaire about their behaviors during the lockdown. Adults have eaten more snacks, especially the ones who had high BMI prior to the lockdown. People who had high BMI also were overeating more frequently and had worse diet quality than people with normal BMI. The diet quality of most participants worsened. Because of panic people bought a lot of highly processed food with long shelf lives and ate it instead of fresh, nutritious and healthy food which was less accessible during the lockdown. Levels of physical activity were also lowered, again especially in the group of people with higher BMI as many people were afraid to exercise outside. Because of that people lost control of keeping their weight within the correct values (63).

In the meta-analysis including adults, weight gain was observed in 12.8–29.9% of cases during the lockdown (64). In one of the Iraqi studies the weight increased in over 30% of people (65). On the contrary, according to Italian authors weight loss was observed in 35.7% of people while weight gain was observed only in 11.1% of people, although, the study group included only people aged 60 and more (66). In another study including only obese people 36.3% of them gained weight during lockdown (67). The lockdown had the greatest impact on those who were already overweight and obese (63). Moreover, it was also found that younger participants gained weight faster than older ones (64).

Most likely lockdown contributed to the acceleration of weight gain, growth of BMI and increased prevalence of both overweight and obesity. This fact is particularly unfavorable because excess body weight is one of the factors associated with the severe course of COVID-19 (9).

Conclusion

Most of the studies reveal that the COVID-19 pandemic and lockdown contributed to the growth of BMI in many countries and in different populations of people and it increased the prevalence of overweight and obesity. Restrictions introduced to prevent infection including lockdown and social distancing caused changes in people's diet and activity forms. The physical activity was decreased while sleep time and screen time were increased and the amount of food consumed increased. More processed food with long shelf life was consumed. However, still it should be remembered that the background of the development of obesity and overweight is complex, and it employs a variety of components such as socioeconomic problems, diet habits, lifestyle, type of work, and even the neighborhood view quality. Additionally, air pollution may be associated with obesity prevalence, especially in children, while the impact of water pollution on obesity is less studied.

Author contributions

IW and MP: conceptualization. IW and KK: resources and writing—original draft preparation. PG, RP, and MP: writing—review and editing. IW and KK: visualization. RP and MP: supervision. PG: funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding

The APC was funded by Wroclaw Medical University (SUBZ. E264.23.039).

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. Chooi YC, Ding C, Magkos F. The epidemiology of obesity. *Metabolism*. (2019) 92:6–10. doi: 10.1016/j.metabol.2018.09.005

2. Płaczkiewicz-Jankowska E, Czupryniak L, Gajos G, Lewiński A, Ruchała M, Stasiak M, et al. Management of obesity in the times of climate change and COVID-19: an interdisciplinary expert consensus report. *Pol Arch Intern Med*. (2022) 132:16216. doi: 10.20452/pamw.16216

3. Swinburn BA, Sacks G, Hall KD, McPherson K, Finegood DT, Moodie ML, et al. The global obesity pandemic: shaped by global drivers and local environments. *Lancet*. (2011) 378:804–14. doi: 10.1016/S0140-6736(11)60813-1

4. Reynolds K, Gu D, Whelton PK, Wu X, Duan X, Mo J, et al. Prevalence and risk factors of overweight and obesity in China. *Obesity (Silver Spring)*. (2007) 15:10–8. doi: 10.1038/oby.2007.527

5. Mean BMI (kg/m2) (crude estimate). WHO's Global Health Observatory data (2017). WHO Global Health Estimates. Available at: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/mean-bmi-(kg-m-)-(crude-estimate)

6. Flores-Dorantes MT, Díaz-López YE, Gutiérrez-Aguilar R. Environment and Gene Association with obesity and their impact on neurodegenerative and neurodevelopmental diseases. *Front Neurosci*. (2020) 14:863. doi: 10.3389/fnins.2020.00863

7. Hruby A, Manson JE, Qi L, Malik VS, Rimm EB, Sun Q, et al. Determinants and consequences of obesity. *Am J Public Health*. (2016) 106:1656–62. doi: 10.2105/AJPH.2016.303326

8. World Health Organization. The top 10 causes of death. (2020). Available at: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

9. de Leeuw AJM, Oude Luttikhuis MAM, Wellen AC, Müller C, Calkhoven CF. Obesity and its impact on COVID-19. *J Mol Med (Berl)*. (2021) 99:899–915. doi: 10.1007/s00109-021-02072-4

10. Wei C, Feng R, Hou X, Peng T, Shi T, Hu X. Nanocolloids in drinking water increase the risk of obesity in mice by modulating gut microbes. *Environ Int*. (2021) 146:106302. doi: 10.1016/j.envint.2020.106302

11. Lee A, Cardel M, Donahoo WT. Social and environmental factors influencing obesity In: KR Feingold, B Anawalt and MR Blackmanet al, editors. *Endotext*. South Dartmouth, MA: MDText.com, Inc. (2019)

12. Gurka MJ, Filipp SL, DeBoer MD. Geographical variation in the prevalence of obesity, metabolic syndrome, and diabetes among US adults. *Nutr Diabetes*. (2018) 8:14. doi: 10.1038/s41387-018-0024-2

13. CDC. More obesity in U.S. rural counties than in urban counties. Available at: https://www.cdc.gov/media/releases/2018/s0614-obesity-rates.html

14. Chen D, Jaenicke EC, Volpe RJ. Food environments and obesity: household diet expenditure versus food deserts. *Am J Public Health*. (2016) 106:881–8. doi: 10.2105/AJPH.2016.303048

15. Dubowitz T, Ghosh-Dastidar M, Cohen DA, Beckman R, Steiner ED, Hunter GP, et al. Diet and perceptions change with supermarket introduction In *A food desert, but not because of supermarket use [published correction appears in health Aff (Millwood)]*. (2015) 34.

16. Maddock J. The relationship between obesity and the prevalence of fast food restaurants: state-level analysis. *Am J Health Promot*. (2004) 19:137–43. doi: 10.4278/0890-1171-19.2.137

17. Boyland EJ, Nolan S, Kelly B, Tudur-Smith C, Jones A, Halford JCG, et al. Advertising as a cue to consume: a systematic review and meta-analysis of the effects of acute exposure to unhealthy food and nonalcoholic beverage advertising on intake in children and adults. *Am J Clin Nutr*. (2016) 103:519–33. doi: 10.3945/ajcn.115.120022

18. Lapierre MA, Fleming-Milici F, Rozendaal E, McAlister AR, Castonguay J. The effect of advertising on children and adolescents. *Pediatrics*. (2017) 140:S152–6. doi: 10.1542/peds.2016-1758V

19. Gómez SF, Rajmil L. Advertising, obesity and child health: the case of Spain. *BMJ Paediatr Open*. (2022) 6:e001482. doi: 10.1136/bmjpo-2022-001482

20. Church TS, Thomas DM, Tudor-Locke C, Katzmarzyk PT, Earnest CP, Rodarte RQ, et al. Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. *PLoS One*. (2011) 6:e19657. doi: 10.1371/journal.pone.0019657

21. Sun M, Feng W, Wang F, Li P, Li Z, Li M, et al. Meta-analysis on shift work and risks of specific obesity types. *Obes Rev*. (2018) 19:28–40. doi: 10.1111/obr.12621

22. Liu Q, Shi J, Duan P, Liu B, Li T, Wang C, et al. Is shift work associated with a higher risk of overweight or obesity? A systematic review of observational studies with meta-analysis. *Int J Epidemiol*. (2018) 47:1956–71. doi: 10.1093/ije/dyy079

23. Bonham MP, Bonnell EK, Huggins CE. Energy intake of shift workers compared to fixed day workers: a systematic review and meta-analysis. *Chronobiol Int*. (2016) 33:1086–100. doi: 10.1080/07420528.2016.1192188

24. Creatore MI, Glazier RH, Moineddin R, Fazli GS, Johns A, Gozdyra P, et al. Association of Neighborhood Walkability with Change in overweight, obesity, and diabetes. *JAMA*. (2016) 315:2211–20. doi: 10.1001/jama.2016.5898

25. Smith M, Hosking J, Woodward A, Witten K, MacMillan A, Field A, et al. Systematic literature review of built environment effects on physical activity and active transport - an update and new findings on health equity. *Int J Behav Nutr Phys Act*. (2017) 14:158. doi: 10.1186/s12966-017-0613-9

26. Rees-Punia E, Hathaway ED, Gay JL. Crime, perceived safety, and physical activity: a meta-analysis. *Prev Med*. (2018) 111:307–13. doi: 10.1016/j.ypmed.2017.11.017

27. Oyeyemi AL, Adegoke BO, Oyeyemi AY, Deforche B, De Bourdeaudhuij I, Sallis JF. Environmental factors associated with overweight among adults in Nigeria. *Int J Behav Nutr Phys Act*. (2012) 9:32. doi: 10.1186/1479-5868-9-32

28. Nicolaidis S. Environment and obesity. *Metabolism*. (2019) 100:153942. doi: 10.1016/j.metabol.2019.07.006

29. Hawkesworth S, Silverwood RJ, Armstrong B, Pliakas T, Nanchahal K, Sartini C, et al. Investigating the importance of the local food environment for fruit and vegetable intake in older men and women in 20 UK towns: a cross-sectional analysis of two national cohorts using novel methods. *Int J Behav Nutr Phys Act*. (2017) 14:128. doi: 10.1186/s12966-017-0581-0

30. Farley AC, Hajek P, Lycett D, Aveyard P. Interventions for preventing weight gain after smoking cessation. *Cochrane Database Syst Rev*. (2012) 1:CD006219. Published 2012 Jan 18. doi: 10.1002/14651858.CD006219.pub3

31. Hartmann-Boyce J, Theodoulou A, Farley A, Hajek P, Lycett D, Jones LJ, et al. Interventions for preventing weight gain after smoking cessation. *Cochrane Database Syst Rev*. (2021) 2021:CD006219. doi: 10.1002/14651858.CD006219.pub4

