

Review

Applications of Machine Learning in Real-Life Digital Health Interventions: Review of the Literature

Andreas K Triantafyllidis^{1,2}, BSc, MSc, PhD; Athanasios Tsanas^{3,4}, BSc, BEng, MSc, PhD

¹Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki, Greece

²Lab of Computing, Medical Informatics and Biomedical Imaging Technologies, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

³Usher Institute of Population Health Sciences and Informatics, Medical School, University of Edinburgh, Edinburgh, United Kingdom

⁴Mathematical Institute, University of Oxford, Oxford, United Kingdom

Corresponding Author:

Andreas K Triantafyllidis, BSc, MSc, PhD
Information Technologies Institute
Centre for Research and Technology Hellas
6th km Charilaou-Thermi Rd
GR 57001 Thermi
Thessaloniki, 60361
Greece
Phone: 30 2310 498100
Fax: 30 2310 498180
Email: atriand@gmail.com

Abstract

Background: Machine learning has attracted considerable research interest toward developing smart digital health interventions. These interventions have the potential to revolutionize health care and lead to substantial outcomes for patients and medical professionals.

Objective: Our objective was to review the literature on applications of machine learning in real-life digital health interventions, aiming to improve the understanding of researchers, clinicians, engineers, and policy makers in developing robust and impactful data-driven interventions in the health care domain.

Methods: We searched the PubMed and Scopus bibliographic databases with terms related to machine learning, to identify real-life studies of digital health interventions incorporating machine learning algorithms. We grouped those interventions according to their target (ie, target condition), study design, number of enrolled participants, follow-up duration, primary outcome and whether this had been statistically significant, machine learning algorithms used in the intervention, and outcome of the algorithms (eg, prediction).

Results: Our literature search identified 8 interventions incorporating machine learning in a real-life research setting, of which 3 (37%) were evaluated in a randomized controlled trial and 5 (63%) in a pilot or experimental single-group study. The interventions targeted depression prediction and management, speech recognition for people with speech disabilities, self-efficacy for weight loss, detection of changes in biopsychosocial condition of patients with multiple morbidity, stress management, treatment of phantom limb pain, smoking cessation, and personalized nutrition based on glycemic response. The average number of enrolled participants in the studies was 71 (range 8-214), and the average follow-up study duration was 69 days (range 3-180). Of the 8 interventions, 6 (75%) showed statistical significance (at the $P=.05$ level) in health outcomes.

Conclusions: This review found that digital health interventions incorporating machine learning algorithms in real-life studies can be useful and effective. Given the low number of studies identified in this review and that they did not follow a rigorous machine learning evaluation methodology, we urge the research community to conduct further studies in intervention settings following evaluation principles and demonstrating the potential of machine learning in clinical practice.

(*J Med Internet Res* 2019;21(4):e12286) doi: [10.2196/12286](https://doi.org/10.2196/12286)

KEYWORDS

machine learning; data mining; artificial intelligence; digital health; review; telemedicine

Introduction

Background

Digital health interventions [1], including modalities such as telemedicine, Web-based strategies, email, mobile phones, mobile apps, text messaging, and monitoring sensors, have enormous potential to support independent living and self-management [2], and reduce health care costs [3]. They have also shown great promise in improving health [4]. With the advent of new tools and algorithms for machine learning, a new class of smart digital health interventions can be developed, which could revolutionize effective health care delivery [5].

The term *machine learning* is widely used across disciplines but has no universally accepted definition [6]. This is in part explained by the breadth of the areas it covers and because researchers from diverse disciplines have historically contributed (and still contribute) to its development. Broadly, it refers to an algorithmic framework that can provide insights into data, while facilitating inference and providing a tentative setting to determine functional relationships.

Machine learning has been applied in multiple health care domains, including diabetes [7], cancer [8], cardiology [9], and mental health [10]. Most of the developed machine learning models and tools in research settings have investigated the potential of prognosis [11], diagnosis [12], or differentiation of clinical groups (eg, a group with a pathology and a healthy control group or groups with pathologies) [13], thus demonstrating promise toward the development of computerized decision support tools [14]. The key requirements for the development of these tools are sufficiently large datasets (in terms of both number of participants and explanatory variables to explore) and accurate labels, typically provided by expert clinicians. The premise is the identification of those data structures or variables (eg, clinical, behavioral, or demographic variables) that are associated with the target outcome (eg, whether a person has cancer). In this regard, useful knowledge can be derived from the available data, which can empower patients to monitor their health status longitudinally and support health professionals in decision making with regard to management, treatment, and follow-up interventions where required.

Despite a considerably growing body of research literature in the use of machine learning in health care applications [15], it is astonishing how few of these suggestions are actually translated into clinical practice [16]. There is remarkably limited empirical evidence of the effectiveness of machine learning applications in digital health interventions. This is rather surprising, since any proposed health care solutions would reach their full potential only if they are embraced by the medical community, becoming integrated within properly designed digital health interventions and tested in real-life studies with patients and health professionals.

Objective

Considering that machine learning models and tools have not been widely and reliably used in clinical practice, whereas the peer-reviewed literature in the field is growing exponentially,

we wanted to assess the progress made in smart data-driven health interventions applied in real-life research settings—that is, the real world in which constraints in available resources or opportunities to collect reliable data may exist, as opposed to simulation or laboratory-based studies [17]. In this direction, we present a systematic literature review of digital health interventions incorporating machine learning algorithms, by identifying and mapping their features and outcomes, with the aim to improve our knowledge of the design and development of impactful intelligent interventions.

