# The Artificial Intelligence (AI)-Based Chatbot for Promoting Physical Activity and Healthy Diet: Protocol for a Systematic Review

**Yoo Jung Oh**
University of California Davis https://orcid.org/0000-0002-7829-8535

**Jingwen Zhang**
University of California Davis

**Min-Lin Fang**
University of California San Francisco

**Yoshimi Fukuoka** ( Yoshimi.fukuoka@ucsf.edu )
University of California San Francisco https://orcid.org/0000-0002-2245-9264

**Protocol**

**Keywords:** Chatbot, Conversational agent, Physical activity, Weight loss, Weight maintenance, Diet, Nutrition, Sedentary behavior, Systematic review, Protocol

**Posted Date:** December 7th, 2020

**DOI:** https://doi.org/10.21203/rs.3.rs-121903/v1

# Abstract

Background

With the rise of artificial intelligence (AI) technologies in recent years, the rapidly expanding fields of AI-supported chatbot lifestyle interventions have offered new solutions to the global epidemic of physical inactivity and obesity. However, to the best of our knowledge no systematic review of chatbot-based lifestyle change intervention exists. The goals of this systematic review are to summarize the characteristics of chatbot interventions and to synthesize and evaluate uses of chatbots to improve physical activity, dietary, and weight management behaviors and to identify knowledge gaps and directions for future studies.

Methods

In collaboration with a medical librarian, six electronic bibliographic databases (PubMed, EMBASE, ACM Digital Library, Web of Science, PsycINFO, and IEEE) will be searched to identify all relevant studies. Main outcomes include changes in self-report and/or objectively measured physical activity, sedentary behavior, diet, and body weight. Additional outcomes include feasibility, acceptability, safety, and user satisfactions of chatbots. Two reviewers will independently screen the title and abstract, conduct a full-text screening to select the qualified studies, extract data from the included studies, and assess the risk of bias using Covidence software. Lastly, we will conduct a qualitative synthesis of the findings. However, if several randomized controlled trials that report similar outcome measurements are identified., quantitative synthesis will be provided.

Discussion

To the best of our knowledge, this is the first systematic review to synthesize and evaluate the existing research that assess the impact of AI chatbots on changing physical activity, dietary, and weight management behaviors. We anticipate our findings to advance knowledge by identifying the key characteristics of effective AI chatbot interventions and by highlighting knowledge gaps and limitations in the literature.

Systematic review registration

International Prospective Register of Systematic Reviews (PROSPERO): CRD42020216761.


# Background

Healthy lifestyle such as engaging in regular physical activity, taking balanced diet, and maintaining healthy weight reduces risks of premature death and chronic diseases and improves metal health and quality of life [1, 2]. Despite these health benefits, maintaining healthy lifestyle is challenging and the prevalence of physical inactivity, sedentary behavior, and obesity has risen dramatically worldwide [3]. Global prevalence of insufficient physical activity was 27.5% in 2016 [4], with high income countries

experiencing increasing rates of insufficient activity. For example, more than 80% of American adults do not meet the current physical activity guidelines [5]. Moreover, the prevalence of obesity is at 42.4% [5] and the prevalence of metabolic syndrome is estimated at more than 30% [6] in the United States (U.S.). Therefore, cost-effective lifestyle interventions that can be disseminated to large and diverse populations are urgently needed.

With the rapid rise of artificial intelligence (AI) technologies in recent years, AI-based lifestyle interventions have offered new possibilities into developing innovative ways to support behavior change. One frontier regards recent advancements in voice and text recognition, natural language processing, and machine learning, which have led to the increasing incorporation of conversational agents, or AI-chatbots, into lifestyle modification programs [7, 8]. AI-chatbots employ dialogue systems to enable both constrained and unconstrained natural language conversations with users. They can be deployed in the forms of mobile apps, thus making programs widely available 24/7. It is worth noting that AI chatbots need to be differentiated from embodied virtual agents that rely on multimodal signals (e.g., images, videos, sounds) to simulate human face-to-face communication. AI chatbots' core feature is natural language conversation that facilitates more flexible communications between human and the chatbot. Through natural conversations, the chatbot can ask questions, understand individual users' backgrounds such as behavior habits, psychological states, social and physical environments, and deliver personalized health information and persuasive messages to facilitate the user to accomplish behavior change goals. More importantly, AI chatbots are always available to respond to users' queries, answer questions, and/or provide social support.

