Health recommender systems for behavior change

HEALTH RECOMMENDER SYSTEMS FOR BEHAVIOR CHANGE:

Exploring their potential for smoking cessation

Santiago Hors-Fraile

HEALTH RECOMMENDER SYSTEMS FOR BEHAVIOR CHANGE:

Exploring their potential for smoking cessation

Dissertation
to obtain the degree of doctor at Maastricht University
and University of Seville
on the authority of the Rector Magnificus Prof. Dr. Pamela Habibović,
in accordance with the decision of the Board of Deans,
to be defended in public on
Thursday, 24th of March 2022 at 13:00 hours
in Maastricht

Santiago Hors-Fraile
He was born on November 19th, 1987 in Seville (Spain).

Supervisors
Prof. Dr. Hein de Vries (Maastricht University)

Co-supervisors
Dr. Francine Schneider (Maastricht University)
Dr. Luis Fernández-Luque (University of Seville)
Dr. Antón Civit-Balcells (University of Seville)

Assessment Committee
Prof. Dr. Stef Kremers, Maastricht University (chairman)
Prof. Dr. Arie Dijkstra, University of Groningen
Prof. Dr. José Luis Sevillano, University of Seville
Dr. Marcel Pieterse, University of Twente

HEALTH RECOMMENDER SYSTEMS FOR BEHAVIOR CHANGE:
Exploring their potential for smoking cessation

PhD thesis, Maastricht University, the Netherlands

The research presented in this thesis was conducted at CAPHRI Care and Public Health Research Institute, Department Health Promotion, of Maastricht University. CAPHRI participates in the Netherlands School of Public Health and Care Research CaRe.
Department of Computer Architecture and Technology. Higher Technical School of Computer Engineering. University of Seville.

ISBN: 978-94-6423-607-1
Cover design: Bregje Jaspers
Layout: Dennis Hendriks || ProefschriftMaken.nl
Printed by: ProefschriftMaken.nl

Table of contents

CHAPTER 1

General introduction and outline of the thesis

The problem of smoking cessation

Smoking is harmful to the health of human beings as it leads to chronic diseases and has detrimental effects on many organs [1]. Its regular consumption is one of the several causes of many types of cancer, heart diseases, stroke, chronic obstructive pulmonary disease (COPD), diabetes, respiratory infections, cataracts, tooth decay, skin stains, premature aging, and a higher risk of tuberculosis [2-6]. As the first avoidable risk factor for death worldwide [7, 8], smoking is a global problem that killed approximately 100 million people in the 20th century [9]. It kills seven million people each year, costing about $1.4 trillion [10]. For instance, in the Netherlands, it is estimated that there are around 2.5 million smokers (14.46% of the population), and nearly 30 thousand smokers die every year from tobacco-related diseases [11]. In Spain, these figures are as high as 9.6 million smokers (20.45% of the population), and nearly 57 thousand die of tobacco-related diseases [11]. Therefore, quitting smoking is of paramount importance to reduce the number of cases of tobacco-related diseases and, consequently, save lives.

Tobacco smoking is an addictive practice [12] that causes many smokers to relapse when they try to quit. Relapsing happens as a result of several factors. First, due to the chemical addictive effects of smoked nicotine [13, 14], smokers deprived of it may develop nicotine abstinence syndrome, which may result in cravings, weight gain, irritability, intestinal disorders, headaches, especially in smokers with heavy dependence [15]. Moreover, the social context of the smokers may act as their trigger to smoke [16-20]. Additionally, the learned behavior of smoking in certain situations, such as smoking at parties and every day, makes it a difficult-to-change habit [21]. Finally, smokers may also have personal factors triggering their desire to smoke, including stress, anger, anxiety, food and drink consumption (especially alcohol), or even events affecting their life [22-26]. All these reasons for relapse make smoking cessation a severe challenge.

Smoking cessation approaches

Due to the adverse health effects of smoking, many smokers want to quit. Just in the US, there are 22.7 million smokers (68% of adult smokers) [27]. Quitting to smoke is a challenging process and may require the involvement of different strategies. Unaided, abrupt cessation or "cold-turkey" is the most commonly used method [28, 29]. However, it is the least effective one [30], with a 3–5% success rate of prolonged abstinence for 6–12 months [31]. Nevertheless, in the past few decades, more successful approaches have been identified - pharmacotherapy on the one hand and psychological or motivational interventions on the other.

Pharmacotherapy requires the utilization of medication in order to reduce symptoms of nicotine withdrawal. One method is the usage of nicotine replacement therapies [32], which are effective in increasing one-year abstinence (OR 1.71, 95% CI 1.55–1.88) [33]. Moreover, drugs can be used to reduce cravings and inhibit the feeling of pleasure when smoking. For instance, the antidepressant Bupropion [34] may double the likelihood of people remaining abstinent compared to a placebo (OR 1.85, 95% CI 1.63–2.10) [35]. The smoking satisfaction reduction drug Varenicline [36] also increases the chances for continuous or sustained abstinence for a six-month or a longer period as compared to a placebo (risk ratio 2.24, 95% CI 2.06–2.43) [36], as found at one-year follow-up (OR 2.96, 95% CI, 2.12–4.12) [33]. Finally, despite the proven benefits of the pharmacological treatments for supporting people to stop smoking, many smokers fail as these treatments are not always effective [37-40]. However, some pharmacological treatments may produce severe side effects, such as terrible nightmares [41-43] or insomnia [42, 44, 45], which compel people to refrain from taking these drugs and relapse. Nicotine replacement and drug usage approaches are intended to reduce the chemical addictive effects of smoking. Despite the effectiveness of these approaches, they do not account for the personal and social addictive nature of smoking that plays a significant role in the acquisition of a smoking habit.

Another means to facilitate smoking cessation is through psychological interventions that aim at increasing a smoker's motivation to quit smoking and developing behavioral skills to cope with difficult situations that may lead to relapse. One way of providing these interventions is through a physician offering brief pieces of advice on smoking cessation. However, this has a small effect, increasing only 2–3% of smoking cessation rates over unaided attempts [46]. A more extensive behavioral intervention is a behavioral therapy that can tackle those addictive factors of smoking, which have been previously described, that neither nicotine replacement therapy nor pharmacological treatments can—the social context and learned behaviors. In general, behavior therapy focuses on changing the smoker's behavior using principles of learning theory [47]. There are different types of behavioral therapy approaches [48]. One of them is motivational interviewing. In this approach, therapists guide their patients toward higher commitment and personal motivation by paying attention to the used language, generating an environment of acceptance and compassion [49-53]. Another approach is contingency management, where the patients' motivation is aimed to be reinforced by offering them a monetary reward for quitting [49-53]. Another popular approach is the cognitive-behavioral theory [54-56]. In this method, the therapist examines, identifies, and targets their patients' thoughts and behaviors that lead them to smoke. Then, they deal with the related factors (e.g., risk perceptions, self-efficacy to quit, beliefs) by providing encouragement, education, and training of coping skills to stop smoking through different techniques (such as role-playing and cognitive restructuring) [57].

One way to deliver behavioral therapy is via personal counseling—the counselor builds a supportive relationship with the smoker patients and helps them gain insights into their own behaviors. For this method, a smoker-focused approach is implemented to guide patients toward making a self-chosen change through communication strategies [58]. Counseling can be done face to face or even remotely (over the phone or in groups). It has been proven to be effective in supporting one-year continuous abstinence with around a 25% success rate [59]. On comparing individual counseling to minimal support, such as brief physician advice and self-help materials, a study showed that the former approach could increase about 40–80% of smokers' chances of quitting smoking after six months, resulting in an approximately 11% success rate [58]. Counseling can also be provided in groups, which has the potential benefit of peer feedback, modeling behaviors after how a peer handled their problems, improving social skills, and saving costs [60]. However, according to the Cochrane systematic review, there is insufficient evidence to support whether group counseling is more or less effective than individual counseling [61].

The described quitting support approaches are not mutually exclusive. Some can be combined for generating beneficial effects [62]. The highest success rates have been achieved in intensive interventions, combining pharmacological treatment and behavioral support [63], with 35% success rates for one-year abstinence [64]. However, the magnitude of the cessation success rates found in the literature for the various types of smoking cessation support approaches may vary as they have all been measured differently [65]. Also, the follow-up periods and definition of abstinence (continuous abstinence or abstinence for a given number of days before the follow-up) are not consistent across studies [66]. A group of experts concluded that the following to be the most relevant measures: a) a seven-day point prevalence abstinence (not smoking for the last seven days), b) continuous abstinence for one and six months (not consuming any tobacco product since the quitting day although some slips may be allowed, depending on the study), and c) prolonged abstinence for 6 and 12 months (not consuming any tobacco product since the quitting day, excluding the grace period after the quitting date, so the relapses in the first few days are not taken into account) [66]. Moreover, the method to check abstinence (self-reported or with biochemical validation) is a variance factor that needs to be considered when interpreting the results that indicate the success rates [66, 67] since biochemical validation provides higher internal validity.

Although optimal rates were achieved combining pharmacological and behavioral treatments, behavioral therapy-based interventions on their own can produce considerably beneficial results. For example, the study by Sykes et al. [68] achieved a seven-day point prevalence abstinence of 17.2%.

Behavior change models

Behavioral therapies rely on evidence-based theories that intend to explain human health behavior [69-71]. These theories are essential because they are used to guide and structure the process of acquiring a new behavior: The theories present different determinants of the behavior, and during therapy, these determinants are addressed and positively changed, which results in the development of desired behavior. Several of these theories have been used in the smoking cessation context. They are the theory of planned behavior [72], social cognitive theory [73], goal setting [74], health belief model [75], and cognitive-behavioral therapy [76]. The transtheoretical behavioral change model [77] is a particular case since it was specifically developed considering smoking cessation in adults as the design reference. All these theories consider the determinants of different behavior changes and explain and aim to predict behaviors differently. As such, each theory has limitations as they exclude factors that influence behaviors and make different assumptions. However, many of these theories can be combined to generate a more comprehensive view of human behavior. One of these combined models is the Integrated Behavioral Change model (I-Change) [78]. It has evolved from the original Attitude–Social Influence–Self-Efficacy model [79-81], which grouped the theory of planned behavior, social cognitive theory, goal setting, health belief model, and transtheoretical behavioral change model.

The I-Change model proposes that the behavior change process consists of four phases: awareness, motivation, action, and actual behavior. People move from the first phase to the fourth one based on different behavioral determinants in each phase, which can be shaped by information factors (channel, source, message, and personal), and preceding factors (biological, psychological, behavioral, environmental). Each phase has determinants of change that influence explain an individual more to transition to the next phase. This empirically tested model provides a comprehensive view of the human behavior change process and can the human behavior of smoking cessation [82-86], making it relevant to be considered in behavioral therapies.

Using these theoretical models, behavioral experts can bring about people's health behavioral change through reflection and counsel a person regarding actions needed to change relevant determinants. However, the availability of behavioral experts is limited, and counseling is also a time-intensive [87-92] and, consequently, costly process.

Computer-based interventions

With the advancements in technology, new approaches have been developed to support healthy behaviors beyond consultation and reduce the time needed to tailor behavioral change support advice. The first steps were taken by personalizing the computer-generated letters and sending them by regular post [93-95]. Some interventions were CD-ROM-based multimedia applications [96, 97], whilst others were included in computer kiosks [98, 99]. The introduction of mobile phones made computer-generated SMS a popular method of delivery [100-102]. Yet, due to the emergence of the Internet, the number of ways to communicate tailored information has increased significantly [103].

In computer-based interventions, the delivery of tailored texts in e-mails [85, 104-106] and videos [107-111] is typical. The potential advantage of reaching a larger audience faster and at lower costs has increased the number of studies that address and test online delivery of behavioral support in recent years. Such usage of computers to provide personalized behavioral feedback to the user is known as computer tailoring [112].

Traditional computer tailoring

In traditional computer tailoring, computers adjust the health materials as per the person to make them relevant and credible for their situation [113], replicating what an actual human counselor would do [114]. In this process, computers use identifiable characteristics of the person from the health information, which would be delivered (e.g., name and references to previously provided data, such as the number of smoked cigarettes), to generate customized information that specific smoker patient would find more relevant. Moreover, the information is adapted to the person's context (e.g., the information provided to an unemployed, pregnant smoker would be different from the one sent to a working man with COPD). It is then provided to make the people aware of their status (e.g., blood pressure decreased from abstinence) [115, 116]. All these adjustments for message personalization facilitate more intensive information processing and make the messages more memorable [113]. Furthermore, more relevant messages lead to higher message processing, elaboration, and attention and are more persuasive. This effect is described in the elaboration likelihood model [117]. The aforementioned personalized message approach has been effective in health communication and promotion interventions [118] in, for instance, smoking cessation [119-124], reducing alcohol drinking [125-127], sunscreen usage [128], physical activity [129-131], nutrition [114, 132, 133], and cancer screening [134, 135].

The first generation of computer-tailored interventions used systems that computed messages using the if-then rules. Users, thus, received messages matching their individual characteristics. This type of system is known as rule-based expert systems, which are a part of the artificial intelligence (AI) field [136]. To create these systems, first, a behavioral model was chosen to build the architecture of behavioral counseling. Then, questions were developed to identify the core variables that influence the motivational processes and the resulting behavioral choices. To enable the creation of personalized messages, users needed to complete a questionnaire. The smokers' responses to these questions were linked to algorithms that matched appropriate messages that fit the responses of the smoker (e.g., gender, health history, age, beliefs, etc.), resulting in a personalized message.

Rule-based expert systems used for computer tailoring have, however, some limitations. Expert systems excel in replicating human knowledge but cannot create new ones. This means that, as we have presented, computer tailoring technology cannot self-learn once the initial working rules are set. Additionally, the high levels of personalization require obtaining data to create a set of rules through many questions, resulting in high user burden, reduced engagement, and elevated dropout rates [137-139].

Recommender systems

The next generation of computer-based systems are recommender systems that use a different type of technology also used by large platforms, such as Netflix [140], Trivago[141], and Amazon [142], to recommend personalized content to watch, hotels to book, and items to purchase, respectively. In the health context, these systems are called health recommender systems (HRS) [143]. HRSs vastly differ in their algorithm philosophy compared to traditional computer-tailoring, going beyond the if-then programming paradigm. HRSs fall into the category of machine learning in AI. In contrast with the more static approach of traditional computer-tailoring interventions, HRSs can learn from and adapt to their users. Instead of using strict rules and knowledge defined by experts based on behavioral models, HRSs use user interactions data with the system as the driving factor for message selection. Typically, HRS users are required to rate the message relevance through a feedback system and help it "learn" how good the recommendation was [144]. In addition, an HRS may use several pieces of information, including the features of the contents of the potential message to recommend, the users' past message preference history, and their demographic and clinical profile. Through this, HRSs can predict the message with the highest relevance for a user among the potential messages in a dataset [145], consequently reducing the need for lengthy questionnaires, as in the case of expert systems.

HRSs can also take advantage of the information related to all users in the system to compute for message selection instead of only using the information for a specific user. This is known as "collective intelligence" [146]. Such systems determine the most relevant content for a given user based on what other similar users previously rated as relevant for them. This reduces the need for lengthy questionnaires because it leverages user feedback provided on the actual recommendations. In addition, it can foster serendipity (unexpected positive outcomes). It contrasts with deterministic expert systems, which provide valuable recommendations that, in principle, would not have been selected by the user [147].

As recommender systems may be used for tailored behavior change interventions, an unresolved challenge is determining what is known about their application in the healthcare context. Since this is still unknown, researchers may find it challenging to design HRSs properly, reducing their ability to impact patients. This is explored in more detail in Chapter 2, where the following research question is answered: "**RQ1: What is the existing knowledge regarding the usage of health recommender systems for patient interventions?**" To answer this question, we will perform a comprehensive analysis of the existing literature and propose a taxonomy for HRSs that can be followed to facilitate the classification of studies and identify their critical elements from a multidisciplinary perspective.

Types of recommender systems

There are different types of recommender systems. We can group these systems under four categories depending on the type of algorithm: content-based filtering, collaborative filtering, case-based recommenders, and hybrid approaches.

Content-based filtering algorithms select messages based on the correlation between the content of the messages and the users' preferences [148]. In other words, they follow the philosophy of "recommend me items based on my past actions and preferences." These preferences can be set by the user directly or be inferred from their actions/ratings. For example, a computer-based smoking cessation support system detects the characteristics of the messages marked to be useful. These characteristics can include length, topic, source, applied health communication techniques, proposed call to action, structure, and complexity. The next time the smoker wants motivation to resist smoking, the system checks what message characteristics the smoker prefers and then recommends a new piece of text that matches that preference as much as possible. This type of algorithm works well for systems with any number of users as they only depend on the user's previous preference history. However, the system must always know all the characteristics of the items to recommend and the items the user has selected or consumed previously [149].

Collaborative filtering algorithms are the type of recommender systems that leverage collective intelligence. They select items based on the correlation between users or items that are similar. This similarity can be defined as a similarity between users or items. In the similarity between users, also known as user-item filtering, the system assumes that the items liked by the people who are similar to a given user will also be liked by that user. In this case, the similarity can be in terms of rating history or demographic details (same age, gender, location, etc.). When the latter type of similarity is used, it is referred to as demographic filtering systems. In the case of similarity between items, the system looks for items similar to those that the user positively rated, considering how other users rated them, which is known as item-based filtering.

Case-based recommender systems retrieve relevant items for the conditions imposed by the user. A case-based recommender system uses a set of already-solved cases to solve the current problem by matching user preferences with item descriptions. For example, a user may request a message for a woman aged 36, mother of two children, unemployed, and with high nicotine dependence and high social support to stop smoking. Then, the system looks for messages already configured to match those specific requirements. They can be conceived as a configurable decision tree but try to maximize user demands if they cannot match one of the existing branches. If one or more criteria are not possible to be met, the system provides the closest recommendation. For instance, in our previous example, the system may not have a combination for all the mentioned variables but may find a message for a woman, a mother, with high nicotine dependence and social support (except for the employment status and age). These recommender systems are very similar to the traditional expert systems used for computer tailoring. Knowledge-based recommenders are a type of case-based recommenders [150]. They have an abstraction layer that makes the interaction with the system less technical and more friendly. For instance, a case-based recommender for smoking cessation support could be: "Hello system. What can you briefly tell me to support me to stay smoke-free on a cold night of wedding party, where my friends are indoors smoking and drinking, where I have been smoke-free for just 23 days already, and where I am a 46-year-old male with medium nicotine dependence and high motivation to quit?" "Well, I can give you a piece of advice that fits your variables the most, but it is intended for people with low nicotine dependence. Alternatively, I can give you a message that is rather lengthy but fills all your other requests. Is any of these two options OK for you?" A similar conversation as a knowledge-based recommender would be: "Hello system. I am in a similar situation to the one I was in on my partners' birthday, and I'm tempted to smoke. Can you help me?" "Well, I can give you a message that is intended for social pressure to smoke in parties."

Hybrid approaches combine any of the three methods above in different ways. For instance, they can be combined in cascade (the output of one is the input of the next one); they can be averaged (the output is the highest-ranked recommendation after combining the scores for all possible recommendations provided by different algorithms), and they can switch and provide recommendations generated from different algorithms each time [151].

Challenges of behavioral interventions using health recommender systems

Despite their many advantages over the previously described expert systems, HRSs also have limitations. First, it is challenging to map the features of the behavioral messages to be delivered and to generate a significant amount of messages, and ensure variety to cover at least one message for all behavioral cases [149, 152]. For example, in the case of a smoking cessation intervention that considers age (teenagers, younger adults, older adults, and senior citizens—four cases), gender (male and female—two cases), nicotine dependence (low, medium, and high—three cases), motivation to quit (low, medium, and high—three cases), there will be 72 possible combinations (4*2*3*3 cases) of user profiles. The designers will need to ensure that each of the 72 possible combinations is considered for addressing each relevant topic. However, no study has been able yet to define the process of generating the messages to cover that many combinations. Second, the lack of initial user data for generating accurate outputs impacts many HRSs that rely only on user history to compute recommendations—known as the cold-start problem [153]. The lack of user data, when the users do not fulfill their user profile, to personalize the recommendations can also affect HRSs that use the user profile variables such as gender, age, smoking behaviors, etc.

Consequently, when user data is limited or not available, the potential for recommendation is also limited. Last, the lack of existing health (behavior) datasets, which exist for other contexts, is a barrier in testing the technical performance of an HRS [154]. For example, MovieLens is a well-known dataset of rated movies [155], which can be used to benchmark algorithms compatible with the movie recommendation context. However, that is not the case in the healthcare domain, as we also need to factor in other metrics such as behavior change and health outcomes. Furthermore, health interventions cannot depend only on benchmarks and, in any case, should be validated by actual human trials. For the second and third challenges, an HRS needs to be built that considers and minimizes these issues.

Due to these limitations, there may have been a scarcity of research in the healthcare domain concerning recommender systems. Although there are several studies recommending recipes and lifestyle activities, they do not focus on health results and do not use behavioral science [146, 156-160]. However, there is a growing interest in applying recommender systems in actual healthcare contexts [161]. Among the studies that focused on supporting patient healthcare, the Perspect intervention [162], the SoLoMo intervention [163, 164], and the SocialPOD [165] were the identified pioneers that use behavioral science—all of them were published in 2016. The first two focused on smoking cessation, while the last one focused on weight loss. Yet, only a minority of the studies used HRS-integrated behavioral change models in their design [166]. This fact needs to be addressed since the impact of HRS-based interventions may be limited due to the lack of behavioral science grounding. To solve this problem, in Chapter 3, we comprehensively disclose a step-by-step process of how to create a complete HRS for a smoking cessation intervention that integrates the I-Change behavioral change model [78] in its design to deliver smoking cessation supportive messages. It presents a solution to the identified needs of describing a process to generate messages for HRSs. Furthermore, it allows the system to compute safe behavioral recommendations using behavioral science evidence, with less user information and without using previous datasets. All this translates into the following research question "**RQ2: How can we combine behavioral science and recommender systems technology to support behavioral change?"**

Relevant success factors: Engagement and appreciation

In recent years, recommender systems have been started to be used in computer tailoring. However, there is scarce evidence about their impact on behavioral change support. Consequently, it is required to study whether HRSs can bring effective positive behavioral change support. However, health communication methods using information systems, such as traditional computer tailoring and HRSs need to engage with people and be appreciated by users to be effective [167-169].

The concept of engagement has different meanings depending on the context [170], and there are several means to measure it. For example, Yardley et al. [171] proposed to differentiate between engagement at the micro-level (with the digital intervention) and macro-level (with the actual pursued behavior). Perksi et al. [172] proposed an integrative conceptual framework where context plays a crucial element in engagement. Short et al. [173] provided a comprehensive overview of the engagement construct measurements in eHealth and mHealth behavioral change interventions. In this study, the researchers reflected on the need to use multiple methods to measure this complex construct using system usage data and psychological aspects, including, for instance, social dimensions. For this thesis, we consider the definition of engagement the micro-level proposed by Yardley et al., which covers user participation and involvement with the intervention to manage health to achieve the desired goals [169, 174, 175]. These goals may include digital metrics such as the spent time for an intervention, the number of counseling sessions, web page visits, etc.

Many studies have shown that higher engagement in smoking cessation interventions may lead to better cessation outcomes [176-182] as low-engaged participants may not need to be exposed to enough "dose" of the intervention to be effective. Smokers need to engage with the internet-delivered digital smoking cessation interventions as a precondition for these interventions to be effective [171]. In the case of Internet-based interventions, the number of web page visits has been identified as a positive predictor of abstinence [183, 184]. A lower engagement and, consequently, a lower usage indicate a negative impact on health behavior change [185]. For example, in the study conducted by Borland et al. [93], higher compliance to a computer-tailored advice intervention by mail was associated with better cessation results. Moreover, in the study by Christofferson et al. [184], it was found that users with higher engagement—the total number of messages sent each week by each user to the system —were significantly more likely to quit smoking after five weeks compared to those with lower engagement levels. The smoking cessation intervention by

Elfaddali et al. [105] has also shown a dose-response effect between the engagement to the smoking cessation program elements and the smoking effect.

Despite the positive relationship between engagement and intervention effects, there are high low usage levels in digital behavioral change interventions [186]. The factors that may influence engagement include the intervention characteristics (e.g., the duration and the number of sessions, actions requested to be done by smokers and prescribers), the characteristics of the user (e.g., age, gender, personality, and technology literacy), and the contextual factors (e.g., social support, and legislation). The more attractive the intervention characteristics are considered, the more engaging the interventions would be. Higher attractiveness can be achieved by, for instance, including videos or animations rather than text [109], adding tailored content relevant to the user, and pushing elements such as automatic notifications [187].

The user demographics is another factor that may impact the engagement in eHealth interventions [188] and the personal user characteristics, such as their effort in the intervention [189]. An analysis of three internet-delivered behavioral trials concluded that user characteristics such as education, self-efficacy, gender, age, and peer influences may be related to higher system usage [190]. In addition, contextual factors such as healthcare providers prescribing the eHealth intervention and offering face-to-face counseling may also contribute toward higher engagement [191].

A positive significative relation can also be found between engagement and better health behavior outcomes when considering the role of HRSs in delivering tailored smoking cessation interventions [164]. In this case, the engagement was calculated by comparing the number of tailored messages rated by the smokers and all the messages they received. This study with HRSs does not consider the spent time on the app, the number of logins, or any other factor related to engagement.

Recommender systems can also be measured by user perceptions of the system qualities and the level of user appreciation [192]. The appreciation of the provided intervention may also be an indicator of changes in behavior [184, 193]. Kroeze et al. [194] defined appreciation as attraction, interest, novelty, reliability, trustworthiness, personalization, comprehensibility, motivation and not irritation, and appropriateness of the information. As highly appreciated interventions may lead to more and longer use [195], they may increase user engagement and overall effectiveness and impact. Appreciation assessment has been extensively used in different interventions [111, 121, 196, 197]; however, it is still beneficial to find a way to increase appreciation of the interventions, as such results are not always consistent. Moreover, the technical metric of “precision” could be considered a proxy for the appreciation of the system as it measures how accurate the recommendations are based on user feedback when they rate a message. Higher ratings imply higher precision because the system managed to select a message well appreciated by the user.

In conclusion, in general, user engagement and appreciation are essential for health communication and, thus, also for eHealth interventions, including those with HRSs, and several factors may impact the engagement levels. In addition, there is relatively little knowledge of user engagement for HRS-based interventions. A more comprehensive perspective of engagement and appreciation in relation to the system may be beneficial in further understanding the HRS-based interventions. Consequently, we describe the results of a study on smoking cessation comparing two HRS algorithms—one with collective intelligence grounded in behavioral science and another without collective intelligence. The main research question to be answered is **RQ3: "Can an HRS grounded in behavioral science benefit from collective intelligence to improve user appreciation, user engagement, and generated smoking cessation outcomes?**" This study has been presented in Chapter 4, and its results are explained in Chapter 5.

Outline of the dissertation

The following chapters will explore the expanding field of recommender systems as a strategy for patient support interventions.
Recommender systems are being increasingly used as a strategy for patient support interventions. However, the previously generated knowledge has not been analyzed from a multidisciplinary perspective, including a behavioral perspective, and the existing research gaps are still unknown, leading us to the following question:

- **Research Question 1 (RQ1):** What is the existing knowledge regarding the usage of health recommender systems for patient interventions?

In Chapter 2, we answer RQ1 by analyzing the state-of-the-art in HRSs to identify the existing research gaps—the aspects that can be improved in the future. In addition, we propose a taxonomy for HRSs to facilitate consistent classification that helps us better comprehend these systems in the healthcare domain and can also be used in future studies.

Derived from the analysis conducted in Chapter 2, we found that HRSs do not use or describe behavioral science to generate recommendations for providing patient behavior support. Hence, reducing their reusability and replicability leads us to the following question:

- **Research Question 2 (RQ2):** How can we combine behavioral science and recommender systems technology to support behavioral change?

In Chapter 3, we answer RQ2. For this, we present the design of a mobile app via a health recommender system by using the promising collective intelligence technology for the specific case of smoking cessation, grounded in the I-Change model for behavior change. We comprehensively detail how the message content, the mobile application, and the recommender system were developed.

Yet, HRSs can have different types of algorithms. They will severely condition the behavioral outcomes of the intervention. In addition, we should also consider how they influence other relevant indicators, such as system appreciation and engagement. Algorithms with collective intelligence seem to be the type of HRS that attract more interest due to their potential benefits. However, it is not clear whether this type of algorithms potentially has advantages over other approaches. Consequently, we face the following question:

- **Research Question 3 (RQ3):** Can an HRS grounded in behavioral science benefit from collective intelligence to improve user appreciation, user engagement, and generated smoking cessation outcomes?

Chapter 4 presents the protocol for conducting an open trial to support smokers quit in a non-clinical context (without face-to-face counseling) by comparing the HRS designed in Chapter 3 with a similar version but without collective intelligence for the achieved appreciation, engagement, and smoking cessation outcomes. Finally, Chapter 5 presents, discusses and reflects on the achieved results.

CHAPTER 2

Analyzing recommender systems for health promotion using a multidisciplinary taxonomy: a scoping review

This chapter has been published as:

Hors-Fraile, S., et al.,
Analyzing recommender systems for health promotion using a multidisciplinary taxonomy: a scoping review.
International journal of medical informatics, 2018. 114: p. 143-155.

Background: Recommender systems are information retrieval systems that provide users with relevant items (e.g., through messages). Despite their extensive use in the e-commerce and leisure domains, their application in healthcare is still in its infancy. These systems may be used to create tailored health interventions, thus reducing the cost of healthcare, and fostering a healthier lifestyle in the population.

Objective: This paper identifies, categorizes, and analyzes the existing knowledge in terms of the literature published over the past 10 years on the use of health recommender systems for patient interventions. The aim of this study is to understand the scientific evidence generated about health recommender systems, to identify any gaps in this field to achieve the United Nations Sustainable Development Goal 3 (SDG3) (namely, "Ensure healthy lives and promote well-being for all at all ages"), and to suggest possible reasons for these gaps as well as to propose some solutions.

Methods: We conducted a scoping review, which consisted of a keyword search of the literature related to health recommender systems for patients in the following databases: ScienceDirect, PsycInfo, Association for Computing Machinery, IEEExplore, and Pubmed. Further, we limited our search to consider only English-language journal articles published in the last 10 years. The reviewing process comprised three researchers who filtered the results simultaneously. The quantitative synthesis was conducted in parallel by two researchers, who classified each paper in terms of four aspects—the domain, the methodological and procedural aspects, the health promotion theoretical factors and behavior change theories, and the technical aspects—using a new multidisciplinary taxonomy.

Results: Nineteen papers met the inclusion criteria and were included in the data analysis, for which thirty-three features were assessed. The nine features associated with the health promotion theoretical factors and behavior change theories were not observed in any of the selected studies, did not use principles of tailoring, and did not assess (cost)-effectiveness.

Discussion: Health recommender systems may be further improved by using relevant behavior change strategies and by implementing essential characteristics of tailored interventions. In addition, many of the features required to assess each of the domain aspects, the methodological and procedural aspects, and technical aspects were not reported in the studies.

Conclusions: The studies analyzed presented few evidence in support of the positive effects of using health recommender systems in terms of cost-effectiveness and patient health outcomes. This is why future studies should ensure that all the proposed features are covered in our multidisciplinary taxonomy, including integration with electronic health records and the incorporation of health promotion theoretical factors and behavior change theories. This will render those studies more useful for policymakers since they will cover all aspects needed to determine their impact toward meeting SDG3.

Introduction

In order to achieve the United Nations Sustainable Development Goals, particularly goal 3, "Ensure healthy lives and promote well-being for all at all ages" (SDG3), it is imperative to invest in health-promotion activities. Over the years, numerous health-promotion interventions have been developed that help people adopt a healthy lifestyle and independently manage their health behaviors. Even though these interventions have been proven to be effective [198], they are not suitable for all as populations tend to present high levels of variability. In order to account for these differences, it is important to tailor the interventions to suit the diverse characteristics of a given population (i.e., economic standards, schedules, and residential location). Given this variability, new technologies can be used to solve geographical-access problems, deliver timely interventions, reduce intervention costs, and to even help users exert better control over the intervention [199].

However, computer-based health interventions suffer from a high user attrition rate [200], which presents a severe problem in public-health actions related to medical informatics. This is why it is relevant to use tailored health interventions [201], which can increase user engagement [202]. Tailored health interventions can also ensure more effective outcomes as compared to non-tailor approaches [118, 120, 203, 204], and the integration of computers can make them scalable and even more cost-effective [119, 205].

As technology evolves, new ways to implement such tailored interventions are being adopted, and researchers and policymakers need access to the correct tools to help them assess their design and usage suitability. One such innovative approach to computer-based tailored health interventions is the use of recommender systems (RS) [206]. RS are machine-learning, information-retrieval software tools, which predict the relevance of an item (e.g., a health resource or a message) for a given user (e.g., a patient) [207]. RS can select, tailor, and send health messages that are relevant to users based on previously retrieved user information. Even though RS have gained popularity in the last decade [208] and have been applied in a wide range of domains, such as e-commerce and leisure, their application in the health-promotion domain—as health recommender systems (HRS)—is still in its infancy. Although some HRS are already in use, there is still a long way to go before they become commonly used in health-related environments [209]. One reason for this could be that the potential of these systems [144] is not clearly defined and known to health professionals. For instance, they could be used as clinical-decision support systems if the end user is a healthcare professional, and as engines to generate relevant healthy lifestyle recommendations when patients are the end users. This latter application could significantly contribute to the field of health promotion. Nevertheless, some challenges should be solved such as legal liability and regulatory compliance. Currently, the legislative frameworks are not fitted to deal with potential errors of HRSs [210].

When sending health-promotion messages to the population by running public health campaigns or, more specifically, by using health-promotion interventions, researchers in social marketing have reported that tailoring the content of these messages to the user's context can improve their efficacy, as compared to the use of general content [211, 212]. The added value of this strategy is that the user will then receive highly tailored messages tailored to his attitudes, social support system, self-efficacy, and the action plans needed to realize a particular health behavior. Yet, eHealth programs, including tailored eHealth programs, suffer from high dropout rates [200]. One strategy aimed to overcome this is to offer messages that are also optimally adapted to user preferences, a strategy used by HRSs. HRSs may optimize the message tailoring for each user by selecting the message contents as per the patient's need, sending them on a timely manner, and adapting the messages with changes in the patients' situation over time. Therefore, HRS may be a useful innovation over the current tailored systems as they may increase user engagement with the intervention and reduce costs.

Considering the immense potential in applying RS to health promotion interventions, it is necessary to present a multidisciplinary overview of the results of using HRSs. To map the existing research literature pertaining to the use of HRS for patients, we conducted a comprehensive scoping exercise by exploring five different databases from different fields (technical, medical, and psychological). A preliminary search for previous scoping reviews that adopt a multidisciplinary approach to the topic of HRS for patients was also conducted in a variety of databases of different fields, but we did not find any relevant occurrences.

The primary objective of this scoping review was to create a body of knowledge about the current state of HRS for patients in the last 10 years, in an attempt to answer the following research questions: What are the actual experiences with HRS for patients? What aspects have been studied? What are the existing research gaps that still need to be covered? These questions will be comprehensively addressed by following a multi-disciplinary approach adopted previously by some authors [213]. We analyzed four aspects—their domain, methodology and procedures, the usefulness of health promotion theoretical factors and behavior change theories, and technical details—in performing an in-depth analysis from all angles, which is required to ensure the success of a tailored, computer-based health intervention. We proposed a scheme of classification for this analysis. It constitutes a new taxonomy which integrates both principles of traditional HRSs, and principles used in computer tailored eHealth approaches. The I-Change Model [78] was used to identify whether the HRSs also address these needed factors for behavior change. This taxonomy intends to facilitate the HRS classification, as there is no other taxonomy covering the those or similar aspects relevant for HRS to our knowledge. Therefore, both policy makers and researchers may easily identify knowledge gaps and common successful patterns in previous studies. For future studies such identification may contribute to increase the study fidelity by minimizing the possibilities of having undisclosed parts or overlooked aspects of the study that reduces their replicability. Future studies that complete the proposed taxonomy will be going through an exercise to include many of the needed requirements to meet SDG3, as it covers not only technical aspects, but also health communication aspects, and domain, and methodologies.

This paper aims to present a clearer picture about how the existing studies can help policymakers make better decisions in terms of public-health actions, including computer-based tailored health interventions, and to help researchers design future studies by building upon the existing knowledge.

Materials and methods

Design

We conducted a scoping review following the PRISMA framework [214] to identify studies relating to HRS in which the end users were patients who received recommendations that may influence their health.

Search approach

The main eligibility criteria were that the studies had to be articles published in journals over the last 10 years (from January 1, 2007, to October 18, 2016, when the search was performed), written in English, and dealing with RS that provided some sort of health recommendations to patients. The information sources selected were five databases, namely, PubMed, PsycInfo, Association for Computing Machinery (ACM), IEEExplore, and ScienceDirect. Electronic searches were conducted using the following keywords: ((“recommender systems”) OR (“recommender system”) OR (“recommendation systems”) OR (“recommendation system”)) AND (health OR patient OR patients). When offered the option, keywords were sought in the entire text (not only in titles, abstracts, and/or metadata). We did not systematically assess the methodological rigor of the articles included as reflected in the convention of scoping reviews [215]. An example of the search process can be found in Appendix A.

Study selection procedure

The study selection was divided into four phases, as described in the PRISMA framework. The first phase (identification) consisted of gathering all the articles retrieved from the database (904 results). This process was done by three researchers (SHF, ACB, FLP) who examined each article in parallel; an article was considered to have passed to the following phase if least one reviewer marked it down as relevant. After removing the duplicates (10 articles) and filtering some publications that were initially retrieved but not published in journals (3 proceedings and 1 book), the three researchers ended up with 890 results. They considered the results indicating the same content in different editions of the same paper to be duplicates. During the second phase (screening), the three researchers screened all the titles of the entries, after which they checked all the studies for eligibility (third phase) using the present inclusion and exclusion criteria. Studies were included if they dealt with HRS and if the end users of the system were patients, irrespective of the type of analysis performed. Studies that did not meet these criteria were excluded. In case of doubt, for example, if the titles were not descriptive enough, the researchers were asked to accept the paper since it could be excluded in the later phases. Accordingly, a result selected by any of the three researchers passed to the next phase, the inclusion phase (84). In this phase, the same three researchers read the abstract of the papers and followed the same acceptance criterion.

Full paper review

The selected publications (42 articles) were fully read to assess their eligibility for the quantitative analysis. Only those publications that all three researchers agreed to pass to the quantitative analysis phase did (19), as shown in Figure 1: Methodology flow diagram.

Data extraction

Our proposed taxonomy intended to cover the relevant information to meet the requirements of SDG3 and was based on the intuitive approach described in the study of Nickerson et al. [216]. However, we followed a two-step approach to ensure that it had the five features that Nickerson et al. proposes for a useful taxonomy: namely, being concise, self-explanatory, robust, comprehensive, and extendible. The first step was to choose the aspects using expert opinion. One of the researchers (SHF) proposed the two first taxonomy aspects and their features, and these were discussed and completed by researchers ACB, ORR, and LFL. The second step was to complete the taxonomy using previous studies, deriving a third aspect from the MIRO study that used the I-Change Model [217], and a technical aspect from previously proposed classifications by Schafer et al. [218] and Montaner et al. [219]. As a result, our taxonomy has four aspects. The first one is the domain aspect, which helps us understand the general features of the study, such as what therapeutic area is being addressed, who the target population is, and what items are being recommended. The second one is the methodological and procedural aspect, which lets us identify the robustness of the study using features such as the number of test users, the system integration with an Electronic Health Record (EHR), and the study cost-effectiveness. The third aspect is the health promotion theoretical factors and behavior change theories, which assesses how much the intervention is grounded in health promotion and psychological techniques. The fourth and final one is the technical aspect, which determines the features of the HRS algorithm such as the used information filtering method, what the recommendation interface is, and what type of feedback users can provide to the HRS.

The details of the taxonomy for each the 19 studies were independently extracted by two researchers (SHF, ORR) in parallel. After their extraction, classification discrepancies were resolved by mutual agreement in a later phase. An "N/A" could also be entered against a given field if analyzing it did not make sense for a given study, as could "Unknown" if a study did not provide information about that field.

Figure 1. Methodology flow diagram.

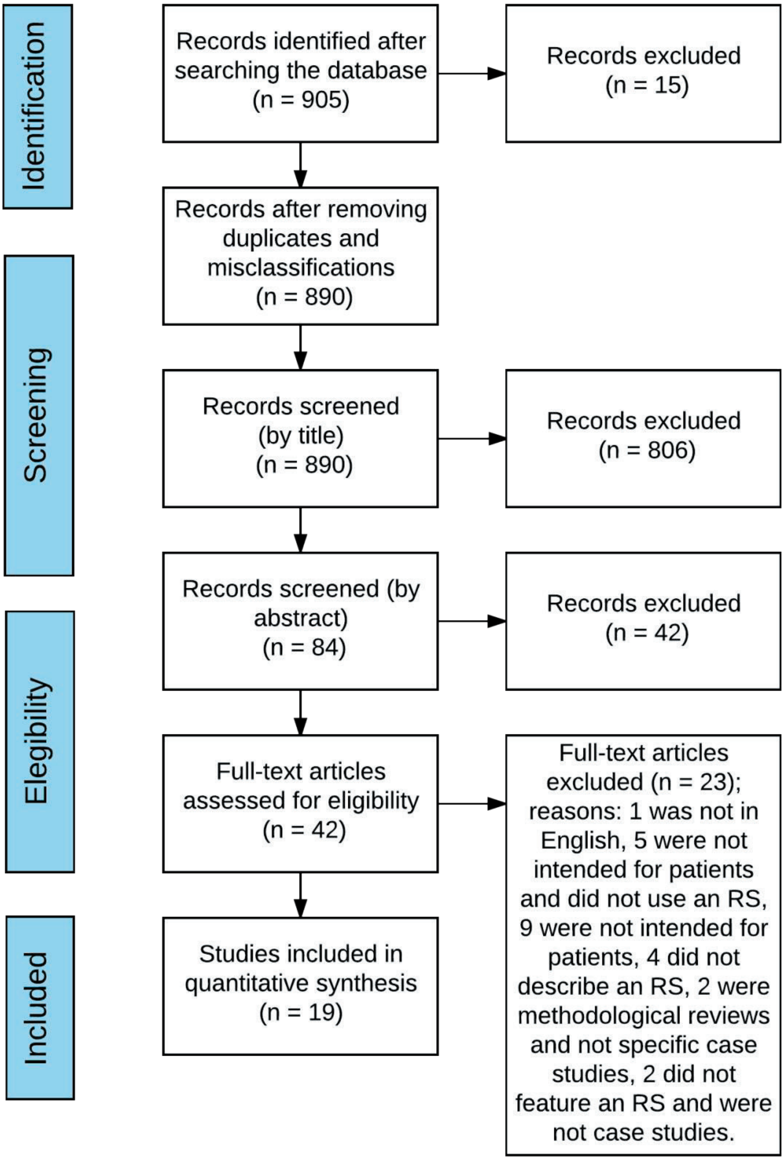


Table 1. Taxonomy of health interventions using HRS.