Methods

Inclusion and Exclusion Criteria

We sought to identify digital health nonpharmacological interventions incorporating machine learning that were assessed in pragmatic studies. In this context, the inclusion criteria for study selection were (1) the study should be conducted with patients or health professionals, or both, in a real-life setting, (2) machine learning algorithms or models were used in the digital health intervention (rather than merely reporting statistical hypothesis testing results or statistical associations), (3) quantitative outcomes of the study were presented, and (4) the article describing the study was written in English. We excluded retrospective studies, case reports, ongoing studies, surveys or reviews, laboratory or simulation studies, studies describing protocols, qualitative studies, and all studies published before 2008 from the review because we wanted to determine the status of recent research developments in the field that have been used in clinical interventional settings.

Literature Search and Screening

We searched the PubMed and Scopus bibliographic databases for studies published after 2008 using the string “(machine learning) OR (data mining) OR (artificial intelligence) AND health” for search within the title, abstract, and keywords of the articles. We limited “Species” in PubMed to humans.

Both authors independently screened the identified articles following the literature search to minimize bias in the selection process. Any disagreements were resolved by discussion between the authors and reaching a consensus. We screened the abstracts of the candidate articles for inclusion and subsequently read the full text of the articles deemed eligible according to the inclusion criteria. Subsequently, we excluded articles not providing sufficient information about the application of machine learning or for being ineligible. We used the Effective Public Health Practice Project (EPHPP) tool to assess the methodological quality of the included studies, which has been found to be reliable [18]. The studies that focused on interventions were synthesized (AKT) according to their target (ie, target condition), study design, number of enrolled participants, follow-up duration, primary outcome and whether this was significantly positive, machine learning algorithms used in the intervention, and outcome of the algorithms (eg, prediction of a target outcome).

The systematic review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses

(PRISMA) guidelines [19]. [Multimedia Appendix 1](#) shows a completed PRISMA checklist.

Results

Literature Search Outcomes

Our last search in November 2018 returned 1386 articles from the PubMed database and 7024 articles from Scopus. We imported all the retrieved records into Mendeley (version 1.19.3) bibliography management software (Mendeley Ltd) [20], which identified 1093 duplicates. We screened the abstracts of the remaining 7317 results according to our inclusion and exclusion criteria and identified 21 eligible articles. The reviewers read the full text of the 21 articles and agreed on 8 for inclusion as eligible articles. The flow diagram in [Figure 1](#) summarizes the

reasons for excluding research articles for study inclusion following the PRISMA format ([Figure 1](#)).

Quality Assessment

On the basis of the EPHPP criteria for selection bias, design, confounders, blinding, data collection, and dropouts, we found the methodological quality to be moderate for 2 of the 8 (25%) studies [21,22] and weak for the remaining 6 (75%) studies [23-28] ([Table 1](#)). Most studies were poorly rated because of selection bias, insufficient care in controlling for confounders, and the high percentage of withdrawals or dropouts (or the absence of their description). The design of a randomized or controlled clinical trial was described in 3 (37%) studies [21,25,28], and 5 (63%) interventions were evaluated in a pilot or experimental single-group study ([Multimedia Appendix 2](#)).

Figure 1. Flow diagram for study inclusion following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) format.

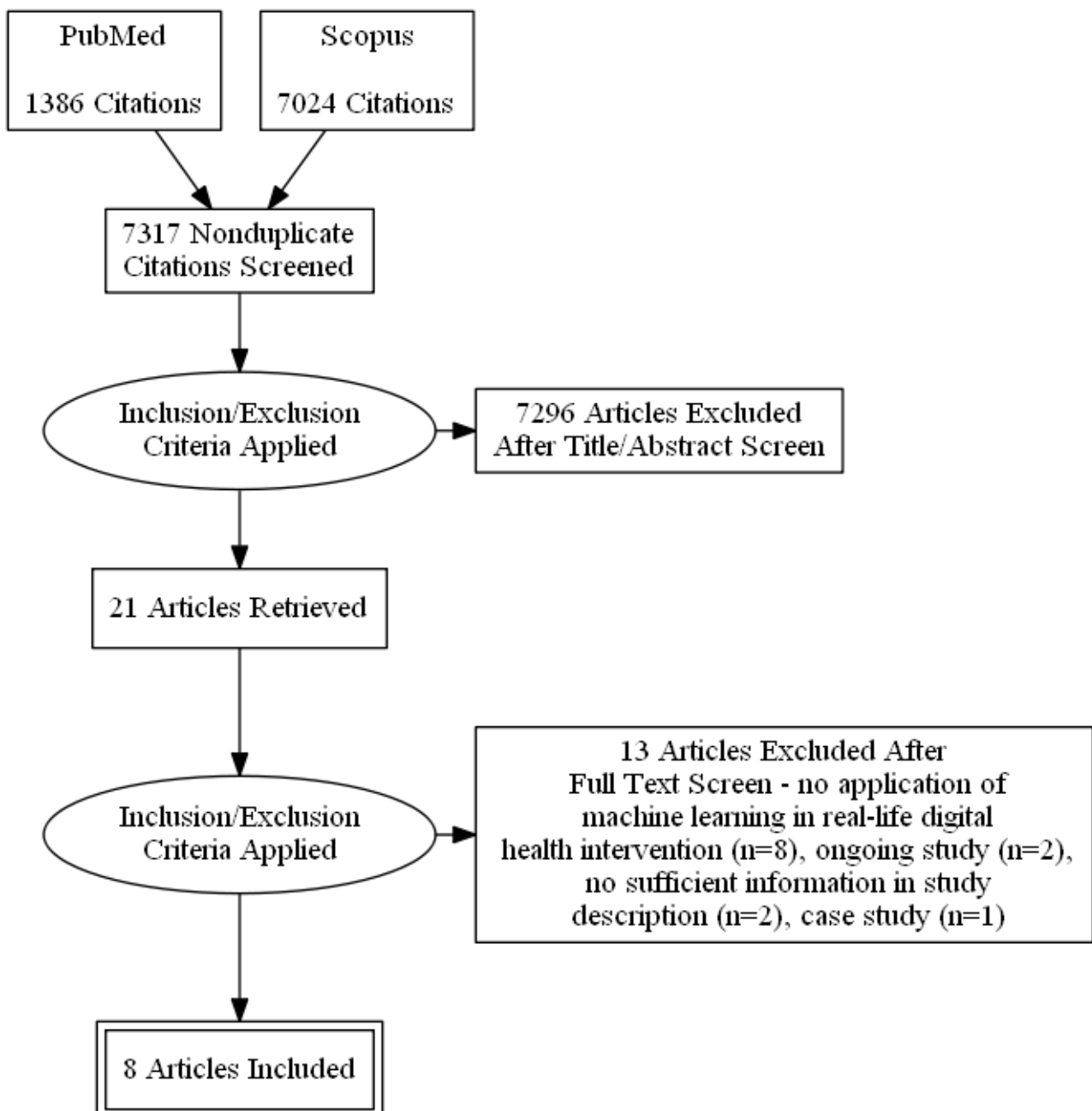


Table 1. Quality assessment of included studies based on Effective Public Health Practice Project (EPHPP) criteria.

Study first author, year, and reference	EPHPP criteria						Global rating
	Selection bias	Study design	Confounders	Blinding	Data collection methods	Withdrawals and dropouts	
Burns, 2011 [23]	Weak	Moderate	Weak	Moderate	Strong	Strong	Weak
Hawley, 2013 [24]	Weak	Weak	Weak	Moderate	Weak	Weak	Weak
Manuvinakurike, 2014 [22]	Weak	Strong	Strong	Moderate	Moderate	Strong	Moderate
Martin, 2012 [25]	Moderate	Strong	Weak	Moderate	Strong	Weak	Weak
Morrison, 2017 [26]	Moderate	Moderate	Weak	Strong	Strong	Weak	Weak
Ortiz-Catalan, 2016 [27]	Moderate	Moderate	Weak	Moderate	Weak	Weak	Weak
Sadasivam, 2016 [21]	Weak	Strong	Strong	Strong	Moderate	Strong	Moderate
Zeevi, 2015 [28]	Weak	Strong	Weak	Moderate	Weak	Weak	Weak

Type of Intervention and Target Population

The interventions targeted depression prediction and management [23], speech recognition for people with speech disabilities [24], self-efficacy for weight loss [22], detection of changes in biopsychosocial condition of patients with multiple morbidity [25], stress management [26], treatment of phantom limb pain [27], smoking cessation [21], and personalized nutrition based on glycemic response [28] (Multimedia Appendix 2).

Of the 8 interventions, 3 (37%) targeted patients: individuals with a diagnosis of depression [23], those with multiple morbidities such as lung disease and cardiovascular disease [25], and those with phantom limb pain [27]. One (13%) intervention targeted people with speech disabilities [24]; 4 (50%) interventions targeted individuals who had no explicit diagnosis of a disease or impairment [21,22,26,28]. All target groups comprised adults. The average number of enrolled participants in the studies was 71 (range 8-214), and the average follow-up study duration was 69 days (range 3-180).

Applications of Machine Learning and Outcomes

Overall, 6 of the 8 (75%) real-life studies of digital health interventions aided by machine learning algorithms showed statistical significance (at the $P=.05$ level) in health outcomes. Different summary measures were used in the identified studies to assess primary outcomes, which reflects the lack of standardization both in methodology and in the metrics used in the research fields. Where possible, we aimed to use the accuracy of the algorithms used and the P value (eg, for showing statistical significance of outcomes in an intervention group compared with a control group) as the principal summary measures. We briefly describe all included studies below in terms of intervention purpose and content, evaluation outcomes, and implications for clinical practice.

Burns et al [23] described a multicomponent mobile-based intervention that used machine learning models to predict the mood, emotions, cognitive and motivational states, activities, and environmental and social context of patients with depression, along with feedback graphs for self-reflection on behavior and coaching provided by caregivers. The predictive models were based on phone sensor-derived variables (eg,

global positioning system, ambient light, phone calls), and regression along with decision trees was used. The accuracy of the models was promising for location prediction (60%-91%), but prediction was very poor for emotions such as sadness. Overall, the 8 participants in the study became less likely to meet the criteria for a diagnosis of major depressive disorder ($P=.03$), and their symptoms of depression and anxiety were decreased by the end of the study ($P<.001$). Patients were also satisfied with the intervention (5.71 average rating on a scale 1 to 7), and 6 of 7 treatment completers (86%) indicated that the intervention was helpful in understanding triggers for negative moods. Despite the benefits of self-reflection on behavior through the use of a multicomponent mobile health monitoring system and the clinical improvements shown in the study, the authors reported that the clinical utility of the prediction models they used should be improved, since the prediction outcomes (eg, location and mood) were merely displayed to the users, and there were no direct interventions based on them.

Hawley et al [24] described the use of a device capable of recognizing the speech of people with dysarthria and generating voice messages. The authors used hidden Markov models to determine the proximity of a spoken word to a personalized speech model for that individual. However, only 67% recognition accuracy was achieved in this real-life observational study with 9 participants. Participants noticed that ease of communication was reduced through the device compared with their usual communication method of either speaking or speaking supported by a conventional voice-output communication aid, mainly due to the low accuracy of speech recognition. Nevertheless, feedback from participants was positive about the device's concept, given that speech recognition was improved.

Manuvinakurike et al [22] focused on changes in self-efficacy for weight loss through the provision of personal health behavior change stories found on the internet. An algorithm based on adaptive boosting was developed to find the most relevant story based on the stage of change and the demographic characteristics of a user, along with the emotional tone and overall quality of the story (accuracy between 84% and 98% for the classification of 5 stages of change). Testing of the algorithm with 103 users revealed significantly greater increases in self-efficacy for weight loss ($P=.02$) and a statistically insignificant effect on

change in decisional balance ($P=.83$). In addition, the medium used to tell the stories, being either text or an animated conversational agent, had no effect on health behavior change. The authors concluded that their approach could maximize participants' engagement in longitudinal health behavior change interventions.

Martin et al [25] used a system in which decision trees could predict unplanned hospital visits of patients with multiple morbidities such as lung disease or cardiovascular disease. Alerts were sent to health professionals, who acted on the alerts according to agreed guidelines. The system was based on information received via patient phone calls with lay care guides. Linguistic and metalinguistic features were extracted, together with the patient's status, to train the prediction models (positive predictive value of 70% for predicting unplanned events). A randomized controlled trial with 214 patients for 6 months (the largest trial we found in the review in terms of number of enrolled participants and duration) showed a reduction of 50% in the number of unplanned hospital events of participants in the intervention group compared with control. The most common response to an alert indicating that a patient needed attention (red alert) was to phone the patient the next day to reassess the situation and contact their general practitioner (3% of calls), suggest or plan a visit to their general practitioner (11% of calls), or call an ambulance (<0.01% of calls). In summary, the authors reported that predictive analytics on an ongoing basis could be used to signify risk of hospitalization and guide the health care system to take appropriate actions.

Morrison et al [26] used push notifications to enhance engagement of smartphone users for stress management. They used a naïve Bayes classifier to predict whether a user would respond to a notification, thereby building a personalized intelligent mechanism for notification delivery, based on the times within a day a user was more likely to view and react to the received messages. However, this exploratory study with 77 participants showed no statistically significant difference between participants receiving the messages sent "intelligently" and those receiving a message daily or occasionally within 72 hours (Cohen $d=0.14$ for intelligent vs daily group and $d=0.5$ for intelligent vs occasional group, for actions taken in response to messages). Although notification delivery based on time had no effect on the study groups (ie, response to notifications was no different), the authors concluded that frequent daily messages may not deter users from engaging with digital health interventions.

Ortiz-Catalan et al [27] applied myoelectric pattern recognition algorithms for the control of a virtual limb in patients with phantom limb pain and used gaming along with augmented and virtual reality for treatment. This single-group study with 14 participants revealed that patients' symptoms of phantom limb pain were significantly decreased (by about 50%) at the end of the provided treatment for 6 months ($P=.0001$ for reduction in intensity and quality of pain). The authors suggested that their novel treatment could be used after failure of evidence-based treatments such as mirror therapy and before proceeding with invasive or pharmacological approaches.

Sadasivam et al [21] used a recommender system to send motivational messages to individuals, targeting smoking cessation. The system was based on Bayesian probabilistic matrix factorization to predict message rating, through the processing of data from the user's previous ratings of messages, along with other users' ratings. This randomized controlled trial with 120 users showed that the system was more effective at influencing people to quit smoking than were standard tailored messages (rule-based system) with proven effectiveness ($P=.02$) and resulted in a similar cessation rate. The authors concluded that their recommender system could be used instead of standard systems for influencing smoking cessation because it was more personalized (it learned and adapted to a person's behavior) and could incorporate a considerably greater number of variables; however, larger trials would be needed to demonstrate the system's effectiveness.

Zeevi et al [28] used gradient boosting regression to predict the postmeal glycemic response of individuals in real life, according to blood parameters, dietary habits, anthropometrics, physical activity, and gut microbiota. The results from this randomized controlled study with 24 participants showed that a personalized diet based on postmeal glycemic predictions could statistically significantly modify elevated postprandial blood glucose ($P<.05$ for predicting low levels of blood glucose ["good diet"] vs high levels of blood glucose ["bad diet"], which was comparable with diets selected by experts). The authors reported that their approach could be used in nutritional interventions for controlling or preventing disorders associated with poor glycemic control, such as obesity, diabetes, and nonalcoholic fatty liver disease. However, evaluation periods of months or even years would be needed first to clearly indicate the effectiveness of the proposed algorithm.

Discussion

Principal Findings

This review is, to our knowledge, the first to systematically examine the features and outcomes of digital health interventions incorporating machine learning that were implemented and assessed in real-life studies [17]. With this aim in mind, we differentiated our review from previous investigations that focused only on the broader use of artificial intelligence in medicine in the context of specific diseases [29,30], machine learning techniques [31,32], or risk prediction models, such as through mining of electronic health records [33,34], and did not consider real-life evaluation of the respective interventions. The need to demonstrate evidence of an intervention's effectiveness in the real world has been highlighted in several other studies [35-37]. Our main finding is that most of the digital health interventions showed significantly positive health outcomes for patients or healthy individuals, which demonstrates the virtue of machine learning applications in actual clinical practice. However, given the small number of studies identified in this review and their considerable limitations highlighted above, further work is warranted to demonstrate the effectiveness of digital interventions relying on machine learning applications in real-life medical care.

Our review found 8 different cases of machine learning applications in a real-life setting: depression prediction and management, speech recognition for people with speech disabilities, self-efficacy for weight loss, detection of changes in biopsychosocial condition of patients with multiple morbidity, stress management, treatment of phantom limb pain, smoking cessation, and personalized nutrition based on glycemic response. The reviewed studies had several implications for clinical practice, such as better engagement of patients with interventions [22], the identification of risk for hospitalization [25], or the introduction of novel treatment methods [27]. Among the studies, those for speech recognition of people with speech disabilities [24] and notification delivery for stress management [26] clearly reported insignificant outcomes, whereas 6 studies showed significant outcomes, but they were of low to moderate methodological quality. Only 3 studies were in the form of a randomized controlled trial, which limited the ability to fully identify the added value of machine learning-enabled interventions compared with standard care. To this end, further rigorous studies with adequately powered samples (recruiting considerably more participants than the average number of 71 participants found in this review) are needed, which would generate the evidence base for the effectiveness of machine learning in clinical practice. To that effect, large trials and publicly accessible databases that have become available over the last few years, such as the UK BioBank and the Physionet database, are providing rich resources that could facilitate insights.

The delivery of motivational messages [21,26] or stories [22] for health behavior change and engagement seems to be an emerging area of digital health interventions incorporating machine learning. These studies also demonstrated the latest efforts to promote individuals' personalized self-management and to put them at the center of health care [38]. Considering the effectiveness of tailored messaging in influencing health behavior change [39], further research in this area is warranted.

The surprisingly small number of identified pragmatic studies in our review might raise some concerns and indicates the substantial challenge of systematically evaluating digital health interventions that incorporate machine learning [40]. In this context, the retrospective validation of algorithms and models, given the availability of one or more datasets, constitutes only the first step in the evaluation process [28]. The second step involves the integration of the algorithms and models within a digital health tool, such as mobile phone-based tools [23], internet-based tools [14], or an aid device [24]. The third step

requires the assessment of the developed tool as a digital health intervention in a real-life research setting (eg, through a randomized controlled trial), together with patients or health professionals, or both [28,41]. The final step would be the monitoring of actual uptake and use of the intervention in real-world settings and outside of a research setting [42], which is, however, rarely reported [43]. Admittedly, this process is challenging and anything but trivial. It requires a significant amount of time and resources, which might not always be available, and multidisciplinary collaboration among experts in different fields, such as engineering, computer science, behavioral science, and medicine, which might not be straightforward. However, such synergistic collaborative approaches are likely necessary in the development of evidence-based, sustainable, and impactful digital health interventions [44,45].

Limitations

We used the term *machine learning*, along with broader terms such as *data mining* and *artificial intelligence*, for our literature search, rather than keywords for specific machine learning algorithms or domains relevant to digital health, such as *telemedicine*. This might have inadvertently omitted studies that could have contributed to the progress made in machine learning applications for digital health. We combined the aforementioned terms with the generic term *health*, aiming to conduct a broad search within the provided boundaries and to include the most pertinent articles relevant to digital health. We searched for articles in a limited number of databases (ie, PubMed and Scopus), which nevertheless are two of the most widely used databases internationally [46]. We did not hand search any studies reported in other reviews or the included studies, and we did not assess the interrater reliability. A meta-analysis was not possible due to the heterogeneity of the included studies.

Conclusion

Our review showed that real-life digital health interventions incorporating machine learning can be useful and effective. Considering the small number of studies examined in this review and their limitations, further evidence of the clinical usefulness of machine learning in health service delivery is needed. We encourage researchers to move beyond the retrospective validation of machine learning models, by integrating their models within appropriately designed digital health tools and evaluating their tools in rigorous studies conducted in real-life settings.

Acknowledgments

The study was supported by the Wellcome Trust through Centre Grant No. 098461/Z/12/Z from the University of Oxford Sleep & Circadian Neuroscience Institute (SCNi). It was also supported by Health Data Research UK, which receives funding from HDR UK Ltd (HDR-5012) funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), the British Heart Foundation, and the Wellcome Trust. The funders had no role in the study or the decision to submit this work to be considered for publication.

Authors' Contributions

AKT was responsible for conducting the study; AKT and AT reviewed the literature and assessed the quality of the included studies; AKT synthesized the literature according to the described methodology; and AKT wrote the first draft of the manuscript and AT contributed to the final version. Both authors read and agreed to submit the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA checklist.

[\[PDF File \(Adobe PDF File\), 500KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Characteristics of included studies and implications for clinical practice.

[\[PDF File \(Adobe PDF File\), 40KB-Multimedia Appendix 2\]](#)

References

1. Widmer RJ, Collins NM, Collins CS, West CP, Lerman LO, Lerman A. Digital health interventions for the prevention of cardiovascular disease: a systematic review and meta-analysis. *Mayo Clin Proc* 2015 Apr;90(4):469-480. [doi: [10.1016/j.mayocp.2014.12.026](https://doi.org/10.1016/j.mayocp.2014.12.026)] [Medline: [25841251](https://pubmed.ncbi.nlm.nih.gov/25841251/)]
2. Triantafyllidis A, Velardo C, Chantler T, Shah SA, Paton C, Khorshidi R, SUPPORT-HF Investigators. A personalised mobile-based home monitoring system for heart failure: The SUPPORT-HF Study. *Int J Med Inform* 2015 Oct;84(10):743-753. [doi: [10.1016/j.ijmedinf.2015.05.003](https://doi.org/10.1016/j.ijmedinf.2015.05.003)] [Medline: [26037921](https://pubmed.ncbi.nlm.nih.gov/26037921/)]
3. Warmerdam L, Smit F, van Straten A, Riper H, Cuijpers P. Cost-utility and cost-effectiveness of internet-based treatment for adults with depressive symptoms: randomized trial. *J Med Internet Res* 2010;12(5):e53 [FREE Full text] [doi: [10.2196/jmir.1436](https://doi.org/10.2196/jmir.1436)] [Medline: [21169166](https://pubmed.ncbi.nlm.nih.gov/21169166/)]
4. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. *Am J Prev Med* 2016 Nov;51(5):843-851. [doi: [10.1016/j.amepre.2016.06.008](https://doi.org/10.1016/j.amepre.2016.06.008)] [Medline: [27745684](https://pubmed.ncbi.nlm.nih.gov/27745684/)]
5. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. *N Engl J Med* 2016 Sep 29;375(13):1216-1219 [FREE Full text] [doi: [10.1056/NEJMp1606181](https://doi.org/10.1056/NEJMp1606181)] [Medline: [27682033](https://pubmed.ncbi.nlm.nih.gov/27682033/)]
6. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. *J Med Internet Res* 2016 Dec 16;18(12):e323 [FREE Full text] [doi: [10.2196/jmir.5870](https://doi.org/10.2196/jmir.5870)] [Medline: [27986644](https://pubmed.ncbi.nlm.nih.gov/27986644/)]
7. 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;15:104-116 [FREE Full text] [doi: [10.1016/j.csbj.2016.12.005](https://doi.org/10.1016/j.csbj.2016.12.005)] [Medline: [28138367](https://pubmed.ncbi.nlm.nih.gov/28138367/)]
8. Forsyth AW, Barzilay R, Hughes KS, Lui D, Lorenz KA, Enzinger A, et al. Machine learning methods to extract documentation of breast cancer symptoms from electronic health records. *J Pain Symptom Manage* 2018 Jun;55(6):1492-1499. [doi: [10.1016/j.jpainsymman.2018.02.016](https://doi.org/10.1016/j.jpainsymman.2018.02.016)] [Medline: [29496537](https://pubmed.ncbi.nlm.nih.gov/29496537/)]
9. Weng SF, Reys J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One* 2017;12(4):e0174944 [FREE Full text] [doi: [10.1371/journal.pone.0174944](https://doi.org/10.1371/journal.pone.0174944)] [Medline: [28376093](https://pubmed.ncbi.nlm.nih.gov/28376093/)]
10. Kelly J, Gooding P, Pratt D, Ainsworth J, Welford M, TARRIER N. Intelligent real-time therapy: harnessing the power of machine learning to optimise the delivery of momentary cognitive-behavioural interventions. *J Ment Health* 2012 Aug;21(4):404-414. [doi: [10.3109/09638237.2011.638001](https://doi.org/10.3109/09638237.2011.638001)] [Medline: [22251028](https://pubmed.ncbi.nlm.nih.gov/22251028/)]
11. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. *Cancer Inform* 2007 Feb 11;2:59-77 [FREE Full text] [Medline: [19458758](https://pubmed.ncbi.nlm.nih.gov/19458758/)]
12. Salvatore C, Cerasa A, Battista P, Gilardi MC, Quattrone A, Castiglioni I, Alzheimer's Disease Neuroimaging Initiative. Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. *Front Neurosci* 2015;9:307 [FREE Full text] [doi: [10.3389/fnins.2015.00307](https://doi.org/10.3389/fnins.2015.00307)] [Medline: [26388719](https://pubmed.ncbi.nlm.nih.gov/26388719/)]
13. Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO. Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. *IEEE Trans Biomed Eng* 2012 May;59(5):1264-1271. [doi: [10.1109/TBME.2012.2183367](https://doi.org/10.1109/TBME.2012.2183367)] [Medline: [22249592](https://pubmed.ncbi.nlm.nih.gov/22249592/)]

