



OPEN ACCESS

EDITED BY

Rubén González-Rodríguez,
University of Vigo, Spain

REVIEWED BY

Pablo Fondo-Ferreiro,
University of Vigo, Spain
Pranavsinh Dhunoo,
Atlantic Technological University, Ireland

*CORRESPONDENCE

Ravi Shankar

✉ Ravi_SHANKAR@nuhs.edu.sg

RECEIVED 30 December 2024

ACCEPTED 17 February 2025

PUBLISHED 28 February 2025

CITATION

Shankar R, Bunde A, Yap A and
Mukhopadhyay A (2025) Development
and feasibility testing of an AI-powered
chatbot for early detection of caregiver
burden: protocol for a mixed
methods feasibility study.
Front. Psychiatry 16:1553494.
doi: 10.3389/fpsy.2025.1553494

COPYRIGHT

© 2025 Shankar, Bunde A, Yap A and
Mukhopadhyay. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Development and feasibility testing of an AI-powered chatbot for early detection of caregiver burden: protocol for a mixed methods feasibility study

Ravi Shankar^{1*}, Anjali Bunde¹, Amanda Yap¹
and Amartya Mukhopadhyay²

¹Research and Innovation, Medical Affairs, Alexandra Hospital, Singapore, Singapore, ²Division of Respiratory and Critical Care Medicine, Department of Medicine, National University Hospital, Singapore, Singapore

Introduction: Caregivers of patients with end-stage kidney disease (ESKD) face significant challenges that contribute to caregiver burden, negatively impacting their physical, psychological, social, and financial well-being. With the growing prevalence of chronic diseases and an aging population, there is an urgent need for accessible and scalable solutions to detect and address caregiver burden. Artificial Intelligence (AI) chatbots using natural language processing (NLP) have shown promise in providing mental health support and monitoring through natural conversations. This study will contribute to research and clinical practice by: (1) validating a novel approach for early detection of caregiver burden through NLP, (2) analyzing the feasibility of AI-powered chatbots for continuous caregiver monitoring, and (3) informing the development of scalable, accessible tools to identify at-risk caregivers.

Methods and analysis: This protocol for the mixed methods aims to evaluate the feasibility, acceptability, and preliminary effectiveness of BOTANIC (Burden Observation and Timely Aid for Navigating Informal Caregiving), an AI-powered chatbot for early detection of caregiver burden. A single-center validation study will be conducted at Alexandra Hospital, Singapore. Twenty primary caregivers of ESKD patients will be recruited to use BOTANIC for 12 weeks. BOTANIC, developed using Python and open-source libraries, will integrate with Telegram and utilize advanced NLP techniques to analyze caregiver conversations and detect signs of burden. The NLP algorithm will analyze conversations to generate burden scores at baseline and at 12 weeks. Participants will also complete baseline and 12-week assessments using validated questionnaires including the Zarit Burden Interview (ZBI), Patient Health Questionnaire-9 (PHQ-9), and Generalized Anxiety Disorder-7 (GAD-7). Primary outcomes include concordance between caregiver burden levels detected by the NLP algorithm and validated assessment scores at both timepoints. Secondary outcomes include user engagement metrics and system satisfaction. Semi-structured interviews will explore participants' experiences with the chatbot. Quantitative data will be analyzed using descriptive statistics and appropriate statistical tests such as paired t-tests or Wilcoxon signed-rank tests, while qualitative data will undergo thematic analysis.

Ethics and dissemination: The study has been approved by the NHG Domain Specific Review Board. Findings will be published in peer-reviewed journals, presented at conferences, and used to inform the development of larger-scale trials of AI-powered caregiver support interventions.

KEYWORDS

caregiver burden, artificial intelligence, chatbot, natural language processing, chronic conditions, feasibility study

Background

Caregiver burden is defined as the multidimensional response to the physical, psychological, social, and financial stressors associated with caregiving (1, 2). With the increasing prevalence of chronic diseases and an aging population, the demand for informal caregiving by caregivers such as family members is on the rise (3, 4). Caregivers play a crucial role in supporting individuals with various chronic conditions, such as dementia, cancer, stroke, and end-stage kidney disease, across different care settings, including home, community, and healthcare facilities (5).

Despite the vital contributions of informal caregivers to the healthcare system and society, caregiving often comes at a significant personal cost. Caregivers are at increased risk of physical health problems, such as cardiovascular disease, immune dysfunction, and premature mortality (6). They also experience high rates of psychological distress, including depression, anxiety, and burnout (7). Moreover, caregiving can lead to social isolation, financial strain, and reduced quality of life (8). These negative consequences, collectively known as caregiver burden, can have far-reaching impacts on caregivers, care recipients, and the healthcare system as a whole.