32. Pistelli F, Aquilini F, Carrozzi L. Weight gain after smoking cessation. *Monaldi Arch Chest Dis*. (2009) 71:81–7. doi: 10.4081/monaldi.2009.367

33. Filozof C, Fernández Pinilla MC, Fernández-Cruz A. Smoking cessation and weight gain. *Obes Rev*. (2004) 5:95–103. doi: 10.1111/j.1467-789X.2004.00131.x

34. Jain P, Danaei G, Manson JE, Robins JM, Hernán MA. Weight gain after smoking cessation and lifestyle strategies to reduce it. *Epidemiology*. (2020) 31:7–14. doi: 10.1097/EDE.0000000000001106

35. Górna I, Napierala M, Florek E. Electronic cigarette use and metabolic syndrome development: a critical review. *Toxics*. (2020) 8:105. doi: 10.3390/toxics8040105

36. Hod R, Mohd Nor NH, Maniam S. Systematic review on e-cigarette and its effects on weight gain and adipocytes. *PLoS One*. (2022) 17:e0270818. Published 2022 Jul 5. doi: 10.1371/journal.pone.0270818

37. Sompa SI, Zettergren A, Ekström S, Upadhyay S, Ganguly K, Georgelis A, et al. Predictors of electronic cigarette use and its association with respiratory health and obesity in young adulthood in Sweden; findings from the population-based birth cohort BAMSE. *Environ Res*. (2022) 208:112760. doi: 10.1016/j.envres.2022.112760

38. Andrade AM, Kresge DL, Teixeira PJ, Baptista F, Melanson KJ. Does eating slowly influence appetite and energy intake when water intake is controlled? *Int J Behav Nutr Phys Act*. (2012) 9:135. doi: 10.1186/1479-5868-9-135

39. Ledikwe JH, Ello-Martin JA, Rolls BJ. Portion sizes and the obesity epidemic. *J Nutr*. (2005) 135:905–9. doi: 10.1093/jn/135.4.905

40. Malik VS, Schulze MB, Hu FB. Intake of sugar-sweetened beverages and weight gain: a systematic review. *Am J Clin Nutr*. (2006) 84:274–88. doi: 10.1093/ajcn/84.1.274

41. Kim TJ, von dem Knesebeck O. Income and obesity: what is the direction of the relationship? A systematic review and meta-analysis. *BMJ Open*. (2018) 8:e019862. doi: 10.1136/bmjopen-2017-019862

42. Cohen AK, Rai M, Rehkopf DH, Abrams B. Educational attainment and obesity: a systematic review. *Obes Rev*. (2013) 14:989–1005. doi: 10.1111/obr.12062

43. Huang H, Wan Mohamed Radzi CW, Salarzadeh JH. Family environment and childhood obesity: a new framework with structural equation modeling. *Int J Environ Res Public Health*. (2017) 14:181. doi: 10.3390/ijerph14020181

44. Blauw LL, Aziz NA, Tannemaat MR, Blauw CA, de Craen AJ, Pijl H, et al. Diabetes incidence and glucose intolerance prevalence increase with higher outdoor temperature. *BMJ Open Diabetes Res Care*. (2017) 5:e000317. doi: 10.1136/bmjdrc-2016-000317

45. Kanazawa S. Does global warming contribute to the obesity epidemic? *Environ Res*. (2020) 182:108962. doi: 10.1016/j.envres.2019.108962

46. Wurtman JJ, Wurtman RJ. Depression can beget obesity can beget depression. *J Clin Psychiatry*. (2015) 76:e1619–21. doi: 10.4088/JCP.15com10380

47. Luppino FS, de Wit LM, Bouvy PF, Stijnen T, Cuijpers P, Penninx BWJH, et al. Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Arch Gen Psychiatry*. (2010) 67:220–9. doi: 10.1001/archgenpsychiatry.2010.2

48. Mannan M, Mamun A, Doi S, Clavarino A. Prospective associations between depression and obesity for adolescent males and females-a systematic review and Meta-analysis of longitudinal studies. *PLoS One*. (2016) 11:e0157240. doi: 10.1371/journal.pone.0157240

49. Milaneschi Y, Simmons WK, van Rossum EFC, Penninx BW. Depression and obesity: evidence of shared biological mechanisms. *Mol Psychiatry*. (2019) 24:18–33. doi: 10.1038/s41380-018-0017-5

50. Kanellopoulou A, Antonogeorgos G, Douros K, Panagiotakos DB. The association between obesity and depression among children and the role of family: a systematic review. *Children (Basel)*. (2022) 9:1244. doi: 10.3390/children9081244

51. Tomiyama AJ. Stress and obesity. *Annu Rev Psychol*. (2019) 70:703–18. doi: 10.1146/annurev-psych-010418-102936

52. Moore CJ, Cunningham SA. Social position, psychological stress, and obesity: a systematic review. *J Acad Nutr Diet*. (2012) 112:518–26. doi: 10.1016/j.jand.2011.12.001

53. van der Valk ES, Savas M, van Rossum EFC. Stress and obesity: are there more susceptible individuals? *Curr Obes Rep*. (2018) 7:193–203. doi: 10.1007/s13679-018-0306-y

54. An R, Ji M, Yan H, Guan C. Impact of ambient air pollution on obesity: a systematic review. *Int J Obes*. (2018) 42:1112–26. doi: 10.1038/s41366-018-0089-y

55. Simkova S, Veleminsky M, Sram RJ. The impact of air pollution to obesity. *Neuro Endocrinol Lett*. (2020) 41:146–53.

56. Parasin N, Amnuaylojaroen T, Saokaew S. Effect of air pollution on obesity in children: a systematic review and Meta-analysis. *Children (Basel)*. (2021) 8:327. doi: 10.3390/children8050327

57. Rundle AG, Park Y, Herbstman JB, Kinsey EW, Wang YC. COVID-19-related school closings and risk of weight gain among children. *Obesity (Silver Spring)*. (2020) 28:1008–9. doi: 10.1002/oby.22813

58. von Hippel PT, Workman J. From kindergarten through second grade, U.S. Children's obesity prevalence grows only during summer vacations. *Obesity*. (2016) 24:2296–300. doi: 10.1002/oby.21613

59. Bekhwani AR, Khan M. Various risk factors of overweight and obesity among children aged 5-16 years. *J Coll Physicians Surg Pak*. (2022) 32:763–7. doi: 10.29271/jcpsp.2022.06.763

60. Pietrobelli A, Pecoraro L, Ferruzzi A, Heo M, Faith M, Zoller T, et al. Effects of COVID-19 lockdown on lifestyle behaviors in children with obesity living in Verona, Italy: a longitudinal study. *Obesity (Silver Spring)*. (2020) 28:1382–5. doi: 10.1002/oby.22861

61. Yang S, Guo B, Ao L, Yang C, Zhang L, Zhou J, et al. Obesity and activity patterns before and during COVID-19 lockdown among youths in China. *Clin Obes*. (2020) 10:e12416. doi: 10.1111/cob.12416

62. Chang TH, Chen YC, Chen WY, Chen CY, Hsu WY, Chou Y, et al. Weight gain associated with COVID-19 lockdown in children and adolescents: a systematic review and Meta-analysis. *Nutrients*. (2021) 13:3668. doi: 10.3390/nu13103668

63. Robinson E, Boyland E, Chisholm A, Harrold J, Maloney NG, Marty L, et al. Obesity, eating behavior and physical activity during COVID-19 lockdown: a study of UK adults. *Appetite*. (2021) 156:104853. doi: 10.1016/j.appet.2020.104853

64. Bakaloudi DR, Barazzoni R, Bischoff SC, Breda J, Wickramasinghe K, Chourdakis M. Impact of the first COVID-19 lockdown on body weight: a combined systematic review and a meta-analysis. *Clin Nutr*. (2022) 41:3046–54. doi: 10.1016/j.clnu.2021.04.015

65. Ahmed HO. The impact of social distancing and self-isolation in the last corona COVID-19 outbreak on the body weight in Sulaimani governorate-Kurdistan/Iraq, a prospective case series study. *Ann Med Surg (Lond)*. (2020) 59:110–7. doi: 10.1016/j.amsu.2020.09.024

66. Di Santo SG, Franchini F, Filiputti B, Martone A, Sannino S. The effects of COVID-19 and quarantine measures on the lifestyles and mental health of people over 60 at increased risk of dementia. *Front Psych*. (2020) 11:578628. doi: 10.3389/fpsyt.2020.578628

67. de Luis Román DA, Izaola O, Primo Martín D, Gómez Hoyos E, Torres Torres B, López Gómez JJ. Effect of lockdown for COVID-19 on self-reported body weight gain in a sample of obese patients. Efecto del confinamiento por COVID-19 sobre la ganancia de peso corporal autorreportada en una muestra de pacientes obesos. *Nutr Hosp*. (2020) 37:1232–7. doi: 10.20960/nh.03307

frontiers | Frontiers in Endocrinology

TYPE Mini Review
PUBLISHED 10 January 2024
DOI 10.3389/fendo.2023.1306290

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Mangesh Tiwaskar,
Shilpa Medical Research Centre, India

*CORRESPONDENCE
Hassan M. Heshmati
hassanheshmati@yahoo.com

RECEIVED 03 October 2023
ACCEPTED 13 November 2023
PUBLISHED 10 January 2024

CITATION
Heshmati HM (2024) Interactions between COVID-19 infection and diabetes.
Front. Endocrinol. 14:1306290.
doi: 10.3389/fendo.2023.1306290

Interactions between COVID-19 infection and diabetes

Hassan M. Heshmati*

Endocrinology Metabolism Consulting, LLC, Hassan Heshmati and Valerie Shaw Endocrine Research, Anthem, AZ, United States

Coronavirus disease 2019 (COVID-19) caused a major pandemic affecting human health and economy around the world since the beginning of 2020. The virus responsible for the disease is "severe acute respiratory syndrome coronavirus 2" (SARS-CoV-2). It invades the target cells by binding to angiotensin-converting enzyme 2 (ACE2). ACE2 is expressed in several organs including endocrine glands. Multiple endocrine and metabolic systems including the endocrine pancreas have been impacted by COVID-19 infection/pandemic. COVID-19 pandemic can promote obesity through alterations in lifestyle (e.g., unhealthy diet and reduced physical activity due to confinement and isolation) leading to type 2 diabetes and/or can directly impair the function of the endocrine pancreas particularly through a cytokine storm, promoting or aggravating type 1 or type 2 diabetes. The increased ACE2 receptors of high adiposity commonly associated with type 2 diabetes and the chronic hyperglycemia of diabetes with its negative impact on the immune system can increase the risk of COVID-19 infection and its morbidity/mortality. In conclusion, there are bidirectional interactions between COVID-19 pandemic and diabetes (e.g., COVID-19 infection can impact diabetes and diabetes can impact COVID-19 infection). The services offered by healthcare systems for the management of diabetes have been adapted accordingly.