14. Triantafyllidis A, Filos D, Buys R, Claes J, Cornelissen V, Kouidi E, et al. Computerized decision support for beneficial home-based exercise rehabilitation in patients with cardiovascular disease. *Comput Methods Programs Biomed* 2018 Aug;162:1-10. [doi: [10.1016/j.cmpb.2018.04.030](https://doi.org/10.1016/j.cmpb.2018.04.030)] [Medline: [29903475](https://pubmed.ncbi.nlm.nih.gov/29903475/)]
15. Contreras I, Vehi J. Artificial intelligence for diabetes management and decision support: literature review. *J Med Internet Res* 2018 May 30;20(5):e10775 [FREE Full text] [doi: [10.2196/10775](https://doi.org/10.2196/10775)] [Medline: [29848472](https://pubmed.ncbi.nlm.nih.gov/29848472/)]
16. Miotto R, Li L, Kidd BA, Dudley JT. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Sci Rep* 2016 Dec 17;6:26094 [FREE Full text] [doi: [10.1038/srep26094](https://doi.org/10.1038/srep26094)] [Medline: [27185194](https://pubmed.ncbi.nlm.nih.gov/27185194/)]
17. Agboola S, Hale TM, Masters C, Kvedar J, Jethwani K. "Real-world" practical evaluation strategies: a review of telehealth evaluation. *JMIR Res Protoc* 2014;3(4):e75 [FREE Full text] [doi: [10.2196/resprot.3459](https://doi.org/10.2196/resprot.3459)] [Medline: [25524892](https://pubmed.ncbi.nlm.nih.gov/25524892/)]
18. Armijo-Olivo S, Stiles CR, Hagen NA, Biondo PD, Cummings GG. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *J Eval Clin Pract* 2012 Feb;18(1):12-18. [doi: [10.1111/j.1365-2753.2010.01516.x](https://doi.org/10.1111/j.1365-2753.2010.01516.x)] [Medline: [20698919](https://pubmed.ncbi.nlm.nih.gov/20698919/)]
19. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)] [Medline: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)]
20. Mohammadi E, Thelwall M, Haustein S, Larivière V. Who reads research articles? An altmetrics analysis of Mendeley user categories. *J Assoc Inf Sci Technol* 2015 Apr 29;66(9):1832-1846. [doi: [10.1002/asi.23286](https://doi.org/10.1002/asi.23286)]
21. Sadasivam RS, Borglund EM, Adams R, Marlin BM, Houston TK. Impact of a collective intelligence tailored messaging system on smoking cessation: the Perspect randomized experiment. *J Med Internet Res* 2016 Nov 08;18(11):e285 [FREE Full text] [doi: [10.2196/jmir.6465](https://doi.org/10.2196/jmir.6465)] [Medline: [27826134](https://pubmed.ncbi.nlm.nih.gov/27826134/)]
22. Manuvinakurike R, Velicer WF, Bickmore TW. Automated indexing of Internet stories for health behavior change: weight loss attitude pilot study. *J Med Internet Res* 2014;16(12):e285 [FREE Full text] [doi: [10.2196/jmir.3702](https://doi.org/10.2196/jmir.3702)] [Medline: [25491243](https://pubmed.ncbi.nlm.nih.gov/25491243/)]
23. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. *J Med Internet Res* 2011;13(3):e55 [FREE Full text] [doi: [10.2196/jmir.1838](https://doi.org/10.2196/jmir.1838)] [Medline: [21840837](https://pubmed.ncbi.nlm.nih.gov/21840837/)]
24. Hawley MS, Cunningham SP, Green PD, Enderby P, Palmer R, Sehgal S, et al. A voice-input voice-output communication aid for people with severe speech impairment. *IEEE Trans Neural Syst Rehabil Eng* 2013 Jan;21(1):23-31. [doi: [10.1109/TNSRE.2012.2209678](https://doi.org/10.1109/TNSRE.2012.2209678)] [Medline: [22875259](https://pubmed.ncbi.nlm.nih.gov/22875259/)]
25. Martin CM, Vogel C, Grady D, Zarabzadeh A, Hederman L, Kellett J, et al. Implementation of complex adaptive chronic care: the Patient Journey Record system (PaJR). *J Eval Clin Pract* 2012 Dec;18(6):1226-1234. [doi: [10.1111/j.1365-2753.2012.01880.x](https://doi.org/10.1111/j.1365-2753.2012.01880.x)] [Medline: [22816797](https://pubmed.ncbi.nlm.nih.gov/22816797/)]
26. Morrison LG, Hargood C, Pejovic V, Geraghty AWA, Lloyd S, Goodman N, et al. The effect of timing and frequency of push notifications on usage of a smartphone-based stress management intervention: an exploratory trial. *PLoS One* 2017;12(1):e0169162 [FREE Full text] [doi: [10.1371/journal.pone.0169162](https://doi.org/10.1371/journal.pone.0169162)] [Medline: [28046034](https://pubmed.ncbi.nlm.nih.gov/28046034/)]
27. Ortiz-Catalan M, Guðmundsdóttir RA, Kristoffersen MB, Zepeda-Echavarría A, Caine-Winterberger K, Kulbacka-Ortiz K, et al. Phantom motor execution facilitated by machine learning and augmented reality as treatment for phantom limb pain: a single group, clinical trial in patients with chronic intractable phantom limb pain. *Lancet* 2016 Dec 10;388(10062):2885-2894. [doi: [10.1016/S0140-6736\(16\)31598-7](https://doi.org/10.1016/S0140-6736(16)31598-7)] [Medline: [27916234](https://pubmed.ncbi.nlm.nih.gov/27916234/)]
28. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized nutrition by prediction of glycemic responses. *Cell* 2015 Nov 19;163(5):1079-1094 [FREE Full text] [doi: [10.1016/j.cell.2015.11.001](https://doi.org/10.1016/j.cell.2015.11.001)] [Medline: [26590418](https://pubmed.ncbi.nlm.nih.gov/26590418/)]
29. Alonso SG, de la Torre-Díez I, Hamrioui S, López-Coronado M, Barreno DC, Nozaleda LM, et al. Data mining algorithms and techniques in mental health: a systematic review. *J Med Syst* 2018 Jul 21;42(9):161. [doi: [10.1007/s10916-018-1018-2](https://doi.org/10.1007/s10916-018-1018-2)] [Medline: [30030644](https://pubmed.ncbi.nlm.nih.gov/30030644/)]
30. Librenza-García D, Kotzian BJ, Yang J, Mwangi B, Cao B, Pereira Lima LN, et al. The impact of machine learning techniques in the study of bipolar disorder: a systematic review. *Neurosci Biobehav Rev* 2017 Sep;80:538-554. [doi: [10.1016/j.neubiorev.2017.07.004](https://doi.org/10.1016/j.neubiorev.2017.07.004)] [Medline: [28728937](https://pubmed.ncbi.nlm.nih.gov/28728937/)]
31. Orrù G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci Biobehav Rev* 2012 Apr;36(4):1140-1152. [doi: [10.1016/j.neubiorev.2012.01.004](https://doi.org/10.1016/j.neubiorev.2012.01.004)] [Medline: [22305994](https://pubmed.ncbi.nlm.nih.gov/22305994/)]
32. Brinker TJ, Hekler A, Utikal JS, Grabe N, Schadendorf D, Klode J, et al. Skin cancer classification using convolutional neural networks: systematic review. *J Med Internet Res* 2018 Oct 17;20(10):e11936 [FREE Full text] [doi: [10.2196/11936](https://doi.org/10.2196/11936)] [Medline: [30333097](https://pubmed.ncbi.nlm.nih.gov/30333097/)]
33. Goldstein BA, Pencina MJ, Montez-Rath ME, Winkelmayer WC. Predicting mortality over different time horizons: which data elements are needed? *J Am Med Inform Assoc* 2017 Dec;24(1):176-181 [FREE Full text] [doi: [10.1093/jamia/ocw057](https://doi.org/10.1093/jamia/ocw057)] [Medline: [27357832](https://pubmed.ncbi.nlm.nih.gov/27357832/)]
34. Xiao C, Choi E, Sun J. Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. *J Am Med Inform Assoc* 2018 Oct 01;25(10):1419-1428 [FREE Full text] [doi: [10.1093/jamia/ocy068](https://doi.org/10.1093/jamia/ocy068)] [Medline: [29893864](https://pubmed.ncbi.nlm.nih.gov/29893864/)]