Current approaches to assessing and addressing caregiver burden have several limitations. Traditional assessment methods, such as the Zarit Burden Interview (ZBI) (9) and the Caregiver Strain Index (CSI) (10), rely on infrequent self-report measures that may not capture the dynamic nature and contextual factors influencing caregiver burden. These assessments are often administered in clinical settings, which can be time-consuming, resource-intensive, and may not reflect caregivers' experiences in real-world settings. Moreover, the lack of regular monitoring and timely support can lead to the escalation of caregiver burden over time, resulting in adverse outcomes such as caregiver burnout, poor mental health, and decreased quality of care (11).

To address these challenges, there is a growing interest in leveraging digital health technologies, particularly artificial

intelligence (AI) chatbots, for caregiver support. AI chatbots are computer programs that simulate human conversation using natural language processing (NLP) techniques (12). NLP enables computers to understand, interpret, and generate human language, allowing for more natural and engaging interactions between users and chatbots (13). By analyzing the linguistic and semantic features of caregivers' responses during chatbot conversations, NLP algorithms can potentially detect signs of caregiver burden, provide personalized support, and facilitate early intervention (14).

Recent studies have demonstrated the significant potential of AI chatbots for mental health support, screening, and behavior change interventions (15–22). Multiple chatbots using different therapeutic approaches have shown promising results. For example, Woebot, a conversational agent based on cognitive-behavioral therapy principles, reduced symptoms of depression and anxiety among college students and demonstrated high engagement rates (15). Wysa, an AI-powered mental health chatbot using a combination of CBT, dialectical behavior therapy, and motivational interviewing techniques, showed improvements in depression, anxiety, and sleep outcomes (16). Similarly, Ai-driven Tess provided on-demand support and coping strategies for individuals with mental health concerns through empathetic responses and demonstrated good feasibility and acceptability (17). Anna, an AI chatbot by Happify Health, successfully increased user engagement in digital mental health interventions through personalized interactions (18). DigiQuit showed success in smoking cessation through personalized behavioral support (19), while CASC demonstrated efficacy in medication adherence monitoring (20).

Studies have shown chatbots can successfully increase adherence to healthy behaviors through theory-based approaches including goal setting, monitoring, real-time reinforcement/feedback, and on-demand support (21). Results demonstrate particular efficacy in promoting healthy lifestyles (40% of studies), smoking cessation (27%), treatment/medication adherence (13%), and reduction in substance misuse (7%) (22).

While AI chatbots have been widely studied for mental health support and monitoring, their specific application to caregiver burden detection and support remains largely unexplored. Previous technological approaches to addressing caregiver burden have primarily focused on digital support groups, educational resources, and care coordination tools. To our knowledge, this is the first study to evaluate an AI-powered chatbot specifically

Abbreviations: AI, Artificial Intelligence; BOTANIC, Burden Observation and Timely Aid for Navigating Informal Caregiving; GAD-7, Generalized Anxiety Disorder-7; NLP, Natural Language Processing; PHQ-9, Patient Health Questionnaire-9; DSRB, Domain Specific Review Board; IRB, Institutional Review Board; ZBI, Zarit Burden Interview.

designed for early detection and monitoring of caregiver burden through natural language conversations.

In the context of caregiving, AI chatbots offer several potential advantages over traditional support methods. First, chatbots can provide 24/7 availability and accessibility, allowing caregivers to seek support at their convenience and in the comfort of their own homes (21, 23). Second, the anonymity and non-judgmental nature of chatbots may create a safe space for users to disclose sensitive information and experiences more openly, overcoming the stigma and barriers associated with seeking help (24, 25). Third, chatbots can deliver personalized and tailored support based on caregivers' individual needs, preferences, and context, using machine learning algorithms to adapt and optimize the conversation flow (26). Fourth, their integration into commonly used platforms like smartphones and messaging apps enables them to reach large and diverse populations (27).

Building on this foundation, we propose to develop and evaluate BOTANIC (Burden Observation and Timely Aid for Navigating Informal Caregiving), an AI chatbot system designed to detect caregiver burden among informal caregivers of patients with various chronic conditions. BOTANIC will leverage advanced NLP techniques and machine learning algorithms to analyze caregivers' responses during chatbot conversations and identify linguistic and behavioral markers of burden. By enabling early detection and continuous monitoring through natural language conversations, BOTANIC aims to help healthcare providers better understand and track caregiver burden patterns over time, potentially informing more timely and appropriate support interventions.