KEYWORDS

coronavirus, angiotensin-converting enzyme 2, COVID-19 infection, pandemic, diabetes, immune system

Introduction

Since the beginning of 2020, COVID-19 infection became a global crisis of the 21st century, causing a major pandemic affecting human health and activities around the world and leading to a major international emergency (1–4). Several endocrine and metabolic systems including the endocrine pancreas have been impacted by this pandemic (5–9). COVID-19 infection caused a major disruption in the management of subjects with endocrine and metabolic disorders, especially those with diabetes, and the services offered by healthcare systems had to adapt accordingly and rapidly.

The purpose of this mini review is to present the interactions between COVID-19 infection and diabetes.

Pandemic

A pandemic is an epidemic that spreads globally, crosses international boundaries, and affects large number of people. Numerous pandemics have occurred throughout the history of mankind (10, 11). The deadliest pandemics were the Plague of Justinian, the Black Death, and the Spanish Flu.

The most recent pandemic was the COVID-19 pandemic (1–3). In January 2020, Chinese authorities announced the isolation of a new type of coronavirus, SARS-CoV-2, following the occurrence in December 2019 of several pneumonia cases of unknown etiology. On March 11, 2020, the World Health Organization declared COVID-19 a pandemic.

Pandemics can influence life at individual, familial, societal, and environmental levels (12). At the individual level, there are health consequences including infection caused by the pathogen, metabolic diseases, mental disorders, impact on pre-existing conditions, and eventually death, financial consequences mainly due to unemployment, and educational consequences caused by remote learning. At the familial level, there is a risk of domestic violence due to prolonged presence of parents and children at home. At the societal level, there are major economic consequences affecting several businesses (e.g., agriculture, restaurant, hotel, store, airline, cruise, convention, concert, sport event, museum, movie, and theater) caused by limitation of social life and activities. At the environmental level, confinement may have some health benefits, at least for short term, due to a reduction in air pollution mainly secondary to decrease in circulating cars and flying planes which can also positively impact life of animals and plants.

COVID-19 infection

Structure and mode of action of COVID-19 virus

The SARS-CoV-2, which is the virus responsible for the disease, is one of the coronaviruses in the family of Coronaviridae. It belongs to genera Betacoronavirus and is the seventh coronavirus known to cause human diseases. The virus is a spherical or pleomorphic enveloped, non-segmented, single-stranded, positive-sense RNA virus (Figure 1). It has four main structural proteins: spike (S), membrane (M), envelope (E), and nucleocapsid (N) proteins (1, 7, 13).

Like other viruses, SARS-CoV-2 can mutate. The mutated virus is referred to as a variant of the original virus. Several variants of SARS-CoV-2 have been reported (e.g., Alpha, Beta, Delta, Epsilon, Eta, Gamma, Iota, Kappa, Lambda, Omicron, Theta, and Zeta). They were initially detected in countries like the United Kingdom, South Africa, Brazil, and India. Some variants are more contagious and aggressive and may show more resistance to the current vaccines. The Delta variant and then the Omicron variant created serious concerns in several countries, where they became the dominant variants affecting adults, adolescents, and children, responsible for spike in hospitalizations. Within a variant, there are also several sub-variants (e.g., BA.4 and BA.5 Omicron sub-variants).

The transmission of human-to-human mainly occurs from direct contact or by droplets spread by infected subjects through cough or sneeze. The survival of the virus in the environment ranges from a few hours to a few days, depending on the conditions. The nose, mouth, and ocular mucosa are the major ways of transmission.

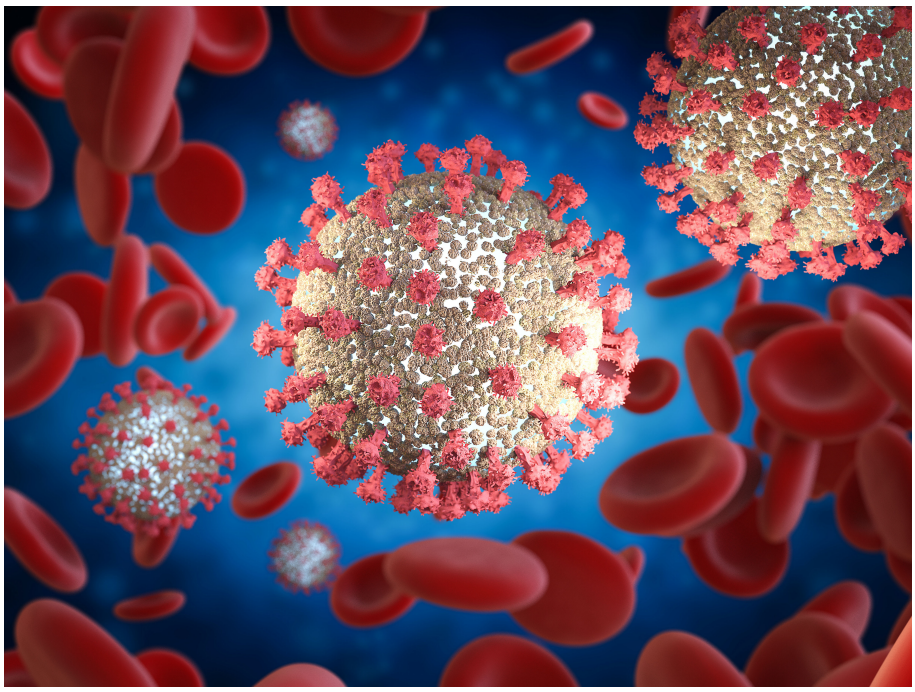


FIGURE 1
SARS-CoV-2. Copyright phonlamai (Kittipong Jirasukhanont)/Depositphotos Inc.

The virus enters and infects the target cells by binding to cell membrane protein receptors. The most well-described cell membrane protein receptor is ACE2, a zinc metalloprotease (7, 9, 13, 14). ACE2 is expressed in multiple organs including pancreas and endothelium. It is abundant in the epithelia of the lung and intestine. After the binding of the virus spike protein to ACE2, there is an internalization of ACE2. The virus uses the genetic system of the host cell and replicates. At the end, the virus leaves the cell through exocytosis. The infected cells undergo apoptosis or necrosis, triggering inflammatory responses.

Health consequences of COVID-19 infection

COVID-19 infection had influenced life at individual, familial, societal, and environmental levels through infection and confinement/isolation and had placed a significant burden upon healthcare worldwide (12). Different systems expressing ACE2 can be impacted by COVID-19 infection (e.g., respiratory, cardiovascular, neurological, gastrointestinal, and endocrine). COVID-19 infection causes alterations in the host immunological status including an increase in pro-inflammatory cytokines (e.g., interleukin-6 and tumor necrosis factor-α). The surge of pro-inflammatory factors (cytokine storm) can cause endothelial dysfunction and host organ damage (9, 15, 16).

Subjects infected with COVID-19 can be asymptomatic, have mild symptoms recovering within 1 to 2 weeks, or be severely affected with the ultimate risk of death. Common symptoms include fever, dry cough, dyspnea, arthralgia, myalgia, ageusia, and anosmia. COVID-19 symptoms can sometimes last several months. The damage to the lungs, heart, and brain increases the risk of long-term symptoms (long haulers). COVID-19 infection interferes with the onset and evolution of multiple endocrine and metabolic disorders (5–9, 17). For some endocrine/metabolic systems and diseases, information on pathophysiology and long-term outcome is relatively limited.

All age groups may be affected by COVID-19 infection. The disease is more severe in men (6, 18). Older subjects (> 65 years), black subjects, smokers, and subjects with immunodeficiency, cardiac and respiratory diseases, cancer, hypertension, diabetes, obesity, and dyslipidemia are considered high-risk populations (13, 19).

Social distancing and home confinement/isolation were the key public health recommendations during the COVID-19 pandemic. More than 4 billion people worldwide experienced the mobility restriction. The home confinement, self-isolation, and unemployment for a long period are responsible for alterations in lifestyle (e.g., unhealthy diet and reduced physical activity) and mental status (e.g., anxiety and depression). The confinement and isolation can impact access to health care (e.g., medications, physicians, and hospital beds) for the management of pre-existing medical conditions (e.g., heart disease, cancer, and diabetes) (12, 20). The changes in lifestyle can lead to insulin resistance and weight gain (overweight or obesity) and death may result from direct consequences of the viral infection, mental complications of confinement/isolation (risk of suicide), or aggravation of pre-existing diseases (4, 8, 20–22).

Diabetes

Prevalence of diabetes

Diabetes is a complex metabolic disease that results from deficiency of insulin secretion and/or action. It affects people regardless of country, age, or gender. Approximately half of the people living with diabetes are unaware of their condition. This proportion is much higher in low-income countries. Diabetes, both type 1 and type 2, is a major cause of morbidity and mortality in the world.

The prevalence of diabetes has risen significantly over the last several decades and is expected to rise dramatically in the years to come. In 2021, the global age-standardized prevalence of diabetes was 6.1% with 529 million people of all ages living with diabetes worldwide; type 2 diabetes accounted for more than 96% of the cases. In the USA, the numbers of subjects with type 1 and type 2 diabetes are around 2 and 35 million, respectively. It is expected that by 2050, the prevalence of diabetes will exceed 10% with more than 1.31 billion people living with diabetes. This increase will be mainly driven by type 2 diabetes which is primarily due to a rise in obesity (23).