Domain	Therapeutic area	*The targeted disease or recommendation topic.*
	Target population	*Description of the users, and other exclusion and inclusion criteria*
	Type of recommendation (items)	*Messages, people, hospitals, paths,…*
	Device interface	*Mobile, web, mobile and web, other (i.e., smartwatch display)*
	Tailoring	*Yes/No*
	Country	*Country or region where the intervention was conducted*
Methodology and procedures	Used metrics to assess performance	*Metrics can be technical (F-score, precision, recall,…) or not (quit smoking,…)*
	Number of test users	*800, 45, 230,… (detail intervention and control groups, if applicable)*
	Effectiveness on patients	*Quantitative measure of the aim of the study (i.e., 30% more physical activity in the intervention than in the control group, average weight loss during the study for obese patients, …)*
	Success percentage	*% of patients that met the objectives of the study (i.e., quit smoking)*
	Duration of the total intervention	*Total length of the period that the users were exposed to the HRS*
	Number of sessions	*Average number of times the users interacted with the HRS during the intervention*
	Electronic Health Record connection	*Yes/No*
	Cost-effectiveness	*Yes/No (If yes, include the details)*
Health promotion theoretical factors and behavior change theories	Attitude	*Yes/No*
	Social influence	*Yes/No*
	Self-efficacy	*Yes/No*
	Action and Coping planning	*Yes/No*
	Supporting Identity change	*Yes/No*
	Rewarding abstinence	*Yes/No*
	Advising on changing routines	*Yes/No*
	Advising on coping	*Yes/No*
	Advising on medication use	*Yes/No*
Technical aspects	Recommendation interface	*
	Recommendation technology	*
	Finding recommendations	*
	Initial profile generation techniques	**
	Profile representation technique	**
	Profile learning technique	**
	Relevance feedback	**
	Profile adaptation technique	**
	Information filtering method	**
	User-profile item matching technique	**

* These technical aspects were retrieved directly from the proposed classification of Schafer et al.
** These technical aspects were retrieved directly from the proposed classification of Montaner et al.

Data analysis

A researcher (SHF) went through all the taxonomy tables created and analyzed the common patterns, contradictory results, and the gaps in all the studies. All the identified elements were presented and discussed with the four other researchers (ORR, LFL, FS, and HDV).

Results

We retrieved 905 initial results from the database search. These included 10 duplicate articles and 5 misclassified results that were actually books and proceedings. From the 890 remaining results, 84 met the inclusion criteria in the title review, 42 met the abstract review criteria, and 19 of them the full-text reading selection [143, 146, 159, 160, 165, 220-233]. We will highlight some of the most relevant findings in the paragraphs below.

The results obtained show that some studies have already used HRS to support patients for different purposes, with different approaches, and using different recommendation techniques. However, there are studies that appear to have misunderstood the concept of an RS. Of the 19 analyzed studies, 3 did not include systems that could be classified as an RS. Instead, they used other kinds of systems that computed recommendations and did not base their recommendations on the user or item feature similarity, or in previous knowledge incorporated by experts.

We present the results for each of the features of our taxonomy. Some features did not apply to certain studies. For example, if a study proposed a theoretical algorithm or conducted a review, we cannot consider whether it has been tested with patients. We highlight these non-applicable studies for each feature analyzed.

A complete description of all the extracted data using our proposed HRS taxonomy (Table 1) can be found in Appendix B.

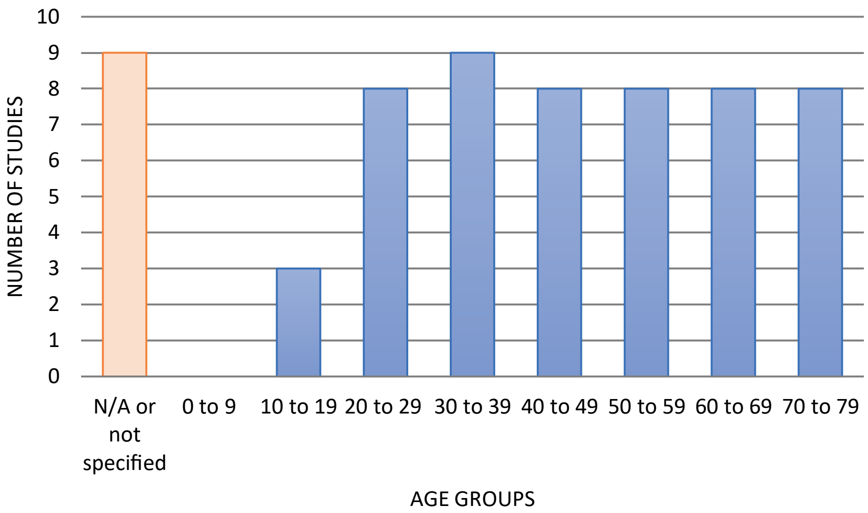
Studied aspects

Domain

Of all 19 eligible studies, 76.32% had the domain aspects we looked for. Of these, most of them focused on generic health promotion rather than recommendations relating to specific diseases (i.e., diabetes). The most frequently covered target population comprised adults (including young and healthy adults).

Derived from the target groups, the age ranges covered can be seen in Figure 2. Please, note that one study can cover several age groups. For 9 studies, either the age was not applicable, or the age range was not specified. At least 10 studies covered the young-adult age group. No study reported coverage for children specifically, and the three studies that covered a population under 19 years of age were designed for diabetics or the overweight population in general; therefore, we included these in the chart.

Figure 2. Number of studies for each age group.



More than half the studies used messages as their recommendation items. Other less frequently used recommendation items were people, and health resources. Similarly, more than half the studies reported at least a mobile-based interface through which the recommendations were delivered.

The studies were conducted in six countries, in the United States, a country in Asia, and four countries in Europe. Further, 60% of the studies in which tailoring was applicable stated that they implemented some type of tailoring technique.

Methodology and procedures

Upon analyzing the methodology and procedures, we found that 23.03% of the results we looked for were applicable and actually reported. The metrics used to assess the performance of the interventions were associated with the technical performance of the HRS (i.e., precision, recall, F-measure, and accuracy). In a lower percentage of the studies, the user perception (i.e., satisfaction, perceived usefulness, value, and trust) and health-related outcomes (i.e., weight loss) were also considered.

Seven studies included tests with users. Two studies measured effectiveness in terms of patient outcomes, one of them not reporting its effectiveness, and the other reporting a positive effect with the control group's average measure for weight loss doubling after the intervention; this outcome applied to 12% of the study population.

Regarding the length of the intervention and the session frequency, the studies reported interventions lasting from 14 days to 4 months long, involving one session where the patient interacted with the HRS and received recommendations. Thirteen studies could have benefited from being connected to an EHR, and two of them reported having a connection with an EHR. No study reported the cost-effectiveness of the intervention.

Health promotion theoretical factors and behavior change theories

In the studies we analyzed, 100% of the results did not find evidence of features of this aspect.

Technical aspects

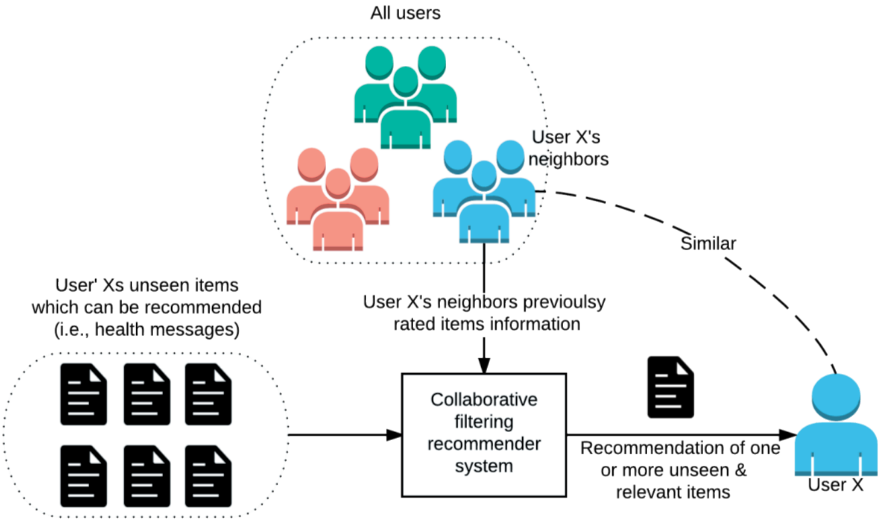
When we analyzed the technical aspects, we found that 45.27% of the studies contained the information we sought. The results of these studies showed that the Top-N interface (a list of the N most probably relevant recommendations) was used the most for the recommendations.

The most frequently used recommendation technology features were people-to-people correlation and user inputs, either standalone or in combination with other recommendation technologies. The "request recommendation list" technique was the most used for finding recommendations.

In 70% percent of the studies, the user profiles were manually generated. The techniques to represent the user-profile analysis were applicable to 12 studies. The most commonly repeated profile representation technique was the vector space model, followed by the history-based model and user-item ratings. In almost 77% of the cases, no profile learning technique was needed because they already had a database with a user profile or had implemented collaborative filtering algorithms. In addition, among all the studies, four reported a profile adaptation technique.

Half the studies analyzed did not included any feedback system, 40% included an explicit feedback system, and 10% implemented an implicit feedback system. The most common method of filtering information was pure collaborative filtering (Figure 3), followed by hybrid methods, content-based filtering, and knowledge-based techniques. Five studies reported their user-profile-item matching technique, and 80% of them had implemented the nearest neighbor approach. This approach recommends new items to a given user among the items other similar users– who are called 'neighbors'. The neighbor similarity can be computed in different ways such as using demographic data, or the users' item rating history.

Figure 3. Collaborative filtering recommender system concept diagram, the most used in the analyzed studies.



Discussion

Using our taxonomy to extract the features of the studies helped us to identify some relevant issues for discussion.

Domain

Although all the therapeutic areas extracted have a direct bearing on the achievement of SDG3, most of them focus on healthier nutrition and generic healthy-lifestyle promotion. More disease-specific HRSs are needed, which address non-generic topics and conditions. In particular, we believe that an excellent area would be substance abuse, one of the issues targeted by SDG3. As tailored messages have proved to be useful in reducing the intake of harmful substances [234], it would be feasible to design and implement an HRS that addresses this issue.

Most of the studies were concentrated in two countries, Spain and Taiwan, which together represented more than 30% of all the studies. In order to achieve a comprehensive vision of the impact of HRS, more countries, especially low- and middle-income ones, should conduct studies on HRS since culture and perceptions of digital elements entering the healthcare loop may affect their actual effectiveness.

Although the most recommended type of item in these studies were messages, none of the studies described anything about them. Consequently, it was not possible to assess them for inclusion of communication and behavior change features. This may have been due to the fact that they overlooked the importance of the message content or that they were not allowed to share the content.

The large proportion of the mobile interfaces used in the reported HRS point in the right direction, toward universal access to healthcare services and resources, especially in low- and middle-income countries.

Methodology and procedures

The effectiveness of HRS on patients was not described in 17 out of the 19 studies. This may be a consequence of the fact that several studies presented theoretical systems, reviews, or descriptions of systems whose results are yet to be achieved in the future. We also noticed that none of the studies reported on the cost-effectiveness of these systems, highlighting the need for further analysis on this feature in health interventions involving HRS. In addition, few studies in our sample used tests in order to assess user acceptance, and suitability of the system to meet its purpose with real users. Finally, sample sizes were low; only one study involved more than 90 users. We were therefore unable to determine the clinical or health outcomes since the statistical power of the samples are very low. Only one study reported health outcomes with a two-fold improvement when using the HRS [165]. However, this result is severely compromised since only three testers completed the study.

EHRs can be used to define the profile of each user such that the recommendations are based on their previous health records. However, only two studies used EHR. We acknowledge that privacy and legal barriers may be the reason for such a small figure. Integration with user data may require additional effort at both the management and technical levels. In addition, the EHR usage is a good way to reduce the manual data entry of user profiles in the HRS, and to increase the extent of information on user characteristics to yield more accurate recommendations. We should take into consideration that it is more common to use alternative platforms and not integrate the experimental system with EHRs until they are mature and at the final phases before being explored. This confirms that the use of HRS is in its infancy and that they are a potential tool to achieve SDGs that have not yet been met.

Health promotion theoretical factors and behavior change theories

We were unable to assess the extent of usage of health behavior theories and factors, because of the complete lack of information about how these messages were designed. Since a description of the messages and the length of intervention are the key elements in replicating the studies and building upon their experience, the utility of the existing evidence is limited. There is a need to develop and analyze additional studies with a more complete description of the intervention and how messages were designed.

Technical aspects

Although the HRS concept was not correctly applied in some situations, only 4 of the 15 studies that were not reviews or theoretical descriptions comprehensively described the technical specifications of the HRS in terms of the classifications and categories used [218, 219]. Consequently, there is little evidence of HRS characteristics that have been tested in the healthcare domain.

An important technical aspect concerns the limited description of applications of any profile adaptation technique. Only 3 out of 10 applicable studies implemented this technique. In order to provide more accurate recommendations over time, HRS need to evolve with the users. This means that these systems should ensure that user information is updated. Similarly, only 5 out of 14 studies implemented some kind of user feedback. Both the profile adaptation and user feedback are key factors for computer health education because the recommendations sent to the users need to be adapted to their current status and updated based on their answers. Otherwise, we will rely on the user's initial status, which will probably not yield accurate results in terms of behavior change interventions that need time to work (i.e., smoking cessation).

Table 2. Gaps in HRS in terms of meeting SDG3.

Gaps in HRS in terms of meeting SDG3			
Domain	**Methodology and procedures**	**Health promotion theoretical factors and behavior change theories**	**Technical**
• Research on sparse therapeutic areas • Lack of studies targeting teenagers and children • No experience in low- and medium-income countries	• Specific cohorts not usually addressed • Lack of reported results • Few patient experiences and limited number of participants • Few cases with EHR integration • Unreported cost-effectiveness	• Completely unreported	• Terminology misconception • Limited profile adaptation techniques implemented • Limited patient feedback systems included • Manual initial user-profile generation • Generic, superficial details used for RS classification

Conclusions

This paper presents a comprehensive scoping review of HRS to explore the current experiences of health interventions for patients using these systems. Due to the lack of a defined taxonomy for these purposes, we also propose a multidisciplinary taxonomy to classify these systems and determine the aspects analyzed and the gaps that should be addressed. We encourage future HRS studies to make sure they follow this taxonomy, assessing domain, methodology, health promotion strategies, and technical aspects. It has been useful to discover some unmet SDG3 needs when using HRS. We consider this taxonomy may be relevant for future use as reporting the domain aspects will contribute an easy context categorization. The methodology and procedures aspects will make easier to understand the robustness and fidelity of the study. Reporting the health promotion theoretical factors and behavior change theories will explain whether how the behavior change the HRS wants to provide is backed by actual theories. Finally, the technical aspects reporting will break down the necessary details to repeat and evolve successful studies. Future HRS studies should cover at least all aspects proposed in our taxonomy when disseminating their results. As a result, policy makers will be able understand their impact towards SDG3.

Although the studies analyzed present interesting approaches that could help meet SDG3, there remain several challenges. In terms of domain, we saw that most of the studies targeted the adult population, were oriented to generic health promotion and nutrition, and were conducted in a reduced number of countries. For the methodological and procedural aspects, we identified a lack of reported results and cost-effectiveness, few and limited patient-testing cases, and that not all studies made use of EHR data. In terms of the health promotion theoretical factors and behavior change theories aspects, we found a complete dearth of information. In terms of the technical aspects, we identified that the studies do not report complete information about the systems; that there are systems mislabeled as RS; and that most of the systems have limitations in terms of generating user profiles, adapting the profiles to changes in the user's circumstances, and collecting feedback from patients.

Consequently, many of the studies may still be considered black boxes whose details about how recommendations are generated are unknown. Although machine learning algorithms are difficult to interpret, and sometimes the dissemination is not aimed towards a full description of the systems, it is necessary to expose their details for both facilitating future research, and providing the information to make informed decisions at a policy maker level. Some institutions are introducing laws to remedy this lack of transparency. For example, the EU have approved the 'General Data Protection Regulation', which will come into force in 2018. It will ban systems generating decisions based solely on automated processing, which may clearly affect HRS that have not doctors in-the-loop [235-238]. That is why we recommend including health care professionals in the design phase of the HRS algorithm and the actual items that are going to be sent, as well as making them part of the intervention with the HRS as some studies are doing [239].

Due to the lack of reported key data in many of the studies of this review, we conclude that it is not possible to provide a guide of specific recommendations in the design of HRS to meet SDG3 yet. Future researchers should strive to innovate in terms of research areas and target groups. They should design HRS-based health promotion interventions by taking into consideration health promotion theoretical factors and behavior change theories, and specifying how the recommended items are made: their contents and wording, the frequency at which they are sent, and the exact tailoring techniques they use. Outlining these factors is also needed in order to be able to understand why certain interventions were or were not effective. In addition, the studies should describe their health-related metrics and test them with a sufficient number of users to achieve statistically significant results. Otherwise, technology-related metrics (i.e., F-score, precision, and accuracy) may prove inadequate to justify the cost and usage in a real-world setting. In this sense, it is necessary to continue reporting results on the evolution of HRS studies, since much existing evidence comes from descriptive theoretical studies or introductory studies. Paying more attention to the technical aspects, such as using correct terminology and comprehensively describing the systems, would benefit other researchers and policymakers willing to build on the previous successful experiences.

Policymakers should facilitate the secure usage of EHR that can feed into HRS and promote new studies that focus on analyzing the cost-effectiveness of these systems. As long as this type of analysis is not conducted, we encourage policymakers to propose and support studies pertaining to HRS in other therapeutic areas apart from nutrition and general well-being. A focus on relevant areas that can help meet SDG3, such as smoking cessation, oncology, mental health, and pregnancy and the early maternity stages, could help population risk prevention and enable users to manage symptoms, thereby having a global impact.

Implications and direct applications for researchers and policymakers: Below are some aspects to consider when applying HRS to computer-based tailored health interventions for public health promotion.

- Implication 1: Policymakers should promote the use of HRS to meet SDG3 because they can potentially act as a tool for scalable health promotion interventions, especially those that use mobile interfaces.
- Implication 2: Other therapeutic areas apart from the ones included in this study are can also benefit from HRS, such as mental health, substance abuse, chronic diseases management, or health education for maternal care and childcare.
- Implication 3. Policymakers should be aware that not all systems that claim to be an HRS are correctly defined. This may be misleading when assessing HRS-related results and making decisions about them. A deeper analysis to validate the correct classification by an IT expert is recommended.
- Implication 4: Wherever possible, policymakers should facilitate EHR integration with HRS for user-profile creation, which will help tailor the system's recommendations to the user's context. This can be done by, for instance, adopting secure computer communications protocols and providing a sample EHR for executing validation tests.
- Implication 5: When using a public computer to run tailored health promotion interventions through HRS, policymakers should ensure that the team leading the

intervention is a multidisciplinary one, including experts in behavior change, tailored health promotion, healthcare professionals, statisticians and technicians, who can collaboratively come up with a detailed design. In tailored interventions, special care should be taken to include a feature where the user profile is updated as the system adapts to the users' changing situation over time.
- Implication 6: Although there is immense potential in the use of HRS in health interventions, there is no information on the effectiveness nor cost-effectiveness thus far, indicating the need for further studies to address these aspects.

In sum, to better identify interventions in computer-based health promotion with HRS that covers all relevant aspects—the domain, methodological and procedural aspects, health promotion theoretical factors and behavior change theories, and technical aspects—policymakers can apply our taxonomy for each intervention.

Limitations

This scoping review analyzed journal articles from five databases, but additional results may be obtained by taking into consideration conference proceedings and grey literature and by using other databases. The methodological rigor of the articles included was not systematically assessed as per the convention of scoping reviews.
None identified.

Summary table

What was already known about the topic:
- HRS can be used to automatically tailor health information.
- There is a growing interest in the scientific community about the use of HRS, and some studies have already been conducted for health promotion.
- The application of tailoring and health communication theories are effective for behavior change.

What this study contributed to existing knowledge:
- HRS adoption to foster healthy lifestyles and promote well-being is currently lacking in terms of scientific evidence and only a few experiences that involve a sufficient number of users. This poses a challenge for policymakers and researchers to make decisions regarding the use of such systems. HRSs have been applied to very few areas that would meet the requirements of SDG, indicating that such systems need to be applied to new unexplored areas.
- Despite the apparent interest in tailoring messages, the data reported is insufficient to determine whether the messages are indeed tailored using health communication

theories. Besides, there is little information about the application of behavior change theories in HRS.
- In order to achieve effective behavior change or to maintain a healthy lifestyle, it is necessary to take into account the current status of the user and the subsequent evolution of their circumstances. The current HRS do not place much emphasis on receiving feedback and adapting according to the user's context.
- This paper has contributed a taxonomy for classifying HRS intended for patients, which can be used by researchers and policymakers in future studies to visualize and understand each HRSs approach.

2.12 Appendix A – Example of the search process

Researchers who wish to repeat the search in Science Direct, will have to click on "expert search" and then introduce the following text without the brackets: [("recommender systems," OR "recommender system," OR "recommendation systems," OR "recommendation system") AND (health OR patient OR patients)]. Next, they should select a year range between 2007 and 2016 and make sure that the checkboxes against journals and books are ticked. Some extra publications may be retrieved, since it is likely that some publications were released from October 14 to December 31. Similarly, the same query can be introduced in the PubMed search bar and filtered by publication date "January 1, 2007 and October 18, 2016." When using to the ACM digital library, this query was adapted to the database library as follows: +("recommendation system" "recommender systems" "recommender system" "recommendation systems") + (patients patient health); next, we filtered all results between 2007 and 2016. Similar searches were performed in the remaining databases.

2.13 Appendix B – Results table

Table 3. Studies domain analysis.

Title	Therapeutic area	Target population	Type of recommendations (items)	Interface	Implemented tailoring elements	Country of the study
A smart mirror to promote a healthy lifestyle	Cardio-metabolic risk	Healthy adults (25-60) & Non-pregnant or breastfeeding & not claustrophobia & no mental disabilities & no overt disease	Messages	Other (mirror)	Yes	Italy and France
Collective-intelligence recommender systems: Advancing computer tailoring for health behavior change into the 21st century.	Generic health promotion	N/A	Messages	N/A	N/A	N/A
Constructing recommendation systems for effective health messages using content, collaborative, and hybrid algorithms.	Generic health promotion	N/A	Messages	N/A	Yes	N/A
Consumers' intention to use health recommendation systems to receive personalized nutrition advice.	Nutrition	Not specified	Messages	Digital (E-mail) vs Fitness Clubs and Doctors	Yes	The Netherlands
Design and evaluation of a cloud-based Mobile Health Information Recommendation system on wireless sensor networks	Generic health promotion	Young adults	Messages	Mobile and Web	Yes	Taiwan
Design of a real-time and continua-based framework for care guideline recommendations.	General chronic patients preventive care	Caregivers of chronic patients	Messages	Mobile	N/A	Taiwan
glUCModel: a monitoring and modeling system for chronic diseases applied to diabetes.	Diabetes	Diabetics	Messages	Web	No	Spain
Health recommender systems: concepts, requirements, technical basics and challenges.	Generic health promotion	N/A	N/A	N/A	Yes	N/A
Mobile peer support in diabetes.	Diabetes	Diabetics	Messages & People and communities	Mobile	Yes	N/A
Multimodal hybrid reasoning methodology for personalized wellbeing services	Generic health promotion	Healthy adults & Non-pregnant & not disabilities & no medical complications	Messages	Mobile	Yes	Unknown
Nutrition for elder care: A nutritional semantic recommender system for the elderly	Nutrition	Elderly	Messages	Web	Yes	Spain
Personalized healthcare cloud services for disease risk assessment and wellness management using social media	None (technical-only)	Not specified	Doctors	Web	No	N/A

Table 3. Continued.

Title	Therapeutic area	Target population	Type of recommendations (items)	Interface	Implemented tailoring elements	Country of the study
Predicting potential side effects of drugs by recommender methods and ensemble learning	Drug side effects	N/A	Drug side effects	N/A	No	N/A
Rethinking Health: ICT-Enabled Services to Empower People to Manage Their Health	Generic health promotion	N/A	N/A	N/A	N/A	N/A
Social networks for improving healthy weight loss behaviors for overweight and obese adults: A randomized clinical trial of the social pounds off digitally (Social POD) mobile app	Weight loss	Overweight and obese adults with Android smartphones/tablets & not psychiatric illness & not receiving treatment for drug or alcohol dependency & not eating disorder & not pregnant & not breastfeeding & not heart condition & not chest pain & lose consciousness	People (Other users)	Mobile	No	USA
Supporting self-management of obesity using a novel game architecture.	Obesity	Overweight	Alternative strategies to coping with factors influencing obesity (i.e. stress)	N/A, but mobile is suggested	N/A	N/A
TPLUFIB-WEB: A fuzzy linguistic Web system to help in the treatment of low back pain problems	Low back pain	Adults	Messages	Web	Yes	Spain
Ubiquitous Multicriteria Clinic Recommendation System.	Generic health services	General public	Clinic and paths	Mobile	No	Taiwan
Which Doctor to Trust: A Recommender System for Identifying the Right Doctors.	Generic health services	General public	Doctor profiles	Mobile and Web (Web app)	No	USA

Table 4. Study methodology and intervention procedure analysis.

Title	Used metrics to assess performance	Tested with users	Effectivity on patients	Percentage of success	Duration of total intervention	Number of sessions	The HRS is connected with a EHR	Cost effectiveness
A smart mirror to promote a healthy lifestyle	N/A	89 in different phases (23 volunteers, 6 for reproducibility, and 60 clinical)	Unknown	Unknown	N/A	Unknown	N/A	N/A
Collective-intelligence recommender systems: Advancing computer tailoring for health behavior change into the 21st century.	N/A	N/A	N/A	N/A	N/A	N/A	No	N/A
Constructing recommendation systems for effective health messages using content, collaborative, and hybrid algorithms.	N/A	N/A	N/A	N/A	N/A	N/A	No	N/A
Consumers' intention to use health recommendation systems to receive personalized nutrition advice.	Effort, Privacy risk, perceived usefulness, perceived value, perceived trust	204 respondents interviews	N/A	N/A	N/A	N/A	No	N/A
Design and evaluation of a cloud-based Mobile Health Information Recommendation system on wireless sensor networks	User satisfaction, perceived usefulness, perceived value, perceived trust	202 participants in a single interviewed group (biased, all under 30)	N/A	N/A	N/A	N/A	Yes	Unknown
Design of a real-time and continua-based framework for care guideline recommendations.	Precision, recall, F-measure	3	N/A	N/A	Unknown	N/A	No	Unknown
glUCModel: a monitoring and modeling system for chronic diseases applied to diabetes.	N/A	No	N/A	N/A	N/A	N/A	No	Unknown
Health recommender systems: concepts, requirements, technical basics and challenges.	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Unknown
Mobile peer support in diabetes.	N/A	N/A	N/A	N/A	Unknown	N/A	No	Unknown
Multimodal hybrid reasoning methodology for personalized wellbeing services	Recall, precision, f-score, Type I and II errors	10	N/A	N/A	14 days	Unknown	No	Unknown
Nutrition for elder care: A nutritional semantic recommender system for the elderly	N/A	N/A	N/A	N/A	N/A	N/A	Yes	Unknown

Table 4. Continued.

Title	Used metrics to assess performance	Tested with users	Effectivity on patients	Percentage of success	Duration of total intervention	Number of sessions	The HRS is connected with a EHR	Cost effectiveness
Personalized healthcare cloud services for disease risk assessment and wellness management using social media	Precision, recall, f-measure, true positive, true negative, false negative, false positive	N/A	N/A	N/A	N/A	N/A	N/A	Unknown
Predicting potential side effects of drugs by recommender methods and ensemble learning	sensitivity (SN), specificity (SP), accuracy (ACC), precision, recall, F-measure (F), area under ROC curve (AUC) and the area under the precision–recall curve (AUPR)	N/A	N/A	N/A	N/A	N/A	No	Unknown
Rethinking Health: ICT-Enabled Services to Empower People to Manage Their Health	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Social networks for improving healthy weight loss behaviors for overweight and obese adults: A randomized clinical trial of the social pounds off digitally (Social POD) mobile app	Weight loss	25	Double the control group average weight loss results.	12%	4 months	N/A	No	Unknown
Supporting self-management of obesity using a novel game architecture.	N/A	N/A	N/A	N/A	N/A	N/A	No	N/A
TPLUFIB-WEB: A fuzzy linguistic Web system to help in the treatment of low back pain problems	N/A	N/A	N/A	N/A	N/A	N/A	No	Unknown
Ubiquitous Multicriteria Clinic Recommendation System.	Utility	10	N/A	N/A	N/A	1	N/A	Unknown
Which Doctor to Trust: A Recommender System for Identifying the Right Doctors.	Precision@10, R-precision, mean average precision	N/A	N/A	N/A	N/A	N/A	N/A	Unknown

Table 5. Technical aspects, part I.

Title	Recommendation interface	Recommendation technology	Finding Recommendations
A smart mirror to promote a healthy lifestyle	Unknown	Unknown	Organic navigation
Collective-intelligence recommender systems: Advancing computer tailoring for health behavior change into the 21st century.	N/A	N/A	N/A
Constructing recommendation systems for effective health messages using content, collaborative, and hybrid algorithms.	N/A	N/A	N/A
Consumers' intention to use health recommendation systems to receive personalized nutrition advice.	Digital (E-Mail)	N/A	N/A
Design and evaluation of a cloud-based Mobile Health Information Recommendation system on wireless sensor networks	Browsing	People to people correlation, user inputs	Organic navigation
Design of a real-time and continua-based framework for care guideline recommendations.	Ordered search results	Unknown*	Unknown
glUCModel: a monitoring and modeling system for chronic diseases applied to diabetes.	Inbox mailing system	Case-based reasoning	Mailing inbox navigation
Health recommender systems: concepts, requirements, technical basics and challenges.	N/A	N/A	N/A
Mobile peer support in diabetes.	Top N	User Input	Organic navigation
Multimodal hybrid reasoning methodology for personalized wellbeing services	Top N	Multimodal Hybrid Reasoning*	Request recommendation list
Nutrition for elder care: A nutritional semantic recommender system for the elderly	Ordered search results	User input and item-to-item correlation	Request recommendation list
Personalized healthcare cloud services for disease risk assessment and wellness management using social media	Top N	People-to-people correlation	Request recommendation list
Predicting potential side effects of drugs by recommender methods and ensemble learning	Top N	Attribute-based recommendations	Request recommendation list
Rethinking Health: ICT-Enabled Services to Empower People to Manage Their Health	N/A	N/A	N/A
Social networks for improving healthy weight loss behaviors for overweight and obese adults: A randomized clinical trial of the social pounds off digitally (Social POD) mobile app	Unknown	Unknown	Unknown
Supporting self-management of obesity using a novel game architecture.	Top N	N/A	Request recommendation list
TPLUFIB-WEB: A fuzzy linguistic Web system to help in the treatment of low back pain problems	Top N	People-to-people correlation, User Inputs	Unknown
Ubiquitous Multicriteria Clinic Recommendation System.	Top N	FINLP-OWA *	N/A
Which Doctor to Trust: A Recommender System for Identifying the Right Doctors.	Top N	Attribute-based recommendations *	Request Recommendation List

Table 6. Technical aspects, part II.

Title	Initial profile generation techniques	Profile representation technique	Profile learning technique	Relevance feedback	Profile adaptation technique	Information filtering method	User profile-item matching technique
A smart mirror to promote a healthy lifestyle	Unknown	Unknown	Unknown	Unknown	Unknown	N/A	Unknown
Collective-intelligence recommender systems: Advancing computer tailoring for health behavior change into the 21st century.	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Constructing recommendation systems for effective health messages using content, collaborative, and hybrid algorithms.	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Consumers' intention to use health recommendation systems to receive personalized nutrition advice.	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Design and evaluation of a cloud-based Mobile Health Information Recommendation system on wireless sensor networks	Empty	History-based model & demographic features & user-item ratings matrix	Not necessary (Collaborative filtering)	Explicit feedback	Unknown	Collaborative filtering	Clustering
Design of a real-time and continua-based framework for care guideline recommendations.	Unknown	Unknown	Not necessary (Database)	Explicit feedback	N/A	N/A	N/A
glUCModel: a monitoring and modeling system for chronic diseases applied to diabetes.	Manual	History-based model	Not necessary (database)	No feedback	Add new information	N/A	N/A
Health recommender systems: concepts, requirements, technical basics and challenges.	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Mobile peer support in diabetes.	Unknown	History-based model & user-item ratings matrix	Not necessary (Database & collaborative filtering)	Explicit feedback	Unknown	Hybrid	Unknown
Multimodal hybrid reasoning methodology for personalized wellbeing services	Manual	Unknown	Not necessary (Database)	No feedback	Unknown	N/A	N/A
Nutrition for elder care: A nutritional semantic recommender system for the elderly	Manual	User-item ratings matrix, Vector space model, History-based model	Not necessary (Database)	Implicit feedback	Add new information	Content-based filtering (+ knowledge-based techniques)	Nearest neighbor
Personalized healthcare cloud services for disease risk assessment and wellness management using social media	Manual	Vector space model	Not necessary (Database)	No feedback	Unknown	Collaborative filtering	Nearest neighbor

Table 6. Continued.

Title	Initial profile generation techniques	Profile representation technique	Profile learning technique	Relevance feedback	Profile adaptation technique	Information filtering method	User profile-item matching technique
Predicting potential side effects of drugs by recommender methods and ensemble learning	Manual	Vector space model	Not necessary (Database)	No feedback	N/A	Collaborative filtering	Nearest neighbor
Rethinking Health: ICT-Enabled Services to Empower People to Manage Their Health	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Social networks for improving healthy weight loss behaviors for overweight and obese adults: A randomized clinical trial of the social pounds off digitally (Social POD) mobile app	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
Supporting self-management of obesity using a novel game architecture.	Manual	N/A	Not necessary (Database)	Unknown	N/A	N/A	N/A
TPLUFIB-WEB: A fuzzy linguistic Web system to help in the treatment of low back pain problems	Manual	User-item matrix, Vector space model	Not necessary (Database)	Explicit feedback	Unknown	Hybrid	Nearest neighbor
Ubiquitous Multicriteria Clinic Recommendation System.	Empty	N/A	N/A	No feedback	N/A	N/A	N/A
Which Doctor to Trust: A Recommender System for Identifying the Right Doctors.	User-item ratings matrix, Vector space model	None, not necessary	Structured information retrieval techniques	Unknown	Manual	N/A	N/A

2

CHAPTER 3

Opening the black box: Explaining the process of basing a health recommender system on the I-Change behavioral change model

This chapter has been published as:

Hors-Fraile, S., et al.,
Opening the black box: Explaining the process of basing a health recommender system on the I-Change behavioral change model. IEEE Access, 2019. 7: p. 176525-176540.

Abstract

Recommender systems are gaining traction in healthcare because they can tailor recommendations based on users' feedback concerning their appreciation of previous health-related messages. However, recommender systems are often not grounded in behavioral change theories, which may further increase the effectiveness of their recommendations. This paper's objective is to describe principles for designing and developing a health recommender system grounded in the I-Change behavioral change model that shall be implemented through a mobile app for a smoking cessation support clinical trial. We built upon an existing smoking cessation health recommender system that delivered motivational messages through a mobile app. A group of experts assessed how the system may be improved to address the behavioral change determinants of the I-Change behavioral change model. The resulting system features a hybrid recommender algorithm for computer tailoring smoking cessation messages. A total of 331 different motivational messages were designed using 10 health communication methods. The algorithm was designed to match 58 message characteristics to each user profile by following the principles of the I-Change model and maintaining the benefits of the recommender system algorithms. The mobile app resulted in a streamlined version that aimed to improve the user experience, and this system's design bridges the gap between health recommender systems and the use of behavioral change theories. This article presents a novel approach integrating recommender system technology, health behavior technology, and computer-tailored technology. Future researchers will be able to build upon the principles applied in this case study.

Introduction

Traditional computer-tailored interventions based on behavioral change models can yield highly personalized motivational messages to help an individual adopt and maintain healthy habits [165, 240-242]. However, these interventions typically provide feedback on 'static' scores for each individual's answers. In contrast, a health recommender system (HRS) can dynamically compute a list of recommended items for each user using artificial intelligence (AI). An HRS is a type of machine learning system that provides users with relevant items (i.e., messages) based on, for instance, their past behavior or similarities they share with other users. Combining HRSs with behavioral change models can yield an innovative approach for the further use and improvement of eHealth interventions [144]. A recent scoping review [243] demonstrates that very few studies that employ HRSs describe their domain, design methodology, procedures, theoretical health promotion factors, behavioral change theories, or technical details. This lack of transparency in the design of an HRS hampers the replication of successful experiments and the identification of factors that contribute to this success, thus limiting the systems' growth potential when applied in healthcare. Also, patients and healthcare professionals may not understand what features are being used to generate the predictions and recommendations computed by AI-based healthcare systems [244]. This is may be due to the difficulty for understanding the nature of the processes and logic followed by AI algorithms. Further, the process for tracing back the origin and reasons for a specific AI-based health recommendation is complex, and sometimes, even impossible. This black box perception may yield trust barriers for its adoption [245], and undesired ethical [246] and legal implications [247].

In order to open this black box regarding how HRSs are designed, developed, and used in combination with computer-tailored technology, we shall describe the creation process followed in one m-Health intervention employed by the European project titled 'SmokeFreeBrain' (www.smokefreebrain.eu) [248]. This intervention, in combination with the standard care provided at the smoking cessation units, aims to help patients stop smoking (i.e., pharmacological treatment or nicotine replacement therapy). This system, tested in the 3M4Chan [249] randomized, controlled trial in Taiwan—whose results will be published in the near future—features an algorithm based on the I-Change behavioral change model [78] and the user context that selects the most relevant motivational messages for each user. The I-Change behavioral change model was chosen among a wide variety of models because it has been extensively used to explain smoking cessation behaviors in previous studies [85, 105, 250, 251].

The primary goal of this paper is to open the black box of our HRS by thoroughly explaining the design process followed to combine it with the I-Change model - although the same reasoning and procedures may be applied for other behavioral models, as well. We present

3

the principles of the design process and its outcomes for the system's key components: motivational messages, a recommendation algorithm, and a mobile app. The transparent description of our HRS is a novel effort in the AI-based healthcare solution domain to increase trustworthiness, fairness, and interpretability on its recommendations.

Our secondary goal is to foster transparency concerning the development of our technology by facilitating their replication, re-use, adaptation, and evolvement for future research and implementation in both similar and different contexts.

Related works

Although recommender systems have been studied since the 1990s, recent and increasing interest has been expressed in applying them in the health domain, as indicated in the study of Schäfer et al. [161]. The first literature review related to this topic was conducted by Sezgin et al. in 2013 [209]. The authors presented the basic types of recommender systems and the challenges they faced as well as identified the first study to discuss HRS, which was published in 2007. Merely seven other studies were considered for inclusion in their review, which reflects the limited number of studies in this field. Sadasivam et al. [146] performed another literature review as part of their discussion of how computer tailoring may be advanced via recommender systems. In 2016, the authors presented the differences between the traditional rule-based tailoring systems and the new trend in data-driven HRSs, although earlier studies recommend food and physical activity, which may be considered characteristics of HRSs. Such systems date back to 2000 [252] and 2006 [253], as identified by Tran et al. in 2018 [254].

One of the first studies to propose the use of recommender systems in the health domain was conducted by Fernandez-Luque et al. in 2009 [144]. The authors suggest that personalized recommendations be generated with the feedback (ratings) users offered to social web content as well as the users' similarity parameters. Recently, Torkamaan et al. proposed a basic model for achieving user satisfaction with HRSs [255] wherein effectiveness, privacy, trust, and transparency are key determinants. The authors additionally analyzed the factors influencing users when rating recommendations generated by HRSs [256], which they determined to be effectiveness, emotional gain, enjoyment, liking, and interest. Tondello et al. suggest that such HRSs can be complemented through indirect methods, such as personalized gamification approaches that support behavior change and engagement [257].

HRSs have been applied to a wide variety of health conditions, such as diabetes, drug-related side effects, lower back pain, generic health promotion, and cardio-metabolic risk. A limited number of studies report effectiveness values following the application of an HRS. For instance, for weight loss, the study conducted by Hales et al. [165] concludes that, by using a mobile app based on an HRS for weight loss, the participants managed to double their weight loss and improve their BMI reduction more so than the participants in the control group who did not benefit from the HRS. Regarding smoking cessation, Ghorai et al. [258] published a study describing an HRS that sent motivational SMSs according to the patients' sex, ethnicity, and craving patterns, the last of which was reported by the users themselves via SMS. Three message intensity categories (normal, moderate, and high intensity) were applied, and the message content from previously successful smoking cessation programs was re-used. This study did not present the results for the HRS efficacy, and no other publication referencing this study has done so thus far. However, a relevant study titled 'PERSPeCT' [162] applied a recommender system to select and send motivational messages to influence users to quit smoking using multiple behavioral theories. This thirty-day study demonstrated that the HRS approach was more influential on users' smoking cessation behaviors than was a traditional rule-based computer-tailoring approach in daily mean ratings and self-reported intervention influence. The PERSPeCT HRS was also helpful in making smokers more ready to set a quit date or to quit altogether compared to the traditional tailoring system. Video messaging rather than text-based messaging was tested in another study [259] whose effects are however not statistically proven in this case. Despite the limited evidence, the results from Hales et al. [165] and Sadasivam et al. [162] demonstrate the added value of an HRS. The collective intelligence generated by the aggregated data in an HRS offers real-time adaptation to users' evolving needs and subsequent feedback. The long-term performance of recommender systems in temporally evolving networks has been investigated with industrial data sets that conclude that adopting heterogenous models is necessary for improving the user experience [260]. This result evidences the need for novel approaches that are able to adapt and evolve over time to match user preferences. Another recent study concludes that combining various models in hybrid recommender systems may improve computer tailoring in digital health interventions [166].

Other digital programs for smoking cessation do no employ HRSs, but rather apply static rules that tailor messages using behavioral change theories. For instance, the Text2Quit program [261] implemented social cognitive theory in a six-month study using tailored SMSs to support university students who wanted to quit smoking and found positive, statistically significant results that favored the participants' receipt of tailored messages in both biochemically confirmed abstinence (11.1% in the experimental group and 5% in the control group) and self-reported abstinence (19.9% in the experimental group and 10.0% in the control group). Xhale.dk [262] also applied social cognitive theory for a message-based smoking cessation program that resulted in no statistically significant differences in the thirty-day abstinence rate checked after twelve months of the delivery of the intervention, compared to the tailored and untailored text message groups. The cognitive-behavioral therapy was also employed by Strecher et al. [263] in an randomized controlled trial of 1,866 smokers, wherein the authors proved that high personalized and tailored messages contributed to an increase in the smokers' six-month abstinence rates. A more recent model titled "Health Action Process Approach" [264] was used in a digital smoking cessation study [265] and resulted in a lower cigarette consumption rate by participants of the experimental group but did not express significant effects for abstinence rates, number of smoking cessation attempts, or states of change. Another frequently used model for smoking cessation is the transtheoretical model for behavior change [266] as identified by Noar et al. [267], and one study to employ that model was conducted by Haugh et al. [102]. The authors aimed to assess a tailored SMS-based program that supported smoking cessation. Although they did prove its feasibility and acceptance, they did not manage to prove differences in significant effects between the three study groups (control, one weekly SMS message, or three weekly SMS messages) because this pilot study was equipped with underpowered data for purposes beyond validating the feasibility and acceptation of such system. Cheung et al. [217] analyzed the success factors of Dutch online smoking cessation interventions; among the six studies identified in the literature, five based their tailoring in socio-cognitive models (e.g., I-Change). By using such tailored digital interventions, smokers were between 1.15 and 2.84 times more likely to quit smoking than smokers in the control groups. Therefore, tailoring messages to motivational characteristics has been demonstrated to result in smokers' increased attention, increased information processing abilities, increased motivation to quit, and successful quitting that can be maintained after 24 months [115, 268].

In conclusion, the use of HRS technology can specifically add value to behavioral change interventions when combined with tailoring, although one pitfall of most HRS studies is that they lack transparency and thorough detail [243]. Consequently, a taxonomy was developed to assist the reporting and classification of HRSs [243]. Valdez et al. proposed a framework for developing an HRS in which three dimensions are considered and covered: domain, evaluation, and inception [213]. Nevertheless, no consensus yet exists on how to describe or design an HRS. HRS technology currently fails to incorporate insights of the successes achieved with behavior change theories for health behavior principles and principles of computer-tailored technology. Beyond traditional digital tailoring programs, grounding HRS in theoretical behavioral change models is not common, as Cheung et al. identified [166] where only 3 of the 19 analyzed articles mentioned the inclusion of a theoretical behavioral model.

The novelty of this paper is that it comprehensively describes how HRSs, health behavior principles, and computer-tailored technology are combined in a single digital health solution for smoking cessation, bridging the gap between AI-based collaborative intelligence for healthcare, human behavior, and personalization. We will describe: (1) the motivational messages' design, (2) the recommender system's design, (3) the mobile app's design, and (4) a clear differentiation between the aspects that must be carefully considered when replicating or evolving this study in future research in addition to the tools required to do so (i.e., tailoring recommendations from the World Health Organization, behavioral change models, data analytics services).

Methods

Setting

Our HRS comprehends firstly a server running an algorithm that selects and sends motivational messages to support patients and encourage them to stay smoke free and secondly a mobile app called Quit and Return (hereafter, 'QaR')—programmed in two versions for both Android and iOS devices—which receives messages from the server – and such messages can then be rated by users according to their perceived usefulness. The system was developed between January 2017 and August 2017.

Motivational messages design process

A group of researchers specializing in behavioral change theories for smoking cessation designed the motivational messages in English. We followed the tailoring recommendations made by the World Health Organization (WHO) [269] and the I-Change model to capture specific behavior change determinants, such as attitude (the perceived advantages and disadvantages of quitting), social support (the support to quit offered by others), skills (the actual capacity of the smoker to manage situations where they are tempted to smoke), self-efficacy (how the smoker perceives his/her ability to successfully quit), and action planning (the various actions that are needed to quit [e.g., mentioning one's desire to quit to others] and to successfully cope with the accompanying challenges), all of which have been identified as key factors for increasing awareness, raising motivation, and changing behaviors in previous studies [270-274].

We chose the WHO's 'Encouraging people to quit smoking' guidelines to complement the I-Change model introduced above because these guidelines constitute a well-known behavioral science publication endorsed by one of the most prestigious health entities in the world that includes contributions from experts both inside and outside the WHO. These guidelines were referenced to design the motivational messages because they provide case scenarios, examples of personalized reasons to stop smoking, common excuses for not quitting—to which we created counter-reactive sentences—and strategies and tips to effectively quit smoking.

Before writing the messages, we defined the message meta-features, which are the details or characteristics related to each user profile (i.e., demographic data, smoking habits, and I-Change-related factors, such as attitudes and self-efficacy for quitting). The team ensured that at least one message covered each meta-feature's value to ensure that participants were provided at least one message for each combination of meta-features, thus avoiding cases wherein the system does not recommend a message. In other words, the system should have a sufficient variety of messages that map all potential user types; for instance, the system should include personalized messages that cover both genders, different age ranges, users with different nicotine dependence levels, users who have the skills to quit smoking and those who do not, users who are supported by others to quit smoking, and those who are quitting without the support of their friends and family, among other groups (for an example of a personalized message, see Appendix B). The messages were first written in English and then translated into Mandarin Chinese, which were then validated by two Taiwanese doctors specializing in smoking cessation.

Recommender system design process

We analyzed the designs of previous behavioral change interventions [123, 275, 276] to become aware of how the authors designed their solutions and applied any conclusions or findings they may have reached. We also analyzed the HRS previously used in the Social Local and Mobile (SoLoMo) intervention [239], which included the users' recommendation ratings after six months. We did so to build upon an existing developed system that would speed up the development process—that is, it would only require adaptations and not a complete from-scratch development—and reduce potential flaws and design pitfalls.

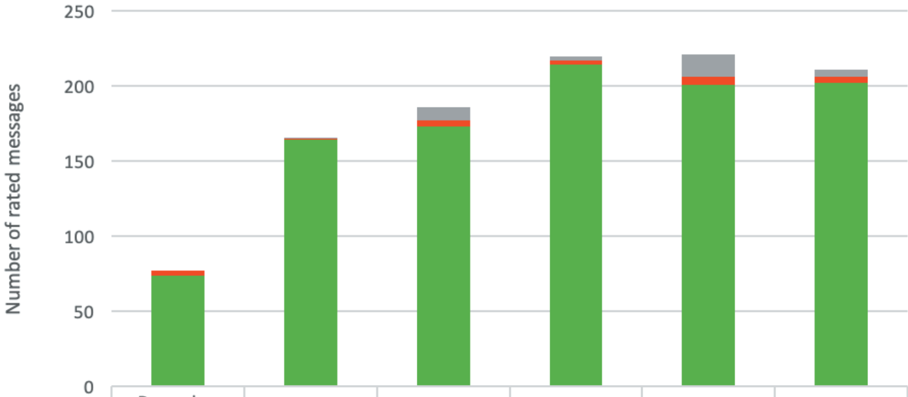
In addition, we assessed the need to increase granularity in feedback options for the received motivational messages, as users were offered merely three options for their responses (positive, negative, or neutral). Although this setup made the rating process simple to understand, most votes were concentrated around the positive option (see Fig. 1), which hindered the potential of the HRS because it did not have information about which messages were neither useful nor liked. Therefore, determining which messages were most useful for each user was difficult as they provided their ratings because most messages were rated as useful regardless of their content. We analyzed which alternatives were used in the existing popular platforms to increase granularity in the feedback options without rendering the rating process too complex for the users. To do so, we checked how the rating process was set up in popular services that use recommender systems, such as Amazon, eBay, and Netflix.

Mobile app design process

The QaR mobile app was evolved from the SmokeFree app used for the SoLoMo study [163, 164]. Our 3M4Chan study required that the QaR app be made available to all users in Taiwan, which resulted in the removal of user data that would be linked to a hospital's electronic health records (as was the case for the SoLoMo study).

To improve the user experience of the SmokeFree app, we analysed its six-month usage statistics using Flurry Analytics, the results of which reveal that users were interested in the messaging section (23.24% of use), the benefits/statistics section (15.84% of use), and their personal profiles (9.95% of use). These three sections resulted in a combined use of 49.03% versus 50.97% of the sixteen other features of the SmokeFree app. Therefore, we simplified and streamlined the SmokeFree app by removing any section other than the messaging, benefits/statistics, and personal profile sections in the QaR app. We did not remove other sections with low usage because they were necessary for making the app work properly and complying with legal restrictions (i.e., identifying the app's creators and giving credit to the funding sources); these sections were: configuration (allowing users to configure their language and notification preferences), tutorial (showing the app's basic functions), and 'about us' (describing the app's authorship).

Figure 1. Six-month HRS message ratings evolution in the SoLoMo study.



	December 2016	January 2017	February 2017	March 2017	April 2017	March 2017
Neutral	0	1	9	3	15	5
Negative	3	1	4	3	5	4
Positive	74	164	173	214	201	202

To make the QaR app available to anyone in Taiwan, we introduced initial questionnaires to complete each user's profile. To keep a low entry barrier, only three questions were required to make the app work, eighteen were required to start a quit attempt, and 52 others were voluntary in an attempt to improve the personalized recommendations.

To complement the SmokeFree app's usage analysis, we conducted a usability report following the MUSiC Performance Measurement method [277] with fifteen participants. From this analysis, we concluded that the tutorial shown when users first used the app could not sufficiently assist them in remembering the 'hamburger menu'—that is, the three horizontal lines in the top left of the blue menu bar that allowed a user to navigate across various sections of the app. Consequently, in the QaR app, we instead developed the main sections to be accessible from a single screen via tabs.

Results

Motivational messages results

Name personalization

Most messages were designed to include the receiver's name with some greeting words.

Meta-features adaptation

Resulting from the message design process and the HRS definition, we divided the meta-features into two groups: 7 basic meta-features, and 51 extended meta-features.

The basic meta-features included the most essential demographic information (gender, and age), and five other smoking-cessation indicators typically required in smoking cessation interventions to determine the patients' smoking habits (quitting date, cigarette consumption, weekly average expenditure in cigarettes, standardized nicotine dependence, and standardized motivation to quit). These 7 variables contain the minimum information required for them to assess a smoking cessation patient. Although these variables were previously used in the SoLoMo intervention, they were also validated by Taiwanese smoking cessation experts coming from Taipei Medical University Hospital, and Wellcome Clinic in Taipei. The 51 other extended meta-features corresponded to patients' comorbidities, living companions, physical activity frequency, and considerations regarding the I-Change model's key factors. These meta-feature topics were selected because the information they provide may be included and referred to in many smoking cessation motivational messages contents. Thus, the information introduced by meta-features allows the personalization and tailoring of these motivational messages."

The users provided such information through questionnaires in the app. The questionnaire related to the 7 basic meta-features was mandatory to start a quitting attempt, whilst the questionnaire for the 51 others was optional. Thus, that extended information allows the system further personalization and tailoring of the motivational messages, and a wider range of relevant topics for smoking cessation. Consequently, users providing a few meta-features only would receive less specific, and probably less relevant tailored motivational messages than users who completed all their extended meta-features.

The I-Change model determinants included attitudes (twelve meta-features), social support (six meta-features), self-efficacy (seventeen meta-features), action planning (eight meta-features), and skills (four meta-features). Section A of the Appendix offers an example message for each determinant.

We created a total of 131 motivational message categories, each of which deals with smoking cessation-specific aspects and is related to one or more user meta-features. As we wrote at least one message to cover all categories and meta-features, the 131 categories unfolded as 331 different messages.

All messages were enounced from a positive point of view and exemplified the benefits of quitting smoking. In section B of the Appendix, we include a sample case of how a category unfolds through various tailored messages. For this example, we took the category 'skin', which is one of the organs that is negatively affected by smoking; in this case, this category is exclusively associated with the 'user age' meta-feature. Based on a user's age, we can determine within which of the three meta-feature types he/she is categorized: < 30 years of age, 30–60 years of age, or 60 years of age or older. Depending on this categorization, the message stresses the importance of maintaining young and healthy skin (< 30), the importance of ceasing the ageing process now (30–60), or the importance of regaining some of the already lost appeal (60+).

Health communication methods

All messages were originally written in plain English (i.e., using active voice, including 'you' pronouns, keeping sentences short) such that educational level was not a factor in understanding them. When translating the messages into Chinese, we requested that the translator maintain that level of simplicity. For instance, we used active rather than passive voice and we did not include superfluous, irrelevant, or distracting information. We kept the length of each message short enough to read in less than one minute (a maximum of 200 words, with an average of 85.5 words per message). We also incorporated several behavior change techniques into the messages; specifically, we covered ten of the eleven groups proposed by Abraham et al [278]. We present an example for each group and underline where the techniques were reflected in section C of the Appendix.

Other techniques we employed include repeating an answer, creating empathy, adding new knowledge, and changing existing misconceptions, for which application examples are described in section D of the Appendix. For the first technique, we both included the user's answer in many messages (as seen in the examples) and allowed the system to repeat sending the same message up to three times, although a message was only repeated if no other user-compatible message existed with fewer repetitions. In this case, we prepended to each message the following piece of text: 'We know we have sent this message already, but we think it is important you remember it'.

We did not ask users to provide what they thought about some typical misconceptions about smoking, as doing so would have required that the initial questionnaire be even longer. Instead, we identified some common misconceptions and included the real facts in some messages, as may be perceived in the above example.

Recommender systems results

Taxonomy

The recommender system algorithm resulting from the design phase responds to the taxonomy proposed by Hors-Fraile et al. [243], as shown in section E of the Appendix.

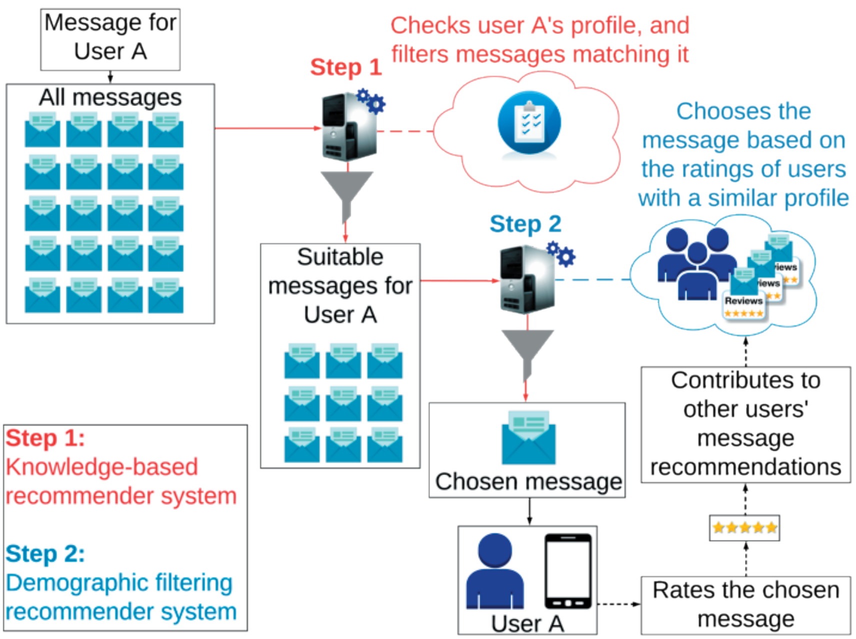
User context selection

The user context is a special type of message category because it is based not on any meta-feature, but rather on the day or moment the message shall be sent. We created thirteen messages based on various possible user contexts (i.e., different weekdays or moments during the day). These messages were only sent if the user was inside the associated context; for instance, a message designed for a Monday could only be sent on Mondays, and section F of the Appendix provides examples of these messages.

Message selection

Messages were selected by a hybrid HRS algorithm in cascade which works in two steps (see Fig. 2).

Figure 2. Implemented HRS diagram for the message selection process.

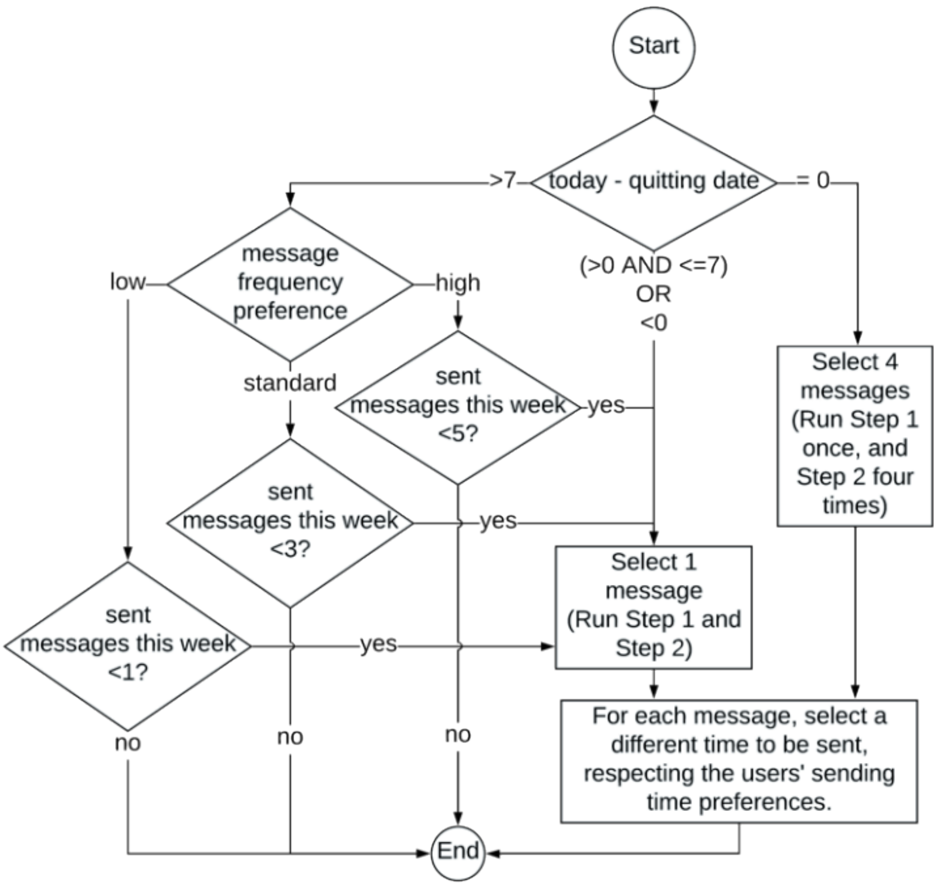


Step 1 (a knowledge-based algorithm) filtered down all the messages that were not compatible with the user's meta-features; that is, it reduced the category possibilities such that any remaining messages were suitable for the user. Step 2 (demographic filtering algorithm) involved selecting from the remaining messages that which the user had not yet received, that which had been received less frequently, and that which was rated more favorably by other users. The impact of the other users' opinions on the final selected message by the HRS to be sent was directly associated with their meta-features' similarity to those of the user to whom the message was sent.

At 00:02 am Taiwan time, the HRS computed the algorithm in the database and prepared the messages to be sent the following day for each user. This process consisted of selecting how many messages should be sent to each user (see Fig. 3), selecting the message according to Step 1, and passing the results to Step 2 (in cascade).

3

Figure 3. Selection of the number of messages for each user decision tree.



HRS Step 1

Step 1 involved checking which messages were potentially suitable for a user profile, which comprised the set of all meta-features associated with the user's questionnaire responses. The algorithm iterated the comprehensive list of messages and discarded those whose meta-features did not match the user profile; this action can be considered the application of a filter to reduce messages that were not applicable for the user. Context-dependent messages introduced in the previous section were treated as if they had a special meta-feature that was compared not against the user profile, but rather to the actual user context. Following the previous example, a message that should be sent in the morning was solely selected as a candidate for sending if its previously calculated time fell within the morning time frame.

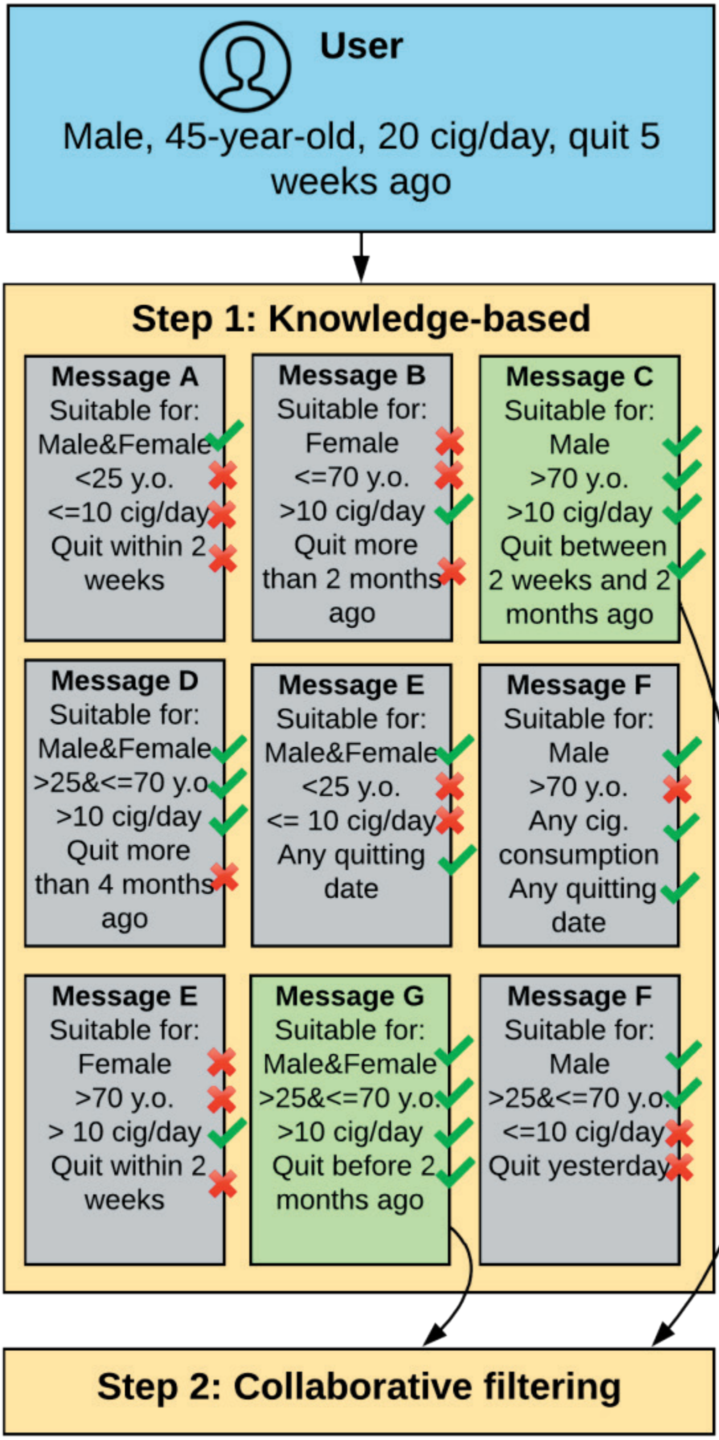
As an illustrative example, Fig. 4 presents a simplified version of a user with only four meta-features: gender, age, cigarette consumption, and amount of time since quitting. If we had a pool of eight messages, only those messages that possessed the four meta-features compatible with the user would be selected to progress to Step 2; in our example, exclusively messages C and G would be candidates for sending.

The HRS grouped the values of our meta-features to reduce the number of possible combinations as described in Table 1. The values for grouping were proposed, and as an initial approximation—to our knowledge—no previous attempts have been made to perform similar categorization.

Table 1. Meta-feature value grouping.

Meta-feature	Grouping	Values
Age	Young	Below 26
	Mature	Between 26 and 70
	Elder	Above 70
Quitting date	Recent	Less than 14 days
	Intermediate	Between 14 days and 59 days
	Late	More than 60 days
Tobacco consumption	Soft smoker	10 or fewer cigarettes a day
	Heavy smoker	11 or more cigarettes a day
Tobacco expenditure	Low	Less than $NT 700 per week
	Medium	Between $NT 700 and 2000 per week
	High	More than $NT 2000 per week

Figure 4. Simplified representation of Step 1 – Knowledge-based algorithm.



HRS Step 2

The messages that reached this step went through the collaborative filtering process, which relied on the hypothesis that one would like what other people with similar user profiles liked. Users who possessed similar profiles to the user under evaluation were called neighbors [279].

By the end of this step, each message was assigned a calculated score that determined its probable relevance to the user. A higher score indicated the given message was more relevant to the user's neighbors and more likely relevant to the user him/herself.

In this step, we aimed to determine the message with the highest score. In order to calculate this score, a matrix was built by the system with all users and their message ratings, excluding the user for whom we were selecting the message (see section G of the Appendix). The message score was the sum of all ratings provided by all users and was scaled by the neighbor similarity score. The next action involved normalizing the message score; if a message had never been rated by a user, its relevance rating score defaulted to the midpoint value of 0.5.

The neighbor similarity score was calculated using the equation illustrated below (1), which considers all meta-features defined by users in their profiles.

$$r_{A,B} = \frac{\sum_{i=1}^{n} \left(\delta_{F_A(i),F_B(i)}\right) + \sum_{j=1}^{m} \left(\frac{\delta_{F_A(j),F_B(j)}}{\max\left(|F_A(j)|, |F_B(j)|\right)}\right)}{\sum_{k=1}^{n+m} \left(\delta_{F_A(k),F_B(k)}\right)} \tag{1}$$

where:

A and *B* are two users;

F_U represents all meta-features completed by user 'u';

F_U (*x*) represents the value of meta-feature 'x' of user 'u';

$|F_U (x)|$ represents the number of values of meta-feature 'x' of user 'u';

n is the total number of single-value meta-features;

m is the total number of multiple-value meta-features; and

$\delta_{y,z}$ is a function that sums the number of matching meta-features between 'y' and 'z'.

Equation (1) returned a value between 0 and 1 because it provided the quotient between the number of matching meta-feature values between user A and B as well as the total number of meta-features they had in common. The numerator was split in two addends to consider the cases of single-value meta-features and multiple-value meta-features.

If a message was rated more than once by a user either because it was sent two or three times or because the user re-opened the message and rated it again, the new rating value overwrote the previous one. If the user had no neighbors, the system then faced the 'cold start' problem [280-282] and picked the first message stored in the list that met the requirements. If the user restarted his/her quitting attempt, the number of times each message had been sent was cleared.

Finally, the algorithm removed the messages sent three times and split the remaining list into three sub-lists depending on how many times messages had been sent to that user: zero, one, or two times, respectively. Then, starting with the sub-list of messages that had never been sent (zero times), the message with the highest score was selected. If this sub-list was empty, the same selection was orderly applied to the one- and two-times sub-lists until a non-empty sub-list was identified. If no message was found, the system did not send any message, as the user had received all relevant messages at least three times. If a message was found, then it became the candidate message for sending.

The maximum number of repetitions was set to three as it was a moderate number in line with the findings by Cacioppo and Petty [283] where they demonstrated that repeated persuasive messages allows greater realization of the meaning, interconnections, and implications of the message arguments. Nevertheless, it was decided that the QaR app would only send repeated messages to a user once no remaining relevant messages sent fewer times to the user were left. Any repeated message would be sent with a complementary text stating that the system knew the message was being repeated but that it intended to refresh users' memory due to its considerable relevance.

Message delivery frequency

The frequency with which the users received the messages was based on the frequency proposed by Abroms et al. [284]; Fig. 3 presents the decision tree used to calculate such frequency. The 3M4Chan intervention was planned to last for six months after a user's quitting day. Assuming users set their quitting day one week in advance, they would receive a maximum of 88 motivational messages during the 6 months at standard frequency, 157 at high frequency, and 42 at low frequency. Users were able to change the message frequency every two weeks; this option was prompted with a push notification and displayed with priority within the QaR app such that users had to respond to continue using the other sections. Message repetition was allowed up to three times based on previous studies of marketing and advertisement [285] as well as psychology [286].

The time frame for sending each message was selected at random within the allotted time range previously configured by each user. This setting eliminated the robotic feeling of messages being sent at the same time although did not bother the users because they limited the day hours during which a message might be sent.

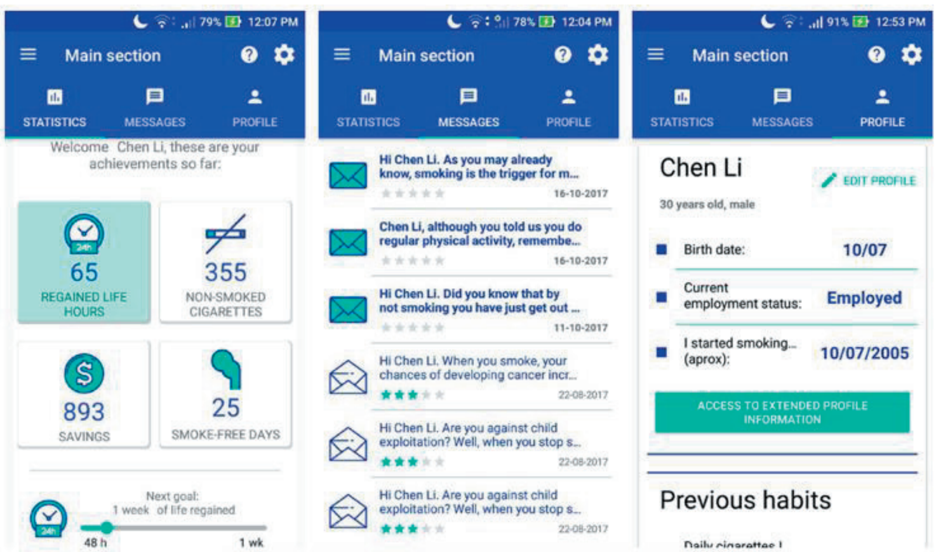
Feedback system

We defined the message rating options to be chosen between 1 and 5 stars, as opposed to the 3-option only message rating in the SoLoMo intervention HRS. The 5-star feedback system redesign aimed to reduce the risk of message ratings concentrating around one option, as can be seen in Fig 1., in order to provide the users with more options to better outline their preferences. This feedback system increased the message rating granularity by 66% compared to the one in the SoLoMo intervention without adding extra cognitive complexity, akin to the widespread application of the system in other fields such as evaluating the quality of hotels or movies."

Mobile app results

The resulting QaR app, designed specifically for the 3M4Chan intervention, is a native app available in two versions with identical functionalities—one for Android API v16 'Jelly Bean' or higher and one for iPhone SE or higher. The three main sections included in the QaR app are: (1) a messaging section with an inbox for the messages users received to support their smoking cessation. Users could rate the relevance of each motivational message and mark messages as 'unread'; (2) a benefits/statistics section containing information and goals about the user's number of regained life hours by not smoking, the number of cigarettes not smoked after quitting, the amount of money saved by not buying cigarettes, and the number of smoke-free days in a manner similar to the original SmokeFree app; and (3) a personal profile section containing all data related to the user's personal details and quitting attempt (see Fig. 5).

Figure 5. QaR app sections.



Discussion

This paper describes how a behavioral change model can be applied to the algorithm logic of HRSs, thus representing a further step towards maturing and consolidating the HRS field, which is considered to be in its infancy [209].

The design we propose covers a commonly missing aspect of HRSs: their grounding in a behavioral change model. In our approach, we embed the I-Change behavioral change model in the algorithm's core, which is however not the only possible solution for the algorithm; rather, other approaches may have been taken and been considered valid. To our knowledge, this study is the first in which a research group proposes and fully describes an HRS to the scientific community by detailing how it may become linked to a psychological model to induce changes in behavior.

We followed a comprehensive design methodology when creating our HRS, although we did not follow any guidelines whilst doing so. However, with the design steps we took, we were able to cover all the aspects proposed by Valdez et al. [213] in their HRS development framework, which was published almost at the same time our HRS was being designed. Covering all the aspects of their framework evidences that we applied valuable design principles in our methodology. To our knowledge, ours is the first HRS design study that aligns with all aspects of this interdisciplinary framework for developing HRSs.

Further, our HRS includes user context elements as did previous studies, such as time and day [287-289]. For instance, we can identify some similarities between the study presented by Lin et al. [288] and our system with regard to the message selection's conceptualization. The authors used constraint rules to ensure the messages were suitable for the user via his/her location, agenda, weather, profile, and time.

Although other studies may have dived more deeply into the complexity and setting of that user context application, none have formally combined these elements with behavioral change models. We might enhance the QaR app by including more context-aware conditions—such as those previously proposed—by defining new meta-features and designing new messages according to such new context-related meta-features. For instance, another option would involve assessing the impact of including meta-features related to the user's culture, religion, or country of origin. These user characteristics have been identified to affect how users may perceive the delivered health recommendations [290, 291]. Consequently, HRSs may be able to more effectively adjust to users by introducing these meta-feature in the similarity computation. In our HRS design, we did not incorporate them because we expected to recruit a homogenous cohort of patients from the Taipei Medical University and Welcome Clinic in Taipei, as described in the published protocol wherein this solution was to be used. Therefore, we anticipated that not including these types of meta-features would not result in a significant loss in the similarity accuracy and would reduce the entry barrier for patients who could already answer seventy other questions that we estimated would help the system differentiate their profiles more thoroughly. Yet, it is unclear what meta-features have higher impact on variables such as message appreciation, user engagement, and smoking cessation. More research is needed to analyze the users' feedback on these variables that would help optimize the system.

The present HRS may be generalized to other health behavioral change topics, such as the promotion of healthy eating or physical activity; however, meta-features, message frequency, and message content would need to be re-designed. We encourage that future researchers continue exploring this approach and tailoring it to their needs.

Nevertheless, our system presents limitations that should be considered when building upon our findings. The design team did not prioritize the algorithm's execution time, which may therefore be a constraint in large-scale environments. In our context, due to the expected number of users and the time frame we were allotted to compute the recommendations, this factor was not an issue. Additionally, this HRS has not yet been tested for metrics such as precision and recall. Finally, message recommendations are sent with push notifications, and each internal smartphone's firmware may handle these push notifications differently. Therefore, some phones may still display notifications on the user's phone, while others will keep them muted until the user opens the QaR app. This variation may affect how users interact with the QaR app and the behavioral impact it may subsequently pose to them.

Conclusions

We detailed the design and implementation process of developing an HRS using the I-Change behavioral change model to help people quit smoking. We reduced the gap between the information technology and psychological behavioral change domains as well as contributed to the research community by making this system's design and implementation principles transparent. Further, this comprehensive description aims to facilitate trust in our proposed solution, as compared to other black box digital health solutions where the artificial intelligence algorithms are totally unknown. We hope our work inspires and serves as a basis for future studies, as more research that combines HRSs and behavioral change models is needed to unveil the full potential of recommender systems in healthcare.

Appendix

A. Examples of messages for each type of included I-Change determinants.

Determinant	**Message description**
Social support	Hi <name>. You told us that all people in your area are smokers. Well, quitting smoking in a situation like that is a real challenge, but it can be done! We suggest you explain the list of reasons why you want to quit smoking to some of those people around you who are smokers and with whom you have a closer contact. You may kindly ask them not to smoke in front of you because you are trying to quit, and also not to offer you any cigarette. Many of them may be also willing to quit, and you can be the spark that fires their determination to start. Be strong, and don't give up!
Action planning	Hi <name>. You told us you don't have a plan to cope with stress. We would like you have one because when you are stressed, your brain is more prone to crave for a cigarette. You know that those cravings last some minutes only. If you get distracted doing some kind of relaxing activity, you will have more chances not to relapse. For instance, some people usually do the following breathing exercise: they take a deep breath, then hold it for 2 seconds, and release it slowly. This is repeated for a minute. Alternatively, other people prefer to drink water, or go for short walk. Any approach is OK as long as you have in mind what you should do in that situation to avoid smoking.
Skills	Hi <name>. You told us that you can relax yourself without cigarettes. That's good to know because some people cannot do it. If your planned strategy doesn't work, consider breathing deeply, holding your breath for 2 seconds, and releasing the air slowly. Repeat the breathing exercise for a minute and you will notice how you feel more relaxed. If possible, you could listen to relaxing music as well.
Attitudes	Hello <name>. I hope everything is going OK. You told us you didn't know if you could have health benefits if you quitted smoking. Actually there is a large list of benefits that you see when you quit smoking. Within 20 minutes of quitting: Your blood pressure and pulse rate drop to normal and the temperature of your hands and feet increases to normal. Within 8 hours of quitting: Your blood carbon monoxide levels drop and your blood oxygen levels increase to normal levels. Within 24 hours of quitting: Your risk of a sudden heart attack goes down. Within 48 hours of quitting: Your nerve endings begin to regrow. Your senses of smell and taste begin to return to normal.... And that's just the beginning!
Self-efficacy	Hi <name>. You told us that you couldn't refuse a cigarette when someone offers it to you. Well, we understand that you may struggle with it. However, you can practice how to kindly refuse a cigarette. A good reinforcing strategy is to add one of the reasons you have to quit smoking to the sentence. For instance: "No. I am quitting because I want to keep my teeth clean and white" Or "No. I am quitting to have a longer life and enjoy with my grandchildren". In this way, not only you say no, but also you remind yourself why you are doing it. If you prefer not saying the reason, you can just think about it in your mind after you say "No". In order to get a natural and almost immediate reaction, you can ask a someone you know to practice an exercise in which this person plays the role of the person who invites you to smoke, and you say no. Even if you think this is not worth doing it, you will feel that this exercise is worth it when you face it in a real situation.

B. Example of message category unfolding in several tailored messages upon meta-features.

Message category: Skin	
Relevant meta-features: User age range	Message description
<30	Hi <name>. You told us you don't have a plan to cope with stress. We would like you have one because when you are stressed, your brain is more prone to crave for a cigarette. You know that those cravings last some minutes only. If you get distracted doing some kind of relaxing activity, you will have more chances not to relapse. For instance, some people usually do the following breathing exercise: they take a deep breath, then hold it for 2 seconds, and release it slowly. This is repeated for a minute. Alternatively, other people prefer to drink water, or go for short walk. Any approach is OK as long as you have in mind what you should do in that situation to avoid smoking.
30-60	Hi <Name of the user>. As a reminder of the benefits of not smoking, since you have stopped smoking you have also stopped a faster wrinkle generation on your face. That's better than any face cream or lotion you can get! Kind regards.
60+	Hi <Name of the user>. It is possible that you have felt your skin better these days. When you stop smoking, it is like you have made your skin look younger, firmer, and with a better healthier color. If not, give it a bit more of time and you will see the results! Kind regards.

C. Examples of applied behavior change techniques.

Behavioral change technique	**Message**
To change beliefs about the benefits and costs of behaviors	Hi <name>. **Remember that by staying smoke free you are preventing seriously ugly permanent stains on your teeth, and also terrible-looking receding gums.** Keep you up smoke free if you want a white smile!
To change risk perception	Hi <name>. **Even for those like you who smoked a few cigarettes, smoking quietly produces negative effects on their body until the consequences are discovered and, it is usually too late.** By quitting smoking you have gained health and avoid the risks of serious illness! Great for you!
To change feelings (or affective attitudes) associated with adopting or ceasing behaviors	Hi <name>. How do you feel today? You told us that you were sure you would be able to resist smoking if you feel sad. We are very happy because many people try to fill that gap when they feel sad by smoking. However, you know that that's not a solution, especially after you have decided quitting smoking. Sometimes life is a bit tougher on us than we would like. Even if you are sure you can do it, a common technique to face sad situations is to remember happy moments to balance out the negative emotions. In addition, **you can imagine yourself within 1 year, as a proud non-smoker who got away the sad moments without smoking. That simple vision of you would take you step closer towards your goal!**
To change (normative) beliefs about other people's behavior and approval of recipients' behavior	Hi <name>. You told us that there are people close to you who are providing you some support. It is clear that **these people care about you.** Remember they are there for you, explain them why you value their support, and **ask them specific things about how they could help you a bit more.** For example, if they are your friends, they could agree on meeting in smoke-free places with you.
To foster a positive behavior-related identity	Hi <name>. You told us that the majority of people in your area are smokers. Well, quitting smoking in a situation like that can be challenging, but it can be done! We suggest you explain those who smoke the list of reasons why you want to quit smoking to some of those people around you who are smokers and with whom you have a closer contact. You may kindly ask them not to smoke in front of you because you are trying to quit, and also not to offer any cigarette. **Many of them may be also willing to quit, and you can be the spark that fires their determination to start.** Be strong, and don't give up!
To enhance self-efficacy	Hi <name>. For many people, driving and smoking goes hand by hand. Although new regulations have come into force to ban smoking while driving under certain circumstances, you may still feel like doing it. You told us this is your case. **The best way to avoid doing it is not carrying tobacco with you, and cleaning your vehicle from any sign of tobacco** - either a car or a motorbike. Especially in cars, where the law can be more flexible if you drive with your windows closed, we encourage you to **remove the lighter and ash-tray, and thoroughly clean the chairs and other textile elements so that they don't smell to tobacco and triggers your desire to smoke.** In addition, think of all the traffic accident risk reduction you will have and possible fines you will avoid by not being distracted by the cigarette while you drive! There are only advantages and **you can get them with these simple steps.**
To change emotional states in readiness for action and during enactment	Hi <name>. You told us you don't have a **plan to cope with stress**. We would like you have one because when you are stressed, your brain is more prone to crave for a cigarette. You know that those cravings last some minutes only. If you get distracted doing some kind of relaxing activity, you will have more chances not to relapse. For instance, **some people usually do the following breathing exercise:** they take a deep breath, then hold it for 2 seconds, and release it slowly. This is repeated for a minute. Alternatively, other people prefer to **drink water, or go for short walk.** Any approach is OK as long as you have in mind what you should do in that situation to avoid smoking.
To enhance social skills	Hi <name>. You told us that you couldn't refuse a cigarette when someone offers it to you. Well, we understand that you may struggle with it. However, you can practice how to kindly refuse a cigarette. **A good reinforcing strategy is to add one of the reasons you have to quit smoking to the sentence. For instance: "No. I am quitting because I want to keep my teeth clean and white" Or "No. I am quitting to have a longer life and enjoy with my grandchildren".** In this way, not only you say no, but also you remind yourself why you are doing it. If you prefer not to tell the reason, you can just think about it in your mind after you say "No". In order to get a natural and almost immediate reaction, you can ask a someone you know to practice an exercise in which this person plays the role of the person who invites you to smoke, and you say no. Even if you think this is not worth doing it, you will feel that this exercise is worth it when you face it in a real situation.

C. Examples of applied behavior change techniques. (Continued)

Behavioral change technique	Message
To facilitate behavior change by prompting environmental change	How are you <name>? You told us that you are not going to remove all your smoking-related stuff yet. Perhaps you've done it by the time this message reaches you. If you haven't, please consider doing it. It may seem obvious, but **some people keep ashtrays, lighters, and cigarettes in their houses and cars. These people are more tempted and will have more probability to relapse. We hope you do it soon if you haven't done it yet to minimize your possibilities of relapse.** We also encourage you to clean all your clothes, and any places where you used to smoke so that they don't smell.
To establish behaviors using rewards	Hi <name>. Although you didn't spend too much on tobacco, you are starting to save some money already. **Have you treated yourself yet? You should!** Even with little money you can provide some nice little price for you for your efforts in smoking cessation. Cheers!
The bold sections highlight where the specific technique is applied.	

D. Other applied behavioral change techniques.

Behavioral change technique	Message
Repeating the answer	"Hi <name>. **You told us that there are people close to you who are providing you significative support.** These people really love you. When you crave for a cigarette, when your determination is shaken to have a smoke, call them and tell them how you feel. They will be there to help you."
Creating empathy	Hi <name>. You told us you don't do physical activity. **There may be many reasons for that and we understand you may not find a good moment to do it** despite the benefits it can provide to your body and your smoking cessation process. However, we are sure you would like to do it if it were easier for you. Well, when you are offered a plan that does not involve physical activity (for example, going to the cinema to spend the evening), you can include minor changes that allow you to do a bit of physical activity while still enjoying the plan. For example, you could go to the cinema on foot, or go for walk to chat about the movie after dinner. Do you promise to introduce little changes like that in your life? Kind regards.
Adding new knowledge	Hello <name>. You told us that you **don't think that quitting smoking would contribute to stop child exploitation. However, many tobacco farms in countries like Pakistan, USA, and Indonesia child labor is used for tobacco farming!** These children and teenagers suffer long and tiring working days, exposed to toxic substances. Dario, 16, who worked in tobacco farms in Kentucky (USA) reported for Human Rights Watch interview reported that "The most difficult crop of all to work is tobacco. You get tired, it takes your energy, you get sick, but you have to go back to the tobacco the next day." Please, consider that smoking is not only bad for your health, but also for the lives of many young people who have to work in this industry
Changing existent misconceptions	Hi <name>. This information may be useful for you. **Did you know that second hand smoke contains up to three times more nicotine and tar, and about five times as much carbon monoxide than first-hand smoke?** Remember that this especially affects to your partner because it is person close to you. If you care about your partner, quitting smoking was a good decision! Kind regards
The bold sections highlight where the specific technique is applied.	

E. Health recommender system taxonomy classification.

Domain	Therapeutic area	Smoking cessation
	Target population	Current smokers willing to quit, speaking Chinese
	Type of recommendation (items)	Messages in text-only format
	Device interface	Android and iPhone mobile phones
	Tailoring	Yes
	Country	Taiwan
Methodology and procedures	Used metrics to assess performance	Smoking cessation rate, days before relapse, user engagement at an individual level, smoking abstinence, Quality Adjusted Life Years (financial aspects), precision of the recommender system, user engagement at an aggregated level, user reliability, user app behavior, user quit attempts, user satisfaction with messages, user mobile app usage, user message ratings.
	Number of tests users	1050 (estimated)
	Effectiveness on patients	Not yet available
	Success percentage	Not yet available
	Duration of total intervention	6 months
	Number of sessions	Minimum: 1 Maximum: 50 (estimated)
	Electronic Health Record connection	No
	Cost-effectiveness	Not yet available
Health promotion theoretical factors and behavior change theories	Attitude	Yes
	Social influence	Yes
	Self-efficacy	Yes
	Action and Coping planning	Yes
	Supporting identity change	Yes
	Rewarding	Yes
	Advising on changing routines	Yes
	Advising on coping	Yes
	Advising on medication use	No
Technical aspects	Recommendation interface	Top-N (N=1)
	Recommendation technology	Attribute based recommendations + People-to-People correlation (Likert)
	Finding recommendations	Selection options + Request recommendation list
	Initial profile generation techniques	Manual
	Profile representation technique	History-based model, user-item rating matrix, demographic features.
	Profile learning technique	Not necessary
	Relevance feedback	Explicit feedback
	Profile adaptation technique	Manual
	Information filtering method	Hybrid: knowledge based + demographic filtering
	User-profile item matching technique	Nearest neighbor (Pearson), Find similar users

F. Example of messages designed for being sent based on the user context (daytime).

Message Category: Special - Context "moment of the day"	
User meta-features: context	Message description
Morning	Good morning <name>! We hope you have a great day. Please let us remind you that the more time you are smoke free, the more vital you will feel. This is because your body will be able to rest better during nights. So, keep up the good work so far!
Afternoon	Good afternoon <name>! If you ever feel like a cigarette after lunch, remember that by smoking your gums receive less oxygen, decreasing the defense mechanisms against bacterial plaque. Now that you don't smoke you have healthier and more beautiful mouth. Coffee and alcoholic drinks may trigger your desire to smoke, so be aware of those triggers to avoid them! Kind regards

G. Demographic filtering process example.

User	**Similarity score with user i**	**Message 1 rating (stars)**	**Message 2 rating (stars)**	**Message 3 rating (stars)**	**...**	**Message M rating (stars)**
# 1	0.4	5	2	-	...	2
# 2	0.8	1	-	3	...	-
# 3	0.1	4	-	5	...	5
...	...	...	...	...	...	...
# N-1	0.9	2	5	-	...	4
Final relevance rating (non-normalized score)		= (0.4*5 + 0.8*1 + 0.1*4 + … + 0.9*2)/(N-1)	12975_Thesis compendium 12_11_2021V02	= (0.4*3 + 0.8*3 + 0.1*5 + … + 0.9*3)/(N-1)	...	= (0.4*2 + 0.8*3 + 0.1*5 + … + 0.9*4)/(N-1)

3

CHAPTER 4

A recommender system to quit smoking with mobile motivational messages: Study protocol for a randomized controlled trial

This chapter has been published as:

Hors-Fraile, S., et al.,
A recommender system to quit smoking with mobile motivational messages:
Study protocol for a randomized controlled trial . Trials, 2018. 19(1): p. 618

Abstract

Background: Smoking cessation is the most common preventative for an array of diseases, including lung cancer and chronic obstructive pulmonary disease. Although there are many efforts advocating for smoking cessation, smoking is still highly prevalent. For instance, in the United States of America in 2015, 50% of all smokers attempted to quit smoking, and only 5-7% of them succeeded - with slight deviation depending on external assistance. Previous studies show that computer-tailored messages which support smoking abstinence are effective. The combination of health recommender systems and behavioral change theories is becoming increasingly popular in computer-tailoring. The objective of this study is to evaluate patients' smoking cessation rates by means of two randomized controlled trials using computer-tailored motivational messages. A group of 100 patients will be recruited in medical centers in Taiwan (50 patients in the intervention group, and 50 patients in the control group), and a group of 1,000 patients will be recruited on-line (500 patients in the intervention group, and 500 patients in the control group). The collected data will be made available to the public in an open-source data portal.

Methods: Our study will gather data from two sources. The first source is a clinical pilot in which a group of patients from two Taiwanese medical centers will be randomly assigned to either an intervention or control group. The intervention group will be provided with a mobile app that sends motivational messages selected by a recommender system that takes the user profile (including gender, age, motivations, and social context) and similar users' opinions. For six months, the patients' smoking activity will be followed-up, and confirmed as "smoke-free" by using a test that measures expired carbon monoxide and urine cotinine levels. The second source will be a public pilot in which Internet users wanting to quit smoking will be able to download the same mobile app as used in the clinical pilot. They will be randomly assigned to a control group that receives basic motivational messages or to an intervention group, that receives personalized messages by the recommender system. For six months, patients in the public pilot will be assessed periodically with self-reported questionnaires.

Discussion: This study will be the first to use the I-Change behavioral change model in combination with a health recommender system and will therefore provide relevant insights into computer-tailoring for smoking cessation. If our hypothesis is validated, clinical practice for smoking cessation would benefit from the use of our mobile solution.

Trial registration: This study has been registered on ClinicalTrials.gov with identifier NCT03108651 on April 11, 2017 - https://clinicaltrials.gov/ct2/show/NCT03108651.

Background

Smoking is one of the leading causes of preventable deaths worldwide [8, 292]. Smoking tobacco is proven to have detrimental effects on different organs and is the root cause of several chronic diseases [293]. Among many conditions, smoking can cause cancer, heart disease, lung disease, chronic obstructive pulmonary disease, and it increases the risk of tuberculosis and other diseases related to the immune system [294, 295]. Additionally, secondhand smoking has also been linked to lung cancer, coronary heart disease, respiratory infections, middle ear disease, and sudden infant death syndrome [4].

Despite efforts to increase awareness about the harmful consequences of smoking, smoking cessation rates remain low [296-300]. Various methods have been used to promote smoking cessation in patients, including behavioral therapy, nicotine replacement therapy, and an array of medications. However, the highly addictive chemical nicotine found in tobacco usually causes intense withdrawal symptoms that make it difficult for smokers to quit [301, 302]. These symptoms include headaches, coughing, fatigue, anxiety, depression, and irritability [303]. Withdrawal symptoms act as triggers that increase the urge to smoke, which otherwise are used to relieve negative emotions such as stress [304]. Consequently, there is a high probability of relapse following smoking cessation [305, 306].

Past Studies

Among new smoking cessation methods, computer-tailored interventions have proven to be effective [85, 261, 307]. These interventions have been tested both in isolation [307-309] and in combination with existing methods [105, 111]. The relatively low cost and universal use of mobile phones make them useful to deliver these computer-tailored interventions to patients [310, 311].

With this information, the SmokeFreeBrain (SFB) project [312] assessed the success rate of different interventions for smoking cessation with respect to health economics. The SFB project studied the cost-effectiveness of these interventions and proposed a plan to implement them.

One part of the SFB project was to develop the Mobile Motivational Messages for Change intervention (from now on, 3M4Chan).

The 3M4Chan intervention measures the effectiveness of motivational messages delivered via smartphone to users who wanted to quit smoking. Utilizing the Social, Local, and Mobile (SoLoMo) study's health recommender system [239] as a basis, the 3M4Chan intervention uses a modified mobile application (app) to improve its usability, and an includes an inventory of messages based on the successful I-Change model of behavioral change [313-316] which better tailor the motivational messages to remain smoke free. The 3M4Chan intervention also differs from the SoLoMo study in the following aspects: duration of the intervention (6 months in 3M4Chan compared to the 12 months in SoLoMo), follow-up frequency, type of performed assessments and outcome variables, and inclusion and exclusion criteria.

Hypothesis

We hypothesize that the patients receiving the motivational messages selected by our health recommender system (and grounded in the I-Change model) will have better smoking cessation rates than those who do not receive any messages and also better rates than those who receive motivational messages that are not based on the I-Change model.

Need for a Trial

Although previous studies have used computer-tailoring [85, 261, 307], the I-Change model [105, 316], mobile phones [310, 311], and the health recommender systems to support smoking cessation [162, 317], this is the first study that will bring together all four features. Our study will use a specially designed app called Quit and Return (有戒有還 in Mandarin; hereafter referred to as “the app”) in two methodologically different pilots – a clinical pilot and a public pilot. They will be assessed differently although will share some common metrics.

If successful, this app could be integrated into treatments for smoking cessation, thereby promoting higher success rates in those attempting to quit. This would result in a better patient health, higher quality of life, lower incidence of smoking-related diseases, and reduced costs in the healthcare sytem.

Methods

Design and setting

The primary objective of this study is to assess whether tailored motivational messages supporting smoking cessation—as selected by a health recommender system and delivered by mobile phones to patients—contribute to higher smoking cessation rates.

Secondary objectives include: (a) comparing smoking cessation rates to other user metrics and technical aspects of the system(b) investigating the relationship between usage of the app used in the 3M4Chan intervention [318] and users’ opinions of messages; and (c) determining if a relation exists between the users’ physical activity levels and their mobile app usage, and their smoking cessation rates.

Both the primary objective and the secondary objectives are common for the clinical and public pilots, although their methodologies and assessments do not completely coincide.

Table 1. Description of the clinical and public pilot.

Name of the pilot	Number of patients	Groups	Treatment	Data to be assessed
Clinical pilot (RCT; 100 patients)	50	Control	Usual care (behavioral therapy and pharmacological treatment)	Smoking cessation rate (clinically validated)
	50	Intervention	Usual care + 3M4Chan app (advanced tailored messages)	Smoking cessation rate (clinically validated)
				Recommender system
				User engagement
				Physical activity
Public pilot (RCT; 1,000 patients)	500	Control	m-health (basic tailored messages)	Smoking cessation rate (self-reported)
				Recommender system
				User engagement
				Physical activity
	500	Intervention	m-health (advanced tailored messages)	Smoking cessation rate (self-reported)
				Recommender system
				User engagement
				Physical activity

The present document shows the protocol version 1.0.4, May 30th, 2018. Major protocol amendments, if any, will be handled by the Taipei Medical University informing any involved stakeholder, and updating the ClinicalTrials.gov registry diligently. Table 1 describes the clinical and public pilot. We followed the Standard Protocol Items: Recommendations for Intervention trials (SPIRIT) guidelines while writing this protocol manuscript (see Additional file 1).

Target population

The target population is any smoker in Taiwan who want to quit smoking.

Recruitment procedure

The participants in this study can be divided into two groups: those who participate in the clinical pilot and those who take part in the public pilot. Participants in the clinical pilot will be recruited from smoking cessation units at Taipei Wellcome Clinic and the Taipei Medical University Hospital in Taiwan between September 1st 2017 and July 31st 2018. They will be asked by a nurse or researcher, whether they want to be part of the study in the waiting room before their first visit to the smoking cessation unit. They will be handed written information about the reasoning of the study, and further, they have to sign an informed consent to participate. The public pilot will be open to any smoker who is willing to quit. In both pilots, each participant must own an Android mobile phone, and accept the terms and conditions of the services provided by the 3M4Chan app. The participants for the public pilot will be recruited online among all the Taiwanese population between September 1st, 2017 and July 31st, 2018. They will be invited to join the study by downloading the app. These invitations will be done with targeted ads on Facebook and Google, as well as with retargeting banners, which have previously been proven to be successful [319]. In addition, posters and informative leaflets will be distributed at the above-mentioned smoking cessation units and at Health Promotion Administration centers in Taiwan to attract candidates in non-digital environments. A total of four posters, two hundred leaflets, and four hundred business card-sized advertisements will be printed to disseminate information about the app.

Inclusion/exclusion criteria

In order to be eligible for the clinical pilot study, each patient has to meet e following inclusion criteria. Participants will be required to be a current smoker aged 20 years or older who owns an Android mobile phone and is able to read Mandarin. Additionally, only those who have smoked at least once per month for the past 2 years and are willing to share information from their medical Electronic Health Record (EHR) with project researchers will be considered. Furthermore, patients will be required to sign an informed consent and agree to be followed up with for 6 months.

For the public pilot study, all interested people who download the 3M4Chan app in their Android smartphone, who are 20-year-old or older, and who can read English or Mandarin will be able to join.

Planned interventions: Control and Intervention groups

In the clinical pilot, participants will be randomly allocated into two groups: Participants in the control group will receive the usual care provided at the medical center, while participants in the intervention group will be provided with a mobile app that allows them to receive tailored motivational messages in addition to the usual care. Similarly, participants in the public pilot (those who download the app but do not visit the medical center) will also be randomly allocated into two groups. Those in the control group will receive basic tailored messages without prioritization (the system sends a random message from a list of messages whose contents fit the user profile). Users in the intervention group will receive messages that are more specifically tailored to them (the message to be sent will be determined by its usefulness as reported by others with a similar user profile).

Data management and quality assurance

Participants' clinical data will be input by doctors into the Realsun, which is a Taiwanese EHR [320], as is routinely done in usual care. To ensure confidentiality, the data will only be accessible by the physicians and researchers involved this project, and the Taipei Medical University Hospital will classify the data as secrets based on legal regulations.

Participant data related to mobile phone activity (including profile details participants of the public pilot, and their registered physical activity time) will be automatically stored in a secure password-protected electronic database located at the ICHIT department of the Taipei Medical University that uses PostgreSQL and is accessible only to researchers involved in this study. Additionally, participants who drop out of the study will be registered along with their reasons for doing so.

The results of this study will be published in peer-reviewed scientific journals and presented at national and international conferences. Moreover, anonymized datasets from the study will be available via a data management system and thereby accessible to third-party researchers. The system will be based on CKAN [321], an open-source data portal platform. It will be possible to use, search, share, and publish these data.

We will assure the quality of the data entry process by checking use cases during the app's testing period: the data that is stored in the database will be evaluated to be the same that is expected values after the use case. If it were not the correct, the system would be debugged until it is fixed. Furthermore, random checks will be conducted throughout the duration of the study at a minimum frequency of once per month to ensure that the data are being stored correctly.

Only participants that meet the study's eligibility criteria, which include signing the informed consent after a full explanation of its consequences, will be enrolled in the clinical pilot. Only participants who accept the terms and services of the app will be enrolled in the public pilot. Data will be kept for 5 years after the research is completed after which all electronic and paper copies will be destroyed.

All consultations will take place at the smoking cessation units at the Wellcome Clinic and Taipei Medical University Hospital. The data analysis will be done at the Taipei Medical University (Taiwan), and the University of Seville (Spain). The data monitoring committee will be formed by researchers of these two institutions with total independence from the sponsors and with no competing interests.

No harm nor additional adverse events from regular smoking cessation care are expected as part of the clinical trials. No harm nor any adverse events are expected as part of the public trial.

The status of the trials is reported to the SmokeFreeBrain project consortium. The data monitoring committee members are part of this consortium. It is the consortium, led by its coordinator, who may decide to terminate the trials if the monthly interim analysis show deviations from the expected plan.

Procedure for generating texts

The motivational messages include a wide range of topics, such as physical activity recommendations, diet tips, reminders about the benefits of being a non-smoker, advice for coping with temptation, little known facts about smoking, suggestions to avoid relapses, how to develop skills required to quit effectively, smoking-related social influences, planning actions to change, quitting smoking self-efficacy expectations, and attitudes towards smoking cessation. The recommender system uses de Vries's I-Change (or Integrated) model [313-315] to tailor messages to users and guide them toward an effective behavioral change.

The messages are designed to cover the different behavioral change factors – determinants- that the I-Change model suggests we should address in order to foster a behavior change. This process consists of two steps. The first one is checking the users' profile, which was previously answered by the own user through a series of in-app questionnaires. These questions depict the I-Change users' profile, providing information about their behavioral determinants (attitudes, social support, skills, self-efficacy, intentions, action planning, etc.). After that information is known, the system iterates over a set of 228 different messages that can be sent to the user. Each message is categorized in one of the 131 to different profiles types created based on the I-Change model. For instance, a message could be targeted to young female users with low social support to quit smoking, whilst other could be targeted to elder male users which have low self-efficacy levels to quit. Then, the system discards those messages which are incompatible with the user we are going to send the message to - for instance, messages targeted to pregnant women if the user is a male -. Finally, the system selects the message which has the highest usefulness rating by other users. The usefulness rating is recorded every time a user receives a message in their phone with a 5-star scale.

Randomization

In the clinical pilot, randomization will be done with computer-generated random numbers in the website *www.randomizer.org*. In the public pilot, the randomization process is handled by the server using the JavaScript random method called Math. Random().

Methods to protect against other sources of bias

The clinical pilot component of this study is designed to be a single-blinded randomized control trial. Care providers at both medical centers will be instructed to treat patients from both groups equally and they will be responsible for patient allocation. Since they do not have any competing interest, we think it is unlikely they influence any group by the lack of blinding. Conversely, the public pilot will be double-blinded as there is no direct interaction with care providers (Table 1). In both cases, outcome assessors will be blinded although the different variables in the data sets will determine the groups.

To avoid bias in the clinical pilot, non-clinical researchers not involved in providing the intervention will be in charge of assessing the results. In the public pilot, although there are no people directly involved providing the intervention, and there is a low risk of bias in the results assessment, the analysis will be done by the same non-clinical researchers not involved in providing the clinical pilot intervention.

Duration of treatment period and follow-up

In the both the clinical and public pilots, patients will be recruited for a period of 7 months, and treated and followed up for 6 months

Measures: Baseline and Outcome and Process measures

The clinical pilot and public pilot are different interventions with shared metrics and also exclusive ones to each case. The primary outcome, the smoking cessation rates, will be measured at the 60, 120, and 180-day follow-up time points, using the urine cotinine and expired carbon monoxide levels in the clinical pilot case, and the self-reported questionnaires in the public pilot case. To assess the secondary outcomes, different metrics are required. Table 2 includes information about how they will be calculated, to which other metrics they will be compared, and to which pilot they are related.

Table 2. Description of metrics to assess the primary and secondary outcome.

Metric	Calculation	Comparisons	Pilot	Related secondary outcome
Primary outcome				
Smoking cessation rate.	Total number of people who relapsed / total number of people in the group at 60 days, 120 days, and 180 days of their quitting date.	• User engagement at an individual level. • User engagement at an aggregated level. • User mobile app usage. • User quitting attempts. • User lifestyle feedback. • Physical activity.	• Clinical • Public	System influence on smoking cessation & App influence on users physical activity levels.
Secondary outcomes				
User engagement at an individual level.	Messages read by the user / total number of messages sent to the user	• Smoking cessation rate.	• Public • Clinical	System influence on smoking cessation.
Engagement at an aggregated level.	Mobile application rolling retention, session length distribution, session frequency, sessions per user, return rate.	• Smoking cessation rate.	• Clinical • Public	System influence on. smoking cessation.
User quitting attempts.	Number and date of quitting attempts.	• Smoking cessation rate.	• Clinical • Public	System influence on. smoking cessation.
User app behavior.	Time spent per app section.	• User message ratings. • User satisfaction with messages.	• Public • Clinical	App usage. and opinions on messages the users received.

Table 2. Continued.

Metric	Calculation	Comparisons	Pilot	Related secondary outcome
User satisfaction with messages.	Satisfaction questionnaire.	• User mobile app usage. • Mobile app behavior. • User message ratings.	• Clinical	App usage and opinions on messages the users received.
User message ratings.	Users' votes for each message in a 5-star scale.	• User app behavior.	• Public • Clinical	App usage and opinions on messages the users received.
User lifestyle feedback.	Comparison of changes in user lifestyle (at baseline and after 6 months) through the questionnaires: EQ-5D-5L, IPAQ for physical activity, and SF-36.	• Smoking cessation rate. • Mobile app usage.	• Clinical	App influence on users physical activity levels.
Physical activity.	Total time (minutes) of activity per user, retrieved by GoogleFit.	• Smoking cessation rate. • Mobile app usage.	• Public • Clinical	App influence on users physical activity levels.

Although they are not necessary to the assessment of the primary and secondary outcomes, the metrics described in Table 3 will also be retrieved and analyzed because they are considered relevant and potentially useful.

Table 3. Description of additional metrics to be measured in the study.

Metric	Calculation	Pilot
User reliability	Comparison of the abstinence self-report at 2-week intervals with the measurements of the CO-oximeter	Clinical
QALY (financial aspects)	Healthcare resources utilization and cost analysis (cost of devices used, pharmacological treatment and time spent for various purposes)	Clinical
Precision of the re-commender system	Messages sent and rated more than four stars / total number of rated messages	Public Clinical
Smoke-free period	Time range between quitting date and the last smoke-free report	Clinical

Statistical analysis

In the clinical pilot, patients will be classified as either smokers or non-smokers according to their positive or negative values in the urine test and the expired CO2 tests (PiCO values between 0-6 will be considered non-smokers) at each follow-up visit. In case of discrepancy between tests, we will consider the patient as a smoker. We will perform repeated analysis of variance ANOVA, and provide p values, to compare the number of non-smokers in the intervention versus control group at 2 months, 4 months and 6 months of their quitting day.

In the public pilot, users will be considered to be non-smokers according to their self-reported answer to the question: "Are you still resisting the temptation or have you smoked? Please, be honest." We will perform repeated analysis of variance ANOVA, and provide p values, to compare the number of non-smokers in the intervention versus control group every 2 weeks.

For secondary outcomes, the numerical variables (User engagement at aggregated level, user mobile app usage, user quitting attempt, user app behavior, and user lifestyle feedback EQ-5D-5L and SF-36, and physical activity) will be assessed with T-tests and variance ANOVA tests. Categorical variables (user engagement at individual level, user satisfaction with messages, user message ratings, and lifestyle answers to the IPAQ questionnaire) will be assessed with Chi-Square tests. In addition, we intend to study the combination of all metrics with a linear regression analysis, and the Kaplan-Meier method to analyze the overall survival rate (non-smokers). All tests will be two-tailed and with a significance level set at 0.05. All analyses will be done according to the intention-to-treat principle, and using the SPSS software version 25. Missing data will be calculated with the last observation carried forward.

Sample size and power calculations

Calculations regarding the needed number of participants to measure the primary outcome – smoking cessation rates – have been done to ensure the statistical significance of the results.

For the clinical pilot, accepting an alpha risk of 0.05 and a beta risk of 0.2 in a two-sided test, 45 subjects will be necessary in both the first group and second groups in order to find a statistically significant proportion difference, which is expected to be 0.2 and 0.5 in the first and second groups, respectively.

For the public pilot, accepting an alpha risk of 0.05 and a beta risk of 0.2 in a two-sided test, 489 subjects are necessary in first group and 489 in the second to find as statistically significant a proportion difference, expected to be of 0.07 in group 1 and 0.14 in group 2.

A dropout rate of 15% and 20% is anticipated for the clinical and public pilot respectively. Although other studies using computer-tailored interventions have reported higher dropout rates, this study will consider participants who do not return for the consultation at month 6 to be failure cases. Consequently, the anticipated dropout rates include only those who withdraw for reasons other than relapse.

Planned recruitment rate

The clinical pilot recruitment rate is expected to be 17 per month in average, with the exception of February 2018 in which we do not expect to have recruit patients due the Chinese new year festivities. In the public pilot, we the planned recruitment rate is 700 users in the first 4 months, and 300 users in the 3 remaining months. This difference in the rate is due to the fact that we plan to conduct a digital marketing campaign to promote the app in the first 4 months, and that we expect a reduced number of new users in

February coinciding with the Chinese new year. Figure 1 summarizes the difference in the recruitment, execution, and assessment periods for both pilots, and Figure 2 details it.

Materials

The clinical pilot will make use of all the materials normally used to treat patients at the smoking cessation units at Taipei Welcome Clinic and Taipei Medical University Hospital.

Carbon monoxide levels will be tested using a PiCO Smokerlyzer CO-oximeter; a value range of 0–6 will be considered normal for a non-smoker. Similarly, a Safecare Biotech COT Rapid Test device, which contains line indicators for positive and negative test results, will be used to measure urine cotinine levels.

Several smoking cessation medications will be used during the trial: Nicotinell TTS 20 and TTS 30 patches, Nicotinell 2 mg chewing gum and 10 mg inhalers, Buporin 150 mg sustained-release tablets, and Champix 0.5 mg and 1 mg tablets. The dosage and duration of each medication will be prescribed by the physician, and the patients will be charged NT$200 per consultation.

In addition, our specially designed mobile app will be used to deliver motivational health messages. Figure 3 shows its interface in both English and Mandarin. The app has been programmed in Android's native language. It uses Mirth Connect as its communication channel [322], Firebase for platform notifications, and PostgreSQL as its database management system.

A nurse will provide patients with instructions for downloading and using the app. Questions related to the patient's basic demographic information and an extended profile (including social influences, plan of action, skills, attitudes, and self-efficacy with regard to smoking) will be assessed in the app.

Users' physical activity data will be collected using their mobile phones. The mobile app will use GoogleFit [323] to track daily physical activity. These data will be available to users through the GoogleFit app but not displayed within the 3M4Chan app, as it will be utilized for research purposes only.

Figure 1. Difference of enrollment, intervention, and assessments periods in the clinical and public pilots.

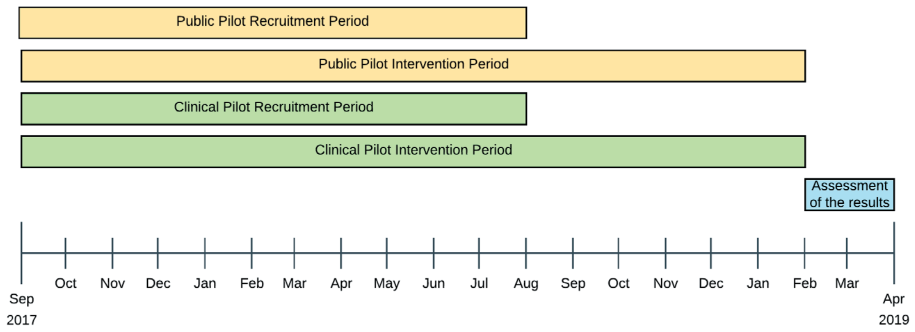


Figure 2. Standard Protocol Items: Recommendations For Intervention Trials (SPIRIT) Schedule of enrollment, interventions, and assessments.

			STUDY PERIOD												
	Enrolment	**Allocation**	**Post-allocation**												**Close-out**
TIMEPOINT**	$-t_1$	*0*	*2w*	*4w*	*6w*	*8w*	*10w*	*12w*	*14w*	*16w*	*18w*	*20w*	*22w*	*24w*	*6 months*
ENROLMENT:															
Eligibility screen	X														
Informed consent	X														
Allocation		X													
INTERVENTIONS:															
Usual care		←	—	—	—	—	—	—	—	—	—	—	—	→	
Usual care and prioritized tailored motivational messages		←	—	—	—	—	—	—	—	—	—	—	—	→	
Non-prioritized tailored motivational messages		←	—	—	—	—	—	—	—	—	—	—	—	→	
Prioritized tailored motivational messages		←	—	—	—	—	—	—	—	—	—	—	—	→	
ASSESSMENTS:															
Urine cotinine and expired carbon monoxide levels	X					X				X				X	
Self-reported smoking cessation questionnaires			X	X	X	X	X	X	X	X	X	X	X	X	
User in-app activity, message ratings, and Google Fit data		←	—	—	—	—	—	—	—	—	—	—	—	→	
EQ-5D-5L, IPAQ, and SF-36 questionnaires	X													X	
Motivational messages satisfaction questionnaire														X	

4

Figure 3. Screenshots of the app (English and Mandarin versions) showing different sections.

The data obtained from the database will be analyzed in a comma separated value (CSV) format using SPSS software.

Patients will be followed up after 60, 120, and 180 days of their initial consultation with a variance margin of 5 days. In the initial consultation, the patient will be assessed in the smoking cessation unit after he or she has been referred from either the pneumology unit or another clinical department. Smoking-related symptoms, treatment adherence, daily cigarette use, and possible adverse events will be discussed and exhaled carbon monoxide, urine cotinine, and other clinical information, such as weight and height, will be assessed. Quality of life (EQ-5D-5L, IPAQ, and SF-36) questionnaires will also be collected. The patient will be instructed on relaxation techniques and on tips to avoid smoking.

During the second consultation, doctors will gather patient information on smoking-related symptoms, treatment adherence, daily cigarette use, possible adverse events, exhaled carbon monoxide and urine cotinine levels, and other clinical information. Coaching for relapse prevention will be performed.

At the third consultation, information about smoking-related symptoms, treatment adherence, possible adverse events, exhaled carbon monoxide and urine cotinine levels, and other clinical information will again be collected. New relaxation techniques will be explained if the previously taught techniques have not been effective. Techniques for relapse prevention will be reviewed. It will also be collected the users' physical and psychological health state, their exhaled carbon monoxide levels, the results from a urine cotinine test, and will be asked clinical information such as weight and height.

At the fourth consultation, doctors will collect information about smoking-related symptoms, treatment adherence, and possible adverse events. Techniques for avoiding the risk of relapse are reinforced. Data about blood pressure, weight, height, and quality of life (using EQ-5D-5L, IPAQ, and SF-36 questionnaires as well as a questionnaire assessing the satisfaction about the motivational messages received through the 3M4Chan app) will also be collected. All input clinical data will be checked for correctness by two healthcare professionals to ensure quality.

The public pilot will use the same mobile app as the clinical pilot. No additional materials will be used.

Discussion

Expected Outputs for Validation of Our Hypothesis

Recently, several recommender systems have been implemented to support smoking cessation [162, 258, 324]. We expect that our recommender-system-aided motivational messages will help increase user engagement at both individual and aggregate levels and lead to prolonged periods of smoking abstinence in accordance with our hypothesis. This will be validated by the metrics measured over the course of our study, such as the health recommender system precision, smoking abstinence, app usage statistics, and user feedback.

Expected Outcomes

We hope to provide a new method to achieve smoking cessation that will be integrated into the usual care received by patients. This can either come from just using the mobile app in the public intervention, or by combining the mobile app with the pharmacological treatment in the clinical intervention setting. Despite being different approaches, achieving an increase to smoking cessation rates will help to reduce the incidence of diseases directly related to smoking as well as those arising from exposure to secondhand smoke. This will lead to increased life expectancy and improved life quality.

Dissemination of the Results

The results of this study will be published, and a subset of the retrieved data will be made public as part of the SmokeFreeBrain project's database. Public datasets will be fully anonymized and protected in accordance with directives 2002/58/EC [325] and 95/46/EC [326], which regulate the protection of privacy in electronic communications and the processing and free movement of personal data.

The data will be made available through an open-source data management portal based on CKAN. Third-party researchers and organizations will have access to both the raw data (where anonymization has made this possible) and to a rich set of metadata for each dataset. In this portal, users will be able to search up to field level different keywords and tags to browse related datasets, and see their format, availability, and licensing type. The published datasets will be a valuable resource for future studies, and we intend this portal become a base reference point for future related studies.

Problems Anticipated

The dropout rate among participants is one of the main challenges we anticipate for this study. Patients might dropout by ceasing to share EHR data, by terminating their use of the app, or by becoming unresponsive to follow-ups over the course of the trial.

Ethics

This study has been approved by the Ethical Committee of Taipei Medical University-Joint Institutional Review Board (TMU-JIRB) and has been registered under Clinical Trials ID: NCT03108651.

Informed Consent

Patients in the clinical pilot will sign an informed consent. All patients using the app, regardless of which pilot they are participating in, will be asked to read and accept the terms and conditions (http://sfb.phr.tmu.edu.tw/page/index.php/privacy-policy-en/), which explain the usage of their data and app analytics, the rights they have during the pilot, and the offered services by the pilot 3M4Chan app in detail.

CHAPTER 5

Applying collective intelligence in health recommender systems for smoking cessation: a comparison trial

This chapter is under review as:

Hors-Fraile, S., et al.,
Applying collective intelligence in health recommender systems for smoking cessation: a comparison trial.
(Submitted to UMUAU SI: Recommender Systems for Health and Wellbeing)

Abstract

Health recommender systems (HRSs) are intelligent systems which can be used to tailor digital health interventions. It is unclear how different types of HRSs may influence users' message appreciation, engagement, and health outcomes. Two HRSs to support smokers stay smoke-free selecting and sending motivational messages to their smartphones were tested in a 6-month trial. Smokers who voluntarily downloaded a mobile app were randomly assigned to two different interventions: one having a knowledge-based algorithm (n=181), and the other having a hybrid algorithm using knowledge-based and demographic filtering (n=190). We assessed participants' message ratings to measure message appreciation, the number of active days, message ratings, number of attempts to quit, and number of abstinence reports to measure engagement; and we used outcomes of those reports to measure smoking abstinence. We also performed subgroup analyses of gender, age, user profile completion, motivation to quit levels, and nicotine dependence levels. Both interventions had similar message appreciation, number of rated messages, and abstinence results when considering the average of the seven-day point prevalence reports across the study, and also considering seven day point prevalence in a pessimistic scenario where dropouts were considered smokers. However, the knowledge-based HRS achieved significantly better seven-day point prevalence abstinence results considering only the last abstinence report of each participant and led to a significantly higher number of active days and higher number of abstinence reports. In addition, participants who voluntarily fully completed their user profile and were sent messages by the hybrid algorithm made significantly more quit attempts.

Introduction

For decades, computers have been used to generate health recommendations [124, 327]. This usage of computers to adjust health materials to each person in order to make them relevant and credible for their situation, replicating what an actual human counselor would do is called computer tailoring [113, 114]. Computer tailoring involves the generation of participant-specific recommendations, typically in the form of messages, by computers, which is done after an assessment of each person to match their characteristics, needs, and interests [116, 328-330]. Traditionally, these systems were designed with if-then-else rules using users' input to some questionnaires to determine the correct feedback from the system, in a fashion similar to following a decision tree. They were designed to be used in single or multiple sessions in which participants had to sit in front of a computer for a large period of time to conduct the intervention - answering many questionnaires and reflecting on the elaborate generated contents derived from their answers. One of the areas where computer-tailoring was effective is in supporting health behavioral change [129, 130]. Recurring areas that apply these computer-tailored behavioral changes are dietary improvements, physical activity promotion, smoking cessation, and mammography screening, as reviewed by Krebs and Neuhauser [118]. Moreover, a review by Cheung et al. (2017) [217] showed that computer-tailored interventions specifically for smoking cessation were the most successful of all eHealth interventions. The impact was higher when there were multiple tailored recommendations [94, 327]. However, the public health impacts were limited [331, 332] when there was low usage [333], as users need to be engaged with the intervention to acquire the desired behavior change.

Health recommender systems (HRSs) are a newfangled way of providing personalized health information using artificial intelligence (AI) [334], which go beyond the traditional if-then programming paradigm used in computer-tailored systems [146]. HRSs can be conceived as adaptive persuasive technologies, capable of learning the most relevant strategy to provide behavior change support [335]. In most common scenarios, they predict the relevance that each user would assign to a given potentially recommendable message, learning and adapting over time from feedback. HRS which use a collaborative filtering approach can take advantage of the 'collective intelligence' of all users' profiles and their feedback, as opposed to traditional computer-tailoring. This can potentially allow the user to avoid having to answer lengthy questionnaires. Instead, users can provide a reduced set of details of their profile, hence reducing entry barriers to receiving personalized support. Consequently, HRSs may be used to offer higher personalization levels at a low-effort cost for users across interventions, potentially leading to increased improvements in health outcomes [276, 336] and higher user engagement rates [263, 337, 338].

One of the most recurrent fields where recommender systems were applied for health-related purposes is food and nutrition. For example, Elahi et al. (2015) introduced and evaluated a food recommender system obtaining user preferences through an interaction design, which showed a positive user feedback [339]. Musto et al. (2020) [340] presented a strategy using knowledge-aware recommender systems to recommend an appropriate diet. Gómez-del-Río et al. (2020) described an activity recommender system to promote healthy habits in obese children while using gamification as a mechanism to enhance personalization and increasing user motivation [341]. In addition, several other studies have investigated the usage of recommender system to help people nourish themselves more healthily [254, 342-344]. For the specific case of smoking cessation support, a previous study using a HRS which featured a hybrid algorithm (a combination of demographic filtering, content-based, and utility-based approaches) showed positive effects of the HRS that supported behavioral changes [164, 239]. This hybrid HRS was embedded into a mobile app incorporated to the routine care workflow of a hospital-based smoking cessation unit, and the app was offered to patients referred from other specialized care units of the hospital.

However, Shäfer et al. (2017) previously explained that advanced HRSs need to achieve high levels of personalization which classic recommender system algorithms (like collaborative filtering) might not have, as they do not take into account key user information such as patient gender, age, comorbidities, context, or ethnicity [161]. Also, other elements, such as being aware of the user context and grounding the recommendations on a behavioral model, are desired [345] and cannot be covered by the collaborative filtering algorithm approach on its own. Yet, demographic filtering is a variant of the collaborative filtering algorithm and uses users' profiles to compute similarities. However, this user profile information is obtained by requesting users to fill in their profile (sometimes with lengthy questionnaires) thereby reverting back to the necessary step in traditional computer tailoring. Despite completion being less time consuming as in a traditional-tailoring approach and not mandatory, even a task of an additional few minutes may be undesirable and lead to lower engagement levels or even increase attrition risks [346], common problems with HRSs [243].

As advanced health recommender systems may offer that needed level of personalization in healthcare, we wanted to compare two HRS in the context of smoking cessation support. The HRSs would select the most relevant motivational messages to be sent to smokers who are making a quitting attempt through a mobile app, aiming to increase their abstinence rates. The first HRS used knowledge-based algorithms (KBA). The second one used hybrid algorithms (HA) which performed a knowledge-based step just as the KBA, but its output was passed to a second step which performed a collaborative filtering – more specifically, a demographic filtering. This demographic filtering step had embedded in its design the principles of the Integrated Model for explaining motivational and behavioral change [345], also known as I-Change. This model was chosen because it was useful for supporting smoking cessation in previous interventions [313]. The I-Change model assumes that people transition through three phases (awareness, motivation, and action) to develop a behavior, which are conditioned by information and predisposing factors. Each of these phases has relevant determinants of change (e.g. attitudes and skills), which are the elements we included as part the algorithm design.

After reading each message, all participants were asked to rate them with a score ranging from 1 to 5 stars for the message relevance and usefulness. Although the ratings were irrelevant to the KBA algorithm, they were useful for improving the collective intelligence of the demographic filtering algorithm used in other participant groups. Also, this provided users of both groups a feeling of control over the system and gave the impression that they could provide feedback, making their experience more homogeneous and comparable.

We expected that the HA would be able to generate more relevant and personalized recommendations than its simpler knowledge-based-only version, leading to less dropouts, higher message appreciation, engagement, and effective behavioral change. This evaluation goes beyond the traditional way of assessing recommender systems, based on how well a given recommendation matches previous interactions between users and items [347-349]. However, some authors already proposed that HRSs need to include some kind of utility function that takes into account both user satisfaction (message appreciation and engagement) and health impacts (smoking cessation behavioral changes) [161, 213], as we do in this study.

At the same time, these higher levels of personalization and relevance in messages might not be needed, for instance, for smokers already highly motivated to quit. Hence, information on specific subgroup effects is also needed. Currently, it is not clear how to best develop more-advanced systems and whether such integration will yield better outcomes in terms of users' message appreciation, engagement levels, and health outcomes compared to simpler HRS algorithms which address a reduced number of user profile elements. This consideration is in line with two recommended future challenges for HRSs identified by Shäfer et al. (2017) [161]: 1) selecting dataset sources and ensuring their quality, and 2) exploring different models to personalize intervention contents based on user's health context, history, and goals.

The first goal of the present study was to compare two smoking cessation support HRSs: one only knowledge based (KBA), and the hybrid approach using knowledge based filtering with collective intelligence using demographic filtering (HA). Both systems were compared on: a) message appreciation, b) engagement with the system, and c) one's own self-reported smoking cessation status as indicated by the last 7-day point prevalence

(7D-PP) report in different time intervals. The second goal was to analyze potential subgroup differences in these three outcomes by gender, age (categorized as either an older or younger generation), profile completion, motivation to quit levels, and nicotine dependence levels.

Methods

Design

A trial was approved by the Ethical Committee of Taipei Medical University – Joint Institutional Review Board (TMU-JIRB), and conducted from November 10, 2017 to January 15, 2020. After voluntarily downloading the mobile app and filling out a baseline measurement, respondents were randomly assigned to either the KBA smoking cessation intervention or the HA smoking cessation intervention. Baseline measurements were taken between November 10, 2017 and July 15, 2019, and post-tests ended 6 months after the baseline. Between the baseline and final follow-up, respondents could freely interact with the system to obtain more motivational message support, as well as provide evaluations on the various messages they received.

The Android version of the app used in this study was launched on November 10, 2017, and the iOS version of the app was launched on August 6, 2018. Different time intervals were defined to perform in-depth analyses to measure the evolution of the two user groups: 0~7, 8~14, 15~21, 22~30, 31~60, 61~120, and 121~180 days. These time intervals were defined so as to have a good understanding of participants' evolution throughout the study, aiming to distribute the collected data in time periods where none would be without data – as time passes, participants are likely to stop using the app as intensively as in the initial days. The registered information would then be sparser across later months, and therefore the time periods should increase their duration. Grouping the data in larger time intervals allowed a more-effective analysis of that period, and those intervals are commonly assessed time points for smoking cessation assessments (7 days, 15 days, 1 month, 2 months, and 6 months). The initial time point of these time intervals was the first day with a valid quitting attempt.

Interventions

Both HRSs selected smoking cessation motivational messages and sent them to a mobile app used in the 3M4Chan study [249] performed within the H2020 Project SmokeFreeBrain [248]. The mobile app interface was the same in both groups, as were the personalized elements included in the motivational messages, such as referencing the name of the user in the message and the initial message delivery frequency between participants. According to Abroms et al. (2015) [284], five messages were sent on the quitting date, one message per day during the first week after the quitting day, and three messages a week after that. This frequency could be changed every 14 days after the second week of the quitting attempt by participants themselves regardless of their group. Users were given the choice of changing the frequency, i.e., increasing it up to one message per day or decreasing it down to one message per week, by answering a weekly question within the app. The time to send each message was set at random given an allotted time range previously configured by each participant.

In total, the system had 311 different messages. The same message could only be sent a maximum of three times to a participant as explained in the system design description by Hors-Fraile et al. (2019) [345]. Also, users could report their abstinence status in the app by answering the following weekly question: "Are you still resisting the temptation or have you smoked? Please, be honest." And possible answers were a) I have not smoked; b) Only one cigarette; c) Two or three cigarettes; and d) Four or more cigarettes.

Messages

The messages used in the study were created by a behavioral science researcher in English. They were then translated into Mandarin Chinese, and validated by two Taiwanese doctors specialized in smoking cessation. The messages followed the tailoring recommendations for smoking cessation support made by the World Health Organization (WHO)[269]. These guidelines included case scenarios of behavioral support sessions, reasons, strategies, and tips for people stop smoking. We reflected that knowledge in the sentences by elaborating on the recurrent topics and following the suggested approaches to support abstinence.

5

In addition, we also introduced determinants of change -psychological constructs which influence a behavior- proposed in the I-Change model [313]. The chosen determinants were used in previous studies to increase awareness, raise motivation, and change behaviors [270-274]. The included determinants were: attitudes towards stop smoking (the perceived advantages and disadvantages of quitting), social support to quit, skills to manage situations when feeling tempted to smoke, self-efficacy to quit –the perception of the smoker's ability to achieve it-, and action planning of the tasks to be successful (e.g., throwing away all ashtrays at home). All messages were enounced from a positive point of view, addressing the reader with the 'you' pronoun, using the active voice and easy to understand vocabulary, avoiding technical terms and complicated words, and with a maximum word count in English of 200, with an average of 85.5 words per message. Ten of the behavior change techniques proposed by Abraham et al. (2011) [278] were included across the messages. The messages also considered health communication methods such as repeating answers, creating empathy, adding new knowledge, and changing existing misconceptions. Additional examples of the applications of these techniques can be found in the Appendixes C and D of the study by Hors-Fraile et al. (2019)[345]. As an illustrative example, the following sentence presents a motivational message intended to enhance the social skills of a participant called 'John', who reported struggling with social pressure to smoking in while completing his profile during the enrollment process in the mobile app: "Hi John. You told us that you cannot refuse a cigarette when someone offers it to you. Well, we understand that you may struggle with it. However, you can practice how to kindly refuse a cigarette. A good reinforcing strategy is to add to the sentence one of the reasons you have to quit smoking. 'No. I am quitting because I want to keep my teeth clean and white' or 'No. I am quitting to have a longer life and enjoy with my grandchildren'. In this way, you say no and also remind yourself why you are doing it. If you prefer not to say the reason, you can just think it. To get a natural and almost immediate reaction, you can ask someone you know to roleplay a person inviting you to smoke, and you have to reject it. Even if you think this exercise is not worth doing it, it can help you succeed in a real situation."

Knowledge-based algorithm (KBA) system

The KBA system computed an adapted KBA approach over a pool of smoking cessation motivational messages. The adaptation consisted of pre-selecting message characteristics based on five user profile features: age, gender, quitting date, number of smoked cigarettes, and weekly expenditure on tobacco products. If the message dealt about time or date-specific contents (e.g., things to do early in the morning, related to the weekend, etc.), then the context when the message was going to be sent (the time and weekday) was also considered - and it was calculated before starting this KBA. Thus, the user requests which are necessary to the run knowledge based algorithms [150] were fixed by design. This was done so that participants would receive only one message according to the message frequency delivery pattern of Abroms et al. (2015) [284]. In addition, the intended intervention behavior was not to give participants the possibility of choosing from different messages, but to send them one relevant message per delivery. In this way, participants could focus their attention on a single message concept, not being flooded with different concepts at the same time which might not be good for remembering them, and which could even overwhelm them, producing them anxiety and becoming a trigger for smoking.

Consequently, the KBA retrieved all messages compatible with the given participant profile. A message was compatible with the participant profile if all the meta-features (the defining attributes) of the messages were also found in the user profile. If a message did not have a specific meta-feature value, it was considered as a valid match. For example, a message that did not have a gender meta-feature was assumed to be compatible for both male and female users. A simple illustrative case is shown in the table below. The last column contains whether the different messages are compatible for the user profile of a participant with the following meta-features: middle-aged, male, less than 10 cigarettes per day, whose quitting date was 9 days ago, with a low weekly expenditure, and who set no time limit to receive motivational messages during the day. The intermediate columns of the table represent the meta-features of each message.

Table 1. Example of how the KBA processes message compatibility for a given user.

	Age	Gender	Number of daily smoked cigarettes	Weekly expenditure	Quitting date	Context	Final compatibility
Message A	Young	Male		High		Only for mornings	No
Message B			<10		Less than 15 days ago	Only for weekends	Yes
Message C	Old	Female	>20			Only for evenings	No
Message D	Middle-aged			High		Only for mornings	Yes
Message E		Male	10-20	Low			No
Message F					More than 30 days ago		No
Message G		Female			Less than 7 days ago		No

The result of applying the KBA algorithm was a list of compatible and potentially relevant messages for that user. Then, as only one message could be sent at a time, the system selected messages from the list that had been sent fewer times to that user and picked one at random in case of a draw.

Hybrid algorithm (HA) system

The HA system computed a two-staged algorithm approach, also known as cascade [350]. The first stage performed the same knowledge-based algorithm as in the KBA described above to filter potential non-compatible messages with the target participant. In the second stage, a user-based demographic filtering algorithm was applied to the output of the first step (as opposed to random selection used in the KBA). This second step selected a motivational message based on the premise of prioritizing those which were found relevant by other similar participants (called neighbors) [279]. The selection was done using a score that is calculated for each message. This score represented the probability that the message was relevant for the given user. The algorithm used for this calculation worked as follows: First, all users and their message ratings were represented in a matrix (not including the user and ratings for which we are going to send the message). Second, a neighbor similarity index was computed for each user in the matrix following the equation below.

Where:

$$sim(A,B) = \frac{\sum_{i=1}^{n}(\delta_{F_A(i),F_B(i)}) + \sum_{j=1}^{m}\left(\frac{\delta_{F_A(j),F_B(j)}}{\max\left(|F_A(j)|, |F_B(j)|\right)}\right)}{\sum_{k=1}^{n+m}(\delta_{F_A(k),F_B(k)})}$$

A and B are two users;
F_u *represents all meta-features completed by user 'u';*
$F_u(x)$ *represents the value of the meta-feature 'x' of user 'u';*
$|F_u(x)|$ *represents the number of potential valid values of meta-feature 'x' of user 'u';*
n is the total number of meta-features which can have only two values (e.g. yes or no, male or female, etc.);
m is the total number of meta-features which can have more than two values (e.g. high, medium, or low motivation level);
$\delta_{y,z}$ *represents a function that sums the number of matching meta-features between the lists of meta-features 'y' and 'z'.*

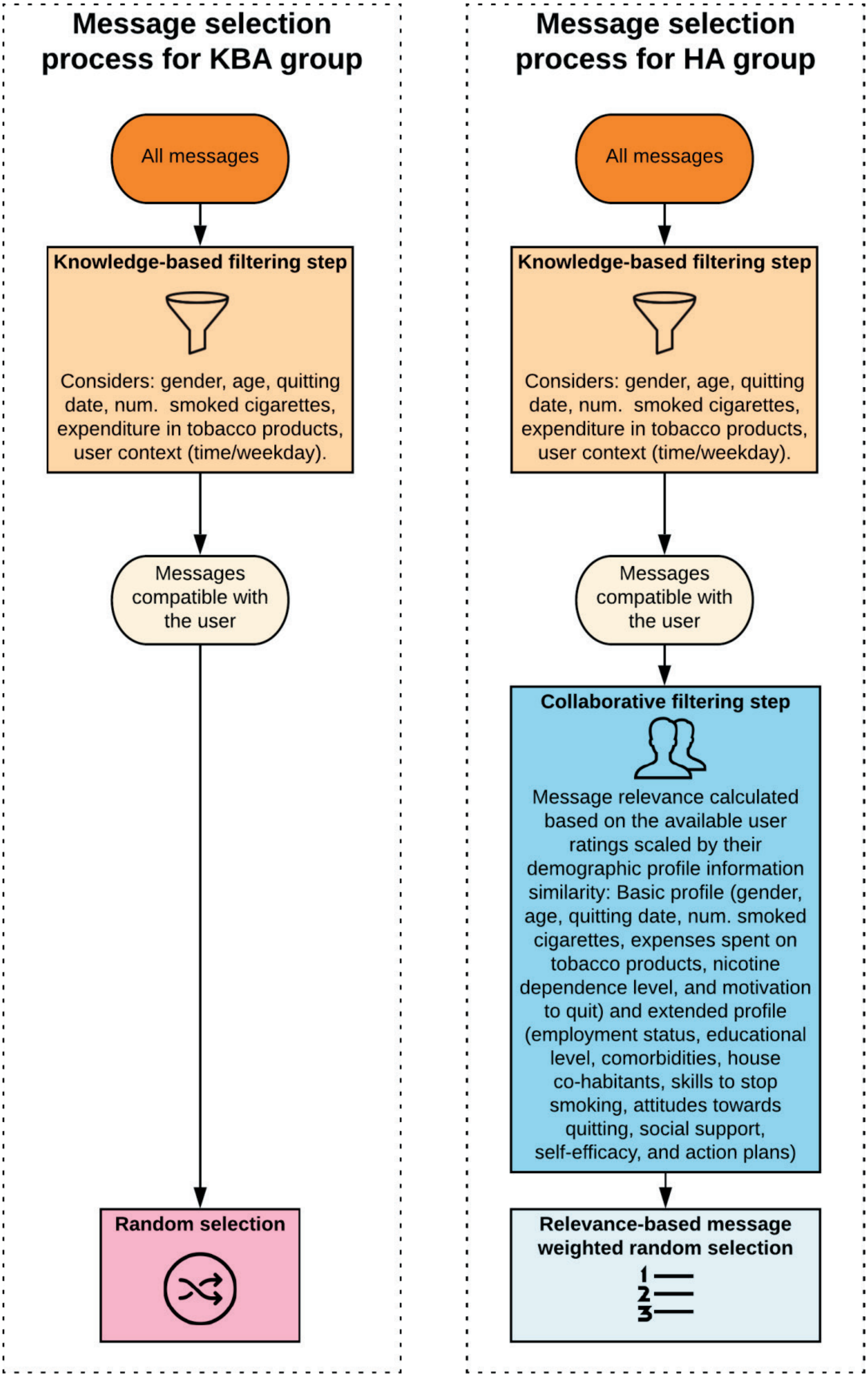
The system used all available information about user meta-features in the neighbor (with a minimum of nine, i.e., the core questions, and a maximum 60 variables, after adding 51 voluntary extended-profile questions). The more meta-features in common between participants, the more similar their profiles were and the higher likelihood of rating messages similarly. Hence, this similarity calculation approach could cover both participants who completed the extended profile and those who did not.

Third, the ratings for a message were multiplied by the corresponding user similarity index and added together. If a message had been re-rated by the same user, only the last rating was considered. Fourth, they were normalized to range between 0 and 1. Messages which were not rated previously, were assigned a 0.5 final score. This represented a 50% chance of being relevant or not. Fifth, the algorithm selected only the subset of messages which had been sent to the user a lower number of times to maximize message diversity and minimize repetitions. Messages which had been sent three times already to that user were discarded – we chose that maximum number because we wanted users realize of their meaning, and implications of the message arguments, following Cacioppo et al. (1989) [283] recommendations. Messages sent once or twice previously, were sent with a complementary sentence in the same message acknowledging the fact that the message had, at some point, been sent already to the participant, but stressing on the importance of the message content and that was why the participant was receiving it again. Finally, a weighted random selection based on the relevance score was run to select the message to be sent.

Figure 1 summarizes the message selection process for the two groups, and the detailed selection process as previously explained by Hors-Fraile et al. (2019) [345].

Participants and recruitment

Figure 1. Message selection summary for the two HRSs used in the study.



Participants considered for this study were smokers over 20 years of age, the legal adult age in Taiwan, willing to quit, who owned a compatible smartphone (Android or iPhone), who downloaded the smoking cessation "Quit and Return" app, who registered on July 15, 2019 or earlier, who accepted the terms and conditions of the mobile app, who made a valid quitting attempt, and who was able to understand any of the two languages in which the mobile app was offered: Mandarin Chinese or English. A quitting attempt was considered valid if it was not cancelled in the first 24 h after it was initiated, when the quitting day was not set in the past, and when the quitting day was set within a maximum of 90 days in the future with respect to the user registration date in the system. The system used a JavaScript random function to randomly allocate registered users between the two groups to receive motivational messages to stop smoking using a different personalization strategy. Participants were blinded to this randomization.

Measurements

Demographics

All participants had to answer nine core questions about their profile to use the app. These included their gender, age, employment status, date on which they began smoking, quitting date, number of cigarettes smoked weekly, and amount of money spent weekly on tobacco. In addition, they had to complete two standardized questionnaires to determine their nicotine dependence [351] and motivation to quit [352]. They could also voluntarily complete a questionnaire with 51 additional extended-profile features about comorbidities, people they share their house with, educational level, physical activity routines, and questions based on the I-Change behavioral change model [313] such as skills for stopping smoking, attitudes towards quitting, social support, self-efficacy, and action plans. This extended-profile questionnaire was used in the intervention by the HA to create a more-comprehensive user profile and calculate similarities in the demographic filtering step.

We considered five user attributes as indicators of potential differences in the study outcome. Four of them were direct variables from the previously described core questions, and the fifth attribute was derived from answering the voluntary questions.

First, we assessed gender (male=0, female=1).

Second, we assessed nicotine dependence based on the Fagerström test [351] (low=0, high=10). Its result was used to classify participants following the categorization included in the Fagerström test: participants with scores of ≤4 were included in the "low-dependence" sub-group (recoded as 0), those with scores of 5 or 6 were included in the "medium-dependence" sub-group (recoded as 1), and those with scores of ≥7 were included in the "high-dependence" sub-group (recoded as 2).

Third, we assessed the level of the motivation to quit based on the Richmond test [352] on a scale from 0 (low) to 10 (high). These test scores were used to classify participants into four motivation sub-groups proposed in the Richmond test using specified cutoff points [352]. We merged the low and medium-low sub-groups to avoid having two groups with too few participants. Thus, smokers with scores of ≤5 were assigned to the "low and medium-low motivated" (recoded as 0) sub-group, those with scores of 6 or 7 were included in the "medium-highly motivated" sub-group (recoded as 1), and those with scores of ≥8 were included in the "highly motivated" sub-group (recoded as 2).

Fourth, we assessed their age on the starting day of their first valid quitting attempt. To do so, we created two participant age sub-groups (generations) based on their potential familiarity with technology as proposed by Berkup [353]: "baby boomers or older" (born before 1964) and "generation X or younger" (born after 1965) (baby boomers or older=0, generation X or younger=1).

Finally, we assessed participants' completion of the extended-profile questionnaire by checking their profile data registry (incomplete=0, complete=1). The latter measurement was only relevant for participants receiving recommendations generated by the HA because the intervention with the simpler knowledge-based algorithm (KBA group) was not influenced by the extended-profile questionnaire.

We did not include all core questions that participants had to complete as measurement variables because a) they were not a reliable indicator of differences on their own, as someone could change that status several times during the smoking cessation process, and participants were unlikely to update that detail in their profile; b) the number of cigarettes smoked weekly and the weekly expenditure on tobacco were related and were already reflected in practical terms within the nicotine-dependence level; c) the date they started smoking was not useful when comparing results because each participant could have had different smoking and abstinence patterns during their lifetime; and d) the quitting date was not useful for comparison as it was a voluntarily set date that was intuitively linked to the date the participant discovered the app. Further, none of the 51 extended features was used to measure outcomes as they were not mandatory and not all participants completed them. Yet, all the excluded variables were only used to generate recommendations.

Outcomes

Message appreciation

Message appreciation was assessed by comparing the average rating provided to messages rated by users of the KBA and HA groups in each time interval. Rating options ranged one to five stars. Users could self-determine whether to rate messages or not, and so the frequency of message appreciation differed between persons.

Engagement with the system

We assessed three metrics in each time interval: (1) the number of active days in the system, with an active day defined as any day when a participant opened the app at least once; (2) the number of rated messages that were sent and rated in a given time interval, and (3) the number of abstinence reports submitted in each time interval. A fourth metric, the number of registered quitting attempts that were not cancelled within the following 24 h after being set, was assessed for the entire intervention period as it did not make sense to subdivide it. We considered those attempts cancelled within 24 h as people just exploring the app with no real commitment to quit smoking.

Smoking behaviors

The 7D-PP smoking status was assessed using self-reported abstinence reports by participants in each time interval. The 7D-PP metric was defined as a self-report of smoking no cigarettes (not even a puff) in the last 7 days. Smokers had to report whether they were currently smoking (1) or non-smoking (0), based on having relapsed or being abstinent at the time of the report (regardless of the number of cigarettes smoked). A maximum of one abstinence report could be sent each week. As most time intervals lasted several weeks, several self-reported status reports could be sent during a given time interval; we always considered the last one submitted within each time interval. Thus, this metric provides the 7D-PP for the last self-reported smoking status submitted in each time interval. The 7D-PP was identified as an important metric to assess self-reported smoking cessation outcomes [66].

To compare smoking cessation outcomes with other studies, smoking behavior changes were calculated as the proportion of eligible participants that sent positive and negative self-reported abstinence reports. Three analyses were conducted. First, an analysis on available data averaging the values of the last 7D-PP abstinence reports of the intermediate time intervals. This was done taking the last report within each time interval, and then using those values to calculate the average across the study for each participant. Second, an analysis on the 7D-PP abstinence taking the very last available abstinence report value ever reported in the study by each participant. Third, a pessimistic analysis of the previous one where non-respondents of the abstinence reports within each time interval were considered as non-abstinent (penalized imputation). For this, we took the very last abstinence report value ever reported, but we considered relapsers those participants who did not submit an abstinence report. We used this latter analysis as a conservative approach to avoid optimistic comparison effects of the intervention [354].

Statistical analysis

To analyze the data, we first performed a descriptive analysis of the sample demographics. Secondly, three kinds of dropout were defined: dropout in terms of no longer sending message ratings, dropout in terms of no longer sending abstinence reports, and dropout in terms of no longer being active in the app. For each of these, logistic regression was performed to identify potential determinants of dropout, which were subsequently used as covariates in the primary analyses on the effects of the type of HRS. The independent variables used in the dropout analysis were: gender, nicotine dependence levels (low, medium, high), motivation level (low and medium-low, medium-high, and high), age (born after 1965 versus born at or before 1965), employment situation (employed versus unemployed), and completed he extended profile (yes versus no).

To take care of the dependencies within observations of each subject, mixed models were used to examine the effects of the app. Depending on the scale level and distribution of the outcome, a different model was used. For message appreciation (after categorizing this variable into 4 levels) an ordinal mixed model, for the engagement metrics (number of active days, number of ratings, number of abstinence reports, and number of quitting attempts) a negative binomial regression mixed model, and for the smoking cessation metric, a logistic mixed model. For each outcome, a suitable model for the (co)variances of the random effects was chosen to adequately capture the dependencies in the outcome across time. Since for smoking cessation, for none of the examined covariance structures of random effects convergence was obtained, the analysis was done by standard logistic regression.

To compare smoking cessation outcomes with other studies, smoking behavior changes were calculated as the proportion of eligible participants that sent the abstinence reports in each time interval. Next, both an analysis on available data according to a logistic regression and a sensitivity analysis were done, assuming a pessimistic scenario (penalized imputation) in which non-respondents within each time interval were considered to be non-abstinent.

For all metrics, we also performed an in-depth analysis of the impact of having completed the extended profile questionnaire, using the same type of regression model for each metric.

Results

Description of the sample and involvement level

In total, 844 participants downloaded the app and registered. Among these, 371 had a valid quitting attempt and were eligible for our study; 290 (78.16%) were male and 81 (21.83%) were female with a mean age of 36.90 (standard deviation (SD) 10.21) years. The mean nicotine dependence-level score was 5.13 (SD 2.59), and the mean motivation-to-quit score was 7.54 (SD 1.91).

In total, 181 users were allocated to the KBA group and 190 users were allocated to the HA group; these numbers slightly differed because of differences in making a valid quitting attempt. No statistically significant differences between the two groups were found in age, gender, nicotine dependence, motivation to quit, employment status, or completion of the extended-profile questionnaire. Table 2 shows a more-detailed analysis. In total, 843 rated messages and 373 abstinence reports were provided by the 371 users of both systems during the 6-month period after making their first valid quitting attempt.

Table 2. Participants' demographic and smoking characteristics distribution between groups.

Variable	Total (N=371)	KBA (N=181)	HA (N=190)	Test statistics p-value
Mean age, years (SD)	36.90 (10.21)	37.71 (10.78)	36.13 (9.60)	$F(1,369)=2.229$ $p=.136$
Generation distribution: Percentage younger (no.)	92.7% (344)	90.6% (164)	94.7% (180)	$\chi^2=2.342$ $p=.126$
Gender: Percentage male (no.)	78.2% (290)	76.8% (139)	79.5% (151)	$\chi^2=.390$ $p=.533$
Mean nicotine dependence score (SD)	5.13 (2.59)	5.08 (2.69)	5.17 (2.50)	$F(1,386)=.127$ $p=.722$
Mean motivation to quit score (SD)	7.54 (1.91)	7.54 (2.03)	7.55 (1.80)	$F(1,369)=.007$ $p=.933$
Employment status: Percentage employed (no.)	83.0% (308)	83.4% (151)	82.6% (157)	$\chi^2=.041$ $p=.839$
Profile competition: Percentage completed (no.)	43.9% (163)	42.0% (76)	45.8% (87)	$\chi^2=.544$ $p=.461$

Dropout analysis

The dropout analysis for messages ratings showed that there was no interaction effect between the app and time period on the dropout rate ($p = .394$), so a potential interaction between the group and time period on the average ratings or the number of rated messages could not be due to differences in dropout rate at the different time intervals. There was also no main effect of the app for this type of dropout ($p = .375$). The variables gender, nicotine dependence, motivation level and completed extended profile turned out to be predictors of this type of dropout and were subsequently used as covariates in the analysis of the outcome variables message ratings and number of rated messages.

For the number of active days there was also no interaction effect of HRS group and period on dropout rate (p = .910). So, the difference between the HRS groups in terms of dropout rate did not differ across time. Averaged across time, there was also no difference in dropout rate between the two HRS groups (p = .583). The variables gender, employment situation, nicotine dependence and completed extended profile were significant predictors of this type of dropout and were subsequently used as covariates in the analysis of the outcome variable number of days active.

For dropout based on smoking cessation reports, no significant HRS group by period interaction (p = .682) and also no main effect of HRS group (p = .158) was found. So the difference between the groups in terms of dropout rate did not differ across time, and there was also, averaged over the time intervals, no difference in dropout rates between the groups. Only completion of the extended profile was a significant predictor of dropout (p = .002).

Overview of outcomes

The results of the main effects of each measure are detailed in Table 3.

Table 3. Results for main effects of type of app on outcome variables for non-dropouts.

Variable	Regression coefficient of HRS group	Test statistic	p-value	Effect[b] size	95% CI for effect size
Message appreciation	.818	F(1,129) = .869	.353	2.265	(.399,12.85)
Engagement: number of rated messages	-.036	F(1,233) =.021	.884	.965	(.595,1.565)
Engagement: number of active days	-.258	F(1,440) = 4.112	.043	0.773	(.602, .992)
Engagement: number of abstinence reports	-.403	F(1,421) = 11.702	.001	0.710	(.530, .843)
Engagement: number of quitting attempts	.105	$\chi^2(1) = 1.190$	.275	1.111	(.920, 1.340)
Smoking cessation-7D-PP reports: available data (on last report available in each time interval)	-.474	$\chi^2(1) = 2.383$	.123	.623	(.341, 1.138)
Smoking cessation-7D-PP reports: available data (on last report available from 0-180 days period)	-1.010	$\chi^2(1) = 5.162$	.023	.364	(.151, .880)
Smoking cessation – 7D-PP reports: pessimistic scenario (on last report available in each time interval)	-.044	$\chi^2(1) = .037$	.847	.957	(.610, 1.501)
Smoking cessation – 7D-PP reports: pessimistic scenario (on last report available from 0-180 days period)	-.301	$\chi^2(1) = .732$	.392	.740	(.371, 1.478)

[a] Coding for HRS: 1= HA; 0 = KBA.

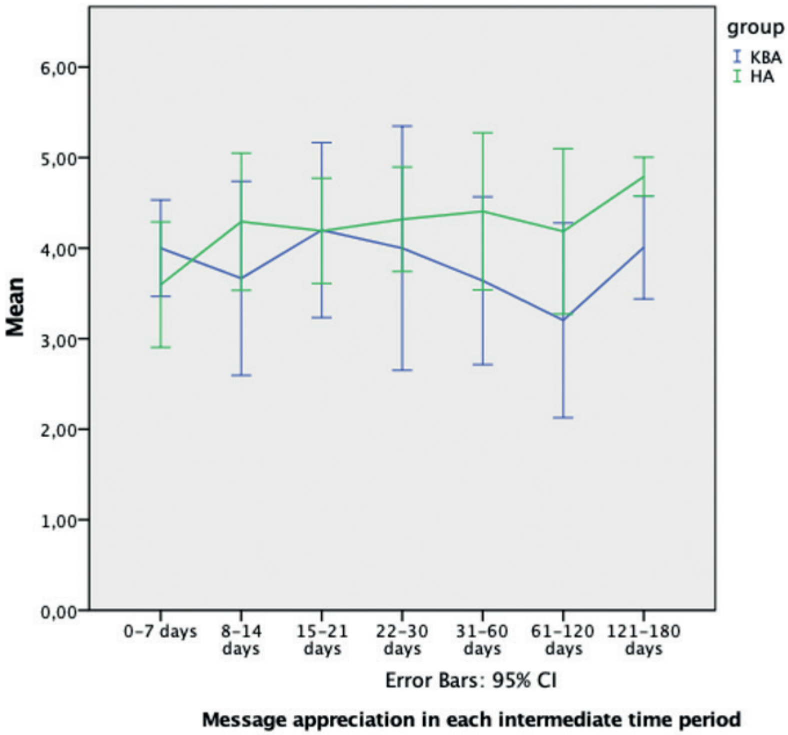
[b] Effect size: for message appreciation (ordinal outcome) and smoking cessation (binary outcome) the odds ratio; for engagement variables: incidence rate ratio.

5

Message appreciation results

The evolution of the mean message appreciation by each of the 2 HRS groups across the different time intervals is shown in Figure 2:

Figure 2. Evolution of the message appreciation.



The ordinal mixed model analysis showed that the difference between the two groups in terms of appreciation did not develop differently across time ($p = .897$). Also, there was no statistically significant difference between the two groups in terms of average ratings across time ($p = .353$). Hence, the difference between the KBA and HA group in terms of the level of message appreciation did not differ across time, and averaged over the study period (0-180 days) there also was no difference between the KBA and HA group.

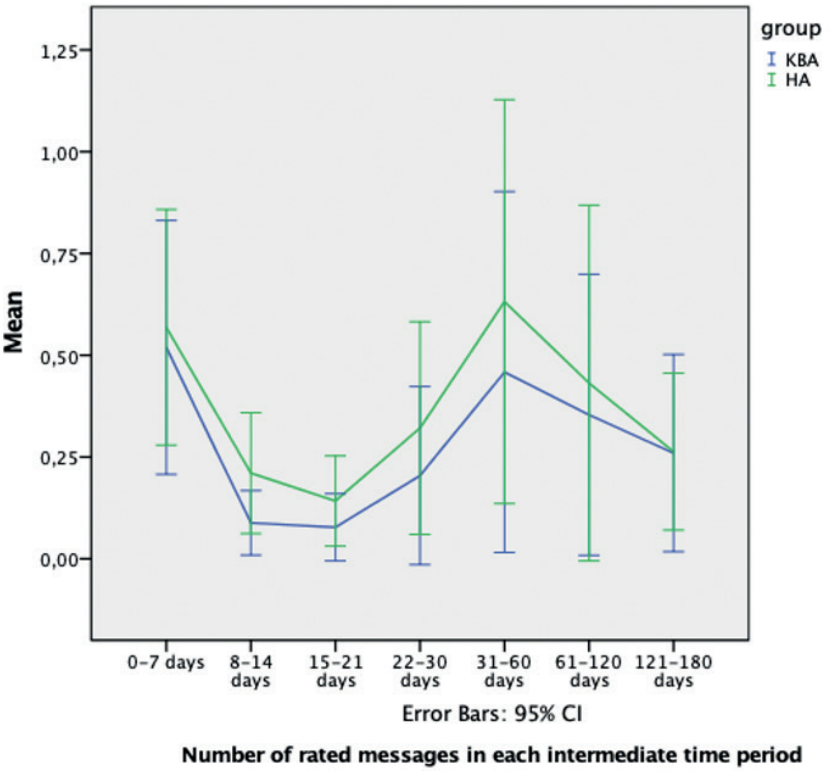
The analysis on the role of completing the extended profile showed that the effect of having completed the extended profile did not differ between the KBA and HA group ($p = .980$). Also, for both groups, completers of the extended profile showed the same level of message appreciation as those who did not complete it ($p = .976$). However, as emerged from the dropout analysis, for both HRS groups completers of the extended profile had a higher probability to stay in the study ($p < .001$).

Engagement resultsº

Number of rated messages

The evolution of the number of rated messages for each of the HRS groups across the different time intervals is shown in Figure 3:

Figure 3. Evolution of the number of rated messages.



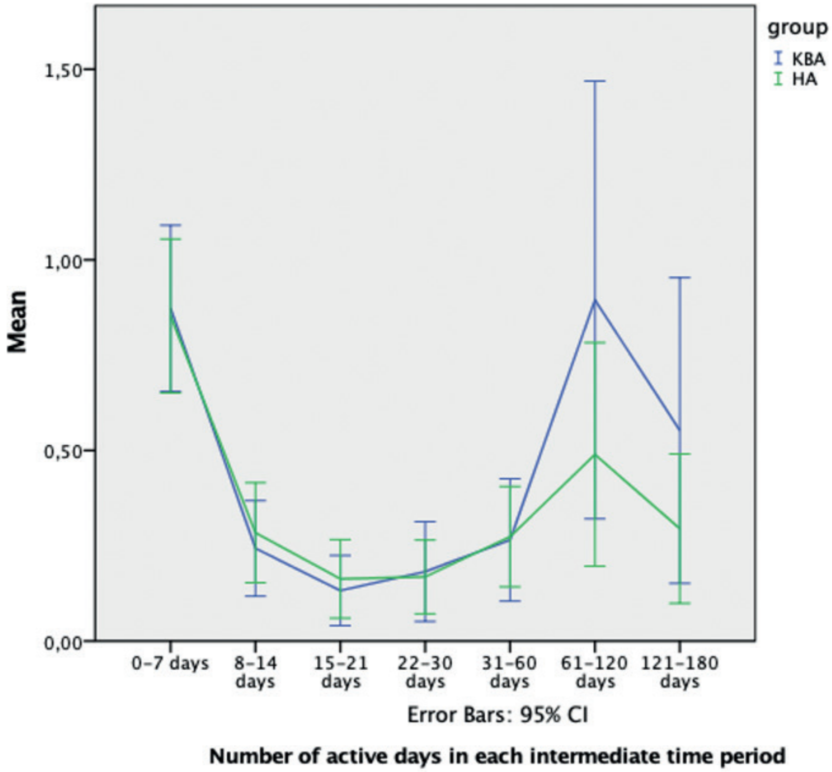
The analysis of number of rated messages, using a negative binomial regression mixed model, showed that the number of rated messages did not develop differently for the two HRS groups across time ($p = .920$). Also, there was no main effect of the type of app when considered across the whole 0-180 days period ($p = .911$).

The analysis of the impact of completing the extended profile questionnaire showed that its effect on the number of rated messages did not differ between the KBA and HA group ($p = .894$) and also, when averaged across both groups, had no significant effect on the number of rated messages ($p = .156$). However, as stated above, there was an effect on dropout: completers of the extended profile had a significantly lower probability to dropout from the study ($p < .001$)

Number of active days

The evolution of the mean number of active days in the mobile app for each of the HRS groups across the different time intervals is shown in Figure 4:

Figure 4. Evolution of the number of active days.



The difference between the HA and KBA in terms of active days did develop differently across time, although it was marginally significant (p = .051). We also found that the covariate employment status of the participants had a marginally significant interaction with the HRS group (p = .070). So, the examination the differences between the two HRS groups at each time interval was done separately for employed and unemployed participants. To control the type I error rate, for each of these groups, the Holm correction was applied when testing the app difference at each of the 7 time intervals. Unemployed participants showed no difference between the two app groups for any time interval in terms of active days. However, employed participants had a significant difference between the two app groups at the last two time intervals (61-120 days (p = .007), and 121-180 days (p= .008)), where the HA led to less active days than the KBA.

Examining the main effect of type of app over the 0-180 days time period also showed that the KBA led to a higher number of active days (p = .043).

There was no interaction effect of HRS group and completion of the extended profile on the number of active days (p = .815). The difference between completers and non-completers of extended profile in the number of active days was the same for the app groups. There was a marginally significant effect of having completed the profile on the number of active days (p = .068), in that completers are on average more days active. Also, completers of the extended profile had a significantly lower probability to dropout from the study (p < .001).

Number of quitting attempts

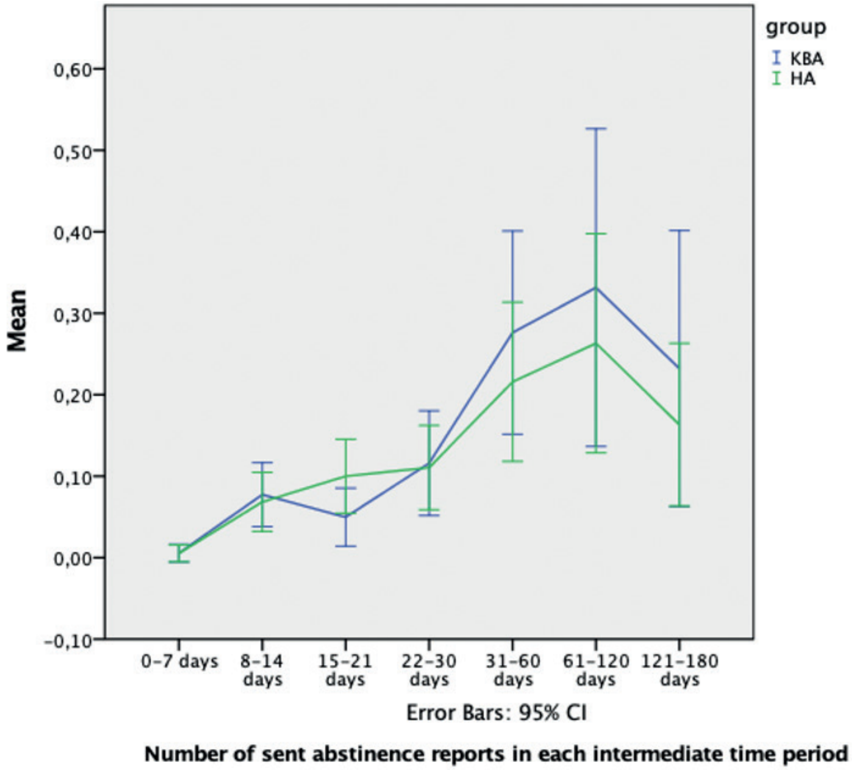
An analysis by the negative binomial regression model of the number of quitting attempts over the whole study period (0-180 days) showed that there was no significant difference between the two groups in terms of the number of quitting attempts (p = .275).

Testing the interaction between HRS group and having completed the extended profile on the number of quitting attempts showed that there was a significant interaction effect (p = .042). The effect is such that the effect of having completed the extended profile on the number of quitting attempts is significantly larger in the HA group. In the HA group there was a very significant effect of having completed the extended profile on the number of quitting attempts (p < .001), with the incidence rate ratio of completers versus non-completers being 1.869. In the KBA group there also was a significant effect of having completed the extended profile on the number of quitting attempts (p = .020), the incidence rate ratio of completers versus non-completers however being lower, 1.342.

Number of abstinence reports

The evolution of the number of sent abstinence reports for each of the HRS groups across the different time intervals is shown in Figure 5:

Figure 5. Evolution of the number of sent abstinence reports.



The difference between the two HRS groups in terms of the number of smoking cessation reports did not develop differently across time (p = .339). However, when testing the difference between the two HRS groups over the 0-180 days period, the average amount of cessation reports (KBA: Mean = 1.11, SD = 1.805; HA: Mean = 0.7, SD = 1.118) was significantly higher for the KBA group (p = .001). Note that this cannot be explained by the dropout rate, as there were no significant differences in dropout rate between the two app groups.

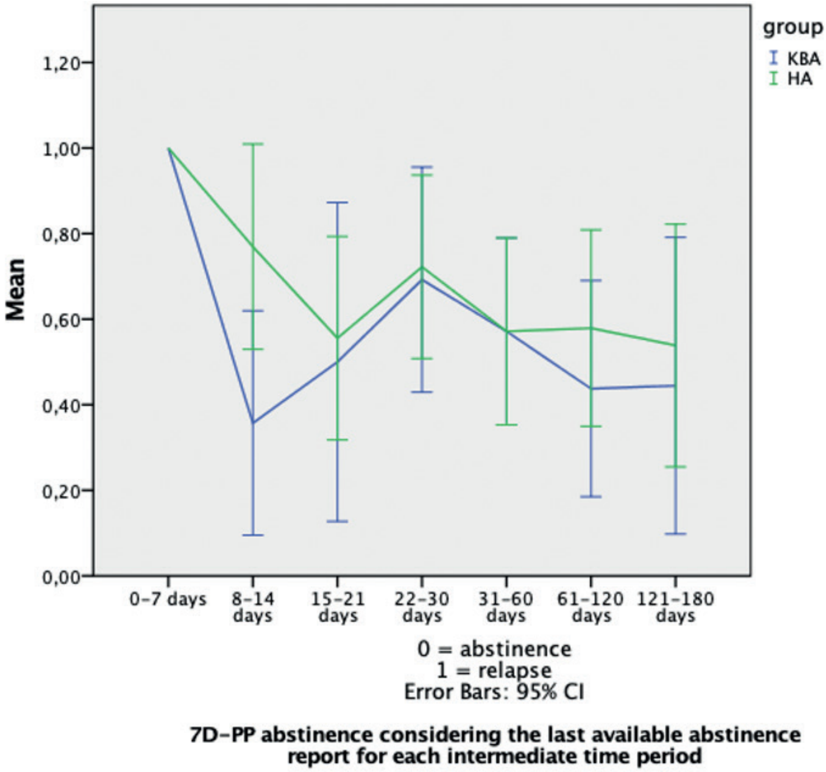
There was no significant effect of having completed the extended profile (p = .870), implying that completers of the extended profile had the same incidence rate of the number of cessation reports as the non-completers (and that was the case for both app groups). However, completers of the extended profile in both groups had a higher probability to stay in the study (p = .002).

Smoking cessation results

7D-PP abstinence: analysis on available data

The evolution of the 7D-PP abstinence for each of the HRS groups across the different time intervals is shown in Figure 6 (the mean was calculated by averaging all the last abstinence reports of each participant in each group, coded as 0 for being abstinent and 1 for relapse):

Figure 6. Evolution of the 7D-PP abstinence (based on the last abstinence report).



The analysis based on the standard logistic regression model revealed no statistically significant interaction between group and time period (p = .737). Also, the difference between the two HRS groups on 7D-PP averaged across 0-180 days was not significant (p = .123). So, no difference between the two groups in terms of 7-day point prevalence of relapse.

An overall analysis using logistic regression only on 7-day point prevalence between 0 and 180 days for the effect of the type of HRS revealed that belonging to the HA group led to a lower 7D-PP of abstinence (OR = 0.364, p = .023). The difference of this analysis with the previous analysis of 7D-PP averaged across 0-180 days is that, in this analysis only each participant's last available abstinence report over the whole time span of 0-180 days was considered, and not for each time interval. In the previous analysis, all participants' intermediate last available abstinence reports were averaged in determining whether there was a difference between the HRS groups.

The in-depth analysis of the impact of completing the extended profile for the two groups showed a marginally significant interaction between app group and completing the extended profile ($p = .065$). In the HA group the completers had a lower probability of being 7-days abstinent (OR = 0.566), whereas in the KBA group the completers had a higher probability of being 7 days abstinent (OR = 1.872), but in none of these groups these differences were significant ($p = .209$ and $p = .175$ respectively).

Also having completed the extended profile, was a significant predictor of dropout ($p = .002$), such that for both groups those who completed the extended profile, had a lower probability to dropout.

7D-PP abstinence: sensitivity analysis under a pessimistic scenario

The standard logistic regression analysis showed no interaction between group and time period ($p = .607$). Also, no main effect of HRS group on 7-day point prevalence was found ($p=.847$). So, there was no significant difference between the two groups in terms of 7-day point prevalence under a pessimistic scenario. Also, the logistic regression analysis for the 0-180 time period, which considered only the last available cessation report, showed no significant effect on the probability of 7-day cessation ($p = .392$).

Discussion

Main findings

The aim of this study was to compare two different health recommender systems for smoking cessation. The first system used a knowledge-based (KBA) algorithm whereas the second used a hybrid algorithm (HA), employing KBA and demographic filtering. Effects were studied concerning participants' message appreciation, engagement with the system, and self-reported smoking cessation outcomes.

Message appreciation

Message appreciation was almost always higher across the intermediate time periods for the HA group as can be seen in the message appreciation evolution chart (see Figure 2). However, this difference was neither significantly different for the intermediate time periods nor averaged across the whole 6-month period. This result suggests that, in our study, although the demographic filtering step in the HA was delivering slightly more relevant recommendations, it did not suffice to be statistically significant compared to a random selection of the filtered messages resulting from the knowledge-based step. This lack of statistical significance contradicts the potential advantage of collaborative filtering that several previous studies showed [355, 356], where messages had higher relevance after going through that step, leading to more and better ratings, which would reflect higher message appreciation [357, 358]. In our case, the lack of statistical significance could be related with the cold-start problem [359] - we estimated during the study design phase that we would recruit a higher number of participants, and those participants would provide more ratings, quickly reducing the impact of the cold-start problem. Another explanation for the lack of statistical significance is the limited sample size: 843 ratings provided by 57 participants, of which 30 were in the HA group. A higher sample size would yield more power to show that the apparent trend of higher appreciation and number of rated messages in the HA is, indeed, due to the demographic filtering step [360, 361].

Engagement

We expected that participants in the HA group showed higher engagement levels than the ones in the KBA group because the HA was supposed to provide more relevant recommendations which would lead to higher engagement. However, we found mixed results. For the number of rated messages, similarly to the message appreciation, we had a higher number of rated messages for all the time intervals, but this was not statistically significant, and the same rationale elaborated for the message appreciation is applicable to this measure. Finding no statistical difference on the number of rated messages can be explained for the same reasons (cold-start, and sample size) than for message appreciation. In addition, as we did not find statistically significant differences in message appreciation previously – which was the metric related to some predecessors of behavioral intentions of use [192], it follows that no statistical difference on the number of rated messages was found either.

Yet, we found differences in other types of engagement metrics: number of active days and number of abstinence reports, both being higher in the KBA group. These were unexpected results because we had expected, at least, the same level of engagement between KBA and HA groups since the HA added an extra filtering layer aiming to provide motivational messages tailored to the preference of the participant rather than at random as in the KBA group. Therefore, in the worst case where the demographic filtering step was not working, message appreciation was to be expected to be similar but not worse, and more in-depth studies may be needed to further understand this finding.

Another potential explanation could be that the factor which made participants engage with the mobile app in terms of active days and submit abstinence reports was not related to the algorithms, but with other variables that we did not consider in our study and which were more favorable for the KBA group. For example, perceived trust in the app could be one of these factors, as it played a relevant role in other health apps before [362, 363]. This explanation is also in line with the study by Dovaliene et al. (2016) [364], who showed that the user satisfaction was not relevant for mobile app engagement – which is what we were aiming to maximize through higher message appreciation as the more a participant appreciates messages, the more it may be satisfied with the app. However, as the mobile app interface was the same for both groups, the perceived trust – or any other non-considered variable – was more prevalent in those messages sent by the KBA.

The identified interaction of employed participants in the KBA group having a higher number of active days in the last two time periods has not an evident rationale behind it either. A potential explanation for this result could be that there was another participant variable that we did not measure, which related to the employment status, and which was the actual factor of differentiation. For example, this could be the participant's income and educational level. The socioeconomic status is a known predictor for smoking prevalence [365, 366] (higher income is typically associated with better health outcomes) and knowledge of smoking effects [367, 368] (higher educational level relates to higher knowledge) and could also be linked to mobile app engagement, as it was shown for other health-related app studies [369-371]. Therefore, the employed participants might present differences among them, and those in the KBA group might have had a higher income or educational level than the participants in the HA group, producing this difference in number of active days. The explanation for this difference being only present in the two last time intervals might be found, again, in the type of participants who did not dropout. If the dropouts in KBA group during the first time periods were those participants with lower socioeconomic status, whilst the dropouts in the HA group were homogenous, the KBA group would end up having a higher proportion of high socioeconomic status participants, which would lead to higher number of active days in the last time periods of the KBA group as we found in our study. Although, this could be a plausible scenario, we did not collect the socioeconomic status of the participants (other than the employment status) and we cannot draw any definitive conclusions.

The results also revealed that completers of the extended profile in the HA group made more quitting attempts. This suggested that, when the HA was provided with more information to personalize the recommendations, it may have influenced the determination to make a quitting attempt after a relapse. The evidence of the rationale for this effect would be more solid if we also found similar effects on message appreciation for the completion of the extended profile in the HA group because participants pleased with the received messages would be willing to re-engage with the app after a relapse. However, we previously identified the non-significant difference of having completed the extended profile in the appreciation metric. Therefore, although the effect on quitting attempts was significant, it will need to be further explored in future studies to reinforce its evidence.

In terms of the completion of extended profile, we expected that those participants who completed it had better results than those who did not complete it in the HA group: the method to calculate the similarity did not change, but participants who had completed the extra 51 variables would benefit from having a more-similar neighborhood, leading to potentially more engaging recommendations. When testing different approaches to calculate the similarity, varying the elements used to compute the similarity also varied the actual outcomes [372, 373]. In our case, during the algorithm design phase, we did not know what the optimal number and type of variables to use were, as each dataset will have a different optimal point [374]. However, we decided that the included variables using the information of the 60 questions were reasonable to define the user for the similarity computation. However, future studies should carefully re-assess what variables may better define user similarities. A lower number of variables may improve the user experience because they will complete shorter but more meaningful questionnaires. Also, we did not know how the number of user variables for computing similarity impacted in other metrics such as smoking abstinence. Increasing the number of questions might begin reducing the effective similarity among neighbors. This can happen if variables that do not really impact smoking and diminish the effect of those which do. Pu et al. (2012) [346] recommended minimizing preference elicitation in profile initialization. Yet, as extended profiles may improve smoking cessation attempts, further experimental research is needed to assess which core questions are needed. We recommend future researchers willing to use demographic filtering in digital health interventions to explore approaches that maximize the number of participants completing user profiles. The trade-off between a mandatory questionnaire and total freedom to complete it should be considered to avoid participants having a poor user experience but also retain the potential engagement benefits found in this study.

Smoking cessation

We found that the HA algorithm was unable to provide significantly better smoking cessation outcomes than the KBA in the 7D-PP analysis averaging the intermediate time periods abstinence report results. The abstinence measurement averaged across time intervals can be considered an approximation of continuous abstinence – which was a relevant measure included in the Russell 2.0 standard for smoking cessation outcomes [66]. However, the lack of abstinence reports at specific time points (e.g., 6 months) reduces the reliability of this approximation as we may be considering the continuous prevalence of participants based on only one report made in the first month, for instance.

The analysis for 7D-PP considering the last abstinence report of each participant as the value for the whole study showed that the KBA algorithm performed better. This 7D-PP can detect delayed quitting and was also included in the Russell 2.0 standard as a relevant smoking outcome measure [66]. However, as our trial was done under real-world conditions where we could not consistently collect abstinence reports at specific time points -it was the participants' decision to voluntarily submit the abstinence reports- this analysis implicitly assumes that the participants remained with the last reported abstinence status until the end of the study. This assumption may have introduced bias, as we cannot know the actual abstinence status of participants at the end of the study if they did not report it.

To further examine the sensitivity of these results, we also considered a pessimistic scenario analysis in which participants who dropped out were considered as smokers (penalized imputation). In this case, we found no differences between groups. This result differs from the corresponding analysis on available data result because in both groups there were participants who never sent an abstinence report and were assumed to be relapsers in this analysis – hence, reducing the differences between groups. Although, some authors showed that the penalized imputation does not necessarily lead to less biased effect estimates or more conservative effect estimates than the complete case analysis [375], and also suggested that it may be too overly critical [376], considering all the previous measures, it seems that the overall impact on smoking cessation outcomes for the participants was not significant between groups.

In addition, it is conceivable that certain messages could result in effects different than intended, and thus lowering the efficacy of the HA system. For example, for a participant called John, one of the messages could read "Hi John. Did you know that as you are no longer a smoker, you are no longer part of a chain that favors child exploitation? Yes, you've read right. In countries like Pakistan, USA, and Indonesia child labor is used in the tobacco industry! Children and teenagers are exposed to toxic substances and hard work conditions. Stay smoke-free for their sake, and for yours!". This message, which was intended to provide knowledge, may not be positively appreciated because, even phrased in positive, thinking about the fact of having contributed to child labor is unpleasant. However, the message contents might cast a deep impact on the participants, leading to non-perceived higher abstinence motivation. Therefore, if users did not like the message, the HA would be unlikely to send it again in favor of other better rated messages, whilst the random selection done by the KBA algorithm would give the same probability of sending the message again to a new user. Unfortunately, in this study we could not traceability back what specific message was rated with what score after the study. Consequently, we cannot validate whether this type of situations happened to be the reason for the unexpected engagement results.

Although the findings of the 7D-PP abstinence considering only the last abstinence report available contradict our initial assumption, the pessimistic scenario and 7D-PP abstinence considering the averaged abstinence reports were in line with a study by Westmaas et al. (2018) [377] who compared two tailoring levels in e-mails to support smoking cessation, and found no differences between the basic and advanced versions. This may mean that having a more complex system for tailoring smoking cessation support does not necessarily provide better abstinence results.

To compare our abstinence results to other similar previous studies, a recent review [378], examined SmartQuit [379], and SmokeFree28 [380], which managed to achieve abstinence rates of 13% at the 2-month and 21% at the 28-day follow-up respectively. Mobile-based interventions for smoking cessation such as Clickotine [381] reported a 7-day abstinence rate of 45.2% and a 30-day abstinence rate of 26.2% in an 8-week study following an intention-to-treat analysis. In more-traditional computer-tailored interventions that had follow-up assessments at 6 months as in our study, abstinence rates were up to 18.3% (10.2% intention to treat) [382] and 20.4% (8.5% following intention to treat) [85]. Those results are similar to results of our study (54.1% in the KBA group and 30.2% in the HA group following the analysis on available data for 7-day point prevalence abstinence considering the last available abstinence report for each intermediate time period, and 8.4% in the KBA group and 11% in the HA group in the pessimistic scenario).

Additional considerations

5

Across all metrics, we identified that the completion of the extended profile was associated with less dropouts. As this was consistent for both the HA and KBA group (in which the completion had no effect in the algorithm), we propose that it could be that the personality of participants who filled out the extended questionnaire is such that they are more curious or willing to interact with the system. This was consistent with the work of Karumur et al. (2018) who explored that user personality can influence user engagement and activity in recommender systems based on a collaborative filtering approach [383].

The external validity of mobile apps for health behavior change has been previously criticized and it is currently been studied in other domains such as physical activity [384]. We conducted this trial under real-world conditions, where we did not reward participants with money nor stimulate them to complete the abstinence reports; additionally, researchers did not recruit or follow-up participants in any way. Hence, the present results of real-world effectiveness in terms of the effects of such an HRS increase the external validity of our study [385], compared to efficacy results as in previously mentioned studies. Thus, the results of this trial study have to be understood within this real-world context and might not be directly comparable to other studies, in which recruitment and follow-up methods may have positively impacted the commitment of participants.

Regarding the recommender system algorithms used in this study, we could not compare their theoretical performance with others found in the literature before running the study because of their differences in design that makes them incompatible with existing datasets which were not focused on health behavior. For instance, in other contexts (e.g. movie recommendations) recommender systems could use existing datasets as benchmarks for performance assessment, such as MovieLens [155]. Hence, commonly used metrics for recommendation systems such as the root mean square error (RMSE), mean average error (MAE), normalized discount cumulative gain (NDCG), and precision could not be applied to assess the algorithms before the study, as either we lacked an initially rated, compatible dataset.

This study focused on metrics which assessed the impacts of the recommendations on participants' behavior instead, as Sahoo et al. (2019) [386] proposed for HRSs, not on the technical performance of the recommender system itself. Schäfer et al. (2017) [161] also concluded that HRSs needed multidimensional user satisfaction measures, which covered message appreciation and engagement metrics. Further, message appreciation can be seen as a proxy for Mean Average Precision at 1. This sets our study in line with previous experimental studies for HRS performance analysis based on hits and total numbers of recommendations or users. This was the case in a study by Rivero-Rodriguez et al. (2013) [387], which followed a similar approach to assess their HRS, using the hit rate (no. of hits / no. of users), as a performance metric. Another example can be found in a study by Bocanegra et al. (2017) [388], in which they used precision to assess their recommender system. Despite applying similar approaches, the direct comparison of performance metrics and reflection on the conclusions generated in other studies are still limited in our case because they are totally different study designs and research questions. A recent scoping review [243] backs the relative scarcity of studies applying HRSs in the health domain and their diversity in therapeutic areas and reported outcomes.

The most similar study we found was the SoloMo study which presented results of another hybrid recommender system for smoking cessation [163]. That algorithm was tested in a clinical context where patients were followed up by healthcare professionals for 1 year [164]. Patients referred to the smoking cessation unit from other specialized care units of the hospital, were invited to the study. The precision of that system (which was directly related to appreciation as previously mentioned) achieved a high score (which would mean high appreciation) – with a minimum value of 0.96 over a total maximum of 1. There were key differences with our present study which explain this difference. First, the end of their messages included the name of the patient's doctor, and this may have increased the perceived quality of the messages, leading to a high precision (and we can therefore assume, also appreciation). Second, patients may have thought that their doctors were going to check their ratings and wanted to please them with higher scores. Third, the potential rating options were only positive, negative, or neutral, as opposed to the 5-star scale included in our study, which provided greater rating granularity. Fourth, the participants enrolled in SoLoMo were referred from a specialized care unit which eventually contributed to enhance their motivation and appreciation of the system potentially driven by a positive framing bias effect. Despite the apparently better results in the SoLoMo study [164], a smaller range in the rating values (only three options compared to the five we had) limits the opportunity of HRSs to learn from users' opinions in digital interventions for smoking cessation.

Further, in a recent scoping review [243], several gaps of the reviewed HRS studies were found. Our study covered many of them, including: 1) reporting the results of our study based on a large user cohort size (n=371) compared to the previous ones found in the literature; 2) using an HRS which was grounded in a behavioral change theory (the I-Change model [313]) recommending messages with behavioral change techniques (following the guide by Abraham et al. (2011) [278]; 3) using advanced profile adaptation and 4) a having clear explicit feedback system (in the HA group).

Limitations

Despite strengths of this study, such as comparing two different HRSs and the fact that no statistically significant differences were found in participants' age, gender, nicotine dependence, employment status, motivation to quit, or completion of the extended profile questionnaire between the KBA and HA groups, our study was subject to some limitations.

First, the HRSs considered all users' feedback for computing recommendations. This implies that the feedback provided by one group affected the generation of recommendations for the other. This design decision was taken to reduce the cold-start effect. Second, between May 22 and June 6, 2018, and between August 1 and 6, 2018, there was a server service interruption that prevented users from registering the app and receiving messages. Third, we could not verify the smoking status self-reports. Although self-reports may provide a valid estimation of cessation rates [389], they were used previously in several previous studies [390], and the Society for Research on Nicotine and Tobacco Subcommittee on Biochemical Verification considered the use of biochemical validation unnecessary in studies with limited face-to-face contact [391], use of bogus-pipeline procedures [392], some biochemical verification methods would have improved the validity of the smoking status reports [393]. Also, the pessimistic scenario analysis we conducted intending to follow a conservative approach may have not accurately reflected the actual behavior of the participants. Fourth, we considered the last status for the abstinence report as the value for each time interval. This way of measuring smoking cessation results hampered direct comparisons with previous studies. Fifth, in this effectiveness study, smokers could report by self-chosen times, resulting in the fact that we could not assess all data for all participants at one specific time (e.g., smoking cessation status after 1 month). Sixth, it is conceivable that specific subgroup effects could have occurred in our analyses, requiring more sophisticated models with more two-way (or even three-way) interactions to explain our results. However, due to the sparsity of the collected data, these more sophisticated models could not be applied. Seventh, user-experience metrics such as perceived quality and satisfaction – which are commonly evaluated nowadays in the field of recommender systems [346, 394] – were not included in this study. Eighth, persuasion profile meta-features to determine what recommendations style (e.g., authority shown in the message, the reflected consensus stated in the message, the message sender liking perception, etc.) [395, 396] would persuade participants the most were not considered. Such type of meta-features could have added extra personalization power to the HRS without needing the participants to complete additional questions in their user profile.

Recommendations

Based on our results, we recommend future studies to keep exploring the usage of different types of HRS to support smoking abstinence, as the results of both algorithms improved the unassisted cessation success rates for 6 months which is around 3-5% [31, 397], and were over some nicotine replacement therapy rates, whose 6-month abstinence rate is around 7% [398, 399]. More research is needed to explore relationships between message appreciation, engagement, and health outcomes. When using HRS with collaborative filtering, new means to determine recommendations relevance other than participant's appreciation should be considered. For instance, including the achieved health outcomes as a complementary rating (e.g., asking when rating a message about abstinence status would provide feedback about how useful for abstinence purposes were previously sent messages so far). In this way, messages that are well appreciated but do not contribute towards supporting abstinence would not be recommended in the future, and messages which are both contributing to support abstinence and highly appreciated would be prioritized to be sent by the HRS.

As we did not find differences between HRS for gender, age, nicotine dependence level, motivation to quit subgroups, we encourage future research is done considering other variables such as trust and socioeconomic status which may help better understand the smokers' behaviors. Still, the gender, age, nicotine dependence level, and motivation to quit could be good meta-features for the HRS similarity computation and suggest keeping them as part of the HRS and message design process.

In addition, future research should consider larger sample sizes with more than 6 months of follow-up time, as results between the KBA and HA in terms of message appreciation and number of rated messages suggest that these differences could become significant. Alternatively, this may be compensated for by increasing the frequency of the sent messages, as the HRS would have more information to process new recommendations faster, and or have larger sample sizes to be better able to detect such differences. To avoid overloading participants with too many messages which may be bothersome and negatively impact their user experience, we consider that a suitable solution could be to offer 'on-demand' messages. This larger sample would facilitate applying more complex statistical models which may help us explain some results about which we could only guess in this study. Also, we suggest pro-actively persuading users to complete their profile if they do not do it voluntarily during the enrollment phase to maximize the impact of the collaborative filtering, giving participants more probability to make a new quitting attempt in case they relapse. This could be achieved, for instance, by making the digital solution ask the participants to complete one or two unanswered question of their user profile every day. It would yield a low entry barrier to start using the solution, whilst it would allow to compute recommendations with increasing user profile information over time.

Conclusions

The first goal of this study was to compare the two presented HRS. We found the KBA led participants to have more active days in the mobile app, to complete more abstinence reports, and to have better 7D-PP abstinence results, despite being non-significant when averaged across time, and also being non-significant in the pessimistic scenario. The additional step of demographic filtering in the HA only improved the number of quitting attempts. However, the HA group seemed to rate a higher number of messages and give better ratings to the messages based on the trends shown on the evolution line charts – yet both measures did not statistically differ between groups. The second goal was to identify potential subgroup differences, and we found that participants who completed their extended profile were more likely to stay in the study and, among employed, those who were in the KBA group had higher engagement in terms of active days than those in the HA group. No other differences were found for any other subgroups (gender, age, nicotine dependence level, motivation to quit).

Our findings provide insights to the usage of health recommender systems in a real-world setting. We conclude that the collaborative intelligence provided mixed results, some of them unexpected, and more research is needed to fully take advantage of it in the context of smoking cessation support. However, this study showed the promising future of health recommender systems: combining behavioral change techniques and models in recommender systems- even with simple algorithms - can lead to higher smoking cessation rates than unsupported quitting and also some nicotine replacement therapies.

CHAPTER 6

General discussion

Aim of this dissertation

In this dissertation, we have provided detailed insight into the usage of recommender system technology for health promotion and behavior change. Chapter 2 contains a formative research study to identify the research gaps in using recommender systems for health promotion through an extensive scoping review of the state-of-the-art HRSs (**RQ1**). Subsequently, in Chapter 3, we have described the multidisciplinary approach to design a mHealth solution for smoking cessation powered by a health recommender system grounded in behavioral science (**RQ2**). Next, in Chapter 4, we have described the protocol to assess the impact of two HRSs that were a part of a mHealth solution in supporting smoking cessation: one HRS used collective intelligence grounded in behavioral science, while the other sent random messages (**RQ3**). Finally, in Chapter 5, we have evaluated and presented the two systems in terms of appreciation, user engagement, and smoking cessation outcomes (**RQ3**).

This chapter summarizes the main findings and explains the added value of the research described in this dissertation and its strengths and limitations. Subsequently, the methodology and the practical consequences of this research are discussed. In addition, the implications and future research recommendations are provided.

Main findings

Chapter 2: Review of the existing HRSs

In Chapter 2, we have presented a review of two state-of-the-art HRSs. This chapter aimed to understand the evidence generated by HRSs, identify the research gaps, and propose reasons and solutions for those gaps.

The scoping review conducted in Chapter 2 analyzed studies using recommender systems for patients in terms of their design, including their domain, assessment methods and procedures, technology, and outcomes. Of the 905 identified studies, 19 met the present study's inclusion criteria.

Our findings have shown that, first, the previous experiences of using recommender systems in healthcare had a sparse coverage, mainly focusing on generic health recommendations for adult users, and were limited to high-income countries. HRSs did not cover specific therapeutic areas, such as diabetes or cancer, particular age ranges, including teenagers and low-income countries. These limitations in scope and geographical context indicate that the efficacy of HRSs needs to be more tested. They are gaps in the field of HRSs, which researchers should cover in the future.

Second, another salient finding was the absence of health promotion theoretical factors and behavioral change theories as part of the design of the systems and messages included as recommendations. Health promotion theoretical factors and behavioral change theories were proven effective in other non-HRS interventions [198, 400], including when the interventions used computer-tailoring [401]. Ghalibaf et al. [329] showed that 50% of the generic computer-tailoring studies used behavioral change methods compared to the 0% present in the studies with HRSs as found in the scoping review. The lack of theoretical support of behavior change factors may reduce the impact of the intended outcomes of the interventions, as previous research found that digital interventions were more effective when they were grounded in behavioral theories [402, 403]. It also indicates that excluding them from HRSs is suboptimal and a missed opportunity to make the HRS-based interventions more effective. The behavioral scientific grounding could be applied to two main aspects of the HRSs: 1) the message selection algorithm, including the frequency of message delivery based on what behavioral change phase the person is at, and 2) the content of the messages—adapting them to include different determinants of behavior change that would benefit the person the most, depending on which behavioral change phase they are at, and also to phrase the messages using health communication techniques. For instance, it could include descriptions of material consequences of their behavior and ensure promotion to assess their own risks [278].

Our findings regarding the absence of health promotion theoretical factors and behavioral change theories in previous HRS-based studies were consistent with another recent scoping review on the same topic [166]. They coincided with the lack of inclusion of theoretical procedures for the design of the messages, and the underreporting of how the messages were created and user feedback was provided. This lack of behavioral change theories in HRSs was also aligned with other studies of digital solutions that proved that a limited number of found cases where behavioral change theories were applied [404-407]. The interventions that did include these theories used the transtheoretical model [408], health belief model [409], social cognitive theory [410], and theory of planned behavior [411], and elaboration of likelihood model [117]. Some just included a few behavioral constructs such as social support and perceived risks [405].

The studies more recent than the ones included in our review started using behavioral models but only partially. For example, the InspiRE system [412] used a recommender system that considered the principles of the transtheoretical model of behavioral change [266], social cognitive theory [413], and the flow concept [414]. In the design of the system, different recommender systems algorithms were considered to be used. Yet, some of them did not support users for behavior change transition (e.g., algorithms that would only recommend actions to reinforce current behaviors, not future healthier habits). The mentioned behavioral theories were then used to determine which algorithm would be more suitable for replicating the human behavior change transition process. However, these behavioral theories were only used as general guidelines for researchers about algorithm selection, and the theories were neither integrated as a part of the algorithms themselves nor the messages. For example, in the intervention of the SocialPOD study [165], the social cognitive theory constructs were included as a part of other elements in the used mobile app but not a part of the recommender system itself.

Third, the reported HRS-related metrics of each study reviewed in Chapter 2 were different and lacked a consistent and complete view of the system performance and impact on the participants. In order to conduct a comprehensive analysis of the HRSs impact, as done in the studies included in the scoping review, the metrics used in those studies were checked for whether they were tested with users, their effectiveness, and their exposure time and frequency to the intervention. The results showed that five studies used only technical performance metrics such as precision and recall [415], three used subjective metrics such as perceived usefulness [192], and only one used a biometric indicator (weight). In addition, only six studies that used some type of metrics conducted actual tests with users. Only one study reported results about the effectiveness of the intervention with regards to patients' health-related outcomes [165]. The level of exposure to the interventions was also not fully discussed in detail in these studies.

Fourth, concerning the aspect of used technology, the studies did not fully report on how the recommendation process was carried out, beginning from the initial profile generation and learning technique to the finding recommendation process and system feedback. This lack of transparency hinders gaining a complete understanding and future replication. Yet, the most frequent information filtering method used by these studies included collaborative filtering. This type of recommendation generation process is known as collective intelligence [146]. The collective intelligence uses information such as user ratings and implicit feedback (e.g., in-app usage behavior) to codify human preferences and behaviors so that computers can understand them for making recommendation predictions [416]. Using algorithms that leveraged other user information to generate recommendations (e.g., collaborative filtering) was the most common approach in our study and that of Cheung et al. [166].

In generic computer-tailoring, only 13% of studies detailed the technical details of the system [329] in comparison with the almost 42% found for HRSs in our review, thus, illustrating a need for more transparency in the description of computer-tailored studies. This discrepancy may be because researchers who used HRSs have technical backgrounds and explored the potential of applying new technologies to the healthcare domain. As a result, they inadvertently highlighted the aspects of their own field in more detail than the healthcare aspects. Similarly, researchers using computer-tailoring had behavioral and health sciences backgrounds and used technologies as a means to deliver behavioral interventions, but not as their research objective. These researchers were more likely to include and state the behavioral aspects of their interventions aligned with their backgrounds and consider the technological aspects less important.

The lack of a comprehensive reporting of key HRS elements in the previous studies in the literature suggests a need for a taxonomy that is adapted to HRSs. Although other taxonomies for generic recommender systems exist, they do not include specific and relevant dimensions for HRSs. Taxonomies aim to structure and organize an area of a discipline or field, allowing transparency and ease of description and evaluation [417]. In our study, we created a taxonomy for HRSs using a multidisciplinary approach, including the following:

1. The domain of the HRS (e.g., the therapeutic area and target population)
2. The implemented methodologies and procedures (e.g., assessed metrics and their effectiveness on patients)
3. The used health promotion theoretical factors and behavioral change theories (e.g., the inclusion of determinants of change such as attitude and self-efficacy)
4. The technical aspects (e.g., information filtering method and recommendation interface)

This taxonomy was to be used as a tool in the reviewing process. Also, it offers future researchers the list of domains (therapeutic area, target population, country of the intervention, type of recommendations, and interface) and methodological and technical aspects that they need to consider for the comprehensive design of an HRS.

One taxonomy for recommender systems was presented by Montaner et al. [219], where profile generation and maintenance (e.g., ratings and gradual forgetting) and profile exploitation (e.g., collaborative filtering) were used as the two main dimensions to classify the systems. However, it did not include any health-related aspects as our proposed taxonomy does.

A more recent taxonomy was published by Lousame and Sánchez [418]. It focused on the collaborative-based recommender systems, adding relevant aspects to understand how they were designed. Their taxonomy considered: several represented entities (e.g., a user is represented by attributes such as age, gender, and country), user and item representation (e.g., purchased items, ratings, comments), and the associations (e.g., implicit or explicit feedback), along with the recommendation methods (e.g., user similarity). However, despite their specialization in collaborative-based systems, Lousame and Sánchez [418] excluded from their taxonomy other types of recommender systems, such as the ones that are knowledge-based, which can also be used for behavioral change. Moreover, they did not include health-related aspects necessary for HRS evaluations.

Friedrich and Zanker [419] and Papadimitriou et al. [420] presented other taxonomies for recommender systems. They focused on categorizing the recommendation explanations. Knowing how users understood why a recommendation was given is insufficient to detail how the recommendations were computed, which is crucial for understanding an HRS. Therefore, these taxonomies proposed a deep analysis of the recommendation explanations to completely understand an HRS, such as the recommendation method, without including other necessary information, while our taxonomy did. Although we did not consider categorizing the explanations of HRSs, all of them are likely to explain how the recommendation was generated to foster trust [421]. Trust is especially relevant in healthcare compared to other fields. For example, a recommendation about an article to be bought or a movie to be watched may not depend on the trust for that recommendation [161]. Thus, these other taxonomies can only be used to complement ours, but they cannot cover the description and classification of an HRS on their own.

In conclusion, although other taxonomies for recommender systems were available, they mainly focused on the technical aspects. However, the most significant difference between these systems and the one described in this thesis was that the latter includes essential elements for systems in the healthcare domain. These specific elements were the methodology and procedures of evaluating the system with patients (e.g., cost-effectiveness, percentage of patient success) and the usage of behavioral science (e.g., usage of determinants of behavioral health change). Therefore, to foster a good understanding of future HRSs and facilitate their replicability and improvement, their healthcare domain and technical and behavioral science aspects should be comprehensively described. Our proposed taxonomy aims to cover these elements.

6

In conclusion, in Chapter 2, we found that HRSs were not comprehensively described in the literature. Several relevant aspects were missing (i.e., the usage of behavioral change techniques). Therefore, a comprehensive multidisciplinary approach was suggested to present and categorize this type of system, along with a taxonomy to foster the development of comprehensive HSRs (**RQ1**).

Chapter 3: Description of our HRS

In Chapter 3, we presented a comprehensive design of a smoking cessation mHealth solution using recommender systems grounded in behavioral science to support smokers to quit smoking. One of the goals of this chapter was to explain the design process to combine an HRS and the I-Change model [78]. The second goal was to foster transparency and reuse the implemented principles in future studies.

The first main topic described in Chapter 3 was combining behavioral science and recommender systems algorithms. Our approach to tailoring using recommender systems followed the idea proposed by Fernandez-Luque et al. [144]. With our HRS, we strived to make the provided contents as personalized and relevant to the needs and interests of the smokers as possible, which was also most important for the smokers using a smoking cessation app [422]. However, we integrated the determinants and phases of the behavior change as part of a recommender system algorithm for the specific case of smoking cessation. We used the I-Change model [78], as it was previously effective in smoking cessation [79]; some of the I-Change model determinants (attitudes, social influence, self-efficacy, action planning, and skills) were considered to model user profiles. The system could factor these I-Change model determinants of change as user profile variables to calculate how similar two users were in the demographic filtering step for sending a recommendation. Therefore, the HRS computed the similarity between two users using demographic elements (e.g., gender and age), smoking-related variables (e.g., nicotine dependence), and also variables related to the determinants of change (user perceptions of success in quitting smoking).

The second main finding was the design of the procedure of generating behavioral change-tailored messages used by the HRS algorithm while covering all the included determinants of change. In the message creation process, content writers manually checked 58 different user profile characteristics (the meta-features), including demographic data, smoking habits, and the included behavioral change determinants. Based on this assessment, they tagged the messages in the system database. Then, they ensured that all messages were addressed, that all determinants of change included at least one message, and that each message had different variations to tailor the different options of the users' demographic profiles, which they can complete while registering. In this new message design process, it was not required to create if-then-else feedback tables exhaustively, as previously done in the traditional message tailoring approach [112], because those relations are created during the execution of the algorithm.

Dijkstra and de Vries [423] published an article on their method to develop traditional computer-tailored interventions. That article was similar to the followed approach in Chapter 3 as both explained the steps to design a behavioral change intervention using computers. However, they differed in terms of the paradigm of the algorithms used. The researchers also used the example of smoking cessation, making their approach even easier to compare to the one in Chapter 3. They both coincided in crucial elements such as message design, the usage of psycho-social determinants of change, the implementation of a questionnaire to learn from the users, and tailoring strategies, including message adaptation. However, the nature of each tailoring method (traditional tailoring use tailoring matrices implementing the if-then-else rules, whereas HRSs use AI to generate the recommendations) made the tailoring methods differ in terms of how each method was explained and described. Besides, the term "feedback" differs among studies, which leads to possible confusion. For example, Dijkstra and de Vries [423] used this term for the answers provided to the user from the system. However, in HRSs, feedback refers to the information the user provided to the system about the relevance of the recommendations [424], as we did in Chapter 3. To avoid future misunderstandings, it is necessary to ensure no ambiguity around the term feedback. It could be simply done by differentiating between the user feedback (the ratings) and the system feedback (the messages).

The third main finding was the set of user interface improvements incorporated into the system. These improvements were based on the usage data of a similar previous system; they aimed to facilitate the participant's access to and interaction with message recommendations. They included a higher feedback granularity level than a previous version where we performed a usability study (offering users the option to rate the messages on a scale of 1–5, instead of 1–3) and facilitated clear sections for a better experience during the system usage. Our solution was built based on a previous system [239] that used, on a superficial level, the transtheoretical model for behavioral change [266] in order to set a message delivery frequency. This previous system's HRS included a message feedback method with only three options (positive, negative, and neutral feedback). However, the new system described in Chapter 3 included five options for feedback (1 to 5 stars). This difference was intended to spread out the users' opinions more, as the feedback received when there were only three options were found to be too homogeneous (over 95% of the feedback was positive). Consequently, it did not help the HRS discriminate good recommendations from the excellent ones.

The fourth main result was that this study used a new comprehensive taxonomy presented in Chapter 2 for the HRS. Other studies also revealed the implementation of behavioral theories in their design. One of those studies was by Baskerville et al. [425], who presented an app for smoking cessation in which the principles of persuasive technology for behavior change [426] were used. Another study was by Bindoff et al. [427], who described the creation of a mobile video game for smoking cessation and considered the need to use persuasive game design [428]; however, any specific behavior change model was not explicitly mentioned. In each of these cases, they did not provide clear step-by-step guidance for replicating their approach integrating behavioral change principles, as we did in Chapter 3. The IDEAS framework for developing more effective digital interventions to change health behavior [429] included such a step-by-step guide. It was a 10-step guide that highlighted the need to involve a multidisciplinary team in the designing process. In retrospect, we used the same steps for our HRS. The IDEAS framework covered the whole process—from ideation to dissemination—at a high level without providing specific recommendations as we had provided in Chapter 3.

In summary, in Chapter 3, we described a process of how a behavioral change model can be integrated with a hybrid HRS using as the example of a mHealth solution for smoking cessation (**RQ2**). We demonstrated the feasibility of such integration and provided a detailed description of its ideation and development processes, including a comprehensive multidisciplinary categorization of the resulting system.

Chapters 4 and 5: Effectiveness of our HRS

Chapter 4 provided a detailed description of the intervention protocol we had conducted to test our HRS. The protocol facilitated pilots among smokers from a clinical setting, as well as from the general public. The public pilot consisted of a trial to evaluate whether the new HRS with collective intelligence presented in Chapter 3 could outperform a simpler version of the same HRS without one by comparing both technical and healthcare outcomes for smoking cessation. These results were detailed in Chapter 5, describing the outcomes of the experimental study through a post-test after six months of comparing both HRSs 1) message appreciation, 2) user engagement, and 3) smoking abstinence rates.

Message appreciation

The first main result was that the HRS using collective intelligence grounded in behavioral science, delivered messages that were appreciated as much as those delivered by the system without collective intelligence. By analyzing the mean message appreciation evolution, we identified that the appreciation generated by the HRS with collective intelligence was consistently higher in all time intervals but the first one. Yet, the message appreciation for the two systems was not statistically significantly different. This lack of significance may suggest that the HRS using collective intelligence took advantage of its demographic filtering step over time. However, it did not suffice to make participants perceive the messages as significantly more relevant for them because it needed more time to generate even higher relevant recommendations due to the cold-start effect [359]. In addition, the low power of our study could have conditioned not finding these results as statistically significant. However, in general, the messages provided by both systems were appreciated better compared to other similar studies. For instance, a computer-tailored intervention for smoking cessation studied by Stanczyk et al. [111] produced an overall appreciation score of 3.31 out of 5. At the same time, Quiñonez et al. [430] reported a score of 3.19 out of 5. The only smoking cessation study using an HRS presenting appreciation results was the SoLoMo [163], which showed an average appreciation score of 4.01 out of 5 [164].

User engagement

The second main result was that the HRS with collective intelligence led to overall similar or even worse engagement outcomes. The HRS without collective intelligence generated a higher number of active days in the mobile app and a higher number of abstinence reports. In the graphical representation of the number of rated messages across time, we identified that the participants rated more messages generated by HRS with collective intelligence than the one with the knowledge-based step only. However, the difference in the number of rated messages was not statistically significant. The same explanation of the cold-start problem and low power in our design presented for the message appreciation could also explain this result.

In addition, no statistically significant differences were found for the overall number of quitting attempts. Yet, participants who completed the extended profile made more quitting attempts using the HRS with collective intelligence than those who also completed it but received messages of the HRS without collective intelligence. Thus, completing the extended profile gave the HRS with collective intelligence more information to personalize the messages. These results add to the theory that higher personalization may lead to higher engagement, as shown by previous studies [431-434].

Although participants who completed their extended profile had significantly less probability of dropping out from the study for all the assessed metrics, this happened for participants using both types of HRSs. As the HRS without collective intelligence did not benefit from having such extended profile information about the participants to personalize the messages further, we could only conclude that this effect was related to their intrinsic motivation of the participants to engage with the system, and probably not to the collective intelligence.

The only other covariate that showed a significant impact on engagement was the employment status: Employed participants receiving messages from the HRS with collective intelligence had fewer active days in the last two time intervals. There is no clear explanation for this effect. However, this lower number of active days in employed participants using the HRS with collective intelligence may be produced by another related socioeconomic variable we did not include in the study. For example, the participants' educational level or income level were associated with different engagement levels in previous studies [369-371]. They showed that, usually, a participant's higher educational level is associated with higher engagement.

The results concerning engagement generated by our system are difficult to be compared with those reported by Stanczyk et al. [111], Quiñonez et al. [430], and in the SoLoMo study [163]. This difficulty is because each intervention was different and proposed a different way to measure the engagement and had different dropout levels. For instance, comparing mobile apps with website-based programs will have to consider different setups for these programs (e.g., three expected 45-minute interactions of a web-based intervention vs. daily one-minute-long interactions for six months with a mobile-based intervention). In our case, we measured engagement by considering the number of active days on the app, the number of rated messages, and the number of abstinence reports. However, other studies used the number of visited sections in a website [178, 435, 436], feature utilization [435], the spent time on the site [435, 436], and the number of visits to the website by a participant [435, 436] and logins, the answers to the system questionnaires, and the number of user messages sent to the system [437]. On the other hand, our data indicate a high positive appreciation evaluation; in our study, many respondents dropped out (92.99–95.15% depending on the variable used to assess the dropout rate), which may have generated biased results. This high dropout rate may have been generated because participants were not invited to the study. Instead, the people who downloaded the mobile app voluntarily were informed that their data would be used in a study and had to fill an online informed consent. Hence, this may have attracted people who were only curious about the app but not really motivated to change their smoking habits. Hence, a thorough usability test and program evaluation, including an in-depth qualitative probing, is necessary to further analyze the level and quality of user appreciation for our HRS system.

Smoking abstinence rate

The third main result was that both versions of the HRS provided support to people to quit smoking. However, the analysis on available data for the 7D-PP abstinence considering only the last available abstinence report showed that the HRS using collective intelligence had a poorer impact on abstinence rates than the one without collective intelligence. However, no significant differences between both versions were found when the 7D-PP abstinence analysis was done averaging the abstinence reports across the time intervals – as in indicator of continuous prevalence-, and in the case of penalized imputation (pessimistic scenario) for both the 7D-PP abstinence considering the last available abstinence report and the 7D-PP abstinence averaging the abstinence reports of the different time intervals. These findings contradicted our initial assumption that the system with collective intelligence would yield better cessation results than the system without it. One potential reason for these weaker benefits might be that the messages that the participants found to be relevant may not always have been the ones they needed to their change their behavior. In this sense, the collective intelligence system might not have sent potentially beneficial messages to a user that were poorly rated by similar other users. However, they might have been beneficial for that particular user. Despite the overall similarity, belief structures may thus still vary among "similar others," which may limit the application of the collaborative filtering approach. If so, this could limit the applicability of collaborative filtering to change particular sets of beliefs and may favor a more personalized tailored approach. One can expect these problems to be more significant during a cold-start period and with a small sample. The collaborative filtering approach needs more research to identify the similarities of belief structures of people considered similar. In the HRS of our study with collective intelligence, when we calculated the similarity between two users, it considered how similar their ratings and profiles (including their belief structures) were. However, averaging these two values was the first approximation to this approach, and the actual weight of each factor should be studied in the future. A second explanation may be that the avoidance of relevant messages may occur if a particular message creates cognitive dissonance [438] in the smokers who are uncomfortable with the message and, consequently, provide a low rating even though the message may have been beneficial. In any case, a limitation to his study is the lack of consistency among the results for smoking cessation for the different analyses, as some present statistically significant differences and others do not. This insufficient significance may be caused for the lack of respondents, resulting in the low power of the study—only 60% instead of the traditional 80%. We had initially planned a sample size of 978 participants to detect significant differences between the control group smoking cessation success rate at 0.07 and the experimental group at 0.14, accepting an alpha risk of 0.05 and a beta risk of 0.2 in a two-sided test. However, we were only able to recruit 371 participants who met the inclusion criteria of our study, implying that our study had a beta risk of 0.4. Thus, we strongly recommended assessing the effects of the new system for a more extended period (for instance, 12 months), as some of the effects of the HRS may only start to become noticeable over a longer period than just six months. Also, it should have a larger sample size to ensure our system has sufficient statistical power.

In the previously mentioned SoLoMo study, a randomized controlled trial was conducted. Participants received extra smoking cessation support from tailored motivational messages generated by a hybrid recommender system (using collaborative filtering). These messages were sent to their mobile phones, in addition to the standard psychopharmacological treatment. The study proved that its HRS combined with the psycho-pharmacological treatment was significantly more effective in supporting the participants to quit smoking for one year than just the psycho-pharmacological treatment. The experimental group had a continued abstinence success rate of 27.5%, while the continued abstinence success rate for the control group was 15.0%, adjusted OR 3.13, P = .002 under the intention to treat analysis [164]. Yet, the smokers had human support from healthcare professionals who enrolled them in the trial and followed up with them throughout the year. This fact could have influenced the participants and, consequently, made them achieve better results [439, 440]; this difference, thus, reduced the comparability of the SoLoMo intervention with our intervention, where smokers voluntarily downloaded the mobile app without any human interaction. Moreover, the inclusion criteria for both these studies were different; the participants had to be willing to quit smoking in the former. In our study, any smoker could have downloaded the app and tried to quit without being committed. Finally, all the SoLoMo participants received free pharmacological medications of Bupropion or Varenicline to increase their chances of abstaining, whereas, in our intervention, we did not provide such support. However, our participants were free to complement their smoking cessation attempts with any other means they wanted (which we did not consider).

Another study using HRSs is the Perspect experiment [162]. It compared a rule-based tailoring system—similar to our knowledge-based system—with a hybrid recommender system that also included collaborative filtering to support smoking cessation. The study participants were followed up for 30 days. It was found out that among those who completed the follow-up, the HRS could increase user ratings but achieved similar results in smoking cessation rates—36% of the participants using the HRS stopped smoking for one day or longer compared to 32% of the participants in the control group. The most similar abstinence metric in our case was the 7D-PP abstinence analysis using the last available abstinence report. For this metric, our systems performed better, as we achieved 54.1% and 30.2% abstinence rates for the HRS without and with collective intelligence, respectively. Our appreciation findings were also in line with the Perspect experiment, as the appreciation of the messages (ratings) increased in both studies. However, no statistical significance was found in our case.

Our abstinence rates (11% for the system without collective intelligence and 8.4% for the one with collecting intelligence, under the penalized imputation analysis) were similar to the other smoking cessation studies of comparable lengths and contexts that did not use HRSs [85, 382]. These rates position our approach of sending multiple and shorter messages using an HRS in the same success range as the previous studies that used fewer (1–3) longer tailored letters. Further studies are needed to assess which method is preferred by which smokers. However, our outcomes were lower than those of the study by Stanczyk et al. [111]. In this study, the computer-tailored text-based intervention achieved a total of 17.9% and 22.6% of six-month seven-day point prevalence. However, higher abstinence levels were achieved with video-based interventions (between 20.9% and 30.6%). Hence, our HRS could be improved by delivering messages in videos instead of text messages.

In conclusion, we found in Chapters 4 and 5 that an HRS grounded in behavioral science using collective intelligence generated mixed results despite its theoretical potential for personalizing and finding relevant messages. For some metrics, the HRS with collective intelligence performed worse than one without it (number of active days, number of abstinence reports, 7D-PP abstinence considering the last abstinence report). However, both HRS produced similar results for appreciation, number of rated messages, number of quitting attempts, 7D-PP averaged for all time intervals, and 7D-PP abstinence under a pessimistic scenario. Only when the collective intelligence had the extended profile information to personalize the recommendations was it able to provide better results in quitting attempts. Given these results, more in-depth research is needed to study how to apply the collective intelligence in HRS for improving these outcomes compared to other approaches.

Strengths

Theoretical strengths

This thesis has several main strengths. Firstly, we integrated behavior change factors in traditional HRSs, which created a new categorization system. This integration is a strength because there was a lack of information about applying behavioral theories in HRSs in the existing literature [166]. With this thesis, we designed and analyzed an HRSs that did include and fully detailed the behavioral change factors integration and provided a way to report such interventions systematically and efficiently in the future.

Secondly, we used HRSs in behavioral change interventions, which is still relatively novel, and limited evidence of these interventions could be found [209][243].

Thirdly, we experimentally compared the addition of collective intelligence to an HRS to demonstrate its added value. The results derived from this comparison provide a better understanding of how HRSs with and without collective intelligence perform. This additional knowledge will help choose the most suitable algorithmic approach in future interventions.

Fourthly, we measured engagement and smoking abstinence to judge the effects of the HRS-based intervention. These parameters related to the participants' health outcomes are not usually assessed and presented in the literature of HRS in favor of other technical measurements [243]. These results allow comparing the health outcomes of HRS for smoking cessation with other types of interventions such as traditional computer tailoring.

Fifthly, throughout this thesis, we provided a multidisciplinary perspective. Researchers with a computer science background and those with a behavioral science background could hopefully understand and benefit from the generated outcomes and the ideas suggested by combining advanced HRS technology with behavioral change technology. This multidisciplinary approach bridges the gap between the academic fields of behavioral science and artificial intelligence, whose combination may improve the outcomes of digital health interventions [441, 442].

Another strength of this thesis is that it fosters replicability and reusability for future studies. It is achieved because the outcomes of this thesis can be used as solid foundations to grow more evidence about HRSs when applied to computer tailoring in the context of smoking cessation. However, it could also be extended to other therapeutic areas since we provided a detailed description of the system (Chapter 3), which was identified as one of the primary needs in the HRS field, as identified in Chapter 2.

Methodological strengths

The design used for the intervention in Chapters 4 and 5 contributed to the external validity of the study. The intervention was done under real-life conditions for the autonomous eHealth solution. It was supposed to be used on its own without any other human interaction. Therefore, no researchers were actively involved in the recruitment of smokers, and smokers were not contacted by the researchers who would encourage them to report their smoking cessation status. In addition, this made that the measure of our results was for effectiveness (assess performance under real-life conditions) instead of efficacy (assess performance under ideal controlled conditions). It provided extra value in understanding how such a solution would benefit smokers if it were to be deployed for their regular use outside of a study.

The methodology followed across chapters 2 to 5 is another strength of this thesis. We matched the three steps to be considered when developing HRSs proposed in the HRS creation framework by Valdez et al. [213]: Understanding the domain, Inception, and Evaluation. Further, we found that the framework lacked an essential element—message generation. This message generation information should be included at the same level as the algorithm design in the framework. Thus, this thesis highlights the value of message content creation. However, this message generation process may not always be necessary in all cases as some HRSs could be used to recommend elements from an existing health database. However, it was a significant component in our case and, thus, should be considered separately as most behavioral change interventions will need to go through this process.

Limitations

In this section, we present the main limitations of this thesis by grouping them into two categories—methodological and technical limitations.

Methodological limitations

Although the scoping review in Chapter 2 aimed to cover the main journal databases, more studies could have been considered by expanding the acceptance criteria with the inclusion of conference proceedings and gray literature. By including these other types of publications, we could have discovered more studies that have used HRSs, which could have contributed to acquiring a more comprehensive overview of the field. In particular, the inclusion of prestigious technical conferences such as ACM RecSys and the IEEE could have increased the coverage of publications.

In Chapter 3, we designed our system by improving the user interface based on the analysis of a previous similar system. However, we should have performed a usability study on our system after implementing these identified improvements.

Regarding the trials in Chapters 4 and 5, the participants who joined the study in the first months of the intervention may have received less relevant messages due to the cold-start problem; this was not considered in our analysis. As a result, these messages may have had a weaker impact on the participants and their measured behavior than those of the cohort who received better-tailored messages.

As the participants did not have a deadline for completing the seven-day point abstinence self-reported smoking cessation questionnaires, it was impossible to set specific timepoints as initially intended in Chapter 4. Instead, we opted to group their answers according to increasing time intervals—0–7 days, 8–14 days, 15–21 days, 22–30 days, 31–60 days, 61–120 days, and 121–180 days. These time frames limited the understanding of how participants behaved, especially in larger time intervals, since only the last report was considered within each time interval. The appreciation and engagement metrics were also assessed following these time intervals to maintain consistency in the analysis results.

In Chapter 5, we did not ask whether smokers were undergoing any other type of smoking cessation treatment (e.g., nicotine replacement therapy) during our analysis. Moreover, the smoking cessation results found in Chapter 5 could not be checked against biochemical validation, and as such, we could not assure their veracity. However, previous studies that used this self-reported method [390] found them to be valid [389], and others considered biochemical validation in studies with limited face-to-face follow-ups not necessary [443] as our case that had none. Next, we did not perform a thorough qualitative and quantitative process evaluation to differentiate between the smokers who liked and benefited from our system and those who did not. Finally, the power of the design was modest, thus limiting possibilities to identify differences between conditions.

Theoretical limitations

Despite the motivational messages created in Chapter 3 and used for the study of Chapters 4 and 5 included behavioral change techniques and were validated by healthcare professionals specialized in smoking cessation, they were neither validated by health communication experts nor by potential participants before the interventions, as other studies did [444, 445]. This lack of validation may have reduced the overall impact of the smoking abstinence messages due to reduced content validity, reducing the effect size [446]. It also implies that the messages sent to the participants may have been of different quality, and the potential impact of this different quality in the measured outcomes was not considered in our study.

Technical limitations

We did not aim to optimize the algorithm computing time to reduce complexity in the development process. This lack of optimization was because it would not affect the results due to the design of the intervention (where all computations were calculated overnight and sent throughout the day) and the expected number of participants in the pilot.

The motivational messages were sent as push notifications to the participants' smartphones. However, each phone might handle notifications differently. For instance, some devices might always show them and report it with a sound, some might notify it when the app is open, and some do not alert them. This variability, which may affect the participant's interaction without a system, was not considered in our analysis as it was impossible to control.

Of the 796 days of the pilot, for 22 days, the servers of the Taipei Medical University department were not available due to a technical virus infection. During this time, motivational messages could not be sent, and the participants were denied access to the app. No specific action was taken to contact the affected users as we did not have their consent to send them any notification.

It was impossible to use the physical activity data in our study as initially planned in the protocol described in Chapter 4. It was due to the differences in the data format for the collection of physical activity (in Android devices, it was the total active time, whereas, in iOS devices, it was the number of steps). Furthermore, the continuous updates on the Google Fit APIs—the intermediary software that provides access to the functionality of another software—to collect the physical activity information made the implemented system invalid without devoting resources for updates, which we did not have. Finally, the users had the option to deny our app access to their physical activity data to protect their privacy, and many of them did it.

Implications and recommendations for future research

The need for full disclosure of the design and working mechanisms of HSRs using a multidisciplinary taxonomy

Understanding how recommender systems are designed is crucial for their reuse and improvement. However, it is nearly impossible to replicate the interventions with the limited information usually disclosed in the published studies. Therefore, HRS-based interventions should start publishing their design, technical, and health-related outcomes. The specific metrics to be assessed will depend on the goals of each study. However, the percentage of patients who met the criteria for fulfilling the objectives of the study can be measured in most studies, thus, become a promising homogeneous indicator for evaluating HRS studies related to different therapeutic areas. Moreover, patient-centered outcomes, such as the perceived quality (e.g., appreciation, usefulness, trust), which other studies have used before [221, 222, 232], can be considered suitable for HRS cross-comparison.

If health-specific metrics are not used in the assessment of HRSs, a pure technical evaluation of HRSs may generate misleading results about their performance. For example, a user may negatively rate a health message that, despite being beneficial for their behavioral change, create cognitive dissonance [438], consequently, is disliked by the user. In such cases, the technical metrics that value the number of positively rated messages would consider the HRS is underperforming when the goal of the system—fostering behavioral change—is being achieved. The relation between what patients find appealing and relevant, and what is actually beneficial for them should be studied in detail in the future.

We recommend that researchers use the HRS taxonomy proposed in Chapter 2 to remember reporting all the relevant details (e.g., the usage of tailoring elements) about the design and evaluation of HRSs. Moreover, it will help them homogenize their description, which will allow for future comparisons and reviews. Finally, in combination with the HRS development framework by Valdez et al. [213], our taxonomy will facilitate the advancement in the usage of HRS as the field of HRS is not consolidated yet [209, 213].

Grounding HRSs in behavioral science

The usage of behavioral science, in combination with the new approach of computer tailoring using HRSs, may have the potential to increase the possibility of tailored eHealth interventions to succeed in bringing about the desired behavioral change in participants. This increase can occur because HRS will take advantage of previous evidence about how individuals can foster behavioral change concerning their current health condition [402, 403]. However, the disadvantage of this combination is that the intervention design becomes increasingly complex, requiring its designers to understand the algorithmic and behavioral aspects.

Although we used the I-Change model to ground our HRS in behavioral science, other behavioral models could have been used, such as the health belief model [447], protection motivation theory [448], and the social cognitive theory [410], self-determination theory [449]. However, whether these approaches will be relevant and useful will depend on the type of intervention required for each therapeutic area, the desired behavioral change that needs to be attained, and the effects achieved by interventions using different models. To decide which theory should be used, the researchers should analyze previous interventions that targeted the desired behavioral change and identify which theories were applied in those interventions to reuse them in the present study as a part of the HRS. For example, they could use principles of intervention mapping [450, 451] as a good planning strategy to design them. In addition, the HRS design also needs to be carefully aligned with the previously found evidence in that field. For instance, short interventions (e.g., one week) with limited contact with the patients (e.g., three times) may require a different HRS design because our system may not learn enough from the participants to provide them with accurate recommendations.

Optimizing the usage of collective intelligence grounded in behavioral science

Appreciation

The results related to appreciation indicated that our approach using HRSs provided high appreciation levels compared to other studies [111, 430]. Moreover, using our two HRSs, we found that the HRS with collective intelligence grounded in behavioral science improved the appreciation for the recommended messages compared to the one without it. Although this difference was not statistically significant, we could identify from the graphic representation of its evolution across time (Figure 2 in Chapter 5) that the differences were consistently higher as time passed. The difference could eventually reach a statistically significant difference if it keeps the same trend over time. Also, the appreciation evolution revealed that the HRS without collective intelligence achieved similar appreciation results at the end of the intervention than the ones achieved during the first days. In contrast, the HRS with collective intelligence achieved higher appreciation levels at the end of the intervention. This result was expected as the hybrid HRS was able to learn from users using collective intelligence. However, the growth was not enough to be detected as statistically significant compared to the ones generated by the knowledge-based HRS. This lack of significance may have been caused because the hybrid HRS started to send more appreciated messages only when users had already rated several other messages - just when the hybrid HRS could take advantage of the collaborative filtering potential. This delayed effect seems to be associated with the cold-start problem that HRSs are known to suffer. However, there are different approximations to solve it. [153, 452-454]. Solving this would reduce the time needed to provide relevant recommendations and increase the appreciation levels. However, it is not certain whether these approaches would contribute to our case and, thus, requires testing and validation.

Two other approaches could be followed to prevent this problem from arising in the future, as presented by Ilarri et al. [455]. One way could be to pre-generate a database with users outside the system to minimize the cold-start problem. For this process, a group of smokers would rate the messages in the dataset for their perceived usefulness. This approach is not ideal since the context in which the raters are not providing feedback is not the same as that in a smoking cessation process, potentially hampering the reliability of the ratings. This approach could work well for pre-rating other types of items where the recommendations do not depend on the context (today's rating will likely be valid tomorrow). This situation may not happen in the smoking cessation process because people experience different smoking-related episodes over time (e.g., cravings), and their perceptions and relationship with tobacco are likely to evolve. However, this could be a feasible approximation that should be considered and studied in the future. The other alternative could be using synthetic data generators [456] to populate the dataset. In this way, an algorithm generates fictitious user data following a probability distribution function to create the necessary users and ratings to reduce the initial lack of information in the system. However, no known generators can replicate the behaviors of smokers. More research to design these generators would be highly significant to the future of AI applied to health as they could be used to benchmark different algorithms. In 2018, the World Health Organization and the International Telecommunication Union created a working group to develop a benchmarking process for AI for health [457]. The creation of this group evidenced the current high interest in expanding the datasets to support the improvement of health AI algorithm evaluations. To foster reusability of the collected data, future HRSs may need to consider following the FAIR (findability, accessibility, interoperability, and reusability) principles [458], and more specifically, the FAIR-Health [459]. This data can then be used for benchmarking and as a sandbox for new research. This type of data would have made the testing of our HRS possible. However, the existing datasets for recommender system benchmarking such as RED [460], MovieLens [155], EachMovie [461], Netflix Price Dataset [462], and Million Song Challenge dataset [463] were not compatible with our HRS design.

Engagement

The collective intelligence (the demographic filtering step) grounded in behavioral science did not increase engagement metrics. We expected that using collective intelligence could provide a higher level of personalization of the messages, making them more relevant and, consequently, engaging for the smokers. However, no statistically significant difference was found in appreciation, which may have led to an insufficient effect in impacting the engagement of the intervention group. Only by considering the completion of the additional information on the user profile as a covariate, we found a significant difference in the engagement metrics favoring the HRS with collective intelligence over the basic knowledge-based one. The messages generated by the HRS with collective intelligence made users have more smoking cessation attempts. However, this does not explain the unexpected results achieved for the number of active days and the number of abstinence reports. These results were theoretically contradictory as we could expect, at least, the same engagement between the two HRS. By design, the collaborative intelligence could not perform worse than the random selection. Therefore, these could only be explained by other reasons, such as using statistical models that did not accurately fit the data.

These engagement findings suggest that demographic filtering in computer-tailored interventions using HRSs can increase some engagement metrics. However, it may not apply to all consistently. Users may need to provide extensive information about themselves with a questionnaire to correctly calculate similarities among users and select relevant messages for each user. This finding contradicted the expected potential benefit of HRSs of not needing to depend on lengthy initial questionnaires like traditional computer tailoring. Thus, we recommend that future HRSs using collective intelligence be designed to enable users to provide extensive information about themselves. We believe that user profile completion can be maximized using techniques such as dividing the questions and spreading their completion across a few days instead of requiring their completion on the first day. In this way, users would not feel intimidated or discouraged from using the app by having to complete a lengthy profile questionnaire at once. The facilitation of patient profile completion would also be possible by combining this progressive user profile completion approach with ecological momentary assessment (EMA) strategies [464, 465] to map the context of the users without requiring a lengthy dedication in one go. In addition, the smoking information may benefit from being updated as smokers evolve [466, 467], and with the EMAs, we could conveniently detect such changes. In combination with the self-learning recommender system technology, it could improve the already existing capacity of recommender systems to adapt to the sub-populations of users with specific needs and characteristics over time [324]. In addition to completing user profiles by reducing the entry barrier, the EMAs could periodically update the user profile and always keep an accurate user model.

In any case, we suggest that the best approach is to reduce the unnecessary user information during the onboarding by syncing the system to the existing trusted databases, such as electronic health records, which was highlighted in the proposed HRS taxonomy (Chapter 2). However, this is not always possible due to the design and nature of the interventions, technical reasons, or even legal and data protection reasons. Nonetheless, there were successful cases such as the SoLoMo study [163], which retrieved all the necessary details from the EHR about the patients—collected by healthcare professionals— to minimize the questions in the mobile app and, consequently, the burden on the patients. Another approach to reducing the burden of completing the questionnaires could be using automatic competition techniques such as input masks or autosuggestion [468]. However, we could not use them in our case because our participant users needed to select their answers from a set of pre-defined answers.

Smoking cessation

The levels of 7D-PP abstinence averaging the results of the different time intervals on available data (as an indicator of continuous prevalence), and the 7D-PP penalized imputation analysis (for both the analysis on the averaged results of the time intervals and using the last abstinence report) achieved by the HRS using collective intelligence grounded in behavioral science did not improve the abstinence levels achieved by the knowledge-based HRS. In contrast, the levels of 7D-PP abstinence, considering only the last available abstinence report for the participants benefiting from collective intelligence, deteriorated. These results contradicted our initial assumption that the combination of the collaborative filtering step, considering the principles of the I-Change model, would generate better recommendations that would induce smokers to quit. We believe that HRSs can be used as a potential means to increase the effects of behavioral and pharmacological treatments. We base this thought on the modest abstinence rates we found in our intervention and on comparing the rates achieved in another very successful intervention for smoking cessation, which also used an HRS [164]. However, given the complexity of collaborative filtering for behavioral change, more research is needed to estimate whether and how eHealth can profit from the HRS technology. We propose that future research should include an analysis of the similarity calculation among users. Such an understanding may help relevance of participants' belief structures in the user similarity computation. This analysis will facilitate a better understanding of how to weigh belief structures, demographic data, and message ratings when designing collective intelligence systems that optimize user matching using only those aspects relevant for bringing about the desired behavior change. This approach contrasts with including aspects that—although making users similar in other contexts—do not help relate users in terms of behavioral change.

Another element to assess is whether these types of eHealth interventions can be used as a standalone or if they should be done in a more blended strategy [469-472], involving counseling to cause a significant positive impact on smoking cessation outcomes. In this case, questions such as how the integration in the real clinical practice should be done and how healthcare professionals can use and benefit from the system to support their patients remain unanswered.

Future studies should overcome the potential deficiencies in our research that may have contributed to the lack of statistical significance in our results for smoking cessation. These deficiencies may include the following: First, the design to assess the smoking cessation may have been suboptimal and, consequently, negatively affected the abstinence analysis. In our study, the participants using both versions of the HRS could answer the seven-day point prevalence reports for smoking cessation at any time after they were made available. Thus, we had to consider the time ranges for grouping and studying the reports since we did not have a homogenous measurement time point. This grouping might have reduced the precision of our smoking abstinence measurements between the two groups and limited comparison accuracy with other studies. Thus, our first recommendation related to smoking cessation outcomes for future HRS-based interventions is to allow smokers to only report their abstinence status on specific dates. Although this may not directly improve the actual outcomes, it will improve their validity as the accuracy of the chosen smoking cessation metric (e.g., seven-day point prevalence) will be higher than the ones in our approach (grouping the reports in periods). However, this approach may reduce the number of smokers reporting their abstinence status as not all smokers may answer on those specific days. We suggest emphasizing the importance of reporting by sending reminders and notifications to the participants.

Another possibility that might explain the achieved abstinence results was that our study did not consider message design as a personalization feature. The messages were designed to be short (a maximum of 200 words), easy-to-understand (using active voice), and clear (removing any distracting or irrelevant information), repeating an answer, creating empathy, adding new knowledge, and changing the existing misconceptions, and including 10 of the 11 behavioral change techniques proposed by Abraham and Kools [278]. However, despite following good practices for healthy communication, the messages were not designed by communication experts. Moreover, none of the factors related to how the messages were communicated (e.g., length, complexity, source, healthy communication techniques, proposed call to action, structure, formality, and included examples) were considered meta-features to compute the selection. For instance, our HRS did not assess whether one smoker would benefit from shorter messages as compared to other longer messages describing the same concept. To consider these communication factors, smokers would have received the same, but differently framed, conceptual message (e.g., one concept in a shorter and succinct message and the same concept in a longer and more elaborated message). However, the smokers' feedback on these messages would not necessarily reflect their preferences. For example, the second message dealing with the same concept might be rated with a low score as it was perceived to be repetitive and irrelevant—although phrased differently—since they had already received the first message. To avoid this, our second recommendation to improve the smoking cessation outcomes in future HRS-based interventions consists of identifying the smokers' preferred style of communication (e.g., short and direct messages, elaborated messages, video messages, or other means). To understand their preference, they could be asked about it in the initial assessment. However, the participants may not know what style they prefer or what would benefit them the most, even if it initially might not have been the most appealing. An approach to solve the problem of smokers not knowing what communication style fits them the best would require a large user dataset. In this alternative, the system could learn the message preferences with AI techniques to check the smokers' feedback for the messages with different communication factors and relate those factors and their ratings to their profiles. In this way, we can identify the smokers who have a specific user profile and determine whether, for instance, they are more likely to benefit from really elaborated messages than their peers who have a different user profile.

Therefore, more studies are needed to assess whether communication factors could be considered relevant variables (and how to optimize them) for the HRS in influencing smoking cessation behaviors.

The third research direction that could increase the chance of future HRSs to improve smoking cessation outcomes is to consider the participants' behaviors as a part of the HRS algorithm. As theorized before, it is possible that the HRS using collective intelligence, excelled in recommending what smokers found pleasant and appreciated more instead of recommending what they needed for behavioral change. To solve this issue, the actual behavior status of the smoker should be inputted into the HRS algorithm. For example, suppose a smoker has rated a few messages as highly relevant because they liked the message content (e.g., found it funny, or it did not create cognitive dissonance) but ends up relapsing. In that case, it must be inferred that although those messages may be pleasant, they were not really useful for behavior change. As a result, the weight of their actual rating should be reduced while computing what message should be sent to other smokers in the same profile. Likewise, if some messages were rated negatively, but the smokers receiving them remain abstinent, then the system should internally rank these messages higher when sending them to other similar smokers in the future. In this way, we may overcome the potential problem of the collective intelligence recommending messages that are liked by the users but not impactful and relevant messages to achieve better health outcomes. However, it may be challenging to know the smokers' behavior. For example, we can ask users frequently about their abstinence status or even inquire about the evolution of their determinants of change status, such as self-efficacy issues and barriers to increase or decrease the internal score of the messages they had previously received. However, they may find such inquisitions bothersome and inconvenient.

A complementary approach to using the data pertaining to the actual behavior for gathering feedback for the HRS is utilizing the smokers' emotions regarding the messages received by them. This approach would enable the HRS to differentiate poorly rated messages, since they are not relevant for the smokers, from the messages that smokers do not like as they create cognitive dissonance, despite being potentially useful. Instead of using the traditional single-dimensional feedback measurement (e.g., rating the relevance of the message using 1 to 5 stars), this multidimensional feedback measurement would add more meaning to the user feedback. For instance, the study by Mizgajski and Morzi [473] explored the usage of emoticons for providing feedback in recommender systems, including options such as "informative," "surprising," "sad," "inspiring," "annoying," and "irrelevant." Moreover, Facebook has already introduced the usage of emoticons as a feedback system instead of the traditional likes or stars to give users more personalized and engaging content. Using such an approach, the users can consider the messages that create cognitive dissonance as annoying or informative instead of irrelevant, supporting the HRS with the help of future recommendations.

Using new ways to deliver HRS-based behavioral change interventions

In Chapter 3, we detailed one way of adapting an HRS to smartphones. Although we used the usage data of a previous similar mobile app to ideate design changes to improve the usability of our app, we did not perform a usability study of the resulting app. Thus, we suggest conducting qualitative and quantitative usability testing of the system before using it; this testing would involve key stakeholders (e.g., smokers and behavioral change experts). In any case, our detailed adaptation may provide new researchers with the design and the usage of new mobile-based HRS technology for computer tailoring. These two elements are relevant as there is a general lack of studies that divulge how HRSs are designed, as shown in the scoping review in Chapter 2. In addition, the provided information may support the usage of large-scale tailored interventions using smartphones as, through the years, the smartphone-related barriers such as costs and accessibility, as identified by Abroms et al. [474], have been almost entirely reduced. Further, we evolved the Libre de humos app [164] used in the SoLoMo study [475]. The mobile app interface changes were made to make the app non-dependent on electronic health records—it, otherwise, could not have been used in our study—and simplify user experience to facilitate usage.

However, the aesthetics of the mobile solution may be directly related to the users' involvement, as explained in the path model of the user engagement scale [476]. However, we do not know to what extent the aspects of the visual appearance of the mobile app, including fonts, colors, proportions, and distributions, could influence the usage of our system as each person may find one style more appealing than others. Therefore, we recommend studying how the user interface graphical elements may influence the usage of the system, which was not done in our study. Using adaptive user interfaces [477] by changing the app background colors, depending on the user's preference, or considering what app section is displayed first, based on what section is used more often, could be another layer of personalization that influences the outcomes of the intervention. These interface changes should be evaluated in the future as the application designed in Chapter 3 had the same type of user interface in our study.

We recommend that researchers keep using mobile apps to deliver HRS-based behavioral change interventions as they may be convenient and accessible. However, researchers should keep exploring how user interfaces can be adjusted to improve user experience, engagement, and, ultimately, health outcomes. However, the variables to adjust the interface can go beyond the screen; non-tangible interfaces such as voice assistant speakers (e.g., Google Nest, Amazon Alexa, and others) [478] could also be used to deliver recommendations. The researchers could compare the results of such interventions with the ones of interventions through the physical screens of computers and smartphones.

Use new data sources for HRSs

There are other approaches to input data in HRSs that should be explored in the future. For example, the wearables monitoring variables such as physical activity, sleep quality, and heart rate could feed meta-features and generate messages related to them. Particularly, the sleep quality is relevant in case the participants are following a pharmacological treatment that has insomnia as a side effect [479, 480]. Physical activity levels can motivate smokers to maintain an active life to prevent potential weight gain related to smoking cessation [481, 482]. Some studies are even considering the usage of wearables to detect smoking activity [483-487]. They can be used as input data for HRSs and as an additional validation element to the self-reported abstinence reports, where biochemical validation is impossible. Therefore, we suggest future researchers assess the potential of biometric data from wearable devices pertaining to variables such as sleep quality or the number of daily steps as a part of the meta-features in HRSs and study their impact on the selection of recommendations.

The usage of implicit feedback (i.e., other elements that are not explicit ratings for learning user preferences) in recommender systems has been explored in the previously discussed studies [424, 488, 489]. However, it is not known whether using it in HRSs can improve the patients' health outcomes. A simple example is if the app has a section with content related to different topics and the smoker accesses that section to read it. Then an HRS can infer that the smoker finds that subject relevant and can factor this information in the message selection process, as done in the SoLoMo study [239]. Another relevant element that can be input in an HRS as implicit feedback could be the location of the smoker. For example, recent studies are using geofencing technology to prevent risks of substance abuse [490]. In our case, if an HRS can locate the smoker in areas where people are likely to relapse more (e.g., football stadiums, restaurants with terraces), the system could send a just-in-time motivational message tailored to provide cues to resist the temptation to smoke in that specific situation by providing an ecological momentary intervention [491-493]. Therefore, we recommend that future researchers explore the usage of implicit feedback and user location contextualization in healthcare collaborative filtering.

Finding alternatives to the message-creation process

The algorithm design process described in Chapter 3 involves an extensive handmade process of writing motivational messages and correctly relating them to the meta-features. This method was tedious and time-consuming. It will be required to be repeated for future interventions unless other design strategies that provide better results are formulated. Although the initial message generation requires human experts in behavioral change to design them, meta-feature categorization of messages could be done using natural language processing semantic analysis [494]. This approach, however, does not seem helpful in our proposed design as designers need to consider the meta-features of the messages in the message generation process. This process is done to ensure comprehensive coverage of the different user profile cases.

Consequently, with the currently available knowledge and technology, the usage of AI to support behavioral science will still require human involvement to, at least, generate the content and the design of the system. Considering the existing gap in reaching that level of automatized recommendation technology, we suggest studying whether categorization techniques (e.g., clustering) could accurately classify the messages to avoid the manual meta-feature relation process. This classification would reduce the time required by message designers and would help them identify clusters that need to have more messages. However, with the emergence of chatbots and their fast development and optimization for coaching [495-501], this problem may not be significant in the future. Suppose HRSs are combined with chatbots to effectively communicate with the users (the HRSs select what should be sent as answers to the user). In that case, their application in generating messages could be more accessible using an algorithmic agent. Ensuring the safety and effectiveness of such messages was discussed in the first review of the studies in the field of psychology with chatbots [502]. The results show that, though promising, the chatbot field is still young, and there is a lack of high-quality evidence ensuring their effectiveness, sustainability, and safety. The potential for health behavior change exists as chatbots can also facilitate new ways of providing feedback to an HRS. For example, it would allow the sending of free texts (or even free speech), considering and allowing the sentiment analysis of these inputs [503-505], which can be used to enrich the decision-making process of an HRS. Consequently, we also encourage researchers to study the usage of HRSs to power chatbots, as the chatbot research field is quickly evolving.

Risks and reflections for practice

Potential risks of using HRSs

The usage of HRSs is not exempt from potential risks and limitations. One of the most common vulnerabilities in recommender systems using collective intelligence is profile injection attacks [506]. Such attacks are produced by ill-intentioned users who provide biased and inaccurate ratings and profiles to alter the behavior of the system. These attacks lead to the system learning incorrectly and, as a result, providing inaccurate recommendations to users. There are solutions to protect recommender systems and minimize the risks of such malicious attacks [507, 508]. However, for our HRS, we did not include any protection measurements against injection attacks. We conducted a risk analysis and concluded a low probability of that happening and a low impact on the outcomes if it did. The two-step design of our hybrid system prevented these attacks have a high impact on the outcome. It was achieved with the first step of the algorithm. It filtered incompatible messages, and as a result of that initial exclusion, even the less relevant messages potentially sent to the user after the second step of demographic filtering would be valid and applicable to the smoker.

When recommender systems take advantage of an established social trust network among users, they are referred to as trust-aware recommender systems [509, 510]. Trust-aware recommender systems can solve the problem of profile injection attacks and bring even more potential benefits to the users [511]. For instance, when a user's friend manages to stop smoking using the system, the user may consider that the messages the friend rated positively are more helpful than those rated positively by unknown others. Further, users' trust in the system is the key for continuance intentions [512], which potentially benefits the engagement with the system. In addition, algorithms using collaborative principles may suffer from comparing items or users with sparse elements in common, and trust can be a valuable way to overcome this problem. However, this problem only occurs when there are millions of messages to be recommended and only dozens of rated messages per user, which was not our case. Although we provided the users of our HRS with a brief explanation of why they had received the messages and why they had to rate them, we did not perform any trust-related analysis. We recommend that future researchers keep exploring the trust element in HRSs. It has great potential to solve many issues, as explained before, and has already been identified as a relevant topic for future recommender systems [161].

Computational costs

In terms of computational complexity –translated to computing time and server costs—the approach of traditional computer tailoring using deterministic decision tree algorithms has a linear computational cost of O(n), where "n" is the maximum length of a branch of the tree. However, recommender systems with collaborative filtering have a computational cost of $O(m^2n)$, where "m" is the number of users, and "n" is the number of messages [513]. This computational cost implies that HRSs may take considerably more time to produce a recommendation than traditional computer tailoring since the number of users in the system increase. The $O(m^2n)$ complexity order may be unsuitable for large-scale interventions where immediate feedback from the system is needed. Our HRS was designed to compute the recommendation selection for each day during the previous night and only sent messages during the day to prevent such a system overload. This preprocessing reduced the possibility of system overload, as done by Amazon [514]; however, in our case, it was not necessary because each recommendation took only a few milliseconds to select. Although this preprocessing approach guaranteed timely recommendations, it presented two main limitations. First, if the system had scaled, the time allocated to preprocess the data may not have been sufficient to perform all the required calculations. Second, the system sent recommendations that were selected during the night. Therefore, if the user provides additional feedback early in the morning before the selected message was delivered, their feedback would not be considered until the next night. This would make the user receive a message based on "old" data. However, the probability and the impact of this risk happening were considered low. These risks were not accounted for because it is assumed that receiving conflicting messages for a participant is not possible if a bit of feedback is not considered. However, we recommend that this type of computational cost risk evaluation be done for each implementation of HRS, as the priority (e.g., fast recommendations vs. constantly updated recommendations) will vary with the intervention context.

Liability and accountability

Liability and accountability are recurrent issues when dealing with automatic systems in healthcare [515, 516]. HRSs are not exempt from this; thus, this aspect needs to be considered in future interventions. In our specific scenario of smoking cessation, a poor recommendation could, at most, be irrelevant for a smoker but would not be harmful. However, in other, more severe scenarios and therapeutic areas, a poor recommendation may be dangerous. The negative effect of wrong messages on the users' health needs to be carefully considered and assessed in ethical and regulatory terms before implementing any intervention using HRSs to protect the users' interests and the professionals behind that intervention. To achieve that, the public health research community will need to be more empowered to design safe interventions. The contributions of this thesis, especially Chapter 2 and Chapter 3 facilitating decision-makers with the understanding of the design and benefits of HRSs, should be a catalyst for that to happen in the future such that the full potential of the HRSs can be unleashed and they can benefit as many people as possible.

In addition, HRSs solution designers need to rely on the legal advice of experts regarding how to optimally comply with the applicable medical liability laws and medical device regulations and make users read and accept the relevant legal disclaimers [517-519]. However, to minimize the risk of any misleading recommendation, we suggest always introducing algorithmic filters, such as our knowledge-based HRS or other technical controls, to ensure that the sent messages, although irrelevant, are never harmful to the users. If the system could send harmful recommendations to the user, it could be considered a medical device depending on the context. Therefore, it will need to comply with the applicable regulations for these types of systems. In these cases, the stated intended use of the system will largely determine the type of regulation that needs to be applied. For instance, an HRS offering lifestyle recommendations will be less prone to harm the user from an incorrect recommendation than an HRS providing tailored guidance about overcoming a disease.

Impact

The main objective of the present research was to extend the existing knowledge of recommender systems applied to the healthcare domain. Recommender systems are AI algorithms that select a recommendation for a user from a pool of different recommendations for a specific purpose using the previously acquired information about that user's preference (e.g., demographic profile and the explicit feedback for previous recommendations). In this thesis, we studied the previously generated evidence related to HRSs. In addition, we combined the field of AI and behavioral science to generate a new approach of personalizing messages to the traditional computer tailoring one and tested this innovative system in a study by including hundreds of participants. Specifically, we focused on supporting behavioral change in people who wanted to quit smoking by delivering them relevant motivational messages to reinforce their determination to remain abstinent.

The first main result is identifying a lack of usage of behavior science change principles concerning HRSs. The previous research with HRSs did not include relevant health behavior determinants and mainly concentrated on the systems' technical performance. Moreover, there was also a lack of consideration of behavioral change theories for HRSs. Hence, the design of the previous HRSs was suboptimal as their participants could not benefit from the behavioral impact derived from the behavioral change models, which had been proven to be effective before. In addition, most of the existing literature dealing with HRSs did not fully reveal how these systems were designed, hampering the reusability of the knowledge generated in their studies. To categorize HRSs and aid the further development and usage of improved HRSs, we developed a taxonomy for HRSs that comprised health, psychological, and technical sections. Our taxonomy can support future researchers at the system design stage not to forget any critical element, such as the behavioral change theories, and facilitate that they elaborately explain their system in future publications.

The second main result is the detailed step-by-step description of combining behavioral science and recommender systems technology to support behavioral change. We created an HRS to select and deliver smoking cessation motivational messages. Its algorithm or "internal architecture" was designed to work according to the phases and determinants of change proposed by the I-Change behavioral change model. Our detailed description of our system ensures transparency for future users and applications and, thus, will also help put the newly acquired scientific knowledge into practice.

The third main result is that we tested an HRS grounded in behavioral science, which used the technology called "collective intelligence" for its impact on message appreciation, user engagement, and smoking cessation outcomes in a study with high external validity. Collective intelligence is a technology based on the principle that "the things relevant to people similar to me will also be relevant to me." The results showed that our HRS could have smoking cessation success rates of around 10% after six months under a pessimistic scenario, which is double the success rates of unaided attempts at smoking cessation, and up to around 30% abstinence under available data analysis. When the collective intelligence technology was included in the system, it was found that the participants neither appreciated the messages more nor had more engagement, contrary to as we initially expected. However, when this system was provided with comprehensive information about the participants (e.g., data associated with the determinants of change of the I-Change behavioral change model), the engagement of the participants in terms of the number of quitting attempts increased. However, the smoking cessation outcomes depended on the type of analysis: Only when the 7D-PP abstinence was analyzed, considering the last available abstinence report did we find statistically significant results. These showed that the HRS using collective intelligence led to lower abstinence rates than the other HSR. The other conducted smoking abstinence analysis did not show statistically significant differences between the two systems. This lack of difference may have been caused by the fact that collaborative filtering may need more time to show additional behavior effects. Hence, although collective intelligence can lead to worse results than more straightforward approaches that do not rely on collaborative algorithms, it can also be a positive element about HRSs for achieving behavioral change. This characteristic can improve some outcomes in the intervention when comprehensive information about the participant is available. In addition, there were metrics – such as appreciation and engagement in terms of the number of rated messages- for which we could not find statistical significance. However, their evolution over time showed a consistent improvement when using collaborative intelligence, promising a positive impact in future studies applying this type of technology.

Overall, our results make significant contributions to the scientific community in many ways:

1. Our results are a foundation for health recommender systems. It is a still-young field lacking comprehensive analysis and categorization.
2. Our thesis provides a fully descriptive HRS design process such that our system can be replicated and adopted in future studies. Making the HRSs description easier will reduce the entry barrier for new researchers to keep exploring and using HRSs in new interventions.
3. Researchers and eHealth designers will know how to optimize their interventions using the best approaches of designing HRSs (e.g., including collaborative filtering, along with extensive participant information for personalizing recommendations).

4. Our findings illustrate and foster the added value of collaboration among different scientific disciplines. It is a multidisciplinary thesis that requires an understanding of psychological and technical concepts for creating effective HRSs.
5. As the core chapters of this thesis have been published in peer-reviewed high-impact journals, they will be accessible in the future to all researchers who are interested in research in the field of HRSs. The design and preliminary results of this study were disseminated in international conferences, such as that of the European Health Psychology Society, and scientific meetings, such as the TMU-IBM Joint Symposium on Innovation in Data Science and Artificial Intelligence in Health Care & JCMIT.

These results will also benefit society as they can be directly applied to the healthcare systems to support smoking cessation. As more people stop smoking, healthcare systems will reduce costs associated with smoking-related conditions developed by smokers. Furthermore, as introduced in Chapter 2, the thesis helps cover the United Nations Sustainable Development Goal #3 - "Ensure healthy lives and promote well-being for all at all ages" - with an innovative approach. Thus, in addition to researchers, policymakers will also benefit from the outcomes enclosed in this thesis as we incorporated the paradigm of precision medicine (AI applied to clinical diagnosis) into the behavioral change domain. However, the combination of AI and behavioral science of this thesis also contributes to society as the company Salumedia Labs, S.L.U. is adapting the knowledge generated in this thesis to other therapeutics areas, such as cancer and diabetes, and following one of our suggestions - exploring how user interfaces can impact recommender systems. To summarize, the relevance of this thesis goes beyond the academic world and is of interest to commercial and non-commercial organizations and other relevant stakeholders in the healthcare ecosystem for providing effective solutions to critical issues and fostering public health.

Conclusion

This thesis contributes to extending the existing scientific body of knowledge in the specific area of smoking cessation support. We determined that recommender systems can complement or substitute traditional computer tailoring for supporting smoking cessation behavioral change. The use of HRSs for health behavior change is an extensive new study field. The HSR methodology may help overcome the limited adaptability and capability of traditional rule-based systems, with its practical advantage of evolving with the users. However, there is a limited number of HRSs that consider behavioral theories in their design. This study described a comprehensive multidisciplinary design of an HRS combining behavioral theories and collective intelligence. We tested the system and found that it provided moderate smoking cessation support, with similar outcomes for continuous prevalence under both available data and penalized imputation analysis.

