Abstract

Background: Fitness devices have spurred the development of apps that aim to motivate users, through interventions, to increase their physical activity (PA). Personalization in the interventions is essential as the target users are diverse with respect to their activity levels, requirements, preferences, and behavior.

Objective: This review aimed to (1) identify different kinds of personalization in interventions for promoting PA among any type of user group, (2) identify user models used for providing personalization, and (3) identify gaps in the current literature and suggest future research directions.

Methods: A scoping review was undertaken by searching the databases PsycINFO, PubMed, Scopus, and Web of Science. The main inclusion criteria were (1) studies that aimed to promote PA; (2) studies that had personalization, with the intention of promoting PA through technology-based interventions; and (3) studies that described user models for personalization.

Results: The literature search resulted in 49 eligible studies. Of these, 67% (33/49) studies focused solely on increasing PA, whereas the remaining studies had other objectives, such as maintaining healthy lifestyle (8 studies), weight loss management (6 studies), and rehabilitation (2 studies). The reviewed studies provide personalization in 6 categories: goal recommendation, activity recommendation, fitness partner recommendation, educational content, motivational content, and intervention timing. With respect to the mode of generation, interventions were found to be semiautomated or automatic. Of these, the automatic interventions were either knowledge-based or data-driven or both. User models in the studies were constructed with parameters from 5 categories: PA profile, demographics, medical data, behavior change technique (BCT) parameters, and contextual information. Only 27 of the eligible studies evaluated the interventions for improvement in PA, and 16 of these concluded that the interventions to increase PA are more effective when they are personalized.

Conclusions: This review investigates personalization in the form of recommendations or feedback for increasing PA. On the basis of the review and gaps identified, research directions for improving the efficacy of personalized interventions are proposed. First, data-driven prediction techniques can facilitate effective personalization. Second, use of BCTs in automated interventions, and in combination with PA guidelines, are yet to be explored, and preliminary studies in this direction are promising. Third, systems with automated interventions also need to be suitably adapted to serve specific needs of patients with clinical conditions. Fourth, previous user models focus on single metric evaluations of PA instead of a potentially more effective, holistic, and multidimensional view. Fifth, with the widespread adoption of activity monitoring devices and mobile phones, personalized and dynamic user models can be created using available user data, including users’ social profile. Finally, the long-term effects of such interventions as well as the technology medium used for the interventions need to be evaluated rigorously.

(JMIR Mhealth Uhealth 2019;7(1):e11098) doi:10.2196/11098

KEYWORDS
review; exercise; physical fitness; automation; mobile apps; web browser; health communication; health promotion
Introduction

Background

Insufficient physical activity (PA) is a worldwide concern as it is a major cause of obesity and the fourth leading risk factor for mortality, accounting for an estimated 3.2 million deaths globally [1]. Maintaining or increasing PA of patients is also an important goal in the treatment for various chronic diseases such as diabetes and cardiovascular illnesses.

Fitness trackers, such as Fitbit and Jawbone, are increasingly being used to monitor personal PA. Activity data collected by associated smartphone apps are being utilized, along with other user-specific or contextual data, to design interventions with the aim of motivating users to increase their PA [2,3]. These interventions take varied forms ranging from activity status reports to personalized fitness-buddy recommendations.

Increasing PA often requires a change in lifestyle or behavior of the user. Feedback based on activity status reports is a common strategy that is often augmented with educational information on the benefits of increased PA. The key limitation of such interventions is the reliance on the self-motivation of users to increase their PA [4,5]. Users may not be motivated for various reasons, for example, they may be inactive by habit or the presented activity goal may be too intimidating for them. Other factors may also play a role in determining the efficacy of interventions. For example, some users may not have the time to perform a recommended PA [6] or there may be constraints imposed by the users’ location, weather, or working environments. Providing information on the benefits of increased PA rarely suffices; effecting behavior change to increase PA additionally requires motivational interventions [7].

The aims, behavior, preferences, context, and lifestyle of users have to be taken into account by apps to design effective interventions [8,9]. A “one size fits all” approach is unable to effectively serve a diverse set of users. Even simple activity recommendations, such as “60 minutes of moderately vigorous physical activity (MVPA)” may be too daunting for a sedentary user or for a cardiac patient. Thus, there is a need for personalization of interventions for promoting PA among users. Personalization implies a modification in the intervention generation or delivery aimed at a specific user. A status feedback does not indicate personalization; personalization implies customized content or advice to help the targeted user in increasing PA.

Previous reviews [10-13] on interventions for increasing PA have studied internet-based or Web-based interventions without focusing exclusively on personalization. They evaluated the success of included studies with respect to intervention delivery (eg, email and website-based) and discussed the utility of theory-based interventions [10]. Other reviews have specifically studied target groups, such as stroke patients [14] or cardiovascular disease (CVD) patients [15]. Another recent review [16] has analyzed the decision support systems used in PA interventions but does not focus on personalization. A survey of tailoring techniques used in real-time PA coaching systems published before August 2013 is presented in the study by op den Akker et al [17].

The term “personalization” has multiple definitions in different domains [18]. We follow the commonly accepted definition in the study by Fan et al [18], which defines personalization as “a process which changes the functionality, interface, content or distinctiveness of system to increase its personal relevance.” According to this definition, if the system is not altered in any of the dimensions mentioned to increase personal relevance, it is not considered as personalization. An earlier study by Hawkins et al [19], defines “tailoring” as a generic term for providing feedback, personalization, and content matching. It uses the term personalization to encompass the tactics of identification, raising expectation, and contextualization. However, following our adopted definition, from the study by Fan et al [18], we also include the category of “content matching” within “personalization.” The review by Akker et al [17] identified 7 categories based on tailoring techniques for activity coaching—feedback, inter-human interaction, adaptation, user targeting, goal setting, context awareness, and self-learning—and discussed relevant studies in these categories. Thus, tailoring has been used as a broad term in the literature and does not necessarily provide the “modification” required for personalization in our adopted definition. In this study, we use the term “personalization” to denote a user-specific modification of an intervention.

Objectives

The purpose of this review was to identify recent literature where the technology-based intervention is personalized with the aim of increasing PA of users. The feedback or recommendation is not just a presentation of the users’ activity status. It is either a personalized feedback based on the history and status of the user to motivate or educate the user or a recommendation to potentially increase PA. A key aspect of such studies, which we focus on, is the user model created, which in turn helps generate personalized recommendations. Findings from this review provide important insights into the current literature and identify significant gaps in the literature. Addressing these gaps could lead to more effective, personalized, and technology-based interventions for promoting PA.

Methods

The Scoping Review

This review aims to identify various interventions, customizations, and user models generated for personalization of technology-based interventions to increase PA of users. We employed a rigorous literature search and chose to conduct a scoping review to analyze our research questions. The research questions we have focused on are as follows: (1) what are the ways of providing technology-based personalized interventions for increasing PA among users and (2) what are the user models used to provide such personalization? For ensuring quality of the included studies, we have used only peer-reviewed articles, including research-in-progress articles, which had full text available. We do not perform additional quality analysis of the studies, as quality assessment does not form part of the scoping review.
study remit [20]. This paper follows the methodology and directions given in the study by Arksey and O’Malley [20].

Search Strategy
PubMed, PsycINFO, Scopus, and Web of Science databases were used to select relevant studies. A comprehensive search was conducted till August 23, 2018, in which articles published since 2013 were targeted. The search string was constructed by considering the criteria required to be satisfied by the studies to be considered: {physical activity} {interventions} having {personalization} provided through some {technology} and identifying or creating a {user model} for the same. The following search string was used: ((fitness OR exercise OR “physical activity” OR “activity level” OR “active living”) AND (intervention OR recommend* OR prescribe OR prescription OR feedback OR message) AND (tailor* OR personaliz* OR personalis*) AND (mobile OR internet OR computer OR device OR “fitness trackers” OR website OR online) AND (profit* OR model)). The search was restricted to papers published in English. This search string ensured our condition of the technology-based intervention having a user model or profile identified for providing the personalization.

In addition to the database searches, we also performed hand searches for additional relevant studies. These studies were found by identifying relevant references from the studies selected. These references were also analyzed for the selection criteria and included in the review if they met the criteria. In addition, a hand search of Journal of Medical Internet Research results for “physical activity interventions” was done to identify several other relevant studies.

Data Extraction
We selected the articles in 2 phases and used Mendeley reference manager to organize them. The first phase involved title, abstract, and keyword review as obtained from the databases searched. This phase was applied to all results obtained from the databases after merging duplicates (a feature provided by Mendeley). The second phase included reviewing the full text of the articles. This was done by obtaining PDF documents for each of the articles that met the inclusion criteria. The full texts were analyzed using the inclusion and exclusion criteria, and studies that were deemed relevant after this phase were included in the scoping review.

Selection Criteria
Studies were eligible for this review if all the following were true: (1) there was an attempt to increase or regulate PA among the target users; (2) the studies had some form of personalized intervention, as recommendations or feedback intended to promote PA of the users; (3) a user model was generated and used for providing the personalized intervention described; (4) the intervention was provided through usage of technology; (5) studies were in English and published in or after the year 2013; and (6) they were not review papers, dissertations, or letters and were published through a peer-reviewed process.

Studies published before 2013 were not included as the popularity of fitness devices and attempts to create trackers and coaches have increased in the last 4 to 5 years, that is, older literature may be less relevant to today’s apps. Moreover, relevant literature until then has already been reviewed in the study by Akker et al [17]. There was no restriction on the study objectives, type of users, or the type of intervention or feedback, other than the focus on personalization for PA promotion. The focus of the review is on methods of personalization and user model generation for technology-based interventions. The interventions where personalization was provided manually were excluded as the user model used for delivering the personalization cannot be identified in a manual process. A comprehensive review of different ways to model users and provide personalized, technology-based interventions for increasing PA in different settings was desired.

The exclusion criteria for this review were as follows: (1) personalization not aimed at increasing PA (eg, personalization in activity tracking or gait detection); (2) personalization provided only in terms of using name or activity status in message, these parameters were filled into standard messages; (3) no user model identified during the intervention; (4) personalization generated manually, even though may be delivered using technology through a website; (5) gender- or culture-based standard tailoring for intervention; and (6) only reports provided without any personalized content for encouraging or educating user, or without any advice.

The inclusion criterion entailed that the technology-based, personalized intervention had to be necessarily aimed at increasing PA. The criterion of increasing PA was not necessarily the main objective of included studies but had to be one of the objectives. For example, in some studies, medication adherence [21] or weight loss [2] was the other objective.

Results

Screening and Study Selection
The screening procedure and study selection was undertaken by 1 researcher and then independently verified by 2 other researchers for adherence to the selection criteria. The initial results were screened for the inclusion criteria and the full-text articles were analyzed using the exclusion criteria. Initial results were obtained by setting the filters of language and duration for all the databases (536 results) and were then searched for duplicates, which resulted in 355 unique studies. The abstracts of these studies were then screened for the criteria of whether the paper tried to increase or regulate PA. In addition, 15 relevant studies were identified by hand searching and cross-references. This led to a selection of 181 papers, which were screened on full text for the remaining selection criteria, resulting in 57 studies.

We found several groups of studies that studied the same system, that is, they were parts or improvements of the same intervention. We also found additional studies through hand searches that belonged to these groups, which helped us understand details of the interventions. We grouped these related studies together and used only 1 representative publication for each of the 23 groups. The groups and the representative studies are listed in Table 1. This step reduced the final number of studies to 49.
Table 1. Studies grouped by the intervention developed or investigated.

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Related studies</th>
<th>Representative study</th>
<th>Representative study</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKTUS</td>
<td>Janols et al and Lindgren et al [23,24]</td>
<td>Janols et al [23]</td>
<td></td>
</tr>
<tr>
<td>Active-O-Meter</td>
<td>Cook et al and De Bourdeaudhuji et al [30,31]</td>
<td>Cook et al [30]</td>
<td></td>
</tr>
<tr>
<td>Food4Me</td>
<td>Marsaux et al, Morales et al, and Marsaux et al [34-36]</td>
<td>Marsaux et al [35]</td>
<td></td>
</tr>
<tr>
<td>I Move</td>
<td>Friederichs et al and Friederichs et al [37,38]</td>
<td>Friederichs et al [38]</td>
<td></td>
</tr>
<tr>
<td>MOPO</td>
<td>Ahola et al, Jauho et al, and Pyky et al [39-41]</td>
<td>Pyky et al [41]</td>
<td></td>
</tr>
<tr>
<td>My Activity Coach</td>
<td>Alley et al and Alley et al [42,43]</td>
<td>Alley et al [43]</td>
<td></td>
</tr>
<tr>
<td>MyBehavior</td>
<td>Rabbi et al [44,45]</td>
<td>Rabbi et al [45]</td>
<td></td>
</tr>
<tr>
<td>myHealthyBehavior</td>
<td>Schulz et al and Schulz et al [46,47]</td>
<td>Schulz et al [47]</td>
<td></td>
</tr>
<tr>
<td>PATH-In</td>
<td>Brooks et al and Williams et al [48,49]</td>
<td>Williams et al [48]</td>
<td></td>
</tr>
<tr>
<td>PATHway</td>
<td>Chatzitofis et al, Claes et al, and Triantafyllidis et al [50-52]</td>
<td>Triantafyllidis et al [52]</td>
<td></td>
</tr>
<tr>
<td>PRO-Fit</td>
<td>Dharia et al [56-58]</td>
<td>Dharia et al [56]</td>
<td></td>
</tr>
<tr>
<td>REACH</td>
<td>Mitchell et al and Mitchell et al [59,60]</td>
<td>Mitchell et al [60]</td>
<td></td>
</tr>
<tr>
<td>RENATA</td>
<td>Reinwad et al and Storm et al [7,61]</td>
<td>Storm et al [7]</td>
<td></td>
</tr>
<tr>
<td>Start to Stand</td>
<td>De Cocker and De Cocker [63,64]</td>
<td>De Cocker [64]</td>
<td></td>
</tr>
<tr>
<td>TaylorActive</td>
<td>Soetens et al and Vandelanotte et al [65,66]</td>
<td>Vandelanotte et al [66]</td>
<td></td>
</tr>
<tr>
<td>TXT2Bfit</td>
<td>Hebden et al and Partridge et al [67,68]</td>
<td>Partridge et al [68]</td>
<td></td>
</tr>
<tr>
<td>YEAH</td>
<td>Kattelmann et al [72,73]</td>
<td>Kattelmann et al [72]</td>
<td></td>
</tr>
</tbody>
</table>

Among the 181 studies assessed for eligibility, most of the studies could be screened using our exclusion criteria. However, a few studies, such as the study by Liu and Chan [22], were identified through the search but were excluded because the definition of personalization used was different. It focused on whether or not to prompt the user based on current and predicted activity status, which differs from the conceptualization adopted.

Figure 1 illustrates the flowchart representing the study selection process.

**Overview of Studies**

We placed no restrictions on the research objective or methodology of the studies to be included in the review, other than following our study criteria. As a result, the studies differ considerably with respect to their research objectives, interventions, data collection methods, and target users. We summarize these diverse settings below before examining the personalized interventions and user models employed in more detail. Moreover, as mentioned earlier, we identified 23 interventions, each of which has been described in more than 1 study. These groups are listed in Table 1. In this review, we study all the articles listed in the table (column 2) but represent each group by a representative study (column 3).

Increasing PA was the research objective in 33 studies. In the remaining studies, the objectives were weight loss, weight management, or obesity prevention (6 studies); maintaining healthy lifestyle that included diet, smoking, alcohol, or exercise management (8 studies); and rehabilitation (2 studies), with PA increase being an auxiliary goal. Furthermore, 1 study had a combined goal of weight management and healthy lifestyle. Among the 33 studies on PA, 31 directly aimed to increase PA of users, 1 study aimed to reduce workplace sitting time [64], and the last study aimed to encourage medical adherence in addition to increasing PA [21]. These studies not only tried to monitor and increase PA but some also focused on helping users overcome barriers to increasing PA and improving their self-efficacy, for example in Oosterom-Calo et al [21]. The 6 studies aimed at weight loss attempted to increase the PA of users to achieve the desired results [eg, 2,74]. One study aimed at providing rehabilitation to patients provided recommendations in consultation with health care experts [75]. The last study aimed at rehabilitation provided real-time as well as weekly adaptation of exercises for the patients [52]. For the 8 studies on healthy lifestyle, their objectives included wellness services [32], exercise and diet recommendation [6,7,47,76-78], and a personalized coaching system [54], which was illustrated through 3 use cases, that is, neck coach, activity coach, and stress coach.
The interventions were presented or delivered to users in multiple ways, for example, through Web apps [7,66,79], mobile phone apps [2,3,74], Kinect devices [52,80], specific activity monitors [81], PDF report [35], text message (short message service, SMS) [9,68,82], printout [27,83,84], or telephone call [60]. In some cases the users were actively pushed by the system toward their PA goal by automatic delivery of interventions periodically [7,9,54]. In other cases, the interventions were relatively more passive and expected higher levels of motivation from the users. They required the user to access the app [45] or answer questionnaires [21,64,79] before they could obtain the personalized intervention.

Data were collected for the intervention (for monitoring users and generating user models) in various ways, for example, through questionnaires [3,21,80], mobile phone sensors [9,32,45,85], specific activity monitors [41,64,66,81], or fitness trackers [2,3,86]. The target population of the studies varied from specific to general users, that is, people with chronic disease [76], elderly adults [79,80], diabetes patients [9,87], cancer patients [78,82,84], people with CVD or at risk of CVD or heart failure [7,21], osteoarthritis patients [48], young adults [3,72,88], and general users [2,32,45,86].

Apart from these differences, the included studies differed in the intervention generation techniques, type of personalization, and user models. In this review, we systematically study these 3 aspects in detail. Note that some studies included in the review may also have incorporated personalization of diet or models for activity detection, but our focus is restricted to only the part of the study concerning PA. The overview presented in Multimedia Appendix 1 highlights the objectives, interventions,
personalization, user models, and theoretical models used in the included studies. These are ordered by their research objectives and by the intervention generation mode within each objective.

Types of Interventions
The interventions included in this review provide various forms of feedback or recommendation. We distinguish between feedback and recommendation, where recommendations are prescriptive in nature, whereas feedback is an informative response to the users’ actions. For example, feedback can be information regarding tips to increase PA such as “exercising with a partner can be a fun and motivating experience,” whereas recommendations are prescriptive suggestions of an activity or goal such as “30 minutes of brisk walk along with your mother” provided to the user. We only consider feedback that is personalized in some aspect (eg, content or timing).

Personalization is achieved in different ways, that is, by personalizing goal, activity, or fitness partner recommendations or by personalizing messages and their timings, as discussed in the following section. On the basis of how this personalized intervention is generated, interventions in our review could be classified into 2 categories: semiautomated and automated. We excluded those studies that had only manual interventions.

Semiautomated Interventions
Semiautomated interventions are those where personalization is not completely automated but includes manual effort from the health care provider. There were 9 studies with semiautomated interventions, and the combinations of manual and automatic elements in them varied.

In the study by Tseng et al [76], automated activity and goal recommendations were provided, which could be modified by a medical expert. Similarly, for the case of SmartLoss [2], the system required a goal to be set in consultation with the nurse, but the platform also automatically provided a set of “SmartTips” in case the user was predicted to be deviating from the weight loss program. Another semiautomated system [35] provided automated educational content to the users using a machine learning method along with manual personalized advice and intervention from an expert. My Activity Coach study [43] used automatic advice recommendation as well as a one-on-one video calling interaction with coach. Similarly, 3 other studies [60,68,82] used telephonic conversations for motivational interviewing, but the participants had website access or automatically delivered messages. The remaining 2 studies [75,80] generated automatic personalized activity or game levels within the limits defined by a health care expert.

Automated Interventions
Automated interventions present in 40 papers in our review used either knowledge-based or data-driven approaches or both to automate the personalization. All the knowledge-based systems relied on either behavior change techniques (BCTs) or PA or clinical guidelines. All the data-driven systems used machine learning techniques to learn user models from historical data.

Knowledge-Based Systems
There were 30 studies using knowledge-based approaches. These systems were rule-based and provided feedback and recommendations based on reasoning modules or rules.

Of these studies, 22 of them attempted to encode knowledge into the system derived from behavior change theories. They provided the personalization intervention by inferring the most suitable user category, where the categories were theory-based. Thus, they provided category-level personalization instead of an individual level of personalization. In addition, 2 studies added an individual level personalization by considering the preference of intervention time [89,90] or by providing walking and cycling routes based on location [27].

There were 8 studies that used PA or clinical guidelines. The study by Ali et al [85] used a hybrid rule- and case-based reasoning model but tried to identify similar cases using the K-nearest neighbor algorithm. Their system was based on the Center for Disease Control and Prevention (CDC) guidelines [91]. The study by Coolbaugh et al [81] used a specific activity monitor for providing the intervention and performed goal adaptation in accordance with American College of Sports Medicine training progression guidelines [92]. This intervention was time-bound and progressed according to the rules laid out in the process flowchart. Moreover, 2 studies [48,93] also used various PA guidelines to recommend step goals. The last 3 studies [52,78,87] used PA guidelines for their specific types of users, that is, cancer patients and diabetes patients. Few studies among these [6,52,78,87] use knowledge-based system constructed using a BCT but provide recommendations based on clinical guidelines.

Data-Driven Systems
Data-driven systems using machine learning approaches were described in 7 studies. The class of machine learning techniques used falls under 2 broad categories, that is, reinforcement learning [9,45] and supervised learning [35,56,74,86,94].

In reinforcement learning, an automated agent learns a policy to optimize a cumulative reward function while sequentially interacting with the environment. At each step, the agent performs an action, obtains a reward, and decides its next action based on the reward with the aim of optimizing the reward function. Thus, in the study by Yom-Tov et al [9], at each step, the agent sent a message, obtained information about the user’s PA, and determined which message to send in the next step. The messaging policy was personalized for the user to maximize cumulative PA. For the other study, multi-armed bandits, a form of reinforcement learning, was used for suggestion generation in the MyBehavior app [45].

Supervised machine learning techniques learn a model from historical data to predict dependent variables from independent variables. The model may be static (such as support vector machine [SVM] or decision trees in the study Marsaux et al [35]) or temporal (such as recurrent neural networks [RNN] in the study by Lim et al [86])—the former does not explicitly model temporal correlations, whereas the latter does. In another study, PRO-fit recommended a fitness partner using geolocation, activity preference, and calendar-based availability on a
It also provided activity recommendation using collaborative filtering and activity prediction from raw accelerometer data. An Internet of Things–based app proposed a context-aware recommendation system to generate a suitable activity for the user based on current fatigue and fitness level. Finally, in the study by Lim et al, a lifestyle model parameterized by heart rate (HR), step count, and burned calories was constructed using RNN, and recommendations for healthy behavior were based on forecasts of these variables.

**Combined Knowledge-Based and Data-Driven System**

There were 3 studies that combined knowledge-based and data-driven approaches. The ATHENA system defined a rule-based recommendation, in which only ranking and validation were done using machine learning. Ensemble-based supervised learning was used for recommendation of food, physical, and mental therapy in study. In the study by Hermens et al, a rule-based system was used for message content along with machine learning to appropriately time the message—an SVM was trained based on historical data to predict effective timing. Another personalized health care system proposed an ontology-based knowledge base, which uses decision trees for providing relevant recommendation to the user.

**Overview of Personalization**

Personalization was found in both recommendations and feedback. In the case of recommendations, personalization was seen with respect to goal setting, activity suggestion, and selection of fitness partners. Feedback was found to be personalized with respect to the content, which could be motivational or educational, or with respect to the timing of its delivery. Status comparative feedback was also considered to be personalized as it was provided to only those for whom it was considered beneficial. Thus, we classify personalized interventions into 6 categories, that is, goal recommendation, activity recommendation, fitness partner recommendation, educational content, motivational content, and intervention timing. These categories are not mutually exclusive, as several studies had more than 1 type of personalization.

**Goal Recommendation**

The category of goal recommendation refers to the prescription of a quantified target goal. This target is in terms of an activity evaluation metric, such as duration of activity, step count, or calorie expenditure. Note that if an activity is prescribed without quantification, then we classify it as an activity recommendation and not a goal recommendation.

Of the studies in the review, 20 of them provided personalized goal recommendation. The specification of the goals differed across the systems and apps. The goals could be specified in terms of game level, training zone and HR, activity duration, step data, or activity level prescription by an expert. These goals were adapted according to the person’s status and did not follow standardized fixed goals (such as “30 minutes of MVPA”). For the case of the REACH intervention, it has been mentioned that personalized step goals are generated based on rate of perceived exertion. However, it is not stated if this is done automatically or by researchers and delivered manually.

In 4 of these studies, personalized goals were indirectly defined or altered after obtaining fixed goals from the user or a guideline. The MyBehavior app used the weekly weight loss goal entered by the user to obtain a personalized target calorie goal using the Harris-Benedict equation. The TaylorActive system also provided goal recommendations and suggestions during a session, but the actual goal was set by the user.

All the above-mentioned studies set a goal for the user before the user activity began. However, in the personalized PA prescription intervention study, the goal was not explicitly known by the user before the activity, although a Web interface allowed the user to check the goal recommendation. It also defined a user goal in terms of target HR and duration of activity, which was sent to the activity monitor. The activity monitor provided visual feedback (blinks on the monitor) to the user when the target goal was achieved or if their HR exceeded the target.

In another study, the feedback was whether the user must increase, strongly increase, or maintain their PA. However, the feedback was not quantified, and thus, this study was not classified as a goal recommendation.

**Activity Recommendation**

The category of activity recommendation includes studies where 1 or more appropriate activities (eg, running and cycling) or behaviors (eg, sleep for X hours) are prescribed to the user. The 22 studies in our review that provided activity recommendation also retained the monitored PA as part of the user model.

Of these, 4 studies offered semiautomated interventions. The activity recommendations were in the form of health care experts’ treatment advice (where the treatment included PA) or activity suggestions. Furthermore, I study did not use a health care expert for the activity plan but suggested activity in the messages sent to the user.

The remaining 18 studies generated activity recommendations based on automated systems. These studies generated the activity or behavior recommendation by considering contextual information such as location or preferences.

In another study, the recommended activities were restricted to standing or walking. It is also important to note that this system encouraged the user to create a goal and activity plan with the aid of the system, which shows the importance of user’s

http://mhealth.jmir.org/2019/1/e11098/
motivation and involvement in planning. In the study by Klein et al [3], the activity and goal recommendations were provided, but the final choice was left to the user. Similarly, in the study by Williams et al [48], the activity plan was generated by the system and adjusted to user level, but the user could request the plan generation. For the personalized coaching system in the study by Hermens et al [54], the activity recommendation was provided in the content of the message [54].

**Fitness Partner Recommendation**

The aim of fitness partner recommendation is to match users of a system who are similar, to motivate them and help them maintain their PA. Of the 3 studies of this type, 2 studies [56,74] used recommender systems for finding a suitable partner. The other study [95] attempted to find a similar user by matching all users who crossed the particular user during a running activity.

**Educational Content**

In the educational content category (21 studies), personalized feedback aimed to increase the knowledge of the users about the importance of or techniques for improving PA.

There is a vast amount of information available on the internet, and providing the user with the most relevant content is the aim of such personalization. A direct way to do this was to provide appropriate links to website content (eg, in Food4Me [35], SmartLoss [2], and the multimodal system in the study by Ali et al [85]). The “My Activity Coach” system [43] and Ninas Saludables intervention [88] provide tailored Web content to users, for example, obese users receive additional content not provided to users with normal weight. The Active2Gether system [3] had an educational phase, where messages that put a user’s insufficient performance into perspective were sent along with the need and benefits of PA. The “start to stand” [64] app provided feedback messages, which also helped impress the harmful effects of too much sitting or sedentary lifestyle, based on the decision rules. Some studies such as those by Storm et al and Short et al [7,84] provided tailored example plans to the user to aid in goal setting. Some studies provided content in terms of tips to increase PA if found to be relevant [30,38,71,97,98]. The I-Move for Life study [84] provided information on the benefits of PA tailored according to the expected outcomes. The Active Plus system [27] provided the user with information on sports opportunities tailored to the location, along with walking and cycling routes. This information is also educational as it provides a feasible method for improving PA of the user.

Educational content may be motivational as well, for example, if the content is provided to help users overcome their specific barriers to performing PA [21,32,66,79,83]. This was in the form of tailored video or textual content.

**Motivational Content**

This category (29 studies) contains personalized feedback that aims to motivate users to improve their PA. What may motivate a user can be inferred from specific rules or BCTs. Note that motivational messages that were not personalized (eg, “Good!” [2]) were not included. Messages in this category targeted users specifically to elicit an action by also utilizing techniques including the users’ name or providing users’ current PA status [3]. However, as mentioned in the exclusion criteria, using only statistics or name in a standard template message is not considered personalization.

A reinforcement learning based study [9] aimed to learn which type of message (negative feedback, positive relative to self, positive relative to others, or no message) best motivates a user. A few studies had both motivational and educational messages as they target the beliefs of users [43,83,84,88]. The studies targeting “stage of change” of the user generally provided personalized motivation [3,41,54,68,72] by determining the stage the user is currently in, for example, precontemplation, preparation, or maintenance. The TaylorActive system [66] provided personally relevant feedback in various categories including what they called the “boosting your confidence” category.

The multimodal system in the study by Ali et al [85] offered motivating content that was not personalized. The Social POD app [74] provided personalized fitness partner recommendation. Personalization of motivational content was done through the fitness partner, who selected a motivational message to be sent to the user.

**Intervention Timing**

This form of personalization takes the context into account and finds the right time to send a feedback or recommendation to the user. Timing of feedback is known to play an important role, for example, a notification reminder sent when the user is busy is likely to be ignored and forgotten.

In our review, 7 studies provided this kind of personalization. Of these, 2 studies [54,86] learned the most appropriate time for intervention from past data using machine learning. The neural network–based model [86] used a greedy policy to determine the best time, which learns from user feedback, after predicting the users’ activity. The personalized coaching system in the study by Hermens et al [54] trained an SVM to determine the appropriateness of a given time to send a message. Moreover, 2 studies (PRO-Fit [56] and Step Up Life [89]) used the calendar context to determine if a given time was suitable for recommendation. Step Up Life intervention [89] additionally used location to determine home and “friendly” locations for providing intervention or reminders. Furthermore, 2 studies [82,90] used the preference of timing obtained from user for providing personalization. The last study [6] mentioned sending reminders at opportune moments, but the exact methodology is unclear.

**Theoretical Background for Personalization**

In our review, 41 studies used a theoretical framework or foundation for providing personalization. Apps with a theoretical background for their personalization either followed guidelines from sports or health care bodies or used BCTs.

Activity training guidelines from the American College of Sports Medicine [92] were followed for recommending activity increments to avoid injuries in the studies by Lee et al and Coolbaugh et al [78,81]. Guidelines from the CDC [91] were
used to generate hybrid rule-based techniques to recommend a suitable activity to users in the studies by Fahim et al and Ali et al [32,85]. BCTs are theory-based methods for changing 1 or several psychological determinants of behavior, such as a person’s attitude or self-efficacy. They aim to create a change among users through appropriate persuasion. Several studies used knowledge-based approaches to incorporate BCTs in providing personalization. The BCTs used were based on Fogg’s Behavior Model [99], Social Cognitive Theory (SCT) [100], Transtheoretical Model (TTM) [101], Theory of Planned Behavior (TPB) [102], I-Change Model [103], Behavioral Change Wheel [104], Activity Theory [104], Protection Motivation Theory [105], Motivational Interviewing [106], and health action process approach (HAPA) [107]. The study by Mukhtar [6] used Fogg’s Behavior Model to create what is termed as a “persuasion strategy” for the user, which takes into account motivation, ability, and trigger as parameters for appropriate recommendation. The Step Up Life intervention also utilizes the Fogg’s Behavior Theory for designing the model [89]. HAPA [107] was used to target different user stages and provide information on behavior risk and intention formation to the user in the study by Storm et al [7]. TTM defines stages of change in users and was used to determine the feedback given to the user, a direct rule-based implementation of the underlying BCT in the study by Pyky et al [41]. A system utilizing TPB and the Stage of Change Model [101] represented the constructs through questions as psychosocial correlates with PA. I-Change Model was used to design the system and questionnaires to effectively motivate users in a few studies [47,71,97]. The SCT has also been used to design the intervention considering that including and addressing social mediators such as family and peer would elicit a positive and sustained response from the user [60,88]. The TaylorActive study [66] used TPB, SCT, and self-determination theory to assess various constructs such as self-efficacy, intrinsic/extrinsic motivation, and action planning during different sessions designed for the user.

Theoretical frameworks were not present in studies using machine learning algorithms for recommending a goal or activity, as the algorithm was used to model user activity and suggest or recommend a better alternative. Some studies [45,74] used BCTs to make design decisions and choices but did not use BCT parameters for user modeling. For instance, the MyBehavior app [44,45] followed BCT to provide low-effort suggestions and used a form of reinforcement learning for activity recommendation. Another study [74] used SCT to design messages and used machine learning to recommend a fitness partner. Table 2 shows the different types personalization provided by the studies in our review.

User Models
Each study in this review created a different user model and defined the user through various attributes. We classify user-related attributes into 5 categories, that is, PA profile, demographics, medical data, BCT parameters, and contextual information.

User models can have a static and/or a dynamic component. The static component is collected only once, typically at the start of the intervention, for example, demographics and preferences. The dynamic component gets updated regularly and includes the monitored quantity describing PA. Some user models also used the personalized quantity as part of the model. All the collected information may not be part of the user model; here, only the data required and used to provide personalization are described under the user model. In cases where it could not be determined how the measured quantity was used, it has been mentioned as part of the profile descriptions.