Health consequences of diabetes

Diabetes causes multiple complications including macrovascular (e.g., cardiomyopathy) and microvascular (e.g., neuropathy, retinopathy, and nephropathy) complications inflicting high cost on healthcare (Figure 2) (24).

Hyperglycemia associated with diabetes causes impaired immunity (e.g., reduced T lymphocytes response, reduced neutrophil function, and disorders of humoral immunity) (25). The frequency and severity of infectious diseases are higher in subjects with diabetes in comparison to those without diabetes. The risk of infection is enhanced in the elderly population (26).

Impact of COVID-19 infection on diabetes

Impact through metabolic changes due to confinement and isolation

COVID-19 infection can impact type 2 diabetes through alterations in adipose tissue mass, cytokines levels, and insulin sensitivity. The confinement and isolation of the COVID-19 pandemic promote unhealthy diet (e.g., overeating) and reduced physical activity, both leading to excess adiposity (overweight or obesity) especially in high-income populations, increased production of cytokines (systemic inflammation), and insulin resistance (8, 20, 21). These alterations can promote or aggravate type 2 diabetes.

Daily exercise (e.g., low to medium-intensity exercise) is essential for preventing the negative impact of inactivity and improving health (20, 22).

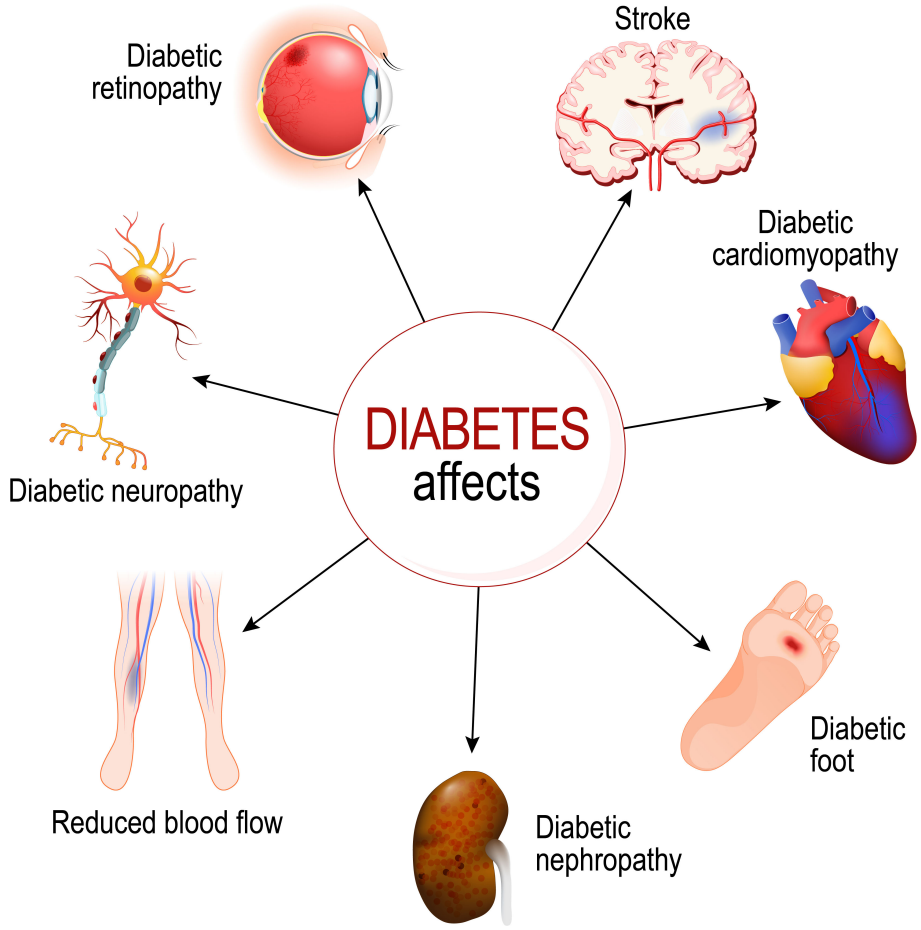


FIGURE 2
Diabetes negatively impacts several organs. Copyright edesignua (Tetiana Zhabska)/Depositphotos Inc.

Impact through cytokine storm

By causing a cytokine storm with the resulting release of multiple pro-inflammatory factors (e.g., interleukin-6 and tumor necrosis factor-α), insulin resistance, endothelial dysfunction, and damage of the pancreatic islets, COVID-19 infection contributes to the promotion or aggravation of type 1 or type 2 diabetes and complications of diabetes (16).

To reduce the risk of severe outcome of diabetes, it is crucial to achieve glycemic control. It is also important that subjects with diabetes be prioritized for COVID-19 vaccination.

Impact through changes in healthcare services

COVID-19 infection has caused a major disruption in the management of subjects with diabetes. The services offered by healthcare systems had to be adapted accordingly. Routine in-person appointments have been minimized to avoid crowds in waiting rooms with the risk of infection. Outpatient management with remote advice and support services using phone calls, video calls, and e-mails have been recommended, promoted, and implemented (19, 27). Elective surgical procedures have been postponed when possible.

Impact of diabetes on COVID-19 infection

Impact through adipose tissue

Type 2 diabetes is closely related to overweight and obesity. Indeed, in subjects with type 2 diabetes, at least 85% have overweight or obesity, and among subjects with obesity, around 30% have type 2 diabetes. ACE2 is expressed in adipose tissue. With higher adiposity commonly associated with type 2 diabetes, more receptors (ACE2) are available for SARS-CoV-2, exposing subjects to COVID-19 infection. In addition, subjects with excess adiposity may experience a more serious COVID-19 infection through several mechanisms (e.g., inflammation, impaired immunity, mechanical lung dysfunction, impact of comorbidities, and vitamin D deficiency) (5, 7, 15, 16, 18, 19, 28–31).

Management of subjects with excess body weight and COVID-19 who require treatment in intensive care units can be challenging (e.g., difficulty for moving, for intubating, and for obtaining diagnostic

imaging) (15). Subjects with overweight or obesity should have weight reduction using appropriate approaches and tools, when indicated (e.g., diet, exercise, behavioral change, drugs, medical devices, gut microbiome modulation, and bariatric surgery) (32–37).

Impact through immune system

In subjects with uncontrolled diabetes, chronic hyperglycemia negatively impacts the immune system and increases the risk of COVID-19 infection and its morbidity/mortality (5–7, 13, 15, 18, 19, 27, 38).

Appropriate glycemic control is essential to reduce the risk of COVID-19 infection. The outpatient plasma glucose goal in case of COVID-19 infection is 72-144 mg/dL with a hemoglobin A1c goal less than 7%. Plasma glucose should be monitored at least twice daily.

Impact through other complications of diabetes

The presence of other complications of diabetes such as endothelial dysfunction, cardiovascular disease, and nephropathy can be responsible for poor COVID-19 outcome (16).

Impact through antidiabetic medications

Some antidiabetic medications are not suitable in severe cases of COVID-19 infection (16). Particularly, sodium-glucose cotransporter-2 inhibitors should be discontinued in subjects severely affected by COVID-19 infection and who are at risk of dehydration. In hospitalized subjects with severe COVID-19 infection, insulin is the preferred treatment for type 2 diabetes with the use of continuous glucose monitoring (6).

Conclusion

There are bidirectional interactions between COVID-19 infection and diabetes. COVID-19 infection can impact the onset and/or the evolution of diabetes and diabetes can impact the onset and/or the evolution of COVID-19 infection.

COVID-19 infection has caused a major disruption in the management of subjects with diabetes. The medical services have been adjusted to the new situation. Routine in-person appointments can be reduced to avoid crowds in waiting rooms with the risk of infection. Outpatient care with remote advice and support services have been promoted. Because of the higher risk of mortality in subjects with diabetes who are infected by SARS-CoV-2, tight glycemic control and proper COVID-19 vaccination are essential in this population.

Author contributions

HMH: Conceptualization, Writing – original draft.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

Author HMH is affiliated with Endocrinology Metabolism Consulting, LLC, Hassan Heshmati and Valerie Shaw Endocrine Research.

Publisher's note

References

1. Sahin AR, Erdogan A, Mutlu Agaoglu P, Dineri Y, Cakirci AY, Senel ME, et al. 2019 Novel coronavirus (COVID-19) outbreak: a review of the current literature. *Eurasian J Med Oncol* (2020) 4:1–7. doi: 10.14744/ejmo.2020.12220

2. Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J Autoimmun* (2020) 109:102433. doi: 10.1016/j.jaut.2020.102433

3. Phelan AL, Katz R, Gostin LO. The novel coronavirus originating in Wuhan, China. Challenges for global health governance. *JAMA* (2020) 323:709–10. doi: 10.1001/jama.2020.1097

4. Woolf SH, Chapman DA, Lee JH. COVID-19 as the leading cause of death in the United States. *JAMA* (2021) 325:123–4. doi: 10.1001/jama.2020.24865

5. Pal R, Banerjee M. COVID-19 and the endocrine system: exploring the unexplored. *J Endocrinol Invest* (2020) 43:1027–31. doi: 10.1007/s40618-020-01276-8

6. Lundholm MD, Poku C, Emanuele N, Emanuele MA, Lopez N. SARS-CoV-2 (COVID-19) and the endocrine system. *J Endocr Soc* (2020) 4:1–13. doi: 10.1210/jendso/bvaa144

7. Kothandaraman N, Rengaraj A, Xue B, Yew WS, Velan SS, Karnani N, et al. COVID-19 endocrinopathy with hindsight from SARS. *Am J Physiol Endocrinol Metab* (2021) 320:E139–50. doi: 10.1152/ajpendo.00480.2020

8. Clemmensen C, Peterson MB, Sørensen TIA. Will the COVID-19 pandemic worsen the obesity epidemic? *Nat Rev Endocrinol* (2020) 16:469–70. doi: 10.1038/s41574-020-0387-z

9. Sorokin AV, Karathanasis SK, Yang ZH, Freeman L, Kotani K, Remaley AT. COVID-19-associated dyslipidemia: implications for mechanism of impaired resolution and novel therapeutic approaches. *FASEB J* (2020) 34:9843–53. doi: 10.1096/fj.202001451

10. Samal J. A historical exploration of pandemics of some selected diseases in the world. *IJHSR* (2014) 4:165–9.

11. Piret J, Boivin G. Pandemics throughout history. *Front Microbiol* (2021) 11:631736. doi: 10.3389/fmicb.2020.631736

12. Chu IYH, Alam P, Larson HJ, Lin L. Social consequences of mass quarantine during epidemics: a systematic review with implications for the COVID-19 response. *J Travel Med* (2020) 27:1–14. doi: 10.1093/jtm/taaa192

13. Muniyappa R, Gubbi S. COVID-19 pandemic, coronaviruses, and diabetes mellitus. *Am J Physiol Endocrinol Metab* (2020) 318:E736–41. doi: 10.1152/ajpendo.00124.2020

14. Mousavizadeh L, Ghasemi S. Genotype and phenotype of COVID-19: their role in pathogenesis. *J Microbiol Immunol Infect* (2021) 54:159–63. doi: 10.1016/j.jmii.2020.03.022

15. Marazuela M, Giustina A, Puig-Domingo M. Endocrine and metabolic aspects of the COVID-19 pandemic. *Rev Endocr Metab Disord* (2020) 21:495–507. doi: 10.1007/s11154-020-09569-2

16. Kazakou P, Lambadiari V, Ikonomidis I, Kountouri A, Panagopoulos G, Athanasopoulos S, et al. Diabetes and COVID-19; a bidirectional interplay. *Front Endocrinol* (2022) 13:780663. doi: 10.3389/fendo.2022.780663

17. Heshmati HM. Consequences of COVID-19 pandemic including endocrine and metabolic impacts. In: Heshmati HM, editor. *Hot Topics in Endocrinology and Metabolism*. London: IntechOpen (2021). p. 89–112. doi: 10.5772/intechopen.100278

18. dos Santos LR, Filho RB, Alves TAFA, Jardine MB, Durate ML. Covid-19 and endocrinology-intersections: a review of current knowledge on physiopathological associations. *Endocrinol Metab Int J* (2020) 8:90–5. doi: 10.15406/emij.2020.08.00286

19. Puig-Domingo M, Marazuela M, Giustina A. COVID-19 and endocrine diseases. A statement from the European Society of Endocrinology. *Endocrine* (2020) 68:2–5. doi: 10.1007/s12020-020-02294-5

20. Narici M, De Vito G, Franchi M, Paoli A, Moro T, Marcolin G, et al. Impact of sedentarism due to the COVID-19 home confinement on neuromuscular, cardiovascular and metabolic health: physiological and pathophysiological implications and recommendations for physical and nutritional countermeasures. *Eur J Sport Sci* (2021) 21:614–35. doi: 10.1080/17461391.2020.1761076

21. Martinez-Ferran M, de la Guía-Galipienso F, Sanchis-Gomar F, Pareja-Galeano H. Metabolic impacts of confinement during the COVID-19 pandemic due to modified diet and physical activity habits. *Nutrients* (2020) 12:1549. doi: 10.3390/nu12061549

22. Clemente-Suárez VJ, Dalamitros AA, Beltran-Velasco AI, Mielgo-Ayuso J, Tornero-Aguilera JF. Social and psychophysiological consequences of the COVID-19 pandemic: an extensive literature review. *Front Psychol* (2020) 11:580225. doi: 10.3389/fpsyg.2020.580225

23. Ong KL, Stafford LK, McLaughlin SA, Boyko EJ, Vollset SE, Smith AE, et al. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* (2023) 402:203–34. doi: 10.1016/S0140-6736(23)01301-6

24. Andersson E, Persson S, Hallén N, Ericsson Å, Thielke D, Lindgren P, et al. Costs of diabetes complications: hospital-based care and absence from work for 392,200 people with type 2 diabetes and matched control participants in Sweden. *Diabetologia* (2020) 63:2582–94. doi: 10.1007/s00125-020-05277-3

25. Casqueiro J, Casqueiro J, Alves C. Infections in patients with diabetes mellitus: a review of pathogenesis. *Indian J Endocrinol Metab* (2012) 16 Suppl 1:S27–36. doi: 10.4103/2230-8210.94253

26. Rajagopalan S. Serious infections in elderly patients with diabetes mellitus. *Clin Infect Dis* (2005) 40:990–6. doi: 10.1086/427690

27. Wake DJ, Gibb FW, Kar P, Kennon B, Klonoff DC, Rayman G, et al. Endocrinology in the time of COVID-19: remodelling diabetes services and emerging innovation. *Eur J Endocrinol* (2020) 183:G67-77. doi: 10.1530/EJE-20-0377

28. Popkin BM, Du S, Green WD, Beck MA, Algaith T, Herbst CH, et al. Individuals with obesity and COVID-19: a global perspective on the epidemiology and biological relationships. *Obes Rev* (2020) 21:e13128. doi: 10.1111/obr.13128

29. Rychter AM, Zawada A, Ratajczak AE, Dobrowolska A, Krela-Kaźmierczak I. Should patients with obesity be more afraid of COVID-19? *Obes Rev* (2020) 21:e13083. doi: 10.1111/obr.13083

30. Gao F, Zheng KI, Wang XB, Sun QF, Pan KH, Wang TY, et al. Obesity is a risk factor for greater COVID-19 severity. *Diabetes Care* (2020) 43:e72–4. doi: 10.2337/dc20-0682

31. Kwok S, Adam S, Ho JH, Iqbal Z, Turkington P, Razvi S, et al. Obesity: a critical risk factor in the COVID-19 pandemic. *Clin Obes* (2020) 10:e12403. doi: 10.1111/cob.12403

32. Gadde KM, Martin CK, Berthoud HR, Heymsfield SB. Obesity. Pathophysiology and management. *J Am Coll Cardiol* (2018) 71:69–84. doi: 10.1016/j.jacc.2017.11.011

33. Heshmati HM. Management of excess adiposity. *Rep Glob Health Res* (2022) 5:144. doi: 10.29011/2690-9480.100144

34. Saxon DR, Iwamoto SJ, Mettenbrink CJ, McCormick E, Arterburn D, Daley MF, et al. Antiobesity medication use in 2.2 million adults across eight large health care organizations: 2009-2015. *Obesity* (2019) 27:1975–81. doi: 10.1002/oby.22581

35. Heshmati HM. Anti-obesity medical devices. In: Himmerich H, editor. *Weight Management*. London: IntechOpen (2020). p. 239–53. doi: 10.5772/intechopen.91697

36. Heshmati HM. Gut microbiome in obesity management. In: Himmerich H, editor. *Weight Management*. London: IntechOpen (2020). p. 255–68. doi: 10.5772/intechopen.91974

37. Radvinsky D, Iskandar M, Ferzli G. Bariatric surgery today: the good, the bad, and the ugly. *Ann Laparosc Endosc Surg* (2017) 2:52. doi: 10.21037/ales.2017.02.26

38. Guo W, Li M, Dong Y, Zhou H, Zhang Z, Chunxia T, et al. Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes Metab Res Rev* (2020) 36:e3319. doi: 10.1002/dmrr.3319

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 08 March 2024
DOI 10.3389/fpubh.2024.1327429

OPEN ACCESS

EDITED BY
Pranav Kumar Prabhakar,
Lovely Professional University, India

REVIEWED BY
Erwin Calgua,
Universidad de San Carlos de Guatemala, Guatemala
Hina Sultana,
University of North Carolina System, United States
Jay Prakash Kumar,
National Institutes of Health (NIH), United States

*CORRESPONDENCE
Tamra Burns Loeb
✉ Tloeb@mednet.ucla.edu

RECEIVED 13 November 2023
ACCEPTED 23 February 2024
PUBLISHED 08 March 2024

CITATION
Loeb TB, Gholami M, Ramm K, Shedd K, Soetenga S, Jackson NJ, Chung UYR, Duru OK, Mangione CM, Hamilton AB and Moin T (2024) Multilevel perceptions of the virtual delivery of the University of California Diabetes Prevention Program on RE-AIM domains due to COVID-19 mandates. *Front. Public Health* 12:1327429. doi: 10.3389/fpubh.2024.1327429

Multilevel perceptions of the virtual delivery of the University of California Diabetes Prevention Program on RE-AIM domains due to COVID-19 mandates

Tamra Burns Loeb[1]*, Maryam Gholami[2], Kate Ramm[1], Kelly Shedd[3], Samantha Soetenga[4], Nicholas J. Jackson[1], Un Young Rebecca Chung[1], O. Kenrik Duru[1], Carol M. Mangione[1], Alison B. Hamilton[1,5] and Tannaz Moin[1,5]

[1]Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at the University of California, Los Angeles, Los Angeles, CA, United States, [2]Altman Clinical and Translational Research Institute (ACTRI), University of California, San Diego, San Diego, CA, United States, [3]UCI Health, University of California, Irvine, Irvine, CA, United States, [4]UCLA Campus Recreation, University of California, Los Angeles, Los Angeles, CA, United States, [5]VA Greater Los Angeles Healthcare System, Los Angeles, CA, United States

Background: The University of California's Diabetes Prevention Program (UC DPP) Initiative was implemented across all 10 UC campuses in 2018. The COVID-19 pandemic and accompanying mandates required swift changes to program delivery, including pivoting from in-person to virtual delivery (i.e., Zoom). Our goal was to assess multilevel constituent perceptions of the use of a virtual platform to deliver UC DPP due to COVID-19 mandates.

Methods: We conducted qualitative interviews with 68 UC DPP participants, coordinators, and leaders to examine the use of virtual platform delivery on the reach, effectiveness, adoption, implementation, and maintenance (RE-AIM) of UC DPP. Transcripts were analyzed using rapid qualitative analysis and emergent themes were categorized using domains corresponding to RE-AIM framework.

Results: Among UC DPP *participants* ($n = 42$), virtual delivery primarily impacted perceptions of UC DPP effectiveness and implementation. Some participants perceived program effectiveness to be negatively impacted, given their preference for in-person sessions, which they felt provided more engagement, peer support, and accountability. Implementation challenges included problems with virtual format (e.g., "Zoom fatigue"); however, several benefits were also noted (e.g., increased flexibility, maintenance of DPP connections during campus closures). UC DPP *coordinators* ($n = 18$) perceived virtual delivery as positively impacting UC DPP reach, since virtual platforms provided access for some who could not participate in-person, and negatively impacting effectiveness due to reduced engagement and lower peer support. UC *leaders* ($n = 8$) perceived that use of the virtual format had a positive impact on reach (e.g., increased availability, accessibility) and negatively impacted effectiveness (e.g., less intensive interactions on a virtual platform). Across constituent levels, the use of a virtual platform had little to no impact on perceptions of adoption and maintenance of UC DPP.

Conclusion: Perceptions of the reach, effectiveness, and implementation of UC DPP using a virtual platform varied across constituents, although all groups noted a potential negative impact on overall program effectiveness.

Unanticipated program adaptations, including virtual delivery, present potential benefits as well as perceived drawbacks, primarily across the effectiveness domain. Understanding differential constituent perceptions of the impact of virtual delivery can help maximize RE-AIM and inform future UC DPP delivery strategies.

KEYWORDS

Diabetes Prevention Program, University of California, virtual delivery, multilevel constituents, RE-AIM

Background

Prediabetes affects 38% of U.S. adults and increases risk of incident type 2 diabetes, a leading cause of morbidity, mortality, and healthcare costs in the U.S. (1, 2). Approximately 1 in 3 U.S. adults had prediabetes in 2019, and without intervention, a significant number are projected to develop incident type 2 diabetes within 5 years (2). The Diabetes Prevention Program (DPP) is a year-long intensive lifestyle intervention which has demonstrated efficacy to lower type 2 diabetes risk among at-risk individuals and those diagnosed with prediabetes (2, 3).

In 2018, the University of California (UC) implemented the DPP across all 10 UC campuses to augment obesity and diabetes prevention efforts, primarily aimed at employees. The UC DPP intensive lifestyle intervention adheres to and is certified by the Centers for Disease Control and National Diabetes Prevention Program (2, 4). The UC Diabetes Prevention Program (UC DPP) is offered free of charge to all UC faculty and staff at risk of developing type 2 diabetes as well as those diagnosed with prediabetes (as defined by the CDC National DPP criteria). The primary outcome of interest for the UC DPP trial was mean percent weight change at 12-month follow-up; secondary outcomes included mean percent weight change at 24-month follow up, challenges and facilitators associated with implementation, and degree of program adoption and maintenance [see (4)]. The evaluation of the UC DPP included diverse UC data sources, including electronic health record (HER) data, administrative claims, campus-based DPP cohort data, site visits, and qualitative interviews (the data source analyzed in the current study) [see (4)]. Our decision to use of qualitative interviews is consistent with the continued need for research utilizing and reporting applications of RE-AIM (5).

In 2020, the COVID-19 pandemic-related public health mandates necessitated unplanned, immediate changes to program delivery, including transitioning from in-person to virtual delivery. All 10 campuses shifted to virtual delivery using the UC Zoom platform due to mandated campus closures. This abrupt shift in program delivery was necessary to continue to offer UC DPP to participants at risk for developing type 2 diabetes and mitigate progression to diabetes among those diagnosed with prediabetes. Research suggests that approximately half of Americans gained weight during the pandemic; this risk was more pronounced among those who reported being overweight before the pandemic (6). Although UC DPP continued to be offered utilizing a virtual platform, there is a lack information about how this change in delivery differentially impacted the perceptions and experiences of UC DPP participants, coordinators, and leaders across RE-AIM domains.

Research comparing in-person to virtual DPP delivery suggests that intensive, multifaceted online DPP programs may be as effective as in-person DPP (7). Offering DPP online can also expand reach to at-risk individuals, although barriers (e.g., lack of internet access, technology, slow internet speed, lack of quiet space) have been noted (8). However, despite the CDC Diabetes Prevention Recognition Program Registry's recognition of online delivery of DPP (9), there are gaps in our knowledge about perceptions of the virtual delivery of preventative health care programs. Accordingly, the purpose of this study is to assess multilevel perceptions of delivering the UC DPP Initiative virtually due to the COVID-19 pandemic utilizing the RE-AIM (reach, effectiveness, adoption, implementation, and maintenance) framework (10, 11).

Methods

A planned component of the overall evaluation of the UC DPP was the use of both quantitative and qualitative data to maximize the use of the RE-AIM model (5) In-depth qualitative interviews, an ideal method to better understand multilevel stakeholder perceptions, experiences, and opinions with respect to the UC DPP, were conducted (4, 12). To accomplish this goal, the RE-AIM framework guided the development of a semi-structured interview guide (5), with versions tailored to DPP participants, coordinators, and leaders. Participants were individuals that participated in UC DPP sessions. Coordinators helped to support the UC DPP at one site, and Leaders were those that provided support across sites; many were affiliated with the University of California, Office of the President.

We sent study invitation emails and letters to UC key constituents between February and July 2021. In-depth qualitative interviews were scheduled with interested constituents. Interviews were conducted by a trained qualitative team member over UC Zoom and lasted approximately 1 h. Participants included those that engaged in UC DPP entirely in-person (pre-pandemic), virtually (during the pandemic), and a combination of in-person and virtually (those that transitioned to virtual from in-person at the beginning of the pandemic, when stay-at-home mandates were imposed).

Three interview guides informed by the RE-AIM framework were developed for the following partner groups: (1) participants, (2)

Abbreviations: UC DPP, University of California Diabetes Prevention Program; RE-AIM, Reach, effectiveness, adoption, implementation, and maintenance; UC, University of California.

coordinators, and (3) leaders. Participant interview guides assessed RE-AIM domains, including Reach (i.e., whether they had received a diagnosis of prediabetes, number of sessions attended), Effectiveness (i.e., what parts of the program worked and did not work for them, their overall satisfaction, suggestions for improvement, whether they participated in DPP on Zoom because of COVID-19 stay-at-home restrictions, and if so, how they would rate the virtual experience and how it could compare to an in-person option), Adoption (i.e., whether UC DPP affected their health and wellness, level of physical activity, eating habits, stress level, and emotional and social health, whether the sessions helped them meet the goals they set for themselves, and how the program helped with accountability), Implementation (i.e., specific program components, including the enrollment process, materials, their group coach, group interactions, and timing and frequency of sessions, as well as facilitators and challenges to participation), and Maintenance (i.e., whether they continued any lifestyle changes they made in the program and whether the program should be offered in-person, virtually, or both). Participants also responded to a question about how COVID-19 impacted participation in UC DPP.

Coordinator and leader guides were similar; coordinators were asked about their specific campus and leaders responded to questions across the UC system. Consistent with the participant guide, coordinator and leader guides also assessed RE-AIM domains, including Reach (i.e., strategies used to raise awareness of UC DPP and outreach/recruitment strategies), Effectiveness (i.e., how they evaluate UC DPP effectiveness, strengths, areas in need of improvement, feedback collected, and specific program components, including materials, training sessions, coaches, and data collection), Adoption (i.e., facilitators and barriers), Implementation (i.e., adaptations made to the program to meet campus or participant needs), and Maintenance (i.e., program needs and obstacles to sustainment). Coordinators and leaders also responded to several questions about the COVID-19 pandemic (i.e., how COVID-19 impacted UC DPP) (see Table 1 for a summary of constituent interview guides).

Qualitative data collection

Interviews were recorded, professionally transcribed, and transcripts were reviewed in detail by the research team. After familiarization with the transcripts, the team used rapid qualitative analysis, a type of manifest content analysis developed for and utilized in health services and implementation research, for example to aid in the rapid identification or expansion of knowledge of intervention components as well as facilitators and barriers of a program, to analyze the data (13, 14). Constituent (i.e., participant, coordinator, and leader) responses that alluded to delivery and use of the virtual platform with respect to RE-AIM domains were reviewed and synthesized. A templated summary of each transcript was created, creating a multilevel inventory of constituent responses to each of the respective interview guide domains. These summaries were combined into matrices to identify and compare themes, as well as to establish thematic saturation [i.e., sufficient, cross-cutting evidence for the multilevel themes presented below; (15)] related to the use of the virtual platform (13). The study was approved by the UCLA Institutional Review Board. All constituents provided verbal consent and were offered a $50 gift card incentive after the interview was completed.

Results

Between April and August 2021, 68 constituents (42 UC DPP participants, 18 coordinators, and 8 leaders) completed interviews. The UC DPP participants' mean age was 46 years (9.8); 33 (79%) were female and 9 (21%) were male. Thirteen (31%) identified as Asian, 8 (19%) Caucasian, 12 (29%) Latino, 1 (2%) Black, 2 (5%) American Indian/Alaskan, or Native Hawaiian/Pacific Islander. Six participants (14%) did not report their racial/ethnic background. Three participants (7%) reported receiving some college education, 27 (64%) had a college degree, and 12 (29%) had an advanced degree. UC DPP coordinators and leaders did not report demographic characteristics for this study.

UC DPP participants

UC DPP participants perceived virtual delivery as having the greatest impact on effectiveness and implementation domains. Few participants commented on reach, but several noted that other group members had dropped out because of the shift to a virtual platform, and one said they were not provided with a virtual option. The majority of participants perceived program *effectiveness* to be negatively impacted by virtual delivery, given their preference for in-person sessions, which provided more engagement and accountability,

> "…I feel like [the virtual option] is not as engaging, if that makes sense. I feel like people are there, but they're not really…there. Like you talk but most of the time, everybody kept their cameras off and sometimes it was just like, anyone there, have any questions or suggestions, you know?"

Another participant echoed this sentiment, stating,

> "I think what was missing…what's missing from the virtual, I think, is that like I said, that in-person camaraderie, the motivation, like I say, weighing in together, just physically being in the room with somebody, you know, it's expressive. There is a lot of thought and emotion tied to this subject. So, I think just that compassion for just having that in- person experience. Over Zoom you can't really gauge someone's…I don't know, facial expressions are everything. I think it just enhances the experience."

Others noted, "So over remote, it was a lot of quiet and everyone mostly—a lot of the time, we all had our cameras off versus being able to see a person," and "The Zoom was just – you are distracted more easily, you know?" Some described reduced interaction with group members and the facilitator, and one stated that it was more difficult to get feedback. Participants also described in-person delivery as easier, presenting fewer obstacles, and included seeing and interacting with others. According to one participant,

TABLE 1 UC DPP participant, coordinator, and leader interview guides.

RE-AIM domain	Participant interview guide	Coordinator and leader interview guide
Reach	Responded to questions about whether they had prediabetes, and if so, who told them, and how the diagnosis made them feel and changed how they thought about their health. They also reported the number of sessions attended.	Asked about strategies used to raise awareness of UC DPP, outreach/recruitment strategies, whether participants reflect the campus population, efforts to ensure a diverse representation of at-risk individuals, how to increase reach, and thoughts about why some faculty and staff participate or decline participation in UC DPP.
Effectiveness	Described what parts of the program worked and did not work for them, their overall satisfaction, suggestions for improvement, whether they participated in virtual DPP on Zoom due to COVID-19 restrictions, and if so, how they would rate the virtual experience and how it would compare to an in-person option.	Asked questions about how they evaluate UC DPP, as well as the effectiveness, strengths, areas in need of improvement, feedback collected, and specific program components (i.e., materials, training sessions, coaches, and data collection).
Adoption	Asked about whether UC DPP affected their health and wellness, level of physical activity, eating habits, stress level, and emotional and social health. Participants recounted the goals they set for themselves, whether it became easier to meet their goals as they attended more sessions, whether the sessions helped them meet their goals, if the DPP recommended goals were attainable, and how the program helped with accountability.	Asked about facilitators and barriers to UC DPP adoption and ways the UC community and participants benefitted from the program.
Implementation	Asked about specific program components, including the enrollment process, materials, their group coach, sessions, group interactions, and timing and frequency of sessions. They also were asked what made it easier to participate, challenges to participation, and whether they had to make changes to participate in the sessions.	Implementation questions focused on adaptations made to the program to meet campus or participant needs, and any unintended consequences of the program.
Maintenance	Asked if they continued any lifestyle changes, they made in the program and whether they thought the program should be offered in-person, virtually, or both.	Questions included program needs and obstacles to sustaining UC DPP.
COVID-19 pandemic	Asked if and how COVID-19 impacted participation in UC DPP.	Asked about how they balanced involvement with UC DPP with campus priorities during the pandemic; how COVID-19 impacted UC DPP.

Coordinators and Leaders answered similar questions; Coordinators were asked about their specific campus and Leaders responded across the UC system. Coordinators were asked to provide their opinion about specific program components.

> "I think once we went remote and obviously, we never went back—I mean we're still not back yet I think—I can't really say because I don't think it's fair to say what didn't work because I don't think it went as anybody planned. I think it worked as best it could in the remote environment. But I don't have anything to compare it to because what was presented in person, I believe would have been super beneficial. Like I said with the weight bands, with the portion plates, it would have been maybe more interactive. We all weighed in together, so I think that part could have really been even more motivational, inspirational, coming together in that sense."

In contrast, several participants noted that virtual was as effective as in-person programming. Participants indicated that "we actually did fine with it," "It was good, and that they did not believe it made a "tremendous difference in the nature of the program." Another participant noted,

> "I think it's a great program, to be honest. I'm excited—if they were to offer it and if I could potentially participate again and if it works with my schedule, I would definitely do it… I think it's great. I think it lost some momentum with the pandemic, but yeah, I really enjoyed the program for the most part."

Another stated,

> "It was fine. We didn't have connectivity issues. We went the whole time. I didn't feel like it was anyhow shortened or less informative or we were missing something because we really still kept talking. I think the group had a lot of commitment to finish this program for the whole year and go through it and continue learning as much, because there was always—there was never a session that did not have a question, a suggestion, or sharing a tip. You could count on that every meeting, so it showed that the other participants were equally invested and involved. It wasn't like it was just one person. Because you know sometimes when you're in Zoom, it's either just one person or it's dead silent or it's hard. That did not happen in these meetings."

With respect to *adoption*, participants noted that DPP was no longer a priority or took a back seat to other pandemic-related concerns, and that they did not own and had to purchase a scale for home use to continue the program. One participant described the shift to virtual having no impact on adoption. With regard to *implementation,* participants' comments focused on implementation challenges, including problems with virtual format (e.g., "Zoom fatigue" or having to choose between taking a lunch break or

participating in the Zoom session). One participant noted, "It was hard to do the virtual after kind of 10 h of non-stop virtual for work;" another stated, I "really did not like the Zoom meeting because my full day was Zoom meetings and I really just kind of let that go." Others described scheduling challenges, being restricted to exercising at home, competing demands (i.e., children at home), technical issues or having to learn how to use Zoom, having to weigh oneself, and having to be honest about lifestyle self-management from home. One participant stated,

> "COVID just shut everything down. We went remote. We didn't receive the materials we were supposed to receive. We didn't receive—we were supposed to get portion plates and I think something else. And you know, it was out of everybody's control, but we still kept meeting."

Others noted that the virtual format was easier or preferred: "I think, yeah, probably the Zoom did make it easier. I think if we were in person on campus…. We would have to walk there or find our way there in person. And so with the Zoom, we just log in and there we are." Another stated, "It (virtual delivery) was good. Yeah. Yeah, it was good. I mean, it worked. It worked out." Some participants noted that the virtual platform helped them to stay connected with others during campus closures. While participants were asked questions about maintenance, they did not describe the impact of virtual delivery on this dimension (see Table 2).

UC DPP coordinators

UC DPP coordinators described the virtual delivery having the greatest impact on the reach and effectiveness domains. Some noted that the virtual platform increased *reach* among those who could not participate in-person: "Retention has been much better in the virtual world." Others stated that participants were lost due to the virtual transition, and that it had no or a similar effect. They also described virtual platforms negatively impacting *effectiveness* due to Zoom being awkward or less engaging, loss of momentum, decreased peer support, and reduced accountability. Similar to participants, coordinators also described DPP as less of a priority in the context of the pandemic, negatively affecting *adoption*. With respect to *implementation*, coordinators discussed negative impacts, including Zoom fatigue, feeling limited in what the program could provide, and difficulties collecting participant data. One noted,

> "Some campuses—again in just this COVID world, they are having a harder time reaching people. So that's been the biggest struggles. And then engagement in this virtual world, in the beginning it was more novel and exciting. They're like, 'Oh, I can still see you and it's Zoom' and…we can still do this. And I think people still like it but they are kind of over it, as well…I think it does work really well for some people, but again, it's finding what works for everyone. Having an option for an in-person and a virtual, depending on what meets the need."

While coordinators were asked questions pertaining to *maintenance*, they did not describe any effects of virtual delivery on this dimension.

UC DPP leaders

Leaders described several benefits of virtual delivery on *reach*, including increased availability and accessibility. One leader noted,

> "I think the teams have been phenomenal in transitioning from in-person to remote meetings… Funding has been especially challenging in the COVID environment, and I also think that the program loses visibility when folks aren't on campus. They're not talking to each—they're not having hallway conversations with others. I do think it's presented challenges, but I think the team has adjusted phenomenally."

Drawbacks were noted for *effectiveness*, including reduced interaction among UC DPP participants on the virtual platform, less visibility, and interruptions (i.e., starting and stopping). One leader stated,

> "A lot of the value of this program is the people getting to know each other and providing support for each other that are participating in the program every week. And you can continue to do that on the Zoom format, but you're not going to create the same kind of personal bonds that you would in person where there's a lot of chatter before and after meetings and stuff like that. I don't think it's the same…. But it's still better than not having the program at all."

With respect to *adoption*, one leader noted that there was a decreased effort to be visible. Leaders did not comment on *implementation*. One leader expressed *maintenance* concerns related to lack of secure funding.

Discussion

This study identified multilevel perceptions of the virtual delivery of a CDC-recognized lifestyle behavior change program, the UC DPP, across all 10 UC campuses. In 2018, the UC system, the third largest employer in California, prioritized diabetes prevention as a system-wide goal, offering a worksite behavior change program, the UC DPP, free of charge to all UC employees with documented prediabetes or who are at risk of developing type 2 diabetes. UC DPP groups are led by UC staff who have completed DPP coach training and are experienced in delivering campus-based wellness programs (4).

Despite significant documented increases in telemedicine (e.g., the provision of clinical services) and telehealth (e.g., health-related services, including administration and continuing medical education) during the pandemic, as well as research focused on patient satisfaction with these services (16), far less is known about multilevel constituents' perceptions surrounding the shift of lifestyle behavior change programs to virtual delivery. Our study found that perceptions of virtual delivery on RE-AIM domains of UC DPP varied across constituent groups, with most reporting a negative impact of virtual delivery on program effectiveness. This study provides evidence that unanticipated program adaptations, including shifting to virtual delivery, present potential benefits as well as perceived drawbacks across RE-AIM domains.

10.3389/fpubh.2024.1327429

TABLE 2 Multilevel constituent perceptions of use of a virtual platform to deliver UC DPP across RE-AIM domains.

RE-AIM Domains	Perceptions (+, –)	Participants (*n* = 42)	Coordinators (*n* = 18)	Leaders (*n* = 8)
Reach	+		Increased retention	Increased • Availability • Accessibility
	–	Increased attrition	Increased attrition	
Effectiveness	+			
	–	Reduced: • Interactions with group members and facilitator • Accountability • Visibility • Feedback Increased: • Distractions • Obstacles	Reduced: • Momentum • Peer support • Accountability • Engagement Increased: • Awkwardness	Reduced: • Interaction among participants • Visibility Increased: • Interruptions
Adoption	+			
	–	• Competing demands during pandemic • Had to purchase scale	• Competing demands during pandemic	Reduced participants' effort to be visible
Implementation	+	• Easier than in-person • Helped to stay connected with others		
	–	Problems with: • Zoom fatigue • Scheduling challenges • Competing demands • Technical issues • Low digital literacy • Accountability	Problems with: • Zoom fatigue • Feeling limited in what UC DPP could provide • Difficulties collecting participant data	
Maintenance	+			
	–			• Concerns about lack of long-term funding

UC DPP participants reported negative effects of virtual delivery across reach, effectiveness, adoption, and implementation domains, with some indicating that virtual delivery had effectiveness and implementation benefits. Future research should focus on facilitating program effectiveness, participant engagement, accountability, interaction, and providing feedback using virtual DPP. The impact of virtual delivery due to COVID-19 on maintenance was limited in this study as it was conducted mid-pandemic; future research should focus on understanding the effects of virtual UC DPP delivery on the maintenance dimension. Barriers described by participants included Zoom fatigue, scheduling, and technical challenges, and competing demands. For participants who face these challenges, in-person DPP delivery may be preferable. Reducing these barriers should increase perceived effectiveness and implementation among other participants.

UC DPP coordinators discussed negative consequences of virtual delivery across reach, effectiveness, adoption, and implementation domains. There were equal numbers of remarks about positive (or neutral) and negative effects of virtual delivery on reach. Coordinators did not perceive other positive benefits to virtual delivery. While UC DPP leaders also discussed the drawbacks of virtual delivery on effectiveness and adoption, they described positive impacts on reach. Research designed to leverage the benefits of UC DPP delivery using virtual platforms and mitigate barriers from the participant, coordinator, and leader perspectives is needed. Understanding the differential impact of these pandemic-related changes can help maximize RE-AIM and inform future strategies for UC DPP delivery.

Limitations of the current study include the inability of each constituent group to plan and prepare for the shift to virtual delivery and to fully anticipate barriers and facilitators to engagement, due to the sudden onset of COVID-19 and accompanying stay at home mandates. Although the abrupt shift to virtual delivery allowed for continuity of UC DPP programming, our understanding of the extent to which socioeconomic factors, lack of technology, and/or low digital literacy affected participants' ability to engage in the program is limited to the remarks provided by participants in these

interviews. These challenges included "Zoom fatigue," scheduling difficulties, competing demands (including having children at home), an inability to exercise outside of one's residence, and low digital literacy. Future research should examine the differential impact of these and other contextual factors on participants' ability to engage in virtual lifestyle change programs. This study sample is comprised entirely of the recollections of UC faculty and staff; the extent to which their perceptions are generalizable to constituents from other institutions of higher education awaits future investigation.

Conclusion

The UC system prioritized diabetes prevention as a system-wide goal, offering UC DPP free of charge to all UC employees at risk for or diagnosed with prediabetes, levering campus wellness resources in diabetes prevention, and shifting to virtual delivery during the COVID-19 pandemic to maintain program continuity. This study examined perceptions of utilizing a virtual platform (UC Zoom) to deliver UC DPP on RE-AIM domains. Perceptions varied across constituent groups, with most describing a negative impact of virtual delivery on program effectiveness. There is a need to develop questions to assess preferences for and potential barriers to virtual delivery, include them in the data routinely collected for the CDC, and refine strategies for UC DPP implementation accordingly. Given that remote and/or hybrid DPP delivery is likely to continue, identifying and addressing the challenges and opportunities of the virtual delivery of UC DPP across the RE-AIM domains is critical for ongoing diabetes prevention programming efforts.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by University of California IRB# 20-000357. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

TL: Writing – review & editing, Writing – original draft, Formal analysis, Data curation. MG: Writing – review & editing. KR: Writing – review & editing, Formal analysis. KS: Writing – review & editing, Formal analysis. SS: Writing – review & editing, Formal analysis. NJ: Writing – review & editing. UC: Writing – review & editing. OD: Writing – review & editing. CM: Writing – review & editing. AH: Writing – review & editing, Formal analysis. TM: Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was funded by NIH/NIDDK (1R01DK124503). TM also receives support from the NIH/NIDDK (R01DK127733 and R18DK122372), NIDDK/CDC (U18DP006535), CDC (U18 DP006708), UCLA, UCOP, and the Department of Veterans Affairs (QUE20-028, CSP#2002 and CSP NODES0017). The funding source had no role in the design, conduct, or reporting of the study or the decision to submit the manuscript for publication.

Acknowledgments

We would like to thank UC DPP participants and constituents who made this study possible.

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. American Diabetes Association. Economic costs of diabetes in the U.S. in 2017. *Diabetes Care*. (2018) 41:917–28. doi: 10.2337/dci18-0007

2. Centers for Disease Control and Prevention (2022). Research behind the national DPP. Available at: https://www.cdc.gov/diabetes/prevention/research-behind-ndpp.htm (Accessed October 19, 2022).

3. Diabetes Prevention Program Research Group. 10-year follow-up of diabetes incidence and weight loss in the diabetes prevention program outcomes study. *Lancet*. (2009) 374:1677–86. doi: 10.1016/S0140-6736(09)61457-4

4. Gholami M, Jackson NJ, Chung UYR, Duru OK, Shedd K, Soetenga S, et al. Evaluation of the University of California Diabetes Prevention Program (UC DPP) initiative. *BMC Public Health*. (2021) 21:1775. doi: 10.1186/s12889-021-11731-7

5. Holtrop JS, Rabin BA, Glasgow RE. Qualitative approached to use of the RE-AIM framework: rationale and methods. *BMC Health Serv Res*. (2018) 18:177. doi: 10.1186/s12913-018-2938-8

6. Khubchandani J, Price JH, Sharma S, Wiblishauser MJ, Webb FJ. COVID-19 pandemic and weight gain in American adults: a nationwide population-based study. *Diabetes Metab Syndr Clin Res Rev*. (2022) 16:102392. doi: 10.1016/j.dsx.2022.102392

7. Moin T, Damschroder LJ, AuYoung M, Macieiewski ML, Havens K, Erl K, et al. Results from a trial of an online diabetes prevention program intervention. *Am J Prev Med*. (2018) 55:583–91. doi: 10.1016/j.amepre.2018.06.028

8. van den Berg AC, Giest SN, Groeneveld SM, Kraaij W. Inclusivity in online platforms: recruitment strategies for improving participation of diverse

sociodemographic groups. *Public Adm Rev*. (2020) 80:989–1000. doi: 10.1111/puar.13215

9. CDC. Centers for disease control and prevention diabetes prevention recognition program standards and operating procedures. (2018). Available at: https://www.cdc.gov/diabetes/prevention/recognition

10. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health*. (1999) 89:1322–7. doi: 10.2105/ajph.89.9.1322

11. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. *Front Public Health*. (2019) 7:64. doi: 10.3389/fpubh.2019.00064

12. Guest G, Namey E, O'Regan A, Godwin C, Taylor J. *Comparing interview and focus group data collected in person and online*. Washington (DC): Patient-Centered Outcomes Research Institute (PCORI) (2020) Available at: https://www.ncbi.nlm.nih.gov/books/NBK588708/.

13. Hamilton AB, Finley EP. Qualitative methods in implementation research: an introduction. *Psychiatry Res*. (2019) 280:112516. doi: 10.1016/j.psychres.2019.112516

14. Lewinski AA, Crowley MJ, Miller C, Bosworth HB, Jackson GL, Steinhauser K, et al. Applied rapid qualitative analysis to develop a contextually appropriate intervention and increase the likelihood of uptake. *Med Care*. (2021) 59:S242–51. doi: 10.1097/MLR.0000000000001553

15. Morse JM. Data were saturated…. *Qual Health Res*. (2015) 25:587–8. doi: 10.1177/1049732315576699

16. Shaver J. The state of telehealth before and after the COVID-19 pandemic. *Prim Care*. (2022) 49:517–30. doi: 10.1016/j.pop.2022.04.002

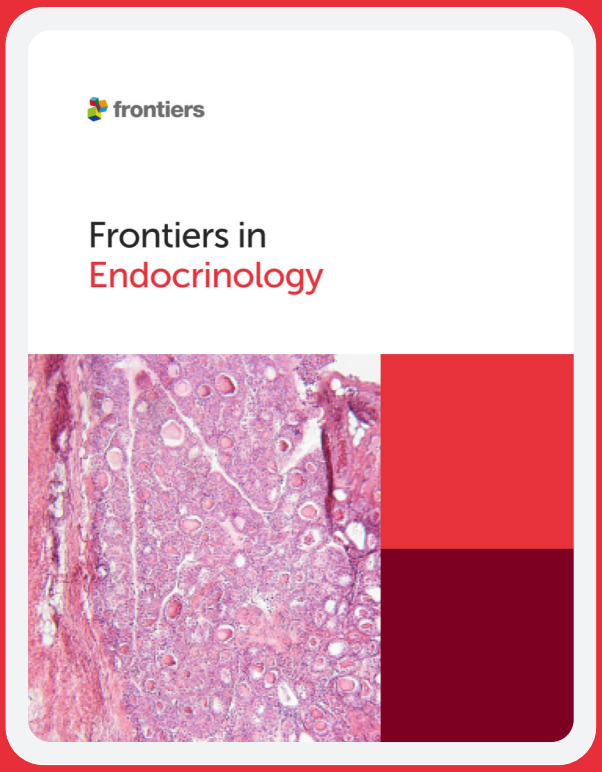
Frontiers in Endocrinology

Explores the endocrine system to find new therapies for key health issues

The second most-cited endocrinology and metabolism journal, which advances our understanding of the endocrine system. It uncovers new therapies for prevalent health issues such as obesity, diabetes, reproduction, and aging.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact