
Review

Predicting Adherence to Behavior Change Support Systems Using Machine Learning: Systematic Review

Akon Obu Ekpezu¹, MPhil; Isaac Wiafe², PhD; Harri Oinas-Kukkonen¹, PhD

¹Oulu Advanced Research on Service and Information Systems, Department of Information Processing Science, University of Oulu, Oulu, Finland

²Department of Computer Science, University of Ghana, Accra, Ghana

Corresponding Author:

Akon Obu Ekpezu, MPhil
Oulu Advanced Research on Service and Information Systems
Department of Information Processing Science
University of Oulu
Pentti Kaiteran Katu 1 Linnanmaa
Oulu, 90570
Finland
Phone: 358 468860704
Email: akon.ekpezu@oulu.fi

Abstract

Background: There is a dearth of knowledge on reliable adherence prediction measures in behavior change support systems (BCSSs). Existing reviews have predominately focused on self-reporting measures of adherence. These measures are susceptible to overestimation or underestimation of adherence behavior.

Objective: This systematic review seeks to identify and summarize trends in the use of machine learning approaches to predict adherence to BCSSs.

Methods: Systematic literature searches were conducted in the Scopus and PubMed electronic databases between January 2011 and August 2022. The initial search retrieved 2182 journal papers, but only 11 of these papers were eligible for this review.

Results: A total of 4 categories of adherence problems in BCSSs were identified: adherence to digital cognitive and behavioral interventions, medication adherence, physical activity adherence, and diet adherence. The use of machine learning techniques for real-time adherence prediction in BCSSs is gaining research attention. A total of 13 unique supervised learning techniques were identified and the majority of them were traditional machine learning techniques (eg, support vector machine). Long short-term memory, multilayer perception, and ensemble learning are currently the only advanced learning techniques. Despite the heterogeneity in the feature selection approaches, most prediction models achieved good classification accuracies. This indicates that the features or predictors used were a good representation of the adherence problem.

Conclusions: Using machine learning algorithms to predict the adherence behavior of a BCSS user can facilitate the reinforcement of adherence behavior. This can be achieved by developing intelligent BCSSs that can provide users with more personalized, tailored, and timely suggestions.

(JMIR AI 2023;2:e46779) doi: [10.2196/46779](https://doi.org/10.2196/46779)

KEYWORDS

adherence; compliance; behavior change support systems; persuasive systems; persuasive technology; machine learning

Introduction

Behavior change support systems (BCSSs) have been effective in improving health and healthier lifestyles. These are persuasive systems that have been designed to change behavior without force or deception [1]. However, the effectiveness of these systems is generally hindered by nonadherence [2-4]. Nonadherence to recommended regimes in BCSSs has the

potential to diminish their long-term benefits [5]. It is associated with the increased prevalence of diseases such as hypertension, diabetes, obesity, dementia, bipolar disorder, and heart failure [2,4,6-8], as well as the increased cost of health care. Yet, there are no standardized factors that can reliably predict adherence [9,10]. Direct adherence monitoring approaches are expensive, burdensome to care providers, and susceptible to distortion by patients, while indirect monitoring approaches such as pill count,

patient questionnaires, electronic medication monitors, or electronic reporting of daily physical activity are susceptible to misinterpretations and overestimation of adherence [11,12]. To implement effective BCSSs and ensure positive behavior change outcomes that can be attributed to the recommended interventions, an accurate assessment of adherence behaviors and their predictors has become imperative. This will guide researchers and health care providers in identifying nonadherent individuals as well as provide measures that will re-engage and help them to adhere [13]. Additionally, an early prediction of user dropout or relapse during interventions may suggest measures that can be used to improve adherence [14].

Existing systematic reviews [2,7,15-19] have sought to examine predictors or determinants of adherence to several BCSSs. They predominately report that there is a lack of consistency regarding reports of adherence, key variables mediating adherence, and reliable measures of adherence. However, findings from these reviews were based on studies that relied solely on self-reported measures of adherence using pharmacological claims and validated questionnaires from behavior change and health psychology theories. Hence, abounding issues of over- and underreporting may limit the validity of the findings.

This review enhances existing knowledge by focusing on predictors of adherence to BCSSs using machine learning techniques. Machine learning techniques have enabled a proficient means of classifying, detecting, and predicting complex phenomena including human behavior. It has also attracted considerable research interest in the development of BCSSs [20-22]. Nonetheless, literature on the use of machine learning techniques as adherence prediction methods in BCSSs is limited [13,23]. Although Bohlmann et al [23] provided literature summaries on machine learning techniques for predicting adherence, they focused on medication adherence only and considered both digital and nondigital interventions. In contrast to previous reviews, this systematic review focuses on the use of machine learning approaches to predict all kinds of adherence problems in BCSSs. In addition, it focuses only on primary studies that used objectively collected data or data generated by the BCSS. Accordingly, this review seeks to answer the following question: What are the existing trends in the use of machine learning techniques to predict adherence to BCSSs? Specifically, this study answers 4 main review questions, as shown in Table 1.

Table 1. Review questions (RQs) and their motivations.

RQ	Question	Motivation
RQ1	What are the targeted adherence problems and their related definitions?	Research on adherence has predominately focused on adherence to medication and pharmacological treatments. However, adherence covers a wider range of health behaviors than medication adherence [9]. This RQ sought to identify other target adherence problems in BCSSs ^a .
RQ2	What are the characteristics of the BCSS including persuasive system features?	Considering the variabilities in adherence problems and BCSSs, this RQ aimed to provide summaries on the characteristics of the BCSS and the persuasive system features that have been used to improve adherence.
RQ3	What are the adopted machine learning approaches in predicting adherence to BCSSs?	This RQ sought to identify the nature of the raw data and predominately used machine learning techniques, feature selection techniques, and performance metrics.
RQ4	What are the limitations or barriers to adherence?	Though various barriers to adherence have been identified in the literature, this RQ sought to identify only those barriers that limit individuals from adhering to the request of the BCSS.

^aBCSS: behavior change support system.

Methods

Literature Search

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) approach, a search on Scopus and PubMed electronic databases was conducted. This search aimed to identify peer-reviewed English conference and journal papers published between January 2011 and August 2022. Scopus indexes a larger number of peer-reviewed scientific journals than the Web of Science and offers results of more consistent accuracy than Google Scholar, while PubMed remains a leading database in biomedical research [24]. Including papers published within the past decade will reveal recent evidence-based research trends [25]. Using the logical OR/AND operators, the search phrases were a combination of keywords related to prediction, adherence, health behavior change interventions, and machine learning (See Multimedia Appendix 1 for the search phrases). Considering the plethora of approaches

to investigate adherence, the search for eligible studies was not limited to a specific study design.