Physical Activity Profile
The user model nearly always included the quantity being monitored—weight, diet, or PA—either recorded automatically or logged by the user. The monitored quantity differs in the included studies because of differing research objectives, intervention systems, and evaluation metrics. PA profile consisted of this monitored quantity along with the historical data of feedback, goals, or activity.

Evaluation of PA was necessary in almost all cases as personalized advice to users would need to consider current PA status of the users. Thus, PA profile was used as part of the user model in 47 studies. However, PA profile data were not used in 2 studies that provided behavior advice to its users based on the assessment and identified problematic beliefs and barriers [21,68].

Most of the studies evaluated PA by calorie or energy expenditure in terms of METs [45,85,86]. Some others estimated it by the time spent at different PA levels, such as vigorous or MVPA [41,87]. There were studies which set the target HR and used specific HR monitors for data collection [80,81,95], whereas I used a smartwatch [75]. Step count was another measure used to evaluate PA, obtained directly from fitness tracking devices [2,3,9]. PA was also evaluated by the time spent in performing an activity [30,74] or the duration. Another study [35] used metrics such as PA level and activity energy expenditure to estimate the level and energy expenditure in performing the PA. Stairs climbed was also used as a measure of PA in the study by Klein et al [3]. The activity level was a common metric used by studies which collected PA-related data through questionnaires [38,43,79,97].

The parameters listed in the PA profile of the studies (see Multimedia Appendix 2) are self-explanatory except for 2 of them. The “start and stand” app [64] had a data attribute named “level of sitting time in 5 domains.” This was obtained through the Workforce Sitting Questionnaire and included time spent in (1) traveling, (2) at work, (3) watching television, (4) using computer at home, and (5) other leisure activities. In the study by Martin et al [2], the “zone of adherence” was a quantity calculated by their mathematical model to predict whether the user needs to be provided special interventions. Furthermore, 1 study [77] used the term “lifestyle” for personalizing the exercise recommendation to a person. This has been categorized as a PA profile metric as lifestyle can be used to deduce the current level of PA of the person. In addition, 1 study [60] used a metric termed as Signal Vector Magnitude to calculate the vector magnitude of acceleration corrected for gravity.
Table 2. Personalization provided.

<table>
<thead>
<tr>
<th>Serial #</th>
<th>Paper reference</th>
<th>Goal recommendation</th>
<th>Activity recommendation</th>
<th>Fitness partner</th>
<th>Educational content</th>
<th>Motivational content</th>
<th>Intervention timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vandelanotte et al [66]</td>
<td>Y a</td>
<td>Y</td>
<td>Y b</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Ahire et al [77]</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Tseng et al [76]</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Storm et al [7]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Schulz et al [47]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Hermens et al [54]</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Lee et al [78]</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Fahim et al [32]</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Dhar et al [56]</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Rabbi et al [45]</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Twardowski et al [94]</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Yom-Tov et al [9]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Lim et al [86]</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Y</td>
</tr>
<tr>
<td>15</td>
<td>Cook et al [30]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Larsen et al [88]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Short et al [84]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Boudreau et al [97]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Moreau et al [87]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Rajanna et al [89]</td>
<td>Y, U</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Irvine et al [83]</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>Y</td>
<td>U</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Friederichs et al [38]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Blake et al [108]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Coolbaugh et al [81]</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Williams et al [48]</td>
<td>Y-initially</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Kwasnicka et al [98]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Janols et al [23]</td>
<td>U</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
</tr>
<tr>
<td>30</td>
<td>Mistry et al [90]</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
</tr>
<tr>
<td>31</td>
<td>Peels et al [27]</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
</tr>
<tr>
<td>33</td>
<td>Ammann et al [79]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Pyky et al [41]</td>
<td>—</td>
<td>U</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
</tr>
<tr>
<td>35</td>
<td>Varadharajan et al [95]</td>
<td>Y</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>Y</td>
<td>—</td>
</tr>
<tr>
<td>36</td>
<td>Codreanu et al [80]</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Marsaux et al [35]</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td>Y</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Alley et al [43]</td>
<td>Y</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
</tr>
<tr>
<td>39</td>
<td>Mitchell et al [60]</td>
<td>Y, U</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Oosterom-Calo et al [21]</td>
<td>—</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
</tr>
<tr>
<td>41</td>
<td>De Cocker et al [64]</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>
Demographics

Demographics formed a part of the user model for 39 studies. Demographic data collected included age and gender [75], body mass index [7,32], employment [64,79], nationality [7], weight [79,95], marital status [47,88,97], and education [64,79]. Several studies collected demographic information but did not use it for providing the personalized intervention. User demographics formed an important part of the user model in 16 papers. Among these, some studies [7,41,54,74,79] did not explicitly state whether demographics was used for personalization or not.

Medical Data

In our review, 16 studies (aimed at rehabilitation, healthy lifestyle, and increasing PA) used medical data as part of their user model. Personalization was based on clinical symptoms [77,80], cholesterol levels [35,93], medical records [6,76,94], pain [23,48,52], and anaerobic threshold (the point between aerobic and anaerobic training of the user) along with HR and HR at rest [75]. It is unclear if the study by Mitchell et al [60] used the medical data for providing personalization. Sleep data were also collected by 2 studies [23,66]; however, whether it was directly used for providing personalization is not clear.

Behavior Change Technique Parameters

In our review, 30 studies used various BCT parameters such as stage of change [79]; subjective PA [3]; motivation [66]; skills, barriers, goals, and outcome expectations [3]; habit strength [7]; and rate of perceived exertion [60]. The MOPO study [41] based its personalization on a data attribute termed “life satisfaction,” which is a self-reported scale on happiness, interest in life, feeling of loneliness, and the ease of living. Various psychosocial parameters such as attitudes, intention, motivation, and confidence are also used along with stage of change [87,97].

Such BCT parameters were inferred using questionnaires such as the 20-item Weight Efficacy-Lifestyle Questionnaire and the 44-item Big Five Inventory Questionnaire that sought answers from users. Studies using BCT parameters had interventions that were knowledge-based, except in the studies by Hermens et al and Hales et al [54,74].

Contextual Information

Contextual information in the user models refers to any additional information that provided cues to the context and/or behavior of the user. The context of the user varied considerably across the 14 studies that used this type of information in our review. This category included user preferences, social media profile, location, time, mood, and energy levels among others.

Activity preferences of a user were generally obtained from the user to recommend a suitable activity to the user. All the 12 studies utilizing preferences also used PA profile in their user models. Preferences were also inferred based on users’ history and adherence to recommendations in the studies by Yom-Tov et al and Lim et al [9,86]. Location and time information was used to determine the feasibility of certain activity recommendations in the studies by Klein et al and Short et al [3,84]. For example, jogging may not be feasible during rainy weather. A study by Codreanu and Florea [80] used the estimated energy level (rested, fatigued, or energetic), defined by the “mood temperature factor.”

In our review, 4 papers used the social media profile to motivate users through activity status posts on websites or by inspiration from friends. Of these, 2 studies [32,95] used the profile to provide better recommendations and persuasion to users. On the other hand, the Active2Gether and PRO-Fit systems used social media in a direct way to generate social comparison [3] and recommend a fitness buddy to the user [56]. The TaylorActive app [66] used various indicators to gauge quality of life, perceived neighborhood environment, learning style, and delivery mode preference. All of these were measured using questionnaires provided to the user. Multimedia Appendix 2 summarizes the parameters across the 5 categories for the user models of the studies in our review.

Results of Individual Studies

The studies included in this review have diverse aims and, thus, different evaluation metrics. For our review, we have considered only the results relevant to PA of users. Not all studies in the review presented evaluations of their proposed interventions and not all of them evaluated a PA metric. In our review, only 27 studies presented evaluations of their proposed interventions.
for PA. Out of these, 15 studies have reported positive statistically significant outcomes. The remaining 12 studies have not shown statistically significant results or shown no improvement at all. The impact and extent of the results vary in the studies as all are not randomized controlled trials (RCTs) and do not try to address similar questions. **Table 3** shows the evaluation and results for the RCTs included in this review.

There are 20 RCTs (listed in **Table 3**), which evaluated their proposed interventions. Of these, some studies were evaluated on the basis of self-reported PA [7,47,83,84] and others used objectively measured PA through devices [35,41,66,93]. A metric used in these studies is MET-minutes or MET hours, which is the metabolic equivalent unit for energy expenditure. These MET minutes have been observed from self-reported data collected through questionnaires. We observe that not all studies report an improvement in PA after intervention as compared with the control group. In the study by De Cocker et al [64], the objectively measured sitting time has no significant difference; however, the self-reported data show significant difference between the intervention and control groups. The MOPO study [41] also reports significant change in self-rated fitness but no significant change in self-reported daily sitting time. On the other hand, studies such as the ones by Vandelanotte et al, Partridge et al, and Irvine et al [66,68,83] reported a significant improvement in PA of users after the intervention. The **Table 4** shows the evaluation of other studies (which are not RCTs), along with the methodology used for evaluation.

Some of the studies evaluated the difference in intervention delivery mediums. In the study by Peels et al [27], 2 kinds of personalized interventions were used—basic and environmental—where environmental intervention provided users with more contextual information, such as walking and cycling routes. In the study by Van Stralen et al [28], it was found that the printed interventions—basic as well as environmental—were significantly effective; however, the Web-based interventions were not. However, in the study by Walthouwer et al [71], no significant difference was observed when participants were provided interventions through the medium of their choice (text or video). Similarly, the study by Blake et al [108] observed no significant difference between delivery modes, email, and SMS. In addition, the study by Schulz et al [47] observed no statistical difference among the sequential intervention module delivery versus the simultaneous module delivery. Another study, the TaylorActive system [65], reported an increase in PA for all groups of intervention delivery—text, video, and combination.

**Table 4** shows the evaluation of other studies (which are not RCTs), along with the methodology used for evaluation. There are 7 studies in the review, which are not RCTs, but present some feasibility or usability analyses [74,81] or are observational [79,88] or single group studies [45,54,82]. Of these, some studies such as the studies by Hermens et al [54] and Coolbaugh et al [81] have very low sample size (8 and 2, respectively). The Personalized Coaching System study [54] conducted many different experiments. We consider the one mentioned in the paper, which aims to improve long-term activity behavior of chronic obstructive pulmonary disease patients. Moreover, 5 out of 8 patients had an improvement in activity level, although exercise capacity and health status show clinical improvement in 3 of these 5 patients. The feasibility of personalized PA prescription intervention [81] was tested on 2 users. Of these, 1 subject showed excellent adherence until week 10, but the other subject had inconsistent participation. These studies do not demonstrate the effectiveness of the interventions due to their low sample sizes. However, they provide directions toward potential feasible interventions for increasing PA.

Other studies such as the studies by Rabbi et al, Ammann et al, and Spark et al [45,79,82] show significant improvements in PA for their users. A different evaluation metric is used by Short et al [84], which evaluated habit strength of performing PA. The self-reported habit strength for PA increased, which has been considered as an effective improvement for PA intervention.

In both RCT and other studies, several studies have shown significant improvement in the PA of the participants due to personalized interventions. The study by Cook et al [30] showed a significant intervention effect with an increase in active commute and leisure time PA as well as PA in schools for the adolescents. The MyBehavior app evaluation study [45] also stated an increase in walking minutes and calories burnt in nonwalking exercises as compared with the baseline. A study for older adults [83] reported a positive impact on PA, with improvements in endurance, strengthening, stretching, and balance improvements. Similarly, I-Move [38] achieved a small but significant improvement in weekly minutes of MVPA. The study by Partridge et al [68] also reported a statistically significant increase in mean MET-minutes per week. In addition, the total walking days increased in the intervention group as compared with the control group. An increase in weekly minutes of MVPA was reported by Larsen et al [88]. They also reported an increase in the diversity of activities undertaken by participants as compared with the baseline.

A total of 2 studies have reported an improvement in self-reported values but have not observed the same for the objectively measured PA values [35,64]. Several of the studies do observe improvements in PA in the intervention groups; however, these are not statistically significant [9,41,47,74]. Furthermore, 1 study [71] tried to analyze the matched delivery preference and reported no intervention effect with a delivery method (video-only, text-only, or combined) of choice. It also reports that the video-only intervention did not see any improvements.
<table>
<thead>
<tr>
<th>Serial #</th>
<th>Paper reference</th>
<th>Dataset size</th>
<th>Variables evaluated</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Soetens et al [65]</td>
<td>803</td>
<td>Effect of time over increase in PA&lt;sup&gt;a&lt;/sup&gt;</td>
<td>PA increases in all groups, time has no significant effect on all completers though has significant effect on those who had low baseline scores for total PA minutes (P&lt;.001)</td>
</tr>
<tr>
<td>5</td>
<td>Storm et al [7]</td>
<td>790</td>
<td>Strength of habit for PA measured with abbreviated version of Self-Reported Habit Index, self-efficacy, and planning</td>
<td>Self-efficacy (P=.1), planning (P=.2), and habit strength (P=.006) improved in the intervention group</td>
</tr>
<tr>
<td>6</td>
<td>Schulz et al [47]</td>
<td>5055</td>
<td>Minutes of PA per day in control, sequential intervention module delivery, and simultaneous module delivery</td>
<td>No statistical difference in sequential and simultaneous delivery for PA or with respect to control group. Sequential delivery could be more effective than simultaneous module delivery after 12 months (P=.7)</td>
</tr>
<tr>
<td>13</td>
<td>Yom-Tov et al [9]</td>
<td>27</td>
<td>PA minutes per week, change in activity with message policy, change from initial to RL&lt;sup&gt;b&lt;/sup&gt;-based learned policy</td>
<td>No statistical difference in treatment and control arm (P=.30) for PA minutes per week. Difference in change of activity between initial and learned message policy statistically significant (P=.004)</td>
</tr>
<tr>
<td>15</td>
<td>Cook et al [30]</td>
<td>555</td>
<td>PA (minutes per week) behavior difference at baseline and postmeasurement for 3 parameters: commuting, leisure time PA, and PA in school</td>
<td>Improvement found in leisure time MVPA&lt;sup&gt;c&lt;/sup&gt; (P&lt;.05), for increase in commute by bicycle (around 30 min) (P&lt;.01) and total MVPA (P&lt;.05)</td>
</tr>
<tr>
<td>17</td>
<td>Short et al [84]</td>
<td>724</td>
<td>Minutes per week of MVPA and resistance training score for all 3 arms–3 module interventions delivered monthly, weekly, or single-module</td>
<td>Significant improvement of MVPA across all groups (P&lt;.05). Significant improvement in resistance score from monthly 3-module intervention to single module (P&lt;.01)</td>
</tr>
<tr>
<td>21</td>
<td>Irvine et al [83]</td>
<td>368</td>
<td>Cardiovascular exercises, stretching exercises, strength exercises, balance exercises (all measured in minutes per week), and number of activities</td>
<td>Improvement in intervention group as compared with control in all (P&lt;.001)</td>
</tr>
<tr>
<td>22</td>
<td>Friederichs et al [38]</td>
<td>4302</td>
<td>Minutes of MVPA per week and number of days ≥30 min activity in I-Move intervention, Active Plus intervention, and control group</td>
<td>I-Move had small but more significant effect than Active Plus in minutes of MVPA per week (P=.03 and P=.07). I-Move had medium sized effect and Active Plus had large size effect for number of days ≥30 min</td>
</tr>
<tr>
<td>23</td>
<td>Blake et al [108]</td>
<td>296</td>
<td>Active travel, moderate activity at work and recreation and vigorous activity at work and recreation in 2 arms for different delivery modes, both with tailored content, one with SMS&lt;sup&gt;d&lt;/sup&gt; and another with email</td>
<td>No significant difference between email and SMS, but significant difference in moderate activity at work (hours per day), with email more effective than SMS (P=.24).</td>
</tr>
<tr>
<td>25</td>
<td>Hargreaves et al [93]</td>
<td>97</td>
<td>Step count</td>
<td>No difference at baseline and 12 weeks. Significant increase in step count of intervention group between week 12 and week 24 (P=.055) but not so significant in comparison group (P=.15)</td>
</tr>
<tr>
<td>30</td>
<td>Mistry et al [90]</td>
<td>337</td>
<td>PA between the 3 groups–standard care, generic message, and intervention group after 4 weeks</td>
<td>No significant difference between groups for change in PA (P&gt;.05)</td>
</tr>
<tr>
<td>31</td>
<td>Peels et al [27]</td>
<td>1729</td>
<td>Number of MET&lt;sup&gt;e&lt;/sup&gt; hours in 4 kinds of tailoring: printed, and Web-based (basic and environment-based in each) and control group</td>
<td>Printed (both basic and environmental) had statistically significant increase in MET hours (P=.025 and P=.31, respectively). No significant increase in both Web-based interventions (P=.59 and P=.887, respectively)</td>
</tr>
<tr>
<td>34</td>
<td>Pyky et al [41]</td>
<td>496</td>
<td>Self-rated health and fitness and leisure time PA</td>
<td>Changes in self-rated fitness and leisure time PA are associated with improved self-rated health (P&lt;.026 and P&lt;.04, respectively). No significant difference between intervention and control for self-reported daily sitting (P=.32) and light housework (but no other leisure time) PA (P=.45)</td>
</tr>
<tr>
<td>37</td>
<td>Marsaux et al [35]</td>
<td>1607</td>
<td>Objective PA in control group, group with personalized advice on diet and PA (L1 group), L1+phenotype (L2 group) and L2+genotype (L3 group)</td>
<td>No significant difference between control and any of the 3 groups in objective PA level measured (P=.73)</td>
</tr>
</tbody>
</table>

Note: <sup>a</sup>PA increases in all groups, time has no significant effect on all completers though has significant effect on those who had low baseline scores for total PA minutes (P<.001).  
<sup>b</sup>RL<sup>b</sup>-based learned policy  
<sup>c</sup>MVPA<sup>c</sup> leisure time (P<.05), for increase in commute by bicycle (around 30 min) (P<.01) and total MVPA (P<.05)  
<sup>d</sup>SMS<sup>d</sup> and another with email.  
<sup>e</sup>MET<sup>e</sup> hours in 4 kinds of tailoring: printed, and Web-based (basic and environment-based in each) and control group.
<table>
<thead>
<tr>
<th>Serial #</th>
<th>Paper reference</th>
<th>Dataset size</th>
<th>Variables evaluated</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>Alley et al [43]</td>
<td>154</td>
<td>PA (min per week) for 3 groups: control, tailoring only, and tailoring+video coaching group</td>
<td>Significant difference in PA between tailoring+video coaching versus control group ($P=.01$) but no significant difference in PA between the 2 intervention groups ($P=.54$)</td>
</tr>
<tr>
<td>39</td>
<td>Mitchell et al [60]</td>
<td>171</td>
<td>Sedentary time, LPA$^f$, and MVPA for intervention group with personalized step goals versus control group with generic advice</td>
<td>Decrease in sedentary time, Improvement in LPA and MVPA for both groups ($P&lt;.005$).</td>
</tr>
<tr>
<td>41</td>
<td>De Cocker et al [64]</td>
<td>312</td>
<td>Sitting time in 3 groups: control, generic intervention, and tailored intervention</td>
<td>Self-reported total sitting time decreased more in tailored group compared with both generic group ($P=.002$) and control group ($P=.002$). But no significant difference in objectively measured data</td>
</tr>
<tr>
<td>47</td>
<td>Kattelmann et al [72]</td>
<td>1639</td>
<td>Total MET-minutes per week estimated from self-reported data</td>
<td>No difference between control and intervention for total MET-minutes per week ($P=.90$). Significant time effect for moderate MET-minutes per week ($P=.002$) and significant time × group × gender effect for vigorous MET-min per week ($P=.05$)</td>
</tr>
<tr>
<td>48</td>
<td>Partridge et al [68]</td>
<td>214</td>
<td>Self-reported PA data analyzed as MET-minutes per week</td>
<td>Significant effect of intervention on average MET minutes per week at 12 weeks ($P=.05$). Total PA days ($P=.003$) and number of walking days ($P=.02$) increased in intervention group</td>
</tr>
<tr>
<td>49</td>
<td>Walthouwer et al [71]</td>
<td>1419</td>
<td>PA duration in text-tailored, video-tailored, and control arm. In the tailoring group, 2 groups were compared, 1 where preference of user to video/text was matched and another without the matching</td>
<td>No significant difference in condition match/mismatch for PA ($P=.33$). Also, no significant difference for video-tailoring × intervention used ($P=.83$) and text-tailoring × intervention used ($P=.81$)</td>
</tr>
</tbody>
</table>

$^a$PA: physical activity.  
$^b$RL: reinforcement learning  
$^c$MVPA: moderately vigorous physical activity.  
$^d$SMS: short messaging service.  
$^e$MET: metabolic equivalent.  
$^f$LPA: light physical activity.
Discussion

Principal Findings

This study provides a review of studies on personalized technology–based interventions for increasing PA. This review adds to the PA literature in several ways. It provides an overview of personalization provided to users in the context of apps that aim at increasing PA. It examines various attributes, which can be personalized for encouraging the user, and identifies the theoretical frameworks used in these studies. This review included all research designs and, thus, provides a comprehensive view of ideas for effectively encouraging PA by means of personalization. We now discuss the review implications with respect to interventions, personalization, user models, theory and guidelines, and results.

Interventions

The widespread adoption of activity monitoring devices, increasing accuracy of data-driven prediction techniques, and ease of automation all facilitate the use of automated interventions. However, PA changes in patients who are under clinical observation may need to be assessed by a health care expert, leading to manual interventions.

Semiautomated systems combine and thereby aim to provide the best of both worlds—automated and manual interventions. Though these are often specialized for patients [75,80], they can also be available for the general user [2,35]. Having a health care expert–based intervention is less scalable but often necessary for patients under specific medical treatments. An interesting case of semiautomation is seen in the study by Dobrican and Zampunieris [75], where the targets were cardiac patients and the aim was rehabilitation. The doctor was involved for medical advice, but adaptive goals were set based on the European Society of Cardiology guidelines [109]. Note that there are arguments suggested against completely automated systems, for example, they have not been effective in weight loss [2].

Commercial fitness apps designed for the general user could take into account specific requirements of users with clinical conditions, including chronic diseases such as diabetes, who may benefit from such interventions. Current systems would need to include adaptive goal recommendation [54] to offer personalization in light of medical constraints and not just preferences of the users (eg, no swimming for elderly patients). From this review, we observe that user models for patients with chronic diseases are similar. PA guidelines, such as European Society of Cardiology’s guidelines [109] for cardiac patients or by Canadian Diabetes Association [110] and BCT–based design could be incorporated to enable effective behavior change.

Personalization

Interventions in the included studies were personalized in one or more ways. Recommendations were personalized with respect to goals, activity, or fitness partners. Feedback was personalized with respect to its educational or motivational content and, in some cases, its timing.

<table>
<thead>
<tr>
<th>Serial #</th>
<th>Paper reference</th>
<th>Method of study design</th>
<th>Dataset size</th>
<th>Variables evaluated</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Hermens et al [54]</td>
<td>Single-case experi-</td>
<td>8</td>
<td>Objectively measured activity behavior (activity level)</td>
<td>5 patients had increased PA(^a) level</td>
</tr>
<tr>
<td>11</td>
<td>Rabbi et al [45]</td>
<td>Single case experiment</td>
<td>16</td>
<td>Minutes of walking per day and calories burnt in nonwalking exercise per day</td>
<td>Intervention had significant effect for walking ($P&lt;.005$) and exercise ($P&lt;.05$)</td>
</tr>
<tr>
<td>16</td>
<td>Larsen et al [88]</td>
<td>Observational study</td>
<td>21</td>
<td>Change in minutes of MVPA(^b) using a semistructured interview among adolescent girls after 12 weeks</td>
<td>Statistically significant increase in weekly minutes of MVPA ($P&lt;.001$). Also reported activity types had larger variation than baseline</td>
</tr>
<tr>
<td>24</td>
<td>Coolbaugh et al [81]</td>
<td>Feasibility study</td>
<td>2</td>
<td>12 weeks of personalized intervention</td>
<td>Feasibility could not be ascertained</td>
</tr>
<tr>
<td>33</td>
<td>Ammann et al [79]</td>
<td>Observational study</td>
<td>803</td>
<td>Weekly total PA minutes across young, middle age, and old age groups</td>
<td>Significant increase in MVPA from baseline for older adults ($P&lt;.5$). All age groups increased weekly PA significantly ($P&lt;.05$) and walking minutes ($P&lt;.01$) over time in intention-to-treat analysis</td>
</tr>
<tr>
<td>44</td>
<td>Hales et al [74]</td>
<td>Pilot study and iterative usability study</td>
<td>9</td>
<td>Calories spent during intentional activity for users as compared with baseline</td>
<td>Calories expended increased from baseline but not statistically significant ($P=57$)</td>
</tr>
<tr>
<td>46</td>
<td>Spark et al [82]</td>
<td>Single group, pre- and post-test study</td>
<td>29</td>
<td>Duration of MVPA for participants in initial intervention (6 months), followed by extended contact information (6-12 months) and no contact follow up (12-18 months)</td>
<td>Significant improvement in minutes/day MVPA to 6 months from baseline ($P=.006$) and to 18 months from baseline ($P=.003$)</td>
</tr>
</tbody>
</table>

\(^a\)PA: physical activity.
\(^b\)MVPA: moderately vigorous physical activity.
Personalization was done either individually or in a category-based manner. The former includes individual models, for example, based on a user’s lifestyle [86], rate of progression [81], and preferences [32] or determined by a health care specialist [80]. In the latter case, category-specific personalization was provided after identifying the most appropriate category for the user. The categories were defined based on BCT [eg, 3] or activity status [eg, 2].