AI chatbots have been tested in realms of medical consultation, disease diagnoses, mental health support [8, 9], and more recently risk communications for the COVID-19 pandemic [10]. There are a few systematic reviews that have summarized characteristics of chatbot interventions in a variety of healthcare domains such as mental health support, disease diagnosis, and sexual health education [8, 11, 12]. However, due to great heterogeneities in research outcomes and study designs, none of the reviews conducted meta-analysis to extract effect sizes. To the best of our knowledge, no systematic review on chatbot lifestyle interventions for promoting physical activity, healthy diet, and weight management has been conducted.

The aims of this systematic review are to 1) describe characteristics, functions, and core conversational capacities of chatbot interventions; 2) investigate whether these chatbot interventions are effective on changing physical activity, diet, weight management behaviors, and other related health outcomes; and 3) identify knowledge gaps and limitations. We will also use the AI-Chatbot Behavior Change Model [7] as a guiding theoretical framework to examine the chatbot interventions' characteristics, theoretical mechanisms, quality and safety checks, and ethical considerations. The model is proposed for guiding the design and evaluation of chatbot-based lifestyle interventions and specifies four conceptual domains regarding (1) designing chatbot characteristics and understanding user backgrounds; 2) building chatbot's relational capacity; 3) building chatbot's persuasive conversational capacity; and 4) evaluating

mechanisms and outcomes. We believe that the findings of this systematic review will assist in designing chatbot-based lifestyle interventions in the future.

# Methods

This systematic review has been registered with the International Prospective Register of Systematic Reviews (PROSPERO): registration number CRD42020216761. The Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) statement [13] has been used in the preparation of this protocol, and the PRISMA-P 2015 Checklist (p. 5-6) is included [see Additional file 1].

### Eligibility criteria

Studies will be selected according to the following criteria (the PICOS framework [14]): populations/participants, interventions and comparators, outcome(s) of interest, and study designs/types. The population of interest will be individuals who use chatbots for their physical activity, diet, and weight management and include both children and adults. Eligible interventions will be text or/and speech based chatbots operating as standalone software or via a web browser. Both contained and unconstrained chatbot designs will be included. Chatbots that are part of virtual reality, augmented reality, embodied agents, and therapeutic robots will be excluded. We will include studies with or without comparators. Applying the AI-Chatbot Behavior Change Model, we will examine main outcomes including changes in self-report and/or objectively measured physical activity, sedentary behavior, diet, and body weight. Additional outcomes include feasibility, acceptability, safety (e.g. adverse events, injury etc.), and user satisfactions of chatbots if available. In terms of study design, quantitative studies, such as randomized controlled trials (RCTs), trials without a control group and/or randomization, cohort studies, case-control studies, and observational studies will be included. However, qualitative studies and case studies will be excluded. Peer-reviewed articles and conference proceedings in English will be included. However, review papers, conference abstracts without the main text, proposals, and editorials will be excluded. No restrictions regarding country of the study conducted or published, study setting, or year of publication will be applied.

### Information sources and search strategy

In collaboration with a medical librarian (MF), a systematic search strategy for six electronic databases (PubMed, EMBASE, ACM Digital Library, Web of Science Core Collection, PsycINFO, and IEEE) was designed using a combination of MeSH/Emtree terms and various keywords to identify peer-reviewed studies related to chatbot for promoting physical activity and healthy diet. In addition, Web of Science will be used to find citing, cited, and relevant references of studies selected for inclusion. An additional file provides detailed sample search strategies [see Additional file 2].

### Study selection and data management process

A flow diagram following the PRISMA guidelines for reporting systematic reviews will be used to illustrate the selection processes and results [13]. Initially, all retrieved studies will be imported into Endnote reference management software to remove duplicates, then the remaining citations will be uploaded to Covidence (Covidence systematic review software, Available at www.covidence.org) which is recommended by Cochrane for use by its systematic review authors. Using Covidence, two review team members will independently screen the title and abstract of the identified articles from the electronic database search and additional sources and then they will independently review the full texts of retrieved articles for eligibility. Any disagreement between the two review team members over the eligibility of particular studies will be resolved through discussion with a third reviewer.