The specific aims of the feasibility study will be:

1. To develop BOTANIC, a Python-based AI chatbot system that uses NLP and machine learning to detect signs of caregiver burden through conversations with caregivers.
2. To assess the preliminary effectiveness of BOTANIC in detecting caregiver burden, as measured by the NLP algorithm and validated caregiver-reported outcome measures, and to explore its impact on secondary outcomes, including user engagement, system satisfaction, and caregiver well-being (e.g., depression, anxiety, quality of life).
3. To evaluate the feasibility and acceptability of BOTANIC among informal caregivers of patients with various chronic conditions at a tertiary public hospital through a validation study design.
4. To gather qualitative feedback from caregivers on their experiences with BOTANIC, perceived benefits and challenges, and suggestions for future enhancements.

We hypothesize that BOTANIC will show promising results in detecting caregiver burden, with significant correlation between NLP-detected burden levels and caregiver-reported outcomes at baseline and follow-up. We also expect BOTANIC to demonstrate feasibility and acceptability among informal caregivers, as indicated by high user engagement, system satisfaction, and positive qualitative feedback.

This paper presents the protocol for developing and evaluating BOTANIC, detailing the planned methodology for both the technical development and feasibility testing phases. No results are presented as this is a protocol paper describing the proposed study design and methods.

Methods

Study design

This study will employ a concurrent validation design to evaluate BOTANIC's burden detection capabilities among informal caregivers of patients with various chronic conditions, using a mixed-methods approach at a tertiary public hospital in Singapore. The study will consist of two main phases: (1) Development of BOTANIC Chatbot, and (2) feasibility testing, using quantitative surveys and qualitative interviews.

Phase 1: development of BOTANIC

Foundational model selection and fine-tuning

For the development of BOTANIC, we will leverage a state-of-the-art foundational large language model (OpenAI's GPT-4), to analyze caregiver responses and detect signs of caregiver burden. The selection of this foundational model is based on factors such as performance, flexibility, and cost-effectiveness for our caregiver chatbot application. Subsequently, we will fine-tune the model using a curated dataset of caregiver conversations and annotations, focusing on specific questions and topics related to caregiver burden. This dataset will be sourced from various online forums, support groups, and existing caregiver chatbot systems, and will be annotated to identify linguistic and semantic patterns associated with caregiver burden. The fine-tuning process will involve adapting the pre-trained language model to the specific domain and style of caregiver conversations, as well as optimizing its performance for the task of caregiver burden detection based on the carefully designed questions.

Telegram bot development

BOTANIC will be developed as a Telegram bot using the Python programming language and the python-telegram-bot library. This library provides a simple and intuitive interface for building chatbots on the Telegram platform, allowing BOTANIC to engage with caregivers through natural conversational interactions. Unlike traditional structured questionnaires, caregivers can express their experiences, feelings, and concerns in their own words through free-text messaging, while the system's underlying architecture processes these natural responses. The bot's questions are carefully designed based on the Zarit Burden Interview, a validated instrument for detecting caregiver burden. The questions are organized into key domains that research has shown to be critical indicators of caregiver burden:

1. Physical and emotional demands of caregiving - These questions assess the direct impact of caregiving activities

and responsibilities on the caregiver's wellbeing, which research shows is a primary indicator of burden levels.

2. Impact on personal relationships and social life - Questions examining how caregiving affects relationships with family and friends, as social isolation is a significant contributor to caregiver burden.
3. Financial strain associated with caregiving - Research demonstrates that economic pressures from caregiving responsibilities are a major source of burden, particularly for those caring for patients with chronic conditions.
4. Expectations and sense of duty - These questions evaluate feelings of obligation and perceived expectations from others, which the Zarit Burden Interview identifies as key burden indicators.
5. Personal life impact and self-care - Assessment of how caregiving affects the caregiver's ability to maintain their own life and wellbeing, which research shows directly correlates with burden levels.

This assessment framework implements the Zarit Burden Interview through natural conversation, with additional validated measures (PHQ-9, GAD-7) included as supplementary indicators to help validate burden detection. The questions are designed to gather burden indicators while maintaining a supportive dialogue that encourages honest responses. Each category's questions are introduced contextually to maintain conversation flow while ensuring systematic assessment of burden dimensions.

The NLP algorithm will employ a multi-stage approach to detect caregiver burden:

1. Text preprocessing including tokenization, lemmatization, and removal of stop words.
2. Feature extraction focusing on linguistic markers associated with burden (emotional words, stress indicators, fatigue-related terms).
3. Semantic analysis using GPT-4's embeddings to understand context and emotional content.
4. Pattern recognition across multiple conversations to identify trends in burden indicators.
5. Integration of structured assessment responses (ZBI questions) with unstructured conversation analysis.
6. Generation of a continuous burden score using a weighted combination of these features.

GPT-4 will serve two distinct functions: handling natural conversation flow and supporting burden detection through semantic analysis. The conversational aspects will be managed through GPT-4's dialogue capabilities, while a separate custom NLP pipeline will analyze the conversations for burden indicators, using GPT-4's embeddings as one component of the analysis.