Only empirical studies that described the development and testing of machine learning models for BCSS adherence prediction were considered. Studies that used only self-reported data, were not reported in English, or did not focus on human participants were ignored.

Study Selection

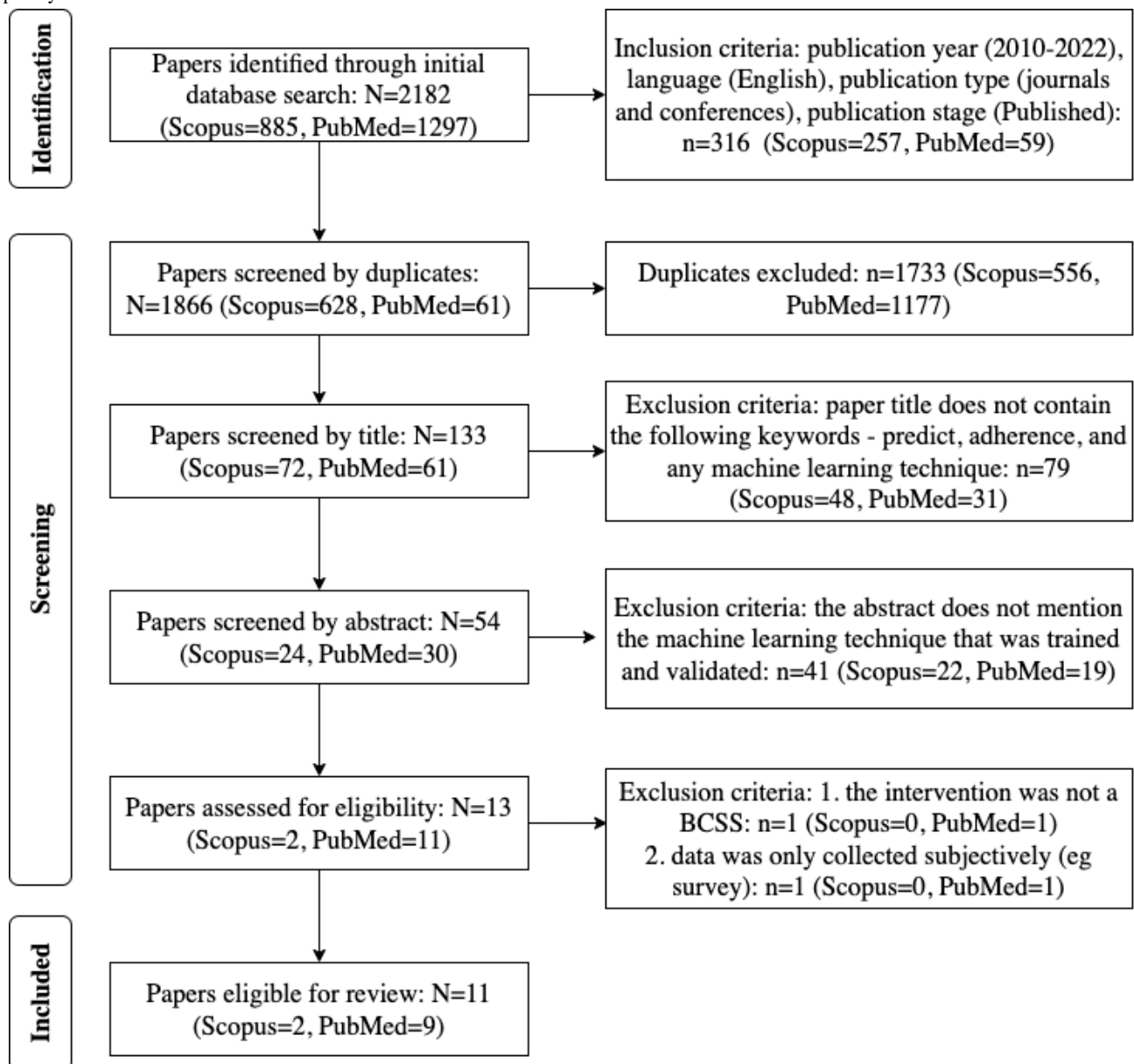
During the initial search of the databases, 2182 papers were retrieved. The results were refined by year, document type, publication stage, source type, and language, resulting in 1866 papers. The exported papers were screened for uniqueness and for titles containing keywords such as adherence, prediction, and any machine learning technique. Of these, 1812 papers were excluded. The remaining 54 papers were screened by abstracts and full texts. Papers were excluded by abstract if the machine learning technique(s) were not mentioned. Furthermore, papers

were excluded by full text if the intervention was not characterized by a BCSS (any form of information system that has been developed to change human behavior voluntarily). Finally, 11 journal papers were considered eligible for this systematic review and thus downloaded for methodological quality assessment.

Figure 1 shows the PRISMA flow diagram of study identification and selection. Since the study selection was not

limited to a specific study design, the Mixed Methods Appraisal Tool by [26] was used to assess the methodological quality of the downloaded papers (see Multimedia Appendix 2 [13,14,26-35]). Accordingly, 11 studies were identified to be of high methodological quality. Pertinent information on the study characteristics and machine learning approaches was extracted using a data extraction form in Microsoft Excel created by the authors. Multimedia Appendix 3 [13,14,27-35] presents a list of included studies.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for study selection. BCSS: behavior change support systems.



Results

All included studies (N=11) were primarily aimed at the development and use of machine learning techniques for the prediction of adherence [27-32], nonadherence, dropout, or relapse [13,14,33-35]. The ensuing sections will elaborate on findings related to the review questions (RQs).

Targeted Adherence Problems and Their Related Definitions (RQ1)

Overview

This review identified 7 health behaviors that BCSSs target: medication adherence [29,31,34], use of health care systems [13,28], physical activity [14,27], diet [33], illicit drug use [30], depression and anxiety [32], and insomnia [35]. Based on the characteristics of these health behaviors, they were grouped into

4 categories of adherence problems, described in the following sections.

Adherence to Digital Cognitive or Behavioral Interventions (n=5)

This category includes health behaviors such as health care system use, illicit drug use, depression, anxiety, and insomnia. This adherence problem focused on predicting adherence or nonadherence to internet-based cognitive behavioral therapy [32,35], and remote health monitoring systems [13,28,30] using machine learning models. Adherence to digital cognitive or behavioral interventions refers to the successful completion of all recommended tasks and achievement of initial set goals. While nonadherence refers to ignoring or not completing the recommended task consecutively after signing up to use the BCSS.

Medication Adherence (n=3)

This adherence problem is linked to medication adherence behavior. Although extant literature posit that medication nonadherence is the most common form of adherence problem, this review identified 3 studies that addressed this problem within BCSSs. It compromises the effectiveness of treatment outcomes in about 85% of patients with chronic and acute medical conditions globally [11,36]. Studies in this category applied machine learning models to remote real-time measurements of medication dosing [29,31,34]. Though these studies had different thresholds for defining medication adherence, it generally referred to a patient's behavior or commitment to taking the medications as prescribed by a physician with an average adherence rate of 80% and above.

Physical Activity Adherence (n=2)

Adhering to physical activity routines has the potential to reduce the risk of chronic diseases irrespective of age or other sociodemographic factors. Whereas some individuals find it difficult to regularly engage in or continue a physical activity routine [14], others discontinue when they have achieved a health or body goal [27]. These studies were observed to have

varying definitions of adherence. For instance, Zhou et al [14] considered an increase in the number of steps over time, while Bastidas et al [27] considered the users' app use patterns. Thus, physical activity adherence may be defined as either a consistent increase in physical activity levels compared to an individual's baseline activity levels or an individual's responsiveness to prompts from the app. These 2 definitions describe behavior compliance and program compliance, respectively [37].

Diet Adherence (n=1)

This was observed in only 1 study [33]. Dietary relapse in a weight loss intervention was predicted. Dietary relapse refers to any instance in which a person exceeded a specified meal or snack point threshold (per meal).

Characteristics of the BCSSs and Persuasive System Features (RQ2)

The BCSSs included mobile apps [13,14,27,28,33,34], web-based apps [30,32,33,35], and sensor-based systems plus mobile apps [29,31]. They were targeted at different groups of people, namely physically inactive women, illicit drug users, obese or overweight people, and patients with a wide range of chronic diseases, such as heart failure, myocardial infarction-anxiety, depression (MI-ANXDEP), insomnia, and Parkinson disease. [Multimedia Appendix 4](#) [14,15,29-44] describes other study-specific characteristics.

The BCSSs leveraged some behavior change techniques and persuasive systems features to improve user adherence. These features were extracted and evaluated using the persuasive systems design (PSD) model [45]. The PSD model has been validated in several studies [37] and is predominately used in the design and evaluation of BCSSs [46]. The model consists of 28 system features that make up 4 categories of persuasive principles (namely primary task support, dialogue support, credibility support, and social support). [Table 2](#) displays the frequency of the PSD features represented in the BCSS. All the studies had primary task support features, 8 studies had dialogue support features, 5 studies had credibility support features, and only 1 study had social support features.

Table 2. Persuasive features identified in the behavior change support systems (BCSSs). Check marks indicate that the feature was identified.

Persuasive features	Studies predicting adherence						Studies predicting nonadherence					Total
	[30]	[28]	[32]	[29]	[31]	[27]	[13]	[35]	[33]	[34]	[14]	
Primary task support												
Personalization	✓		✓	✓	✓	✓	✓	✓		✓	✓	8/11
Self-monitoring		✓				✓	✓				✓	4/11
Reduction				✓		✓			✓	✓		4/11
Rehearsal					✓							1/11
Tailoring	✓											1/11
Dialogue support												
Reminders		✓		✓				✓	✓	✓		5/11
Suggestions							✓				✓	2/11
Praise	✓										✓	2/11
Rewards									✓			1/11
Similarity	✓											1/11
Liking								✓				1/11
Credibility support												
Real world feel	✓					✓		✓		✓	✓	5/11
Expertise	✓					✓		✓				3/11
Verifiability	✓							✓		✓		3/11
Third party endorsement	✓											1/11
Social support												
Social facilitation	✓						✓					2/11

Primary task support simplifies and motivates users to perform recommended tasks (eg, exercise). A total of 5 features of the primary task support principle were used: personalization, self-monitoring, reduction, rehearsals, and tailoring. Personalization was the most used feature in this category. It delivers personalized content to the users. For example, artificial intelligence generated workouts according to an individual's characteristics and preferences. Self-monitoring enables app users to view and track their activity levels and health status in real-time (eg, the app enables users to monitor, visualize, and track activity levels and calories burned in real time). Reduction breaks down tasks such as daily point goals into specific meal or snack targets. The least used features in this category were rehearsal (practicing the target behavior, eg, gait movements) and tailoring (eg, the app provided content that was distinct to users of specific age groups and health goals, such as alcohol or smoking cessation, weight loss, or mental health).

Dialogue support provides a means to help users to achieve their goals via human-computer interactions. The dialogue support features included reminders (eg, medication prompts), suggestions (eg, the app advises users based on their input to the app), and praise (automated feedback on the completion of a task). The least used features included rewards (eg, point-based incentives), similarity (eg, therapy resembling traditional cognitive behavioral therapy), and liking (eg, user-friendly and appealing design).

System credibility support provides a means for users to trust the system. Features identified in this category included expertise (app provided theory-based information and were designed to improve engagement, effectiveness, and security), verifiability (app provided links to related sites), third-party endorsement (from the National Institute of Health), and real-world feel. The real-world feel feature was implemented as "Contact Us" (a means to communicate with the developers of the app) and "About Us" (providing information about the developers of the app).

Social support provides a means of supporting users via social influence. However, it was the least used principle. Social facilitation was the only identified social support feature and it was implemented by allowing user participation in online app forums. Perhaps, the minimal use of social support features may be attributed to the negative sentiments associated with some of its features [38].

It is important to note that the effectiveness of these persuasive system features in improving adherence was not explicitly evaluated in the included studies. However, these features may have directly or indirectly improved adherence rates. Some of the studies used behavior change techniques such as goal-setting, web-based human coaching, face-to-face counseling sessions, and feedback from psychologists and expert program providers to improve adherence behavior. However, this review could not identify behavioral theories or models upon which these techniques were based.

Machine Learning Approaches (RQ3)

Developing machine learning models for predicting user adherence or nonadherence was the general aim of all the included studies. This process conventionally consists of 4 main stages: data collection, feature selection, model training, and model validation.

Data Collection

Apart from the gait-related data, which were collected in a controlled environment (laboratory), the data used to train or test the machine learning models in the majority of the studies were objectively collected by the health app while the study participants were performing the behavior of interest in an uncontrolled environment. For example, data such as log data, training behavior, walking steps, and 3D movement scans were automatically extracted in a contactless manner without any form of self-report from the users.

Studies on medication adherence were observed to use images and videos to capture participants' medication adherence behavior. The apps used technologies such as computer vision [34], internet-connected smart sharp bin [29], and flight sensors [31] for the real-time monitoring of self-administered injections and medication ingestion. Time-stamped data of injection needles discarded into the smart sharp bin, time-stamped skeletal joint data, and images of the participants taking the medication were retrieved, validated, and then used to generate the data set for training or testing the model. Similarly, studies on physical activity adherence [14,27] used continuously collected time-dependent physical activity data from app users to develop machine learning models.

Studies on adherence to digital cognitive or behavioral interventions [13,28,30,32,35] and dietary adherence [33] used a combination of objectively collected data and participants' responses to questions on self-assessment delivered by the health app. Goldstein et al [33] used a BCSS that asked predefined questions related to triggers of dietary relapse for the analysis.

Considering that objectively collecting trigger-related data or self-assessment data on health symptoms from a mobile app may currently be challenging as triggers (such as food cravings and hunger) are physiologically motivated, future studies on BCSSs that seek to extract trigger-related data may consider using physiological sensor data. This is a noninvasive approach to detecting hunger and cravings using wearable body sensors [39]. Such sensor data may also be integrated into the health app to enable self-monitoring.

Feature Selection or Engineering

Among the 11 studies, 5 performed feature selection, 5 performed feature engineering, and 1 adopted features based on existing literature (see [Multimedia Appendix 5 \[14,15,29-37\]](#)). Feature selection and feature engineering were both aimed at enhancing model performance by eliminating irrelevant features and generating new features from raw data respectively. Due to the complexities of combining 2 or more machine learning techniques (ie, ensemble learning), some studies [28] applied more than 1 feature selection method. However, there were no differences in the selected features.

The flat features algorithm (including filter, wrapper, and embedded methods) were the predominately used feature selection method. This algorithm assumes all features to be independent [40]. Interestingly, each study had its own set of unique predictors irrespective of the category of adherence problem. [Multimedia Appendix 5 \[14,15,29-37\]](#) highlights the various feature selection approaches.

Model Training

The learning problem was a binary classification. Thus, there were 2 class labels (outcomes), namely adherence/nonadherence, adherers/nonadherers, relapse/nonrelapse, and dropout/nondropout. An overview of the adopted techniques and the outcomes of the best-performing techniques is provided in [Table 3](#).

Table 3. Identified machine learning and model validation techniques.

Ref	Machine learning techniques	Evaluation metrics	Predicted outcome
[30]	Logistic regression and random forest ^a	AUROC ^{a,b} , specificity, sensitivity, PPV ^c , NPV ^d , and confusion matrix	Successful or early dropout
[13]	Logistic regression, random forest ^a , and decision trees	Accuracy, precision ^a , and AUROC	Dropout or nondropout
[28]	Decision tree, MLP ^e , and KNN ^{a,f}	Precision, sensitivity, F_1 -score, TPR ^g , FPR ^h , and AUROC ^a	Adherers or nonadherers
[32]	Random forest ^a	Accuracy	Adherence or nonadherence
[35]	Logistic regression, SVM ⁱ , and decision trees (boosted) ^a	AUROC ^a , TPR, FPR, and PRAUC ^j	Dropout or nondropout
[33]	Ensemble methods ^a	Accuracy, sensitivity ^a , specificity, and AUROC ^a	Relapse or not
[29]	XGB ^k , extra trees, random forest, MLP, gradient tree boosting, recurrent neural network, and LSTM ^{a,l}	Accuracy, specificity ^a , sensitivity, precision, F_1 -score, and AUROC	Adherence or nonadherence
[34]	XGB ^a	Accuracy, precision ^a , sensitivity, AUROC, TPR, and FPR	Adherence or nonadherence
[31]	Decision trees ^a , KNN, naive Bayes, SVM, and random forest	Confusion matrix	Adherence or nonadherence
[27]	LSTM ^a and SVM	Accuracy, sensitivity, F_1 -score, and confusion matrix	Adherent or nonadherent
[14]	Logistic regression ^a and SVM	Accuracy, sensitivity, specificity, and AUROC	Relapse or not

^aBest performing machine learning technique or most relevant metric for the outcome prediction.

^bAUROC: area under the receiver operating characteristic curve.

^cPPV: positive predictive value.

^dNPV: negative predictive value.

^eMLP: multilayer perceptron.

^fKNN: k-nearest neighbor.

^gTPR: true positive rate.

^hFPR: false positive rate.

ⁱSVM: support vector machine.

^jPRAUC: precision-recall curve.

^kXGB: extreme gradient boost.

^lLSTM: long short-term memory.

A total of 13 supervised machine-learning techniques were used across the included studies. Logistic regression, support vector machines, and random forest were the most used techniques cutting across all 4 categories of the adherence problems. The machine learning techniques mapped to specific adherence problems included support vector machines for physical activity adherence; extreme gradient boosting, extra trees, recurrent neural network, naive Bayes, and gradient tree boosting for medication adherence; and ensemble methods for dietary adherence. Random forest was observed to be the predominant best-performing model in studies on adherence to digital cognitive or behavioral interventions, while long short-term memory (LSTM) was a common best-performing model between medication adherence and physical activity adherence. Overall, the predominant best-performing models across all included studies were random forest, decision trees, logistic regression, k-nearest neighbor, LSTM, and ensemble learning.

Model Validation

Owing to the relatively small and imbalanced data sets used in some of the included studies, cross-validation methods were adopted to eliminate bias that may occur during data split. The following cross-validation methods were identified: K(5)-fold cross-validation [29,34]; leave-one-out cross-validation [28,31,33]; and stratified K(10)-fold cross-validation [13,30,35].

Besides cross-validation methods, several performance metrics were used to compare and evaluate the performance of the various machine learning models. It was observed that the choice of performance metrics was dependent on the context of the study and more than 1 metric was used to evaluate the performance of a model (see Table 3). The predominately used metrics in order of frequency included area under the receiver operating characteristic curve (7/11), accuracy (6/11), sensitivity (6/11), specificity (4/11), precision (3/11), F_1 -score (3/11), true

positive rate (3/11), false positive rate (3/11), confusion matrix (3/11), positive predictive value (1/11), negative predictive value (1/11), and precision-recall curve (1/11).

Generally, it was observed that the prediction models had good classification accuracies. This was an indication that the features or predictors used in each of these studies were a good representation of the intervention domain. Nonetheless, due to the plethora of digital platforms, the interaction between technology and behavior may affect the generalizability of the results [33,34]. Thus, Koesmahargyo et al [34] posit replication and integration of data from various digital platforms.

Barriers to Adherence (RQ4)

Studies suggest that the rate of adherence may be affected by the following:

1. Achievement of set health or body goals
2. Issues associated with trust and the tolerability of the technology
3. The complexity of the system, and the mismatch between the system design and the needs and preferences of its users
4. Inappropriate timing for delivering or sending notifications or suggestions to the users; since these timings are usually not well chosen, they may either inconvenience the users when delivered or may not be effective in getting their attention
5. The insufficient open collaborative relationship between health app providers and the users. Prior studies [37] refer to this as a lack of accountability in adherence prediction models

These barriers may be classified into two nonadherent groups: intentional nonadherence (1 and 2), or unintentional nonadherence (3, 4, and 5).

Discussion

This systematic review provides an overview of existing trends in the use of machine learning techniques to predict adherence to different BCSSs. This was achieved by finding answers to a set of review questions using data extracted from the 11 included studies. The rest of this section will summarize and discuss findings based on the review results.

This review identified 4 categories of adherence problems: adherence to digital cognitive or behavioral interventions, medication adherence, physical activity adherence, and diet adherence. These problems collectively represent what Middleton et al [4] describe as an “adherence challenge.” However, when considering the taxonomy of key health behaviors [41], it was observed that the behaviors identified in this systematic review were not exhaustive. Consequently, the prediction of adherence to other health behaviors is an open research area for further investigation.

On the use of persuasive system design features, it was observed that primary task support features were the most used, while social support features were the least used. This finding is consistent with that of related systematic reviews [37]. Though the included studies claimed that either the implementation of the BCSS or the selected features (predictors) for the machine

learning algorithm was theory-based, this systematic review could not identify the behavior change theories adopted by the studies. Hence, it was not clear if the operating mechanisms of behavioral theories were considered in most of the included studies. Nonetheless, prior studies have provided evidence of the effectiveness of theory-based interventions and persuasive system features in promoting adherence behavior in BCSSs. Future studies should therefore be intentional about the use of these mechanisms as measures of improving adherence behavior.

The relevant predictors identified align with findings from existing literature. Similar to existing reviews [4,42], exercise history, intensity, and frequency emerged as relevant predictors of physical activity adherence. Exercise, fatigue, cognitive load, and confidence were the most relevant predictors of diet adherence, affirming previous findings (eg, [42]). While communication with or feedback/advice from the physician or health provider, fear, and patients’ cognitive capacity were the most relevant predictors of medication adherence as also found in past reviews (eg, [3,42]). Furthermore, some of the identified predictors of physical activity adherence and medication adherence affirm 2 viewpoints from social learning theory [43]: that individuals develop beliefs that they can perform the necessary tasks to obtain the desired outcome based on prior accomplishment of similar behaviors and verbal persuasion by credible sources. This systematic review identified the completion of homework assignments as a predictor of cognitive or behavioral intervention adherence, while Heesch et al [44] identified the same predictor for physical activity adherence. Furthermore, this review suggests that not all initially selected predictors or features of adherence are subsequently considered most relevant by the feature selection algorithms (see [Multimedia Appendix 5](#) [14,15,29-37]). Using multiple feature selection methods yields the same feature set. Future studies should consider incorporating the feature selection or engineering techniques identified in this review to enable a comparison of their results with the existing literature.

Most of the included studies used traditional machine learning techniques, with limited use of advanced learning techniques such as ensemble learning, reinforcement learning, and deep learning. Among the 13 supervised machine learning algorithms, only 2 were deep learning techniques (multilayer perceptron and recurrent neural network–LSTM), 1 ensemble, and zero reinforcement learning. Perhaps the sparing use of deep learning techniques may be attributed to the small sample sizes of these studies, considering that deep learning is more efficient in the analysis of huge amounts of data. For instance, LSTM may have been a more appropriate algorithm for the study by Evangelista et al [28], because the data collected captured changes in conditions that evolved slowly over time. However, it was not used probably due to a sample of only 14 participants. Interestingly, in a study with a large data set (342,174 injection historic drop data) [29], LSTM outperformed traditional machine learning models like random forest. Future studies should therefore consider using advanced learning methods instead of traditional learning techniques. Deep learning techniques can automate feature engineering or selection and extract complex and nonlinear patterns from data. Reinforcement learning is well suited for systems with inherent time delays where

decisions are evaluated by a long-term future reward and not an immediate knowledge of the effectiveness of a system [47]. In addition, since they learn by observing the results of their actions, they are applicable in study settings with scarce or varying data as found in BCSSs [22,47].

This systematic review observed that quite a small amount of data were used in most of the included studies. With each study participant treated as a single data point, data were extracted from as little as 12 study participants to as many as 7697 study participants. Although training a machine learning model requires a reasonable amount of data to train the model, the required sample size for training and producing a model with good generalizability is not well established [33,48]. Regardless, the included studies adopted suitable machine learning techniques, dimensionality reduction techniques, and evaluation methods that are designed to improve model performance irrespective of the sample size. For instance, Zhou et al [14] performed data augmentation on the training data.

Multiple metrics can be used to evaluate model performance (see Table 3). However, the choice of which metrics best measure the model performance depends on the nature of the problem, the researcher's understanding of the domain or problem, and the expected outcome of the study. For instance, Gu et al [29] prioritized predicting nonadherers (those who will not perform the recommended task on time), hence specificity (true negative rate) was a preferred metric. Considering that a wrong prediction may lead the health app provider to develop unnecessary persuasive strategies for the user, Pedersen et al [13] and Bastidas et al [27] aimed to reduce false negatives (ie,

participants at high risk of dropout are not identified as such), hence a high precision was a preferred metric. However, if the study's goal is to validate the hypothesis that machine learning methods can be used in predicting adherence [14,31] rather than to compare machine learning methods, then choosing the most appropriate metric becomes irrelevant.

A major research challenge reported in 9 of 11 of the included studies was the scarce and small-sized data sets and their effect on the generalizability and reliability of the research results. A specific study limitation pertained to collecting data in a controlled environment [31]. This method of data collection does not represent the entire range of user behavior in a free-living environment.

Conclusions

Findings from this systematic review indicate that though the use of machine learning techniques in the prediction of adherence to BCSSs is scarce and is only beginning to gain research interest, it has the potential to accurately predict adherence behavior in real time using objectively collected data. This systematic review is unique as it has not yet been reported in the literature, and it provides an overview of machine learning approaches in determining predictors of specific adherence problems in BCSSs. A grasp of these trends across different BCSSs will guide researchers in choosing appropriate features and machine learning techniques that favor the prediction of specific adherence problems. In summarizing findings from 11 journal papers, this systematic review highlights research gaps and areas for future research. It also acknowledges limitations that may exist due to the selection strategy for eligible studies.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search syntax and results.

[\[DOCX File , 18 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Multimedia quality assessment.

[\[DOCX File , 22 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

List of reviewed primary studies.

[\[DOCX File , 19 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Characteristics of the studies and the BCSSs.

[\[DOCX File , 18 KB-Multimedia Appendix 4\]](#)

Multimedia Appendix 5

Input data and feature selection approaches.

[\[DOCX File , 31 KB-Multimedia Appendix 5\]](#)

References

1. Oinas-Kukkonen H. A foundation for the study of behavior change support systems. *Pers Ubiquit Comput* 2012;17(6):1223-1235 [doi: [10.1007/s00779-012-0591-5](https://doi.org/10.1007/s00779-012-0591-5)]
2. Burgess E, Hassmén P, Pumpa KL. Determinants of adherence to lifestyle intervention in adults with obesity: a systematic review. *Clin Obes* 2017;7(3):123-135 [doi: [10.1111/cob.12183](https://doi.org/10.1111/cob.12183)] [Medline: [28296261](https://pubmed.ncbi.nlm.nih.gov/28296261/)]
3. Aziz F, Malek S, Mhd Ali A, Wong MS, Mosleh M, Milow P. Determining hypertensive patients' beliefs towards medication and associations with medication adherence using machine learning methods. *PeerJ* 2020;8:e8286 [FREE Full text] [doi: [10.7717/peerj.8286](https://doi.org/10.7717/peerj.8286)] [Medline: [32206445](https://pubmed.ncbi.nlm.nih.gov/32206445/)]
4. Middleton KR, Anton SD, Perri MG. Long-term adherence to health behavior change. *Am J Lifestyle Med* 2013;7(6):395-404 [FREE Full text] [doi: [10.1177/1559827613488867](https://doi.org/10.1177/1559827613488867)] [Medline: [27547170](https://pubmed.ncbi.nlm.nih.gov/27547170/)]
5. Aswad O, Lessard L. Service design methods' ability to personalize telehealth: a systematic literature review. 2021 Presented at: The 27th Annual Americas Conference on Information Systems (AMCIS 2021); August 9-13, 2021; Montreal, QC
6. Abdul-Razak S, Daher AM, Ramli AS, Ariffin F, Mazapuspavina MY, Ambigga KS, et al. Prevalence, awareness, treatment, control and socio demographic determinants of hypertension in Malaysian adults. *BMC Public Health* 2016;16:351 [FREE Full text] [doi: [10.1186/s12889-016-3008-y](https://doi.org/10.1186/s12889-016-3008-y)] [Medline: [27097542](https://pubmed.ncbi.nlm.nih.gov/27097542/)]
7. Patoz M, Hidalgo-Mazzei D, Pereira B, Blanc O, de Chazeron I, Murru A, et al. Patients' adherence to smartphone apps in the management of bipolar disorder: a systematic review. *Int J Bipolar Disord* 2021;9(1):19 [FREE Full text] [doi: [10.1186/s40345-021-00224-6](https://doi.org/10.1186/s40345-021-00224-6)] [Medline: [34081234](https://pubmed.ncbi.nlm.nih.gov/34081234/)]
8. Lie SS, Karlsen B, Oord ER, Graue M, Oftedal B. Dropout from an eHealth intervention for adults with type 2 diabetes: a qualitative study. *J Med Internet Res* 2017;19(5):e187 [FREE Full text] [doi: [10.2196/jmir.7479](https://doi.org/10.2196/jmir.7479)] [Medline: [28559223](https://pubmed.ncbi.nlm.nih.gov/28559223/)]
9. World Health Organization. In: Sabaté E, editor. *Adherence to Long-term Therapies: Evidence for Action*. Geneva, Switzerland: World Health Organization; 2003.
10. Vrijens B, De Geest S, Hughes DA, Przemyslaw K, Demonceau J, Ruppert T, et al. A new taxonomy for describing and defining adherence to medications. *Br J Clin Pharmacol* 2012;73(5):691-705 [FREE Full text] [doi: [10.1111/j.1365-2125.2012.04167.x](https://doi.org/10.1111/j.1365-2125.2012.04167.x)] [Medline: [22486599](https://pubmed.ncbi.nlm.nih.gov/22486599/)]
11. Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med* 2005;353(5):487-497 [doi: [10.1056/NEJMra050100](https://doi.org/10.1056/NEJMra050100)] [Medline: [16079372](https://pubmed.ncbi.nlm.nih.gov/16079372/)]
12. Flynn M, Hall EE. Prediction of adherence to a 9-week corporate wellness walking program. *Health (Irvine Calif)* 2018;10(12):1734-1748 [FREE Full text] [doi: [10.4236/health.2018.1012131](https://doi.org/10.4236/health.2018.1012131)]
13. Pedersen DH, Mansourvar M, Sortsø C, Schmidt T. Predicting dropouts from an electronic health platform for lifestyle interventions: analysis of methods and predictors. *J Med Internet Res* 2019;21(9):e13617 [FREE Full text] [doi: [10.2196/13617](https://doi.org/10.2196/13617)] [Medline: [31486409](https://pubmed.ncbi.nlm.nih.gov/31486409/)]
14. Zhou M, Fukuoka Y, Goldberg K, Vittinghoff E, Aswani A. Applying machine learning to predict future adherence to physical activity programs. *BMC Med Inform Decis Mak* 2019;19(1):169 [FREE Full text] [doi: [10.1186/s12911-019-0890-0](https://doi.org/10.1186/s12911-019-0890-0)] [Medline: [31438926](https://pubmed.ncbi.nlm.nih.gov/31438926/)]
15. Holmes EAF, Hughes DA, Morrison VL. Predicting adherence to medications using health psychology theories: a systematic review of 20 years of empirical research. *Value Health* 2014;17(8):863-876 [FREE Full text] [doi: [10.1016/j.jval.2014.08.2671](https://doi.org/10.1016/j.jval.2014.08.2671)] [Medline: [25498782](https://pubmed.ncbi.nlm.nih.gov/25498782/)]
16. Essery R, Geraghty AWA, Kirby S, Yardley L. Predictors of adherence to home-based physical therapies: a systematic review. *Disabil Rehabil* 2017;39(6):519-534 [doi: [10.3109/09638288.2016.1153160](https://doi.org/10.3109/09638288.2016.1153160)] [Medline: [27097761](https://pubmed.ncbi.nlm.nih.gov/27097761/)]
17. Areerak K, Waongengarm P, Janwantanakul P. Factors associated with exercise adherence to prevent or treat neck and low back pain: a systematic review. *Musculoskelet Sci Pract* 2021;52:102333 [doi: [10.1016/j.msksp.2021.102333](https://doi.org/10.1016/j.msksp.2021.102333)] [Medline: [33529988](https://pubmed.ncbi.nlm.nih.gov/33529988/)]
18. Di Lorito C, Bosco A, Booth V, Goldberg S, Harwood RH, Van der Wardt V. Adherence to exercise interventions in older people with mild cognitive impairment and dementia: a systematic review and meta-analysis. *Prev Med Rep* 2020;19:101139 [FREE Full text] [doi: [10.1016/j.pmedr.2020.101139](https://doi.org/10.1016/j.pmedr.2020.101139)] [Medline: [32793408](https://pubmed.ncbi.nlm.nih.gov/32793408/)]
19. Ormel HL, van der Schoot GGF, Sluiter WJ, Jalving M, Gietema JA, Walenkamp AME. Predictors of adherence to exercise interventions during and after cancer treatment: a systematic review. *Psychooncology* 2018;27(3):713-724 [FREE Full text] [doi: [10.1002/pon.4612](https://doi.org/10.1002/pon.4612)] [Medline: [29247584](https://pubmed.ncbi.nlm.nih.gov/29247584/)]
20. Spanakis G, Weiss G, Boh B, Lemmens L, Roefs A. Machine learning techniques in eating behavior e-coaching. *Pers Ubiquit Comput* 2017;21(4):645-659 [FREE Full text] [doi: [10.1007/s00779-017-1022-4](https://doi.org/10.1007/s00779-017-1022-4)]
21. 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 [FREE Full text] [doi: [10.2196/12286](https://doi.org/10.2196/12286)] [Medline: [30950797](https://pubmed.ncbi.nlm.nih.gov/30950797/)]
22. Yom-Tov E, Feraru G, Kozdoba M, Mannor S, Tennenholtz M, Hochberg I. Encouraging physical activity in patients with diabetes: intervention using a reinforcement learning system. *J Med Internet Res* 2017;19(10):e338 [FREE Full text] [doi: [10.2196/jmir.7994](https://doi.org/10.2196/jmir.7994)] [Medline: [29017988](https://pubmed.ncbi.nlm.nih.gov/29017988/)]
23. Bohlmann A, Mostafa J, Kumar M. Machine learning and medication adherence: scoping review. *JMIRx Med* 2021;2(4):e26993 [FREE Full text] [doi: [10.2196/26993](https://doi.org/10.2196/26993)] [Medline: [37725549](https://pubmed.ncbi.nlm.nih.gov/37725549/)]
24. Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *FASEB J* 2008;22(2):338-342 [FREE Full text] [doi: [10.1096/fj.07-9492LSF](https://doi.org/10.1096/fj.07-9492LSF)] [Medline: [17884971](https://pubmed.ncbi.nlm.nih.gov/17884971/)]

25. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *Int J Evid Based Healthc* 2015;13(3):132-140 [[FREE Full text](#)] [doi: [10.1097/XEB.000000000000055](https://doi.org/10.1097/XEB.000000000000055)] [Medline: [26360830](#)]
26. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Educ Inf* 2018;34(4):285-291 [doi: [10.3233/EFI-180221](https://doi.org/10.3233/EFI-180221)]
27. Jossa-Bastidas O, Zahia S, Fuente-Vidal A, Férez NS, Noguera OR, Montane J, et al. Predicting physical exercise adherence in fitness apps using a deep learning approach. *Int J Environ Res Public Health* 2021;18(20):10769 [[FREE Full text](#)] [doi: [10.3390/ijerph182010769](https://doi.org/10.3390/ijerph182010769)] [Medline: [34682515](#)]
28. Evangelista LS, Ghasemzadeh H, Lee JA, Fallahzadeh R, Sarrafzadeh M, Moser DK. Predicting adherence to use of remote health monitoring systems in a cohort of patients with chronic heart failure. *Technol Health Care* 2017;25(3):425-433 [[FREE Full text](#)] [doi: [10.3233/THC-161279](https://doi.org/10.3233/THC-161279)] [Medline: [27886024](#)]
29. Gu Y, Zalkikar A, Liu M, Kelly L, Hall A, Daly K, et al. Predicting medication adherence using ensemble learning and deep learning models with large scale healthcare data. *Sci Rep* 2021;11(1):18961 [[FREE Full text](#)] [doi: [10.1038/s41598-021-98387-w](https://doi.org/10.1038/s41598-021-98387-w)] [Medline: [34556746](#)]
30. Ramos LA, Blankers M, van Wingen G, de Bruijn T, Pauws SC, Goudriaan AE. Predicting success of a digital self-help intervention for alcohol and substance use with machine learning. *Front Psychol* 2021;12:734633 [[FREE Full text](#)] [doi: [10.3389/fpsyg.2021.734633](https://doi.org/10.3389/fpsyg.2021.734633)] [Medline: [34552539](#)]
31. Tucker CS, Behoora I, Nembhard HB, Lewis M, Sterling NW, Huang X. Machine learning classification of medication adherence in patients with movement disorders using non-wearable sensors. *Comput Biol Med* 2015;66:120-134 [[FREE Full text](#)] [doi: [10.1016/j.combiomed.2015.08.012](https://doi.org/10.1016/j.combiomed.2015.08.012)] [Medline: [26406881](#)]
32. Wallert J, Gustafson E, Held C, Madison G, Norlund F, von Essen L, et al. Predicting adherence to internet-delivered psychotherapy for symptoms of depression and anxiety after myocardial infarction: machine learning insights from the U-CARE heart randomized controlled trial. *J Med Internet Res* 2018;20(10):e10754 [[FREE Full text](#)] [doi: [10.2196/10754](https://doi.org/10.2196/10754)] [Medline: [30305255](#)]
33. Goldstein SP, Zhang F, Thomas JG, Butryn ML, Herbert JD, Forman EM. Application of machine learning to predict dietary lapses during weight loss. *J Diabetes Sci Technol* 2018;12(5):1045-1052 [[FREE Full text](#)] [doi: [10.1177/1932296818775757](https://doi.org/10.1177/1932296818775757)] [Medline: [29792067](#)]
34. Koesmahargyo V, Abbas A, Zhang L, Guan L, Feng S, Yadav V, et al. Accuracy of machine learning-based prediction of medication adherence in clinical research. *Psychiatry Res* 2020;294:113558 [[FREE Full text](#)] [doi: [10.1016/j.psychres.2020.113558](https://doi.org/10.1016/j.psychres.2020.113558)] [Medline: [33242836](#)]
35. Bremer V, Chow PI, Funk B, Thorndike FP, Ritterband LM. Developing a process for the analysis of user journeys and the prediction of dropout in digital health interventions: machine learning approach. *J Med Internet Res* 2020;22(10):e17738 [[FREE Full text](#)] [doi: [10.2196/17738](https://doi.org/10.2196/17738)] [Medline: [33112241](#)]
36. Nasseh K, Frazee SG, Visaria J, Vlahiotis A, Tian Y. Cost of medication nonadherence associated with diabetes, hypertension, and dyslipidemia. *Am J Pharm Benefits* 2012;4(2):e41-e47
37. Ekpezu A, Wiafe I, Oinas-kukkonen H. Technological factors that influence user compliance with behavior change support systems: a systematic review. *Proc Annu Hawaii Int Conf Syst Sci Hawaii* [doi: [10.2196/preprints.46779](https://doi.org/10.2196/preprints.46779)]
38. Nutrokpor C, Ekpezu A, Wiafe A, Wiafe I. Exploring the impact of persuasive system features on user sentiments in health and fitness apps. 2021 Presented at: Proceedings of the Ninth International Workshop on Behavior Change Support Systems, BCSS 2021; April 12-14, 2021; Bournemouth, UK p. 25-39
39. Irshad MT, Nisar MA, Huang X, Hartz J, Flak O, Li F, et al. SenseHunger: machine learning approach to hunger detection using wearable sensors. *Sensors (Basel)* 2022;22(20):7711 [[FREE Full text](#)] [doi: [10.3390/s22207711](https://doi.org/10.3390/s22207711)] [Medline: [36298061](#)]
40. Tang J, Alelyani S, Liu H. Feature selection for classification: a review. In: Aggarwal CC, editor. *Data Classification: Algorithms and Applications*. New York, NY: CRC Press; 2014:37-64
41. Ratz T, Lippke S. 8.06—health behavior change. *Clin Psychol Sci (Second Ed)* 2022;8:95-117 [doi: [10.1016/b978-0-12-818697-8.00068-6](https://doi.org/10.1016/b978-0-12-818697-8.00068-6)]
42. Culos-Reed SN, Rejeski WJ, McAuley E, Ockene JK, Roter DL. Predictors of adherence to behavior change interventions in the elderly. *Control Clin Trials* 2000;21(5 Suppl):200S-205S [doi: [10.1016/s0197-2456\(00\)00079-9](https://doi.org/10.1016/s0197-2456(00)00079-9)] [Medline: [11018576](#)]
43. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall; 1986.
44. Heesch KC, Mâsse LC, Dunn AL, Frankowski RF, Mullen PD. Does adherence to a lifestyle physical activity intervention predict changes in physical activity? *J Behav Med* 2003;26(4):333-348 [doi: [10.1023/a:1024205011001](https://doi.org/10.1023/a:1024205011001)] [Medline: [12921007](#)]
45. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. *Commun Assoc Inf Syst* 2009;24(1):485-500 [doi: [10.17705/1CAIS.02428](https://doi.org/10.17705/1CAIS.02428)]
46. Merz M, Augsburg U, Ackermann L, Ackermann L. Design principles of persuasive systems—review and discussion of the persuasive systems design model. 2021 Presented at: AMCIS 2021 Proceedings; August 9-13, 2021; Montreal, Canada URL: https://aisel.aisnet.org/amcis2021/sig_hci/sig_hci/3
47. Yu C, Liu J, Nemati S, Yin G. Reinforcement learning in healthcare: a survey. *ACM Comput Surv* 2021;55(1):1-36 [[FREE Full text](#)] [doi: [10.1145/3477600](https://doi.org/10.1145/3477600)]

48. Balki I, Amirabadi A, Levman J, Martel AL, Emersic Z, Meden B, et al. Sample-size determination methodologies for machine learning in medical imaging research: a systematic review. *Can Assoc Radiol J* 2019 Nov;70(4):344-353 [FREE Full text] [doi: [10.1016/j.carj.2019.06.002](https://doi.org/10.1016/j.carj.2019.06.002)] [Medline: [31522841](https://pubmed.ncbi.nlm.nih.gov/31522841/)]

Abbreviations

BCSS: behavior change support system

LSTM: long short-term memory

MI-ANXDEP: myocardial infarction-anxiety, depression, or both

PSD: persuasive systems design

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

RQ: review question

Edited by K El Emam, B Malin; submitted 02.03.23; peer-reviewed by R Marshall, C Manlhiot; comments to author 31.07.23; revised version received 20.09.23; accepted 28.10.23; published 22.11.23

Please cite as:

Ekpezu AO, Wiafe I, Oinas-Kukkonen H

Predicting Adherence to Behavior Change Support Systems Using Machine Learning: Systematic Review

JMIR AI 2023;2:e46779

URL: <https://ai.jmir.org/2023/1/e46779>

doi: [10.2196/46779](https://doi.org/10.2196/46779)

PMID:

©Akon Obu Ekpezu, Isaac Wiafe, Harri Oinas-Kukkonen. Originally published in JMIR AI (<https://ai.jmir.org>), 22.11.2023. 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 JMIR AI, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.ai.jmir.org/>, as well as this copyright and license information must be included.