35. Dunkley AJ, Bodicoat DH, Greaves CJ, Russell C, Yates T, Davies MJ, et al. Diabetes prevention in the real world: effectiveness of pragmatic lifestyle interventions for the prevention of type 2 diabetes and of the impact of adherence to guideline recommendations: a systematic review and meta-analysis. *Diabetes Care* 2014 Apr;37(4):922-933. [doi: [10.2337/dc13-2195](https://doi.org/10.2337/dc13-2195)] [Medline: [24652723](https://pubmed.ncbi.nlm.nih.gov/24652723/)]
36. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR Ment Health* 2016;3(1):e7 [FREE Full text] [doi: [10.2196/mental.4984](https://doi.org/10.2196/mental.4984)] [Medline: [26932350](https://pubmed.ncbi.nlm.nih.gov/26932350/)]
37. Aziz Z, Absetz P, Oldroyd J, Pronk NP, Oldenburg B. A systematic review of real-world diabetes prevention programs: learnings from the last 15 years. *Implement Sci* 2015 Dec 15;10:172 [FREE Full text] [doi: [10.1186/s13012-015-0354-6](https://doi.org/10.1186/s13012-015-0354-6)] [Medline: [26670418](https://pubmed.ncbi.nlm.nih.gov/26670418/)]
38. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. *Ann Behav Med* 2003 Aug;26(1):1-7. [Medline: [12867348](https://pubmed.ncbi.nlm.nih.gov/12867348/)]
39. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol Bull* 2007 Jul;133(4):673-693. [doi: [10.1037/0033-2909.133.4.673](https://doi.org/10.1037/0033-2909.133.4.673)] [Medline: [17592961](https://pubmed.ncbi.nlm.nih.gov/17592961/)]
40. World Health Organization. *Monitoring and Evaluating Digital Health Interventions: A Practical Guide to Conducting Research and Assessment*. Geneva, Switzerland: WHO; 2016.
41. Claes J, Buys R, Woods C, Briggs A, Geue C, Aitken M, et al. PATHway I: design and rationale for the investigation of the feasibility, clinical effectiveness and cost-effectiveness of a technology-enabled cardiac rehabilitation platform. *BMJ Open* 2017 Jun 30;7(6):e016781 [FREE Full text] [doi: [10.1136/bmjopen-2017-016781](https://doi.org/10.1136/bmjopen-2017-016781)] [Medline: [28667228](https://pubmed.ncbi.nlm.nih.gov/28667228/)]
42. Michelson D, Davenport C, Dretzke J, Barlow J, Day C. Do evidence-based interventions work when tested in the “real world?” A systematic review and meta-analysis of parent management training for the treatment of child disruptive behavior. *Clin Child Fam Psychol Rev* 2013 Mar;16(1):18-34. [doi: [10.1007/s10567-013-0128-0](https://doi.org/10.1007/s10567-013-0128-0)] [Medline: [23420407](https://pubmed.ncbi.nlm.nih.gov/23420407/)]
43. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *J Med Internet Res* 2018 Jun 06;20(6):e199 [FREE Full text] [doi: [10.2196/jmir.9275](https://doi.org/10.2196/jmir.9275)] [Medline: [29875089](https://pubmed.ncbi.nlm.nih.gov/29875089/)]
44. Buijink AWG, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. *Evid Based Med* 2013 Jun;18(3):90-92. [doi: [10.1136/eb-2012-100885](https://doi.org/10.1136/eb-2012-100885)] [Medline: [22923708](https://pubmed.ncbi.nlm.nih.gov/22923708/)]
45. Cao L, Ramesh B. Agile requirements engineering practices: an empirical study. *IEEE Softw* 2008 Jan;25(1):60-67. [doi: [10.1109/MS.2008.1](https://doi.org/10.1109/MS.2008.1)]
46. Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *FASEB J* 2008 Feb;22(2):338-342. [doi: [10.1096/fj.07-9492LSF](https://doi.org/10.1096/fj.07-9492LSF)] [Medline: [17884971](https://pubmed.ncbi.nlm.nih.gov/17884971/)]

Abbreviations

EPHPP: Effective Public Health Practice Project

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by CL Parra-Calderón; submitted 21.09.18; peer-reviewed by K Cato, J Claes, M Larsen; comments to author 12.11.18; revised version received 07.01.19; accepted 26.01.19; published 05.04.19

Please cite as:

Triantafyllidis AK, Tsanas A

Applications of Machine Learning in Real-Life Digital Health Interventions: Review of the Literature

J Med Internet Res 2019;21(4):e12286

URL: <https://www.jmir.org/2019/4/e12286/>

doi: [10.2196/12286](https://doi.org/10.2196/12286)

PMID: [30950797](https://pubmed.ncbi.nlm.nih.gov/30950797/)

©Andreas K Triantafyllidis, Athanasios Tsanas. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 05.04.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.