Unlike traditional digital questionnaires, BOTANIC is designed to provide a more engaging and naturalistic interaction experience through conversational AI. The chatbot will offer several advantages over standard digital forms: 1) ability to ask follow-up questions based on user responses, 2) detection of subtle linguistic markers of burden that may not be captured in structured questionnaires,

3) provision of empathetic responses and supportive interactions, and 4) ongoing monitoring through regular conversations rather than point-in-time assessments. Additionally, BOTANIC will provide educational resources about caregiving, stress management techniques, and links to support services based on detected burden levels.

The bot will use the fine-tuned foundational model to analyze the caregiver's responses and detect potential indicators of caregiver burden. The conversation flow will be structured to gather relevant information and provide a safe space for caregivers to express their experiences and feelings. The bot will be designed to be empathetic and understanding while focusing on the primary goal of detecting caregiver burden. We will implement error handling and fallback mechanisms to ensure that the bot can gracefully handle unexpected user inputs or technical issues, and provide appropriate guidance and support to the caregiver throughout the conversation.

Testing and evaluation

Throughout the development process, we will conduct rigorous testing and evaluation to ensure the quality, reliability, and effectiveness of BOTANIC. This will involve unit testing and integration testing of the bot's functionality, as well as performance testing to ensure it can handle a large number of concurrent users and conversations. We will also conduct user acceptance testing (UAT) with a diverse group of caregivers to gather feedback on the bot's usability, clarity of questions, and overall user experience. This feedback will be used to identify areas for improvement and to iteratively refine the bot's performance.

Phase 2: feasibility testing

Participants

A purposive sample of 20 primary informal caregivers of patients with various chronic conditions will be recruited from the outpatient clinics, inpatient wards, and community services affiliated with Alexandra Hospital. The inclusion criteria are: (1) aged 21 years or above, (2) providing primary care for a patient with a chronic condition (e.g., dementia, cancer, stroke, heart disease, end-stage kidney disease) for at least 3 months, (3) able to communicate in English, (4) owning a smartphone with a compatible messaging app installed, and (5) willing to use BOTANIC for 12 weeks. Caregivers with severe cognitive impairment, active psychosis, or substance abuse will be excluded. A sample size of 20 is considered sufficient for feasibility testing, allowing for a diverse representation of caregivers and chronic conditions (28–30).

Procedures

Eligible caregivers will be approached by a research coordinator during their care recipients' clinic visits, hospital stays, or community events. The study purpose, procedures, risks, and benefits will be explained, and written informed consent will be obtained. Participants will complete a baseline assessment, including demographic information, the Zarit Burden Interview

(ZBI), the Patient Health Questionnaire-9 (PHQ-9), and the Generalized Anxiety Disorder-7 (GAD-7). They will then receive training on how to use BOTANIC and be encouraged to engage with the chatbot regularly throughout the 12-week study period. The NLP algorithm will analyze initial conversations to establish baseline burden scores, and continue analyzing conversations throughout the study period, with final burden scores calculated at week 12. The validated assessment questionnaires (ZBI, PHQ-9, GAD-7) will also be repeated at 12 weeks. This will allow comparison between the NLP-detected burden levels and standard assessment scores at both timepoints.

Outcomes

The primary outcome is the concordance between burden levels detected by BOTANIC's NLP algorithm and standard assessment measures (ZBI, PHQ-9, GAD-7). The algorithm will assign a continuous burden score based on the linguistic and semantic features of caregivers' responses, with higher scores indicating greater burden severity. Secondary outcomes include user engagement metrics (conversation frequency, duration, response rate), and system usability and satisfaction (System Usability Scale, User Experience Questionnaire).

Data analysis

Descriptive statistics will be used to summarize participants' demographic and clinical characteristics, engagement metrics, and usability and satisfaction ratings. Means, standard deviations, and proportions will be reported as appropriate. The feasibility and acceptability of BOTANIC will be evaluated based on predefined criteria, such as the proportion of eligible caregivers enrolled, the proportion of participants completing the study, average engagement levels, and mean usability and satisfaction scores.

The preliminary effectiveness of BOTANIC in detecting caregiver burden will be assessed using paired t-tests or Wilcoxon signed-rank tests, comparing the NLP-generated burden scores and ZBI, PHQ-9, GAD-7 scores at baseline and 12 weeks. Effect sizes (Cohen's *d*) will be calculated to determine the magnitude of change.

To evaluate the bot's effectiveness in detecting caregiver burden, we will employ both quantitative and qualitative methods. Quantitatively, we will analyze the caregiver responses to the bot's questions and compare the detected burden levels with the ZBI scores. This will help us assess the accuracy and validity of the bot's burden detection capabilities. Qualitatively, we will conduct interviews with caregivers who have interacted with the bot to gather their feedback on the relevance and clarity of the questions, the bot's ability to understand their experiences, and the overall user experience of interacting with the bot.

The caregivers who completed the feasibility testing phase will be invited to participate in semi-structured interviews. Participants will be selected to achieve maximum variation in terms of demographic characteristics, chronic conditions, caregiving experiences, engagement levels, and changes in caregiver burden scores. Recruitment will continue until data saturation is reached, i.e., no new themes or insights emerge from additional interviews.

Qualitative data collection

Individual interviews will be conducted by a trained qualitative researcher using a semi-structured interview guide. The guide will explore caregivers' motivations for using BOTANIC, perceived benefits and challenges, impact on caregiving experiences and well-being, and suggestions for improving the chatbot's content, functionality, and user experience. Interviews will be conducted in English, either in-person or via videoconferencing, based on participants' preferences and public health guidelines. Interviews will be audio-recorded, transcribed verbatim, and anonymized. Field notes will be taken to capture non-verbal cues, contextual information, and researcher reflections.

The qualitative component of this study will follow the Consolidated Criteria for Reporting Qualitative Studies (COREQ) guidelines to ensure comprehensive and transparent reporting of qualitative research findings.

Thematic analysis

The interview transcripts will be analyzed using thematic analysis, following Braun and Clarke's six-phase approach (31). The analysis will involve familiarization with the data, generating initial codes, searching for themes, reviewing and refining themes, defining and naming themes, and producing the final report. The analysis will be conducted iteratively, with constant comparison within and across cases, to ensure the coherence and distinctiveness of the themes. Coding will be performed independently by two researchers, with regular meetings to discuss and resolve discrepancies. The qualitative findings will be integrated with the quantitative results using a mixed-methods matrix (32), facilitating a comprehensive understanding of the feasibility, acceptability, and impact of BOTANIC from the caregivers' perspectives.

Figure 1 presents the study flow diagram illustrating both the technical development of BOTANIC and its clinical feasibility testing pathway.

Patient and public involvement

Informal caregivers contributed to this study's conceptualization and design. During development, two caregiver representatives assessed intervention feasibility and helped refine outcome measures. These representatives will participate in results interpretation. Study findings will be disseminated to participants through written summaries and a presentation session.

Discussion

The proposed study aims to evaluate the feasibility, acceptability, and preliminary effectiveness of BOTANIC, an AI-powered chatbot for early detection of caregiver burden. If successful, BOTANIC could provide a scalable and accessible solution to address the growing issue of caregiver burden across various chronic conditions and care settings.

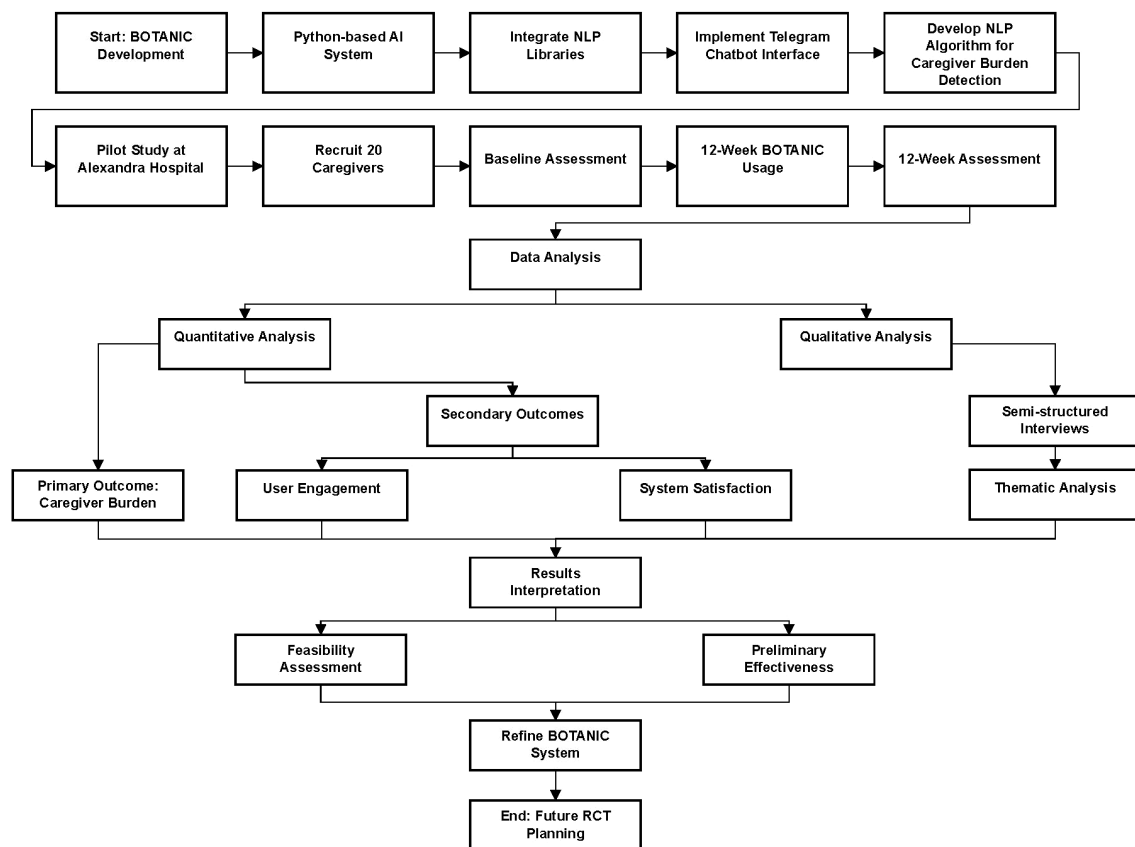


FIGURE 1
Study flow diagram showing the two main phases.

The study's strengths lie in its innovative approach, combining advanced NLP techniques and machine learning algorithms to analyze caregiver conversations and detect signs of burden. The NLP algorithm employs a multi-stage approach that includes text preprocessing, feature extraction of linguistic burden markers, semantic analysis using GPT-4's embeddings, pattern recognition across conversations, integration of structured assessments, and generation of continuous burden scores. GPT-4 serves dual functions - managing natural conversations while supporting burden detection through semantic analysis, with a separate NLP pipeline analyzing conversations for burden indicators. This approach differs from previous chatbot interventions, which have primarily focused on providing general mental health support or adherence to specific health behaviors (15–22). By tailoring the chatbot's content and functionality to the specific needs and experiences of caregivers, BOTANIC has the potential to offer more targeted and personalized support. The mixed-methods design, incorporating both quantitative and qualitative data, will provide a comprehensive understanding of the chatbot's performance, user experience, and impact on caregiver outcomes.

The study's focus on a diverse sample of caregivers in the multi-racial Singaporean context is another strength, as it will provide insights into the chatbot's acceptability and effectiveness across different cultural and linguistic backgrounds. This is particularly important given the growing recognition of the need for culturally

sensitive and inclusive caregiver interventions (33, 34). The study's findings will contribute to the limited literature on digital health interventions for caregiver support in Asian populations and inform the adaptation of BOTANIC for other cultural contexts.

However, several limitations should be acknowledged. The small sample size ($n=20$) and single-center design may limit the generalizability of the findings to wider caregiver populations and healthcare settings. The 12-week study duration, while sufficient for feasibility testing, may not capture the long-term engagement, effectiveness, and sustainability of the chatbot. Future studies should include larger, multi-center samples and longer follow-up periods to establish the robustness and scalability of BOTANIC.

The effectiveness of BOTANIC will largely depend on the quality and diversity of the training data used to develop the NLP models. Ensuring the representativeness, validity, and cultural sensitivity of the data is crucial to avoid bias and improve the chatbot's performance across different caregiver populations and contexts. Collaboration with caregivers, healthcare professionals, and community organizations will be essential to create large, diverse, and high-quality datasets that capture the complexity and nuances of caregiver experiences.

Another critical consideration is the accuracy and reliability of the algorithm. Rigorous testing, validation, and continuous monitoring of the NLP models and dialogue management system will be necessary to ensure the chatbot's performance, consistency,

and appropriateness. Establishing clear protocols for handling high-risk situations and integrating human oversight and escalation mechanisms will be essential to ensure the chatbot's safety and effectiveness.

The successful implementation of BOTANIC in real-world settings will require careful consideration of the technical, organizational, and human factors that influence its adoption, usability, and sustainability. Integration with existing healthcare systems, such as electronic health records and caregiver support services, will be necessary to ensure seamless data sharing, care coordination, and continuity of support. Engaging key stakeholders, including caregivers, healthcare professionals, administrators, and policymakers, in the co-design, evaluation, and dissemination of the chatbot will be crucial to ensure its relevance, acceptability, and alignment with their needs and priorities.

Ethical and social implications of using AI chatbots for caregiver support should also be thoroughly examined and addressed. Ensuring data privacy, security, and confidentiality is paramount to protect caregivers' sensitive information and maintain their trust in the system. Obtaining informed consent, respecting caregivers' autonomy and preferences, and providing clear information about the chatbot's capabilities and limitations are essential to ensure ethical and transparent use of the technology.

Future research should focus on larger, multi-center trials with longer follow-up periods to establish the effectiveness, cost-effectiveness, and sustainability of BOTANIC. Economic evaluations, such as cost-utility and cost-benefit analyses, will be necessary to demonstrate the value proposition and return on investment of the chatbot compared to usual care or alternative interventions.

Conclusion

This protocol describes the development and planned evaluation of an AI-powered chatbot system (BOTANIC) designed to detect and monitor caregiver burden through natural conversations. BOTANIC represents an innovative approach to address the urgent need for accessible and scalable solutions to detect and support caregivers at risk of burden. By harnessing AI and NLP technologies, BOTANIC aims to provide personalized and timely support to caregivers through natural conversations, potentially improving their well-being and the quality of care they provide.

This feasibility study is a crucial first step in evaluating the potential of BOTANIC and informing its future development and implementation. The findings will contribute to the evidence base on digital health interventions for caregiver support and guide the design of larger, definitive trials.

However, the success of BOTANIC will ultimately depend on its ability to accurately detect burden, provide safe and appropriate support, and seamlessly integrate with existing healthcare systems and caregiver support services. Engaging caregivers, healthcare professionals, and other stakeholders in the co-design and evaluation process will be essential to ensure the chatbot's relevance, acceptability, and adoption in real-world settings.

As the global burden of chronic diseases and the demand for informal caregiving continue to rise, innovative solutions like BOTANIC offer hope for a more sustainable and compassionate model of caregiver support. With further research and refinement, AI-powered chatbots have the potential to transform the landscape of caregiver interventions and contribute to better outcomes for caregivers, care recipients, and healthcare systems worldwide.

Ethics statement

The studies involving humans were approved by National Healthcare Group Domain Specific Review Board. 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

RS: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. AB: Conceptualization, Methodology, Writing – review & editing. AY: Methodology, Writing – review & editing. AM: Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study is funded by the National University of Singapore (NUS) Academic Health Programme (AHP) Fund (Reference: HPHSR-Enabling-Sep24/AHP/C01/RS). The funding body had no role in the study design, data collection, analysis, interpretation, writing of the manuscript, or decision to publish. The study protocol was developed independently by the research team.

Acknowledgments

The authors would like to thank the healthcare professionals and caregivers at Alexandra Hospital, Singapore, for their valuable input during the conceptualization of BOTANIC. Special thanks to all the caregivers who will participate in this study.

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

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.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsy.2025.1553494/full#supplementary-material>