However, it did not provide higher 7D-PP abstinence (considering only the last abstinence report of the participants) than the other HRS without collective intelligence. Also, the system using collective intelligence did not result in a statistically significantly higher participant's appreciation for the messages. Mixed results were found for the generated engagement. The HRS with collective intelligence was only statistically significantly fostering more quitting attempts than the system without it when the users' profiles were fully completed.

This thesis took the first steps toward using HRSs for tailored eHealth interventions from a multidisciplinary perspective. Future research will need to explore other types of recommendation paradigms, therapeutic areas, behavioral theories, and delivery interfaces concerning HRSs and compare this type of technology directly with traditional computer-tailored eHealth to determine in which cases it might be more suitable and convenient. In this way, researchers can select the type of technology for improving the outcomes of their interventions and, ultimately, provide better support to patients willing to adhere to healthier behaviors.

References

1. Service, U.S.P.H., et al., *How tobacco smoke causes disease: the biology and behavioral basis for smoking-attributable disease: a report of the Surgeon General*. 2010: US Government Printing Office.
2. Organization, W.H. *Tobacco Factsheet*. 2017; Available from: http://www.who.int/mediacentre/factsheets/fs339/en/.
3. (CDC), C.f.D.C.a.P. *Smoking and Tobacco Use*. 2017; Available from: https://www.cdc.gov/tobacco/basic_information/index.htm.
4. Health, U.D.o. and H. Services, *The health consequences of smoking—50 years of progress: a report of the Surgeon General*. 2014, Atlanta, GA: US Department of Health and Human Services, Centers for Disease ….
5. Kelly, S.P., et al., *Smoking and cataract: review of causal association.* Journal of Cataract & Refractive Surgery, 2005. 31(12): p. 2395-2404.
6. Alkhatib, M.N., R.D. Holt, and R. Bedi, *Smoking and tooth discolouration: findings from a national cross-sectional study.* BMC Public Health, 2005. 5(1): p. 27.
7. Organization, W.H., *WHO global report on mortality attributable to tobacco.* 2012.
8. Vineis, P. and C.P. Wild, *Global cancer patterns: causes and prevention.* The Lancet, 2014. 383(9916): p. 549-557.
9. Mathers, C.D. and D. Loncar, *Projections of global mortality and burden of disease from 2002 to 2030.* PLoS medicine, 2006. 3(11): p. e442.
10. Goodchild, M., N. Nargis, and E.T. d'Espaignet, *Global economic cost of smoking-attributable diseases.* Tobacco control, 2018. 27(1): p. 58-64.
11. Metrics, I.f.H. and Evaluation, *Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2016 (GBD 2016) Results*. 2017, Institute for Health Metrics and Evaluation Seattle.
12. Loddenkemper, R., M. Kreuter, and F. Herth, *The tobacco epidemic*. 2015: Karger.
13. Tiwari, R., et al., *TO NICOTINE ADDICTION: AN OVERVIEW.* 2012.
14. Cinciripini, P.M., et al., *Tobacco addiction: implications for treatment and cancer prevention.* Journal of the National Cancer Institute, 1997. 89(24): p. 1852-1867.
15. Krasnegor, N.A., *Cigarette smoking as a dependence process*. Vol. 79. 1979: Department of Health, Education, and Welfare, Public Health Service, Alcohol ….
16. Yang, T., et al., *Attitudes to smoking cessation and triggers to relapse among Chinese male smokers.* BMC Public Health, 2006. 6(1): p. 65.
17. Andrews, J.A., et al., *The influence of peers on young adult substance use.* Health psychology, 2002. 21(4): p. 349.
18. Shiffman, S. and M. Prange, *Self-reported and self-monitored smoking patterns.* Addictive Behaviors, 1988. 13(2): p. 201-204.
19. McBride, C.M., et al., *Partner smoking status and pregnant smoker's perceptions of support for and likelihood of smoking cessation.* Health Psychology, 1998. 17(1): p. 63.
20. Shiffman, S. and M.E. Jarvik, *Situational Determinants of Coping in Smoking Relapse Crises 1.* Journal of Applied Social Psychology, 1987. 17(1): p. 3-15.

R

21. Hunt, W.A., *Learning mechanisms in smoking*. 2017: Routledge.
22. Pisinger, C., et al., *Motives to quit smoking and reasons to relapse differ by socioeconomic status.* Preventive medicine, 2011. 52(1): p. 48-52.
23. Buczkowski, K., et al., *Motivations toward smoking cessation, reasons for relapse, and modes of quitting: results from a qualitative study among former and current smokers.* Patient preference and adherence, 2014. 8: p. 1353.
24. Cummings, C., J.R. Gordon, and G.A. Marlatt, *Relapse: Prevention and prediction.* The addictive behaviors, 1980: p. 291-321.
25. Shiffman, S.M., *Determinants of Relapse Following Smoking Cessation.* 1980.
26. Marlatt, G.A., *Relapse prevention: A self-control program for the treatment of addictive behaviors*. 1980: University of Washington.
27. Babb, S., *Quitting smoking among adults—United States, 2000–2015.* MMWR. Morbidity and mortality weekly report, 2017. 65.
28. Chaney, S.E., S.W. Sheriff, and L. Merritt, *Gender differences in smoking behavior and cessation.* Clin Nurs Stud, 2015. 3(3): p. 17-22.
29. Fiore, M.C., et al., *Methods used to quit smoking in the United States: Do cessation programs help?* Jama, 1990. 263(20): p. 2760-2765.
30. Raupach, T., R. West, and J. Brown, *The most "successful" method for failing to quit smoking is unassisted cessation.* Nicotine & Tobacco Research, 2013. 15(3): p. 748-749.
31. Hughes, J.R., J. Keely, and S. Naud, *Shape of the relapse curve and long-term abstinence among untreated smokers.* Addiction, 2004. 99(1): p. 29-38.
32. Stead, L.F., et al., *Nicotine replacement therapy for smoking cessation.* Cochrane database of systematic reviews, 2012(11).
33. Wu, P., et al., *Effectiveness of smoking cessation therapies: a systematic review and meta-analysis.* BMC Public Health, 2006. 6(1): p. 300.
34. Hughes, J.R., et al., *Antidepressants for smoking cessation.* Cochrane database of systematic reviews, 2014(1).
35. Excellence, N.I.f.C. and G. Britain, *Guidance on the use of nicotine replacement therapy (NRT) and bupropion for smoking cessation*. 2002: National Institute for Clinical Excellence London.
36. Cahill, K., et al., *Nicotine receptor partial agonists for smoking cessation.* Cochrane database of systematic reviews, 2016(5).
37. Hajek, P., et al., *Varenicline in prevention of relapse to smoking: effect of quit pattern on response to extended treatment.* Addiction, 2009. 104(9): p. 1597-1602.
38. Agboola, S.A., et al., *Abstinence and relapse among smokers who use varenicline in a quit attempt—a pooled analysis of randomized controlled trials.* Addiction, 2015. 110(7): p. 1182-1193.
39. Hays, J.T., et al., *Sustained-release bupropion for pharmacologic relapse prevention after smoking cessation: a randomized, controlled trial.* Annals of internal medicine, 2001. 135(6): p. 423-433.

40. Hurt, R.D., et al., *Nicotine patch therapy based on smoking rate followed by bupropion for prevention of relapse to smoking.* Journal of Clinical Oncology, 2003. 21(5): p. 914-920.
41. Moore, T.J., J. Glenmullen, and C.D. Furberg, *Thoughts and acts of aggression/violence toward others reported in association with varenicline.* Annals of pharmacotherapy, 2010. 44(9): p. 1389-1394.
42. Williams, K.E., et al., *A double-blind study evaluating the long-term safety of varenicline for smoking cessation.* Current medical research and opinion, 2007. 23(4): p. 793-801.
43. Thompson, D.F. and D.R. Pierce, *Drug-induced nightmares.* Annals of Pharmacotherapy, 1999. 33(1): p. 93-98.
44. Jiménez-Ruiz, C., I. Berlin, and T. Hering, *Varenicline.* Drugs, 2009. 69(10): p. 1319-1338.
45. Wilkes, S., *Bupropion.* Drugs of today (Barcelona, Spain: 1998), 2006. 42(10): p. 671.
46. Stead, L.F., et al., *Physician advice for smoking cessation.* Cochrane database of systematic reviews, 2013(5).
47. Zinbarg, R.E. and J.W. Griffith, *Behavior therapy.* 2008.
48. Carroll, K.M., *Behavioral therapies for co-occurring substance use and mood disorders.* Biological psychiatry, 2004. 56(10): p. 778-784.
49. Petry, N.M., *Contingency management: what it is and why psychiatrists should want to use it.* The psychiatrist, 2011. 35(5): p. 161-163.
50. Ledgerwood, D.M., *Contingency management for smoking cessation: where do we go from here?* Current drug abuse reviews, 2008. 1(3): p. 340-349.
51. Krishnan-Sarin, S., et al., *Contingency management for smoking cessation in adolescent smokers.* Experimental and clinical psychopharmacology, 2006. 14(3): p. 306.
52. van den Brand, F.A., et al., *Effect of a workplace-based group training programme combined with financial incentives on smoking cessation: a cluster-randomised controlled trial.* The Lancet Public Health, 2018. 3(11): p. e536-e544.
53. van den Brand, F.A., et al., *How financial incentives increase smoking cessation: a two-level path analysis.* Nicotine & Tobacco Research, 2020.
54. Dobson, K.S. and D.J. Dozois, *Historical and philosophical bases of the cognitive-behavioral therapies*. 2010: Guilford Press.
55. Webb, M.S., et al., *Cognitive–behavioral therapy to promote smoking cessation among African American smokers: A randomized clinical trial.* Journal of consulting and clinical psychology, 2010. 78(1): p. 24.
56. Hofmann, S.G., et al., *The efficacy of cognitive behavioral therapy: A review of meta-analyses.* Cognitive therapy and research, 2012. 36(5): p. 427-440.
57. Schnoll, R.A., et al., *A randomized pilot study of cognitive-behavioral therapy versus basic health education for smoking cessation among cancer patients.* Annals of Behavioral Medicine, 2005. 30(1): p. 1-11.

58. Lancaster, T. and L.F. Stead, *Individual behavioural counselling for smoking cessation.* Cochrane database of systematic reviews, 2017(3).
59. Zhu, S.-H., et al., *Telephone counseling for smoking cessation: effects of single-session and multiple-session interventions.* Journal of consulting and clinical psychology, 1996. 64(1): p. 202.
60. Herkov, M., *About group therapy.* Psych Central, 2020.
61. Stead, L.F., A.J. Carroll, and T. Lancaster, *Group behaviour therapy programmes for smoking cessation.* Cochrane database of systematic reviews, 2017(3).
62. Hartmann-Boyce, J., et al., *Additional behavioural support as an adjunct to pharmacotherapy for smoking cessation.* Cochrane database of systematic reviews, 2019(6).
63. Cahill, K., et al., *Pharmacological interventions for smoking cessation: an overview and network meta-analysis.* Cochrane database of systematic reviews, 2013(5).
64. Alterman, A.I., P. Gariti, and F. Mulvaney, *Short-and long-term smoking cessation for three levels of intensity of behavioral treatment.* Psychology of addictive behaviors, 2001. 15(3): p. 261.
65. Hughes, J.R., et al., *Measures of abstinence in clinical trials: issues and recommendations.* Nicotine & Tobacco Research, 2003. 5(1): p. 13-25.
66. Cheung, K.L., et al., *Exploring consensus on how to measure smoking cessation. A Delphi study.* BMC Public Health, 2017. 17(1): p. 890.
67. Glasgow, R.E., et al., *Biochemical validation of smoking status: pros, cons, and data from four low-intensity intervention trials.* Addictive behaviors, 1993. 18(5): p. 511-527.
68. Sykes, C.M. and D.F. Marks, *Effectiveness of a cognitive behaviour therapy self-help programme for smokers in London, UK.* Health Promotion International, 2001. 16(3): p. 255-260.
69. Montaño, D.E. and D. Kasprzyk, *Theory of reasoned action, theory of planned behavior, and the integrated behavioral model.* Health behavior: Theory, research and practice, 2015. 70(4): p. 231.
70. Glanz, K., B.K. Rimer, and K. Viswanath, *Health behavior and health education: theory, research, and practice*. 2008: John Wiley & Sons.
71. Bell, C.C., B. Flay, and R. Paikoff, *Strategies for health behavior change*, in *The health behavioral change imperative*. 2002, Springer. p. 17-39.
72. Norman, P., M. Conner, and R. Bell, *The theory of planned behavior and smoking cessation.* Health psychology, 1999. 18(1): p. 89.
73. Zheng, P., et al., *A randomized controlled trial of group intervention based on social cognitive theory for smoking cessation in China.* Journal of epidemiology, 2007. 17(5): p. 147-155.
74. Lorencatto, F., et al., *Assessing the quality of goal setting in behavioural support for smoking cessation and its association with outcomes.* Annals of Behavioral Medicine, 2016. 50(2): p. 310-318.

75. Reisi, M., et al., *Factors affecting cigarette smoking based on health-belief model structures in pre-university students in Isfahan, Iran.* Journal of education and health promotion, 2014. 3.
76. Perkins, K.A., C.A. Conklin, and M.D. Levine, *Cognitive-behavioral therapy for smoking cessation: a practical guidebook to the most effective treatments*. 2008: Taylor & Francis.
77. Sutton, S., *Transtheoretical model applied to smoking cessation.* Understanding and changing health behaviour: From health beliefs to self-regulation. Amsterdam: Harwood Academic Publishers, 2000: p. 207-25.
78. De Vries, H., *An integrated approach for understanding health behavior: The I-Change Model as an example.* Psychology and Behavioral Science International Journal, 2017. 2(2).
79. De Vries, H., A. Mudde, and A. Dijkstra, *The attitude-social influence-efficacy model applied to the prediction of motivational transitions in the process of smoking cessation.* Understanding and changing health behaviour: From health beliefs to self-regulation, 2000: p. 165-187.
80. De Vries, H., et al., *Differential beliefs, perceived social influences, and self-efficacy expectations among smokers in various motivational phases.* Preventive medicine, 1998. 27(5): p. 681-689.
81. De Vries, H., *Planning and evaluating health promotion.* Evaluating health promotion, 1998.
82. Kim, J.H. and Y.H. Choi, *Effects of a Strength Based I-Change Smoking Cessation Program for Smoking Middle School Boys.* Journal of Korean Academy of Nursing, 2017. 47(2).
83. Panday, S., et al., *Determinants of smoking cessation among adolescents in South Africa.* Health education research, 2005. 20(5): p. 586-599.
84. Smit, E.S., et al., *Predictors of successful and unsuccessful quit attempts among smokers motivated to quit.* Addictive behaviors, 2014. 39(9): p. 1318-1324.
85. Te Poel, F., et al., *Efficacy of a single computer-tailored e-mail for smoking cessation: results after 6 months.* Health Education Research, 2009. 24(6): p. 930-940.
86. Hilberink, S.R., et al., *Characteristics of patients with COPD in three motivational stages related to smoking cessation.* Patient Education and Counseling, 2006. 61(3): p. 449-457.
87. Kaplan, C.P., et al., *Smoking cessation counseling with young patients: the practices of family physicians and pediatricians.* Archives of pediatrics & adolescent medicine, 2004. 158(1): p. 83-90.
88. Helgason, A.R. and K.E. Lund, *General practitioners' perceived barriers to smoking cessation-results from four Nordic countries.* Scandinavian journal of public health, 2002. 30(2): p. 141-147.
89. Yarnall, K.S., et al., *Primary care: is there enough time for prevention?* American journal of public health, 2003. 93(4): p. 635-641.

90. Smit, E.S., H. de Vries, and C. Hoving, *The PAS study: a randomized controlled trial evaluating the effectiveness of a web-based multiple tailored smoking cessation programme and tailored counselling by practice nurses.* Contemporary clinical trials, 2010. 31(3): p. 251-258.
91. Polosa, R., et al., *Feasibility of using E-mail counseling as part of a smoking-cessation program.* Respiratory care, 2009. 54(8): p. 1033-1039.
92. Searight, H.R., *Efficient counseling techniques for the primary care physician.* Primary Care: Clinics in Office Practice, 2007. 34(3): p. 551-570.
93. Borland, R., J. Balmford, and D. Hunt, *The effectiveness of personally tailored computer-generated advice letters for smoking cessation.* Addiction, 2004. 99(3): p. 369-377.
94. Dijkstra, A., et al., *Tailoring information to enhance quitting in smokers with low motivation to quit: three basic efficacy questions.* Health Psychology, 1998. 17(6): p. 513.
95. Dijkstra, A., H. De Vries, and J. Roijackers, *Computerized tailored feedback to change cognitive determinants of smoking: a Dutch field experiment.* Health Education Research, 1998. 13(2): p. 197-206.
96. Elwood, W.N. and M. Vega, *Process evaluation results from a condom use intervention with substance abusers in treatment: from "Cinderella?" to "Show me!".* Alcoholism Treatment Quarterly, 2005. 23(1): p. 47-62.
97. Schinke, S.P., T.M. Schwinn, and A.J. Ozanian, *Alcohol abuse prevention among high-risk youth: computer-based intervention.* Journal of Prevention & Intervention in the Community, 2005. 29(1-2): p. 117-130.
98. Kreuter, M.W., et al., *Use of computer kiosks for breast cancer education in five community settings.* Health Education & Behavior, 2006. 33(5): p. 625-642.
99. McDonald, E.M., et al., *Evaluation of kiosk-based tailoring to promote household safety behaviors in an urban pediatric primary care practice.* Patient education and counseling, 2005. 58(2): p. 168-181.
100. Kong, G., et al., *Text messaging-based smoking cessation intervention: a narrative review.* Addictive behaviors, 2014. 39(5): p. 907-917.
101. Naughton, F., J. Jamison, and S. Sutton, *Attitudes towards SMS text message smoking cessation support: a qualitative study of pregnant smokers.* Health education research, 2013. 28(5): p. 911-922.
102. Haug, S., et al., *Continuous individual support of smoking cessation using text messaging: a pilot experimental study.* Nicotine & Tobacco Research, 2009. 11(8): p. 915-923.
103. Lustria, M.L.A., et al., *Computer-tailored health interventions delivered over the Web: review and analysis of key components.* Patient education and counseling, 2009. 74(2): p. 156-173.

104. Oenema, A., et al., *Efficacy and use of an internet-delivered computer-tailored lifestyle intervention, targeting saturated fat intake, physical activity and smoking cessation: a randomized controlled trial.* Annals of Behavioral Medicine, 2008. 35(2): p. 125-135.
105. Elfeddali, I., et al., *Preventing smoking relapse via Web-based computer-tailored feedback: a randomized controlled trial.* Journal of medical Internet research, 2012. 14(4): p. e109.
106. Schneider, F., et al., *Periodic email prompts to re-use an internet-delivered computer-tailored lifestyle program: influence of prompt content and timing.* Journal of medical Internet research, 2013. 15(1): p. e23.
107. Alley, S., et al., *Web-based video-coaching to assist an automated computer-tailored physical activity intervention for inactive adults: a randomized controlled trial.* Journal of medical Internet research, 2016. 18(8): p. e223.
108. Walthouwer, M.J.L., et al., *Comparing a video and text version of a web-based computer-tailored intervention for obesity prevention: a randomized controlled trial.* Journal of medical Internet research, 2015. 17(10): p. e236.
109. Soetens, K.C., et al., *Using online computer tailoring to promote physical activity: a randomized trial of text, video, and combined intervention delivery modes.* Journal of Health Communication, 2014. 19(12): p. 1377-1392.
110. Cheung, K.L., et al., *Effectiveness of a video-versus text-based computer-tailored intervention for obesity prevention after one year: a randomized controlled trial.* International Journal of Environmental Research and Public Health, 2017. 14(10): p. 1275.
111. Stanczyk, N., et al., *Comparison of text and video computer-tailored interventions for smoking cessation: randomized controlled trial.* Journal of medical Internet research, 2014. 16(3): p. e69.
112. Kreuter, M.W., et al., *Tailoring health messages: Customizing communication with computer technology*. 2013: Routledge.
113. de Vries, H. and J. Brug, *Computer-tailored interventions motivating people to adopt health promoting behaviours: introduction to a new approach.* 1999.
114. Brug, J., A. Oenema, and M. Campbell, *Past, present, and future of computer-tailored nutrition education.* The American journal of clinical nutrition, 2003. 77(4): p. 1028S-1034S.
115. Dijkstra, A., *Working mechanisms of computer-tailored health education: evidence from smoking cessation.* Health Education Research, 2005. 20(5): p. 527-539.
116. Dijkstra, A., *The psychology of tailoring-ingredients in computer-tailored persuasion.* Social and personality psychology compass, 2008. 2(2): p. 765-784.
117. Petty, R.E. and J.T. Cacioppo, *The elaboration likelihood model of persuasion*, in *Communication and persuasion*. 1986, Springer. p. 1-24.

118. Krebs, P., J.O. Prochaska, and J.S. Rossi, *A meta-analysis of computer-tailored interventions for health behavior change.* Preventive medicine, 2010. 51(3-4): p. 214-221.
119. Smit, E.S., et al., *Cost-effectiveness and cost-utility of Internet-based computer tailoring for smoking cessation.* Journal of Medical Internet Research, 2013. 15(3): p. e57.
120. Smit, E.S., H. de Vries, and C. Hoving, *Effectiveness of a Web-based multiple tailored smoking cessation program: a randomized controlled trial among Dutch adult smokers.* Journal of medical Internet research, 2012. 14(3): p. e82.
121. Stanczyk, N.E., et al., *Influence of delivery strategy on message-processing mechanisms and future adherence to a Dutch computer-tailored smoking cessation intervention.* Journal of medical Internet research, 2013. 15(2): p. e28.
122. Elfeddali, I., et al., *A randomized controlled trial of Web-based Attentional Bias Modification to help smokers quit.* Health Psychology, 2016. 35(8): p. 870.
123. Etter, J.-F. and T.V. Perneger, *Effectiveness of a computer-tailored smoking cessation program: a randomized trial.* Archives of Internal Medicine, 2001. 161(21): p. 2596-2601.
124. Strecher, V.J., *Computer-tailored smoking cessation materials: a review and discussion.* Patient education and counseling, 1999. 36(2): p. 107-117.
125. Jander, A., et al., *Effects of a web-based computer-tailored game to reduce binge drinking among Dutch adolescents: a cluster randomized controlled trial.* Journal of medical Internet research, 2016. 18(2): p. e29.
126. Bingham, C.R., et al., *Efficacy of a web-based, tailored, alcohol prevention/intervention program for college students: initial findings.* Journal of American College Health, 2010. 58(4): p. 349-356.
127. Martinez-Montilla, J.M., et al., *A web-based, computer-tailored intervention to reduce alcohol consumption and binge drinking among Spanish adolescents: cluster randomized controlled trial.* Journal of medical internet research, 2020. 22(1): p. e15438.
128. de Vries, H., et al., *Internet-based computer tailored feedback on sunscreen use.* Journal of Medical Internet Research, 2012. 14(2): p. e48.
129. Kroeze, W., A. Werkman, and J. Brug, *A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors.* Annals of behavioral medicine, 2006. 31(3): p. 205-223.
130. Neville, L.M., B. O'Hara, and A. Milat, *Computer-tailored physical activity behavior change interventions targeting adults: a systematic review.* International Journal of Behavioral Nutrition and Physical Activity, 2009. 6(1): p. 30.
131. Smeets, T., J. Brug, and H. de Vries, *Effects of tailoring health messages on physical activity.* Health Education Research, 2008. 23(3): p. 402-413.
132. Brug, J., et al., *The impact of a computer-tailored nutrition intervention.* Preventive medicine, 1996. 25(3): p. 236-242.

133. Springvloet, L., et al., *Short-and medium-term efficacy of a Web-based computer-tailored nutrition education intervention for adults including cognitive and environmental feedback: randomized controlled trial.* Journal of medical Internet research, 2015. 17(1): p. e23.
134. Rawl, S.M., et al., *Computer-delivered tailored intervention improves colon cancer screening knowledge and health beliefs of African-Americans.* Health Education Research, 2012. 27(5): p. 868-885.
135. van der Aalst, C.M., et al., *The effectiveness of a computer-tailored smoking cessation intervention for participants in lung cancer screening: a randomised controlled trial.* Lung Cancer, 2012. 76(2): p. 204-210.
136. Giarratano, J.C. and G. Riley, *Expert systems: principles and programming*. 1989: Brooks/Cole Publishing Co.
137. Kreuter, M.W. and R.J. Wray, *Tailored and targeted health communication: strategies for enhancing information relevance.* American journal of health behavior, 2003. 27(1): p. S227-S232.
138. Kukafka, R., *Tailored health communication*, in *Consumer Health Informatics*. 2005, Springer. p. 22-33.
139. Schulz, D.N., et al., *Program completion of a web-based tailored lifestyle intervention for adults: differences between a sequential and a simultaneous approach.* Journal of medical Internet research, 2012. 14(2): p. e26.
140. Gomez-Uribe, C.A. and N. Hunt, *The netflix recommender system: Algorithms, business value, and innovation.* ACM Transactions on Management Information Systems (TMIS), 2015. 6(4): p. 1-19.
141. Sunartio, R., H.N. Palit, and A. Gunawan, *Hotel Recommender System Menggunakan Metode Pendekatan Graph pada Dataset Trivago.* Jurnal Infra, 2020. 8(1): p. 215-220.
142. Schafer, J.B., J.A. Konstan, and J. Riedl, *E-commerce recommendation applications.* Data mining and knowledge discovery, 2001. 5(1-2): p. 115-153.
143. Wiesner, M. and D. Pfeifer, *Health recommender systems: concepts, requirements, technical basics and challenges.* International journal of environmental research and public health, 2014. 11(3): p. 2580-2607.
144. Fernandez-Luque, L., R. Karlsen, and L.K. Vognild. *Challenges and opportunities of using recommender systems for personalized health education*. in *MIE*. 2009.
145. Melville, P. and V. Sindhwani, *Recommender systems.* Encyclopedia of machine learning, 2010. 1: p. 829-838.
146. Sadasivam, R.S., et al., *Collective-intelligence recommender systems: advancing computer tailoring for health behavior change into the 21st century.* Journal of medical Internet research, 2016. 18(3): p. e42.
147. Kotkov, D., S. Wang, and J. Veijalainen, *A survey of serendipity in recommender systems.* Knowledge-Based Systems, 2016. 111: p. 180-192.

R

148. Van Meteren, R. and M. Van Someren. *Using content-based filtering for recommendation*. in *Proceedings of the Machine Learning in the New Information Age: MLnet/ECML2000 Workshop*. 2000.
149. Thorat, P.B., R. Goudar, and S. Barve, *Survey on collaborative filtering, content-based filtering and hybrid recommendation system.* International Journal of Computer Applications, 2015. 110(4): p. 31-36.
150. Burke, R., *Knowledge-based recommender systems.* Encyclopedia of library and information systems, 2000. 69(Supplement 32): p. 175-186.
151. Burke, R., *Hybrid recommender systems: Survey and experiments.* User modeling and user-adapted interaction, 2002. 12(4): p. 331-370.
152. Aggarwal, C.C., *Recommender systems*. Vol. 1. 2016: Springer.
153. Bobadilla, J., et al., *A collaborative filtering approach to mitigate the new user cold start problem.* Knowledge-based systems, 2012. 26: p. 225-238.
154. Zaier, Z., R. Godin, and L. Faucher. *Evaluating recommender systems*. in *2008 International Conference on Automated Solutions for Cross Media Content and Multi-Channel Distribution*. 2008. IEEE.
155. Harper, F.M. and J.A. Konstan, *The movielens datasets: History and context.* Acm transactions on interactive intelligent systems (tiis), 2015. 5(4): p. 1-19.
156. Agapito, G., et al. *DIETOS: A recommender system for adaptive diet monitoring and personalized food suggestion*. in *2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. 2016. IEEE.
157. Trattner, C. and D. Elsweiler, *Food recommender systems: important contributions, challenges and future research directions.* arXiv preprint arXiv:1711.02760, 2017.
158. Trattner, C. and D. Elsweiler. *Investigating the healthiness of internet-sourced recipes: implications for meal planning and recommender systems*. in *Proceedings of the 26th international conference on world wide web*. 2017.
159. Cappella, J.N., S. Yang, and S. Lee, *Constructing recommendation systems for effective health messages using content, collaborative, and hybrid algorithms.* The ANNALS of the American Academy of Political and Social Science, 2015. 659(1): p. 290-306.
160. Espín, V., M.V. Hurtado, and M. Noguera, *Nutrition for Elder Care: a nutritional semantic recommender system for the elderly.* Expert Systems, 2016. 33(2): p. 201-210.
161. Schäfer, H., et al. *Towards health (aware) recommender systems*. in *Proceedings of the 2017 international conference on digital health*. 2017. ACM.
162. Sadasivam, R.S., et al., *Impact of a collective intelligence tailored messaging system on smoking cessation: the Perspect randomized experiment.* Journal of medical Internet research, 2016. 18(11): p. e285.
163. Hors-Fraile, S., et al., *Tailoring motivational health messages for smoking cessation using an mHealth recommender system integrated with an electronic health record: a study protocol.* BMC public health, 2018. 18(1): p. 1-10.

164. Carrasco-Hernandez, L., et al., *A Mobile Health Solution Complementing Psychopharmacology-Supported Smoking Cessation: Randomized Controlled Trial.* JMIR mHealth and uHealth, 2020. 8(4): p. e17530.
165. Hales, S., et al., *Social networks for improving healthy weight loss behaviors for overweight and obese adults: a randomized clinical trial of the social pounds off digitally (Social POD) mobile app.* International journal of medical informatics, 2016. 94: p. 81-90.
166. Cheung, K.L., et al., *How recommender systems could support and enhance computer-tailored digital health programs: A scoping review.* Digital health, 2019. 5: p. 2055207618824727.
167. Leon, A., *The respective roles of user participation and user involvement in information system implementation success.* 1991.
168. Lee, T.-H. and W.R. King, *The effects of user participation on system success: toward a contingency theory of user satisfaction.* Information management research, 1996. 1(1): p. 55-80.
169. Sawesi, S., et al., *The impact of information technology on patient engagement and health behavior change: a systematic review of the literature.* JMIR medical informatics, 2016. 4(1): p. e1.
170. O'Brien, H.L. and E.G. Toms, *What is user engagement? A conceptual framework for defining user engagement with technology.* Journal of the American society for Information Science and Technology, 2008. 59(6): p. 938-955.
171. Yardley, L., et al., *Understanding and promoting effective engagement with digital behavior change interventions.* American journal of preventive medicine, 2016. 51(5): p. 833-842.
172. Perski, O., et al., *Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis.* Translational behavioral medicine, 2017. 7(2): p. 254-267.
173. Short, C.E., et al., *Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies.* Journal of medical Internet research, 2018. 20(11): p. e292.
174. Bano, M. and D. Zowghi, *A systematic review on the relationship between user involvement and system success.* Information and Software Technology, 2015. 58: p. 148-169.
175. Hwang, M.I. and R.G. Thorn, *The effect of user engagement on system success: a meta-analytical integration of research findings.* Information & Management, 1999. 35(4): p. 229-236.
176. Schwarzer, R. and L. Satow, *Online intervention engagement predicts smoking cessation.* Preventive medicine, 2012. 55(3): p. 233-236.
177. Pike, K.J., et al., *American Cancer Society's QuitLink: randomized trial of Internet assistance.* Nicotine & Tobacco Research, 2007. 9(3): p. 415-420.

R

178. Strecher, V., et al., *The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial.* Journal of medical Internet research, 2008. 10(5): p. e36.
179. Danaher, B.G., et al., *Mediators of a successful web-based smokeless tobacco cessation program.* Addiction, 2008. 103(10): p. 1706-1712.
180. Rabius, V., et al., *Comparing internet assistance for smoking cessation: 13-month follow-up of a six-arm randomized controlled trial.* Journal of medical Internet research, 2008. 10(5): p. e45.
181. Japuntich, S.J., et al., *Smoking cessation via the internet: a randomized clinical trial of an internet intervention as adjuvant treatment in a smoking cessation intervention.* Nicotine & Tobacco Research, 2006. 8(Suppl_1): p. S59-S67.
182. Perski, O., et al., *Smokers' and drinkers' choice of smartphone applications and expectations of engagement: A think aloud and interview study.* BMC medical informatics and decision making, 2017. 17(1): p. 25.
183. Richardson, A., et al., *Engagement promotes abstinence in a web-based cessation intervention: cohort study.* Journal of medical Internet research, 2013. 15(1): p. e14.
184. Christofferson, D.E., et al., *Engagement and abstinence among users of a smoking cessation text message program for veterans.* Addictive behaviors, 2016. 62: p. 47-53.
185. Donkin, L., et al., *A systematic review of the impact of adherence on the effectiveness of e-therapies.* Journal of medical Internet research, 2011. 13(3): p. e52.
186. Kohl, L.F., R. Crutzen, and N.K. de Vries, *Online prevention aimed at lifestyle behaviors: a systematic review of reviews.* Journal of medical Internet research, 2013. 15(7): p. e146.
187. Murray, E., et al., *Attrition revisited: adherence and retention in a web-based alcohol trial.* Journal of medical Internet research, 2013. 15(8): p. e162.
188. Hardiker, N.R. and M.J. Grant, *Factors that influence public engagement with eHealth: A literature review.* International journal of medical informatics, 2011. 80(1): p. 1-12.
189. Lentferink, A.J., et al., *Key components in eHealth interventions combining self-tracking and persuasive eCoaching to promote a healthier lifestyle: a scoping review.* Journal of medical Internet research, 2017. 19(8): p. e277.
190. Wangberg, S.C., T.S. Bergmo, and J.-A.K. Johnsen, *Adherence in Internet-based interventions.* Patient preference and adherence, 2008. 2: p. 57.
191. Lie, S.S., et al., *Dropout from an eHealth intervention for adults with type 2 diabetes: a qualitative study.* Journal of medical Internet research, 2017. 19(5): p. e187.
192. Pu, P., L. Chen, and R. Hu. *A user-centric evaluation framework for recommender systems*. in *Proceedings of the fifth ACM conference on Recommender systems*. 2011.
193. Schneider, F.M., et al., *Learning from entertaining online video clips? Enjoyment and appreciation and their differential relationships with knowledge and behavioral intentions.* Computers in Human Behavior, 2016. 54: p. 475-482.

194. Kroeze, W., et al., *Comparison of use and appreciation of a print-delivered versus CD-ROM-delivered, computer-tailored intervention targeting saturated fat intake: randomized controlled trial.* Journal of Medical Internet Research, 2008. 10(2): p. e12.
195. Schulz, D.N., et al., *Effects of a web-based tailored multiple-lifestyle intervention for adults: a two-year randomized controlled trial comparing sequential and simultaneous delivery modes.* Journal of medical Internet research, 2014. 16(1): p. e26.
196. Lochbuehler, K., et al., *Influence of smoking cues in movies on children's beliefs about smoking.* Pediatrics, 2012. 130(2): p. 221-227.
197. Cremers, H.-P., et al., *A web-based computer-tailored smoking prevention programme for primary school children: intervention design and study protocol.* BMC public health, 2012. 12(1): p. 277.
198. Kumar, S. and G. Preetha, *Health promotion: an effective tool for global health.* Indian journal of community medicine: official publication of Indian Association of Preventive & Social Medicine, 2012. 37(1): p. 5.
199. Griffiths, F., et al., *Why are health care interventions delivered over the internet? A systematic review of the published literature.* Journal of medical Internet research, 2006. 8(2): p. e10.
200. Eysenbach, G., *The law of attrition.* Journal of medical Internet research, 2005. 7(1): p. e11.
201. Brug, J., et al., *Computer-tailored nutrition education: differences between two interventions.* Health Education Research, 1999. 14(2): p. 249-256.
202. Couper, M.P., et al., *Engagement and retention: measuring breadth and depth of participant use of an online intervention.* Journal of medical Internet research, 2010. 12(4): p. e52.
203. Solenhill, M., et al., *The effect of tailored web-based feedback and optional telephone coaching on health improvements: a randomized intervention among employees in the transport service industry.* Journal of medical Internet research, 2016. 18(8): p. e158.
204. Broekhuizen, K., et al., *A systematic review of randomized controlled trials on the effectiveness of computer-tailored physical activity and dietary behavior promotion programs: an update.* Annals of Behavioral Medicine, 2012. 44(2): p. 259-286.
205. Smith, M.Y., et al., *Determining the cost-effectiveness of a computer-based smoking cessation intervention in primary care.* Manag Care, 2007. 16(7): p. 48-55.
206. López-Nores, M., et al., *Property-based collaborative filtering for health-aware recommender systems.* Expert Systems with Applications, 2012. 39(8): p. 7451-7457.
207. Resnick, P. and H.R. Varian, *Recommender systems.* Communications of the ACM, 1997. 40(3): p. 56-58.
208. Park, D.H., et al., *A literature review and classification of recommender systems research.* Expert systems with applications, 2012. 39(11): p. 10059-10072.
209. Sezgin, E. and S. Özkan. *A systematic literature review on Health Recommender Systems.* in *2013 E-Health and Bioengineering Conference (EHB)*. 2013. IEEE.

R

210. Reed, C., E. Kennedy, and S. Silva, *Responsibility, Autonomy and Accountability: legal liability for machine learning.* Queen Mary School of Law Legal Studies Research Paper, 2016(243).
211. Strecher, V.J., et al., *The effects of computer-tailored smoking cessation messages in family practice settings.* Journal of Family Practice, 1994. 39(3): p. 262-270.
212. Schmid, K.L., et al., *Targeting or tailoring? Maximizing resources to create effective health communications.* Marketing health services, 2008. 28(1): p. 32.
213. Valdez, A.C., et al., *Recommender systems for health informatics: state-of-the-art and future perspectives*, in *Machine Learning for Health Informatics*. 2016, Springer. p. 391-414.
214. Liberati, A., et al., *The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration.* Journal of clinical epidemiology, 2009. 62(10): p. e1-e34.
215. Arksey, H. and L. O'Malley, *Scoping studies: towards a methodological framework.* International journal of social research methodology, 2005. 8(1): p. 19-32.
216. Nickerson, R.C., U. Varshney, and J. Muntermann, *A method for taxonomy development and its application in information systems.* European Journal of Information Systems, 2013. 22(3): p. 336-359.
217. Cheung, K.L., B. Wijnen, and H. de Vries, *A review of the theoretical basis, effects, and cost effectiveness of online smoking cessation interventions in the Netherlands: A mixed-methods approach.* Journal of medical Internet research, 2017. 19(6): p. e230.
218. Schafer, J.B., J. Konstan, and J. Riedl. *Recommender systems in e-commerce*. in *Proceedings of the 1st ACM conference on Electronic commerce*. 1999.
219. Montaner, M., B. López, and J.L. De La Rosa, *A taxonomy of recommender agents on the internet.* Artificial intelligence review, 2003. 19(4): p. 285-330.
220. Colantonio, S., et al., *A smart mirror to promote a healthy lifestyle.* Biosystems Engineering, 2015. 138: p. 33-43.
221. Wendel, S., et al., *Consumers' intention to use health recommendation systems to receive personalized nutrition advice.* BMC health services research, 2013. 13(1): p. 126.
222. Wang, S.-L., et al., *Design and evaluation of a cloud-based Mobile Health Information Recommendation system on wireless sensor networks.* Computers & Electrical Engineering, 2016. 49: p. 221-235.
223. Lin, Y.-F., et al., *Design of a real-time and continua-based framework for care guideline recommendations.* International journal of environmental research and public health, 2014. 11(4): p. 4262-4279.
224. Hidalgo, J.I., et al., *glUCModel: A monitoring and modeling system for chronic diseases applied to diabetes.* Journal of biomedical informatics, 2014. 48: p. 183-192.
225. Chomutare, T., E. Arsand, and G. Hartvigsen, *Mobile peer support in diabetes.* Studies in health technology and informatics, 2011. 169: p. 48-52.

226. Ali, R., et al., *Multimodal hybrid reasoning methodology for personalized wellbeing services.* Computers in biology and medicine, 2016. 69: p. 10-28.
227. Abbas, A., et al., *Personalized healthcare cloud services for disease risk assessment and wellness management using social media.* Pervasive and Mobile Computing, 2016. 28: p. 81-99.
228. Zhang, W., et al., *Predicting potential side effects of drugs by recommender methods and ensemble learning.* Neurocomputing, 2016. 173: p. 979-987.
229. Honka, A., et al., *Rethinking health: ICT-enabled services to empower people to manage their health.* IEEE reviews in biomedical engineering, 2011. 4: p. 119-139.
230. Giabbanelli, P.J. and R. Crutzen, *Supporting self-management of obesity using a novel game architecture.* Health informatics journal, 2015. 21(3): p. 223-236.
231. Esteban, B., et al., *TPLUFIB-WEB: A fuzzy linguistic Web system to help in the treatment of low back pain problems.* Knowledge-Based Systems, 2014. 67: p. 429-438.
232. Chen, T., *Ubiquitous multicriteria clinic recommendation system.* Journal of medical systems, 2016. 40(5): p. 113.
233. Guo, L., et al., *Which doctor to trust: a recommender system for identifying the right doctors.* Journal of medical Internet research, 2016. 18(7): p. e186.
234. Weitzel, J.A., et al., *Using wireless handheld computers and tailored text messaging to reduce negative consequences of drinking alcohol.* Journal of Studies on Alcohol and drugs, 2007. 68(4): p. 534-537.
235. Holzinger, A., *Interactive machine learning for health informatics: when do we need the human-in-the-loop?* Brain Informatics, 2016. 3(2): p. 119-131.
236. Holzinger, A., et al. *Towards interactive Machine Learning (iML): applying ant colony algorithms to solve the traveling salesman problem with the human-in-the-loop approach*. in *International Conference on Availability, Reliability, and Security*. 2016. Springer.
237. Holzinger, A., et al., *A glass-box interactive machine learning approach for solving NP-hard problems with the human-in-the-loop.* arXiv preprint arXiv:1708.01104, 2017.
238. Goodman, B. and S. Flaxman, *European Union regulations on algorithmic decision-making and a "right to explanation".* AI magazine, 2017. 38(3): p. 50-57.
239. Hors-Fraile, S., et al., *Design of two combined health recommender systems for tailoring messages in a smoking cessation app.* arXiv preprint arXiv:1608.07192, 2016.
240. Abroms, L.C., et al., *A randomized trial of Text2Quit: a text