**User Models**

User models were created using a variety of different measurements, that is, PA profiles, demographics, medical data, BCT parameters, and contextual information.

Various parameters were used to evaluate PA, and all the profiles aimed to measure 1 or more “dimensions” of PA. An interesting visualization of multidimensional PA was proposed in 1 study [111]. The premise is that PA cannot be judged only on 1 criterion, for example, number of steps or time of vigorous activity, and has multiple dimensions including sedentary time. All the interventions for PA were restricted in their dimensions, and a multidimensional profile would be useful to obtain a holistic view of the user.

User models based on social profiles used the least amount of other contextual parameters. They promoted behavior change through social influence and are promising for both effective persuasion and user modeling. Among the included studies, social profiles were used for buddy matching [74] and also to post status data on social media to promote PA.

Personalized and dynamic user models can be created using the wealth of multimodal user data available from smartphones. Most of the existing apps do not use all the available data. PRO-fit utilized some of the available data sources—the phones’ geolocation, the users’ social network, and the users’ calendar—effectively [56]. By integrating all the available data, a richer profile can be created, and when combined with reinforcement learning techniques, the most effective interventional policies for each user can be learnt. As user behavior may change over time, it is important to employ online learning algorithms that can continuously monitor user models, adapt to their changing lifestyle patterns, and accordingly modify interventions as well.

**Theory and Guidelines**

Theory-based studies used BCTs to only make design decisions. Furthermore, 1 study did not completely define all the phases of TTM during the design process but utilized the readiness parameter defined by the model [45]. In addition, incorporation of BCTs was usually done via questionnaires in these studies, which may be infeasible or obtrusive to the user. Thus, automated learning of BCT parameters may be worth exploring. There is preliminary work in this direction. A user’s awareness of daily life for profiling users and their behavior [86] is another approach for estimating user behavior. User preferences could also be learnt through greedy approaches [86] or through inherent model design [45].

Another problem with methodologies based on BCTs is that they generally set a fixed ideal goal for a user. In contrast, PA guidelines suggest PA progression to prevent fatigue or muscular injuries. The generic goals of 60 min of PA or 10,000 steps may be too difficult and hence demotivating to a user who is sedentary or has clinical complications. Such users often require help, in the form of intermediate goals, to reach the final goal. PA guidelines can be utilized in such cases. There are attempts in studies [87] to use PA guidelines while using BCT for motivating users. Another study [78] also encourages its users, that is, cancer patients, to follow guidelines set by American Cancer Association [110] while planning their PA.

As identified across the PA literature, an “intention-behavior” gap exists among users. This poses the classic problem that although users are motivated and have intentions to increase their PA, they are not sufficiently active. Many studies were based on BCTs. However, healthy lifestyle induced during the intervention does not ensure that the user does not go back to a sedentary lifestyle after the intervention [54]. The sustained effects of interventions were not evaluated by all the studies but only by a few studies (e.g. [45,60,82]). Habit strength and formation has been addressed and evaluated in the study by Storm et al [7]. It is important that the sustained long-term effects of intervention are analyzed, as it would help to identify effective methods of promoting PA.

**Results of Individual Studies**

Direct positive results demonstrating the effectiveness of personalized interventions have been observed in a diverse set of studies—there are studies implementing data-driven automated systems [45], which recommend activities, whereas there are also studies which provide only personalized educational and motivational content [30]. These results indicate that activity or definite goal recommendation is not required for an effective personalized PA intervention. Effectively personalized motivational and educational content can help induce behavior change among the participants as well. It is also interesting to note that most of the studies with significant improvements are based on theoretical models (e.g. [7,30,54,88]). However, most of these studies also use self-reported values and collect data through questionnaires (e.g. [79,84,108]).

The self-reported PA values need to be considered with caution. As observed in the studies by Marsaux et al and De Cocker et al [35,64], there can be different results when self-reported and objectively measured data are compared. Thus, positive results obtained by interventions based on self-reported data need to be evaluated with objectively measured data through accelerometers and sensors. However, it can be inferred by the positive results obtained through BCT that their incorporation could help users, even if the users’ perceived PA level is incorrect. This makes it worthwhile to find ways of incorporating BCTs and theoretical guidelines in other data-driven–based interventions. However, it also needs to be noted that even the studies that do not show significant improvements use BCTs [47,74,90].
Studies have evaluated not just the PA metrics but also the intervention delivery mediums. The intervention delivery was found to not matter in the cases of video versus text [71] and SMS versus email [108]. However, a difference was found in the case of print versus Web-based intervention [27]. It is possible that the print medium was found to be effective as the participants were adults over the age of 50 years. However, further studies need to be performed to analyze the differences between intervention delivery mediums and their effects on the users.

The sample sizes of the studies reviewed vary considerably. The studies have also not been analyzed for quality to recommend future directions. However, our review indicates that there is scope for more rigorous evaluation in terms of intervention delivery, personalization, and intervention method. Many studies in the review perform pilot studies or feasibility studies or identify RCT protocols, which are yet to be completely evaluated. Evaluation of various systems to identify the effectiveness of intervention medium (along with the personalization aspect) in motivating users could be useful.

Limitations
This review was restricted to specific databases and an appropriate search query. It is possible that some studies may have been left out due to their journal or indexing bias. In addition, the search was restricted to a time frame that was considered relevant for the personalization aspect of the study and could again have led to studies being left out of the review. Moreover, as this is a scoping review, we have included studies without quality analysis and also studies without any evaluation. Though it helps identify the breadth of research, as the quality of studies is not assessed, the gaps identified may not be completely accurate.

Conclusions
This study provides a comprehensive review of personalized technology–based interventions, as recommendations or feedback, for promoting PA. Overall, the studies show that these interventions for increasing PA are more effective when they are personalized, compared with a “one size fits all” generic advice. Gaps have been identified in several aspects, such as in the development of a multidimensional user model and the use of behavioral theory in automated personalization. On the basis of these gaps, research directions for improving the efficacy of personalized technology–based interventions have been suggested.

Acknowledgments
This work was supported by the Ministry of Education, Singapore (grants R-253-000-129-114 and R-253-000-138-133).

Conflicts of Interest
None declared.

Multimedia Appendix 1
Overview of studies in the review.

[PDF File (Adobe PDF File), 97KB - mhealth_v7i1e11098_app1.pdf]

Multimedia Appendix 2
User model parameters for studies in the review.

[PDF File (Adobe PDF File), 83KB - mhealth_v7i1e11098_app2.pdf]

References


75. Dobrican R, Zampunieris D. A Proactive Solution, using Wearable and Mobile Applications, for Closing the Gap between the Rehabilitation Team and Cardiac Patients. : IEEE; 2016 Presented at: IEEE International Conference on Healthcare Informatics (ICHI); 4-7 Oct, 2016; Chicago, IL, USA p. 146. [doi: 10.1109/ICHI.2016.23]


82. Spark LC, Fjeldsoe BS, Eakin EG, Reeves MM. Efficacy of a text message-delivered extended contact intervention on maintenance of weight loss, physical activity, and dietary behavior change. JMIR Mhealth Uhealth 2015;3(3):e88 [FREE Full text] [doi: 10.2196/mhealth.4144] [Medline: 26373696]


109. McMurray JJ, Adamopoulos S, Anker SD, Auricchio A, Böhm M, Dickstein K, Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology, et al. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. Eur J Heart Fail 2012 Aug;14(8):803-869. [doi: 10.1093/eurjhf/hfs105] [Medline: 22828712]


Abbreviations

- **BCT**: behavior change technique
- **CDC**: Centers for Disease Control and Prevention
- **CVD**: cardiovascular disease
- **HAPA**: health action process approach
- **HR**: heart rate
- **LPA**: light physical activity
- **MET**: metabolic equivalent
- **MVPA**: moderately vigorous physical activity
- **PA**: physical activity
- **RCT**: randomized controlled trial
- **RL**: reinforcement learning
- **RNN**: recurrent neural networks
- **SCT**: social cognitive theory
- **SMS**: short messaging service
- **SVM**: support vector machine
- **TPB**: Theory of Planned Behavior
- **TTM**: Trans-Theoretical Model