### Data items and data collection process

We will develop data collection forms based on the Cochrane Handbook for Systematic Reviews of Interventions [15] and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) strategy [13]. The data extraction form includes domains of 1) study characteristics (e.g. author, published year, study design, setting, duration), 2) participant characteristics (e.g. age, gender, education, sample size, socio-economic status), 3) chatbot intervention characteristics (e.g. dialogue system infrastructures, media, anthropomorphic cues, features of relational capacity, features of persuasive conversational capacity), 4) outcomes and mechanisms (e.g. self-report and/or objectively measure physical activity), and 5) safety and ethical assessment. Data extractions for domains from 2) to 5) will be guided by the AI-Chatbot Behavior Change Model [7]. If there is any missing data, we may contact study authors to obtain the data. One reviewer will use Covidence software to extract the five data items specified above. A second reviewer will cross-check the accuracy of the extracted data. The data items to be extracted will be pilot tested using a template spreadsheet and a sample of potential studies obtained through the initial iterative search process.

### Risk of bias assessment

Selected studies will be assessed by two independent reviewers using the Newcastle Ottawa Scale for non-randomized studies [16]. Any included randomized controlled trials will be assessed for methodological quality using the Cochrane Collaboration's Risk of Bias (RoB) 2.0 tool for randomized controlled trials [17].

### Data synthesis

We will provide a qualitative synthesis of the findings of the included studies (i.e., characteristics of chatbot interventions, study designs and target population characteristics, and measurements of outcomes and other indicators for mechanisms of changes). If several RCTs that report similar outcome measurements are identified, we will provide quantitative synthesis. Intervention effects will be summarized for each study by calculating risk ratios (for dichotomous outcomes) or standardized mean differences (for continuous outcomes). Differences in effect sizes will be combined using a random-effects meta-analysis. If the heterogeneity statistics ($I^2$ and Q) suggest no significant heterogeneity in

studies, a fixed-effects model will be fitted to obtain parsimony. All estimates will be reported with 95% confidence intervals.

# Discussion

AI-based chatbot interventions are increasing in popularity and use, contributed by the recent advancements in artificial intelligence, natural language processing, and machine learning. This review presents several strengths and benefits to contribute to the field of AI-based chatbot interventions in the health domain. To the best of our knowledge, this is the first systematic review to integrate and synthesize the existing research that assessed the impact of AI chatbots on changing physical activity, diet, and weight management behaviors as well as related health outcomes.

This systematic review has the potential to advance our knowledge about whether AI chatbot interventions could be considered as an effective means of changing one's physical activity and dietary, and weight management behaviors by identifying key characteristics of the study participants (e.g., gender, socio-economic status), chatbot characteristics (e.g., media, dialogue designs), and outcomes (e.g., changes in relational perceptions, changes in behaviors). Through qualitative synthesis of the findings, the proposed systematic review will highlight which characteristics of AI chatbot interventions are effective in contributing to behavior and health outcome changes in the target populations. In addition, this review may help to identify research gaps concerning the feasibility, acceptability, and effects of AI chatbots on changing lifestyle behaviors. Combined with the theoretical and ethical discussions presented in Zhang et al. [7], our review will point out the blind spots and limitations in existing research and promising future directions for advancing chatbot-based interventions and programs.

Despite strengths and benefits of this systematic review, several potential limitations need to be acknowledged. First, because natural language chatbots are relatively new and are rapidly evolving, there is a possibility that there will be a paucity of quantitative data to answer some of our research questions of interest. In this case, qualitative and narrative synthesis of the findings will address our questions. Second, by limiting our search to English language papers, we may miss relevant non-English papers which could potentially introduce a language bias. However, the previous research studies suggest that restrictions on non-English language introduce little to no systematic bias [18, 19]. Thus we believe that this language restriction will not significantly affect the validity and precision of our findings. Reasons of the excluded papers will be noted in detail.

In sum, the findings from this systematic review will provide a comprehensive summary of chatbots-based lifestyle interventions. The insights will also inform health care providers and researchers to continue the development and implementation of effective AI chatbot-based interventions for physical activity, diet, and weight management. We believe that our review results will be able to assist in designing chatbot interventions and research studies in the future in several ways. First, the research gaps identified in this review will guide novel research questions for future studies. Studies with new and

diverse perspectives have the potential to push the theoretical discussion of chatbots forward. Second, since we will cover a broad range of outcomes, such as user perceptions towards the chatbot systems, user characteristics, and objectively measured behavioral outcomes, the findings of this systematic review will provide practical recommendations on how to adapt and modify the chatbot system to diverse user characteristics and their behaviors.

# Abbreviations

AI: Artificial Intelligence; PROSPERO: International Prospective Register of Systematic Reviews; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis; NRS: Non-randomized study; RCT: Randomized controlled trial

# Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent forpublication

Not applicable.

### Availability of data and materials

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Funding

This project was supported by a grant (K24NR015812) from the National Institute of Nursing Research (Dr. Fukuoka). Publication made possible in part by support from the UCSF Open Access Publishing Fund. The study sponsors had no role in the study design; collection, analysis, or interpretation of data; writing the report; or the decision to submit the report for publication.

### Author’s contributions

YF, JZ, and YO contributed to the conception and design of the review; MF and YO developed the search strategies; YF, JZ, and YO wrote sections of the protocol. All authors contributed to manuscript revision, read and approved the submitted version. YF is the guarantor of the review.

### Acknowledgements

Not applicable.

# References

[1] Ekelund U, Steene-Johannessen J, Brown WJ, Fagerland MW, Owen N, Powell KE, et al. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. The Lancet. 2016;388:1302–10. doi:10.1016/s0140-6736(16)30370-1.

[2] Bize R, Johnson JA, Plotnikoff RC. Physical activity level and health-related quality of life in the general adult population: A systematic review. Preventive Medicine. 2007;45:401–15. doi:10.1016/j.ypmed.2007.07.017.

[3] Ding D, Lawson KD, Kolbe-Alexander TL, Finkelstein EA, Katzmarzyk PT, van Mechelen W, et al. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. The Lancet. 2016;388:1311–24. doi:10.1016/s0140-6736(16)30383-x.

[4] Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. The Lancet Global Health. 2018;6:e1077–86. doi:10.1016/s2214-109x(18)30357-7.

[5] Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity and severe obesity among adults: United States, 2017–2018. NCHS Data Brief, no 360. Hyattsville, MD: National Center for Health Statistics . 2020. https://www.cdc.gov/nchs/data/databriefs/db360-h.pdf.

[6] Moore JX, Chaudhary N, Akinyemiju T. Metabolic Syndrome Prevalence by Race/Ethnicity and Sex in the United States, National Health and Nutrition Examination Survey, 1988–2012. Prev Chronic Dis. 2017;14. doi:10.5888/pcd14.160287.

[7] Zhang J, Oh YJ, Lange P, Yu Z, Fukuoka Y. Artificial Intelligence Chatbot Behavior Change Model for Designing Artificial Intelligence Chatbots to Promote Physical Activity and a Healthy Diet. J Med Internet Res. 2020;22:e22845. doi:10.2196/22845.

[8] Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. 2018;25:1248–58. doi:10.1093/jamia/ocy072.

[9] Pereira J, Díaz Ó. Using Health Chatbots for Behavior Change: A Mapping Study. J Med Syst. 2019;43. doi:10.1007/s10916-019-1237-1.

[10] Miner AS, Laranjo L, Kocaballi AB. Chatbots in the fight against the COVID-19 pandemic. npj Digit Med. 2020;3. doi:10.1038/s41746-020-0280-0.

[11] Gentner T, Neitzel T, Schulze J, Buettner R. A Systematic Literature Review of Medical Chatbot Research from a Behavior Change Perspective. 2020. doi:10.1109/compsac48688.2020.0-172.

[12] 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;:070674371982897. doi:10.1177/0706743719828977.

[13] Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2016;:i4086. doi:10.1136/bmj.i4086.

[14] Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. BMC Med Inform Decis Mak. 2007;7. doi:10.1186/1472-6947-7-16.

[15] Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). Cochrane, 2020. Available from www.training.cochrane.org/handbook.

[16] Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa: Ottawa Hospital Research Institute; 2018. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed 21 Oct 2020.

[17] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;:l4898. doi:10.1136/bmj.l4898.

[18] Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. Int J Technol Assess Health Care . 2012;28:138–44. doi:10.1017/s0266462312000086.

[19] Jüni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. 2002;31:115–23. doi:10.1093/ije/31.1.115.

## Supplementary Information

Additional file 1: PRISMA-P 2015 Checklist (file format: Microsoft Word)

Additional file 2: Sample search strategies for PubMed and EMBASE (file format: Microsoft Word)

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx
- Additionalfile1.docx
- Additionalfile2.docx

- Additionalfile2.docx