References

- Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. *Jama*. (2014) 311:1052–60. doi: 10.1001/jama.2014.304
- Schulz R, Sherwood PR. Physical and mental health effects of family caregiving. *Am J Nurs*. (2008) 108:23–7. doi: 10.1097/01.NAJ.0000336406.45248.4c
- Lindt N, van Berkel J, Mulder BC. Determinants of overburdening among informal carers: a systematic review. *BMC Geriatrics*. (2020) 20:304. doi: 10.1186/s12877-020-01708-3
- Zhylybekova A, Koshmaganbetova GK, Zare A, Mussin NM, Kaliyev AA, Bakshshalizadeh S, et al. Global research on care-related burden and quality of life of informal caregivers for older adults: a bibliometric analysis. *Sustainability*. (2024) 16:1020. doi: 10.3390/su16031020
- Reinhard SC, Given B, Petlick NH, Bemis A. Advances in patient safety supporting family caregivers in providing care. In: Hughes RG, editor. *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*. Agency for Healthcare Research and Quality (US, Rockville (MD) (2008).
- Schulz R, Beach SR. Caregiving as a risk factor for mortality: the Caregiver Health Effects Study. *Jama*. (1999) 282:2215–9. doi: 10.1001/jama.282.23.2215
- Sallim AB, Sayampnanathan AA, Cuttulan A, Ho R. Prevalence of mental health disorders among caregivers of patients with alzheimer disease. *J Am Med Dir Assoc*. (2015) 16:1034–41. doi: 10.1016/j.jamda.2015.09.007
- Pinquart M, Sörensen S. Differences between caregivers and noncaregivers in psychological health and physical health: a meta-analysis. *Psychol Aging*. (2003) 18:250–67. doi: 10.1037/0882-7974.18.2.250
- Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist*. (1980) 20:649–55. doi: 10.1093/geront/20.6.649
- Robinson BC. Validation of a caregiver strain index. *J Gerontol*. (1983) 38:344–8. doi: 10.1093/geronj/38.3.344
- Chiao CY, Wu HS, Hsiao CY. Caregiver burden for informal caregivers of patients with dementia: A systematic review. *Int Nurs Rev*. (2015) 62:340–50. doi: 10.1111/inr.2015.62.issue-3
- Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. *J Am Med Inform Assoc*. (2018) 25:1248–58. doi: 10.1093/jamia/ocy072
- Hirschberg J, Manning CD. Advances in natural language processing. *Science*. (2015) 349:261–6. doi: 10.1126/science.aaa8685
- Kocaballi AB, Berkovsky S, Quiroz JC, Laranjo L, Tong HL, Rezazadegan D, et al. The personalization of conversational agents in health care: systematic review. *J Med Internet Res*. (2019) 21:e15360. doi: 10.2196/15360
- Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Ment Health*. (2017) 4:e19. doi: 10.2196/mental.7785
- Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. *JMIR Mhealth Uhealth*. (2018) 6:e12106. doi: 10.2196/12106
- Fulmer R, Joerin A, Gentile B, Lakerink L, Rauws M. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. *JMIR Ment Health*. (2018) 5:e64. doi: 10.2196/mental.9782
- Boucher EM, Harake NR, Ward HE, Stoeckl SE, Vargas J, Minkel J, et al. Artificially intelligent chatbots in digital mental health interventions: a review. *Expert Rev Med Devices*. (2021) 18:37–49. doi: 10.1080/17434440.2021.2013200
- Carrasco-Hernandez L, Jódar-Sánchez F, Núñez-Benjumea F, Moreno Conde J, Mesa González M, Civit-Balcells A, et al. A mobile health solution complementing psychopharmacology-supported smoking cessation: randomized controlled trial. *JMIR Mhealth Uhealth*. (2020) 8:e17530. doi: 10.2196/17530
- Chaix B, Bibault JE, Pienkowski A, Delamon G, Guillemassé A, Nectoux P, et al. When chatbots meet patients: one-year prospective study of conversations between patients with breast cancer and a chatbot. *JMIR Cancer*. (2019) 5:e12856. doi: 10.2196/12856
- van Baal ST, Le STT, Fatehi F, Verdejo-Garcia A, Hohwy J. Testing behaviour change with an artificial intelligence chatbot in a randomized controlled study. *J Public Health Policy*. (2024) 45:506–22. doi: 10.1057/s41271-024-00500-6
- Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial intelligence-based chatbots for promoting health behavioral changes: systematic review. *J Med Internet Res*. (2023) 25:e40789. doi: 10.2196/40789
- Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and conversational agents in mental health: A review of the psychiatric landscape. *Can J Psychiatry*. (2019) 64:456–64. doi: 10.1177/0706743719828977
- Khawaja Z, Bélisle-Pipon JC. Your robot therapist is not your therapist: understanding the role of AI-powered mental health chatbots. *Front Digit Health*. (2023) 5:1278186. doi: 10.3389/fdgh.2023.1278186
- Abd-Alrazaq AA, Alajlani M, Alalwan AA, Bewick BM, Gardner P, Househ M. An overview of the features of chatbots in mental health: A scoping review. *Int J Med Inform*. (2019) 132:103978. doi: 10.1016/j.ijmedinf.2019.103978
- Shum H-y, He X-d, Li D. From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Front Inf Technol Electronic Engineering*. (2018) 19:10–26. doi: 10.1631/FITEE.1700826
- Milne-Ives M, de Cock C, Lim E, Shehadeh MH, de Pennington N, Mole G, et al. The effectiveness of artificial intelligence conversational agents in health care: systematic review. *J Med Internet Res*. (2020) 22:e20346. doi: 10.2196/20346
- Julious SA. Sample size of 12 per group rule of thumb for a pilot study. *Pharm Statistics*. (2005) 4:287–91. doi: 10.1002/pst.v4:4
- Wutich A, Beresford M, Bernard HR. Sample sizes for 10 types of qualitative data analysis: an integrative review, empirical guidance, and next steps. *Int J Qual Methods*. (2024) 23:16094069241296206. doi: 10.1177/16094069241296206
- Hertzog MA. Considerations in determining sample size for pilot studies. *Res Nurs Health*. (2008) 31:180–91. doi: 10.1002/nur.20247
- Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. (2006) 3:77–101. doi: 10.1191/1478088706qp0630a
- O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. *BMJ*. (2010) 341:c4587. doi: 10.1136/bmj.c4587
- Tran JT, Theng B, Tzeng HM, Raji M, Serag H, Shih M, et al. Cultural diversity impacts caregiving experiences: A comprehensive exploration of differences in caregiver burdens, needs, and outcomes. *Cureus*. (2023) 15:e46537. doi: 10.7759/cureus.46537
- Napoles AM, Chadiha L, Eversley R, Moreno-John G. Reviews: developing culturally sensitive dementia caregiver interventions: are we there yet? *Am J Alzheimers Dis Other Dement*. (2010) 25:389–406. doi: 10.1177/1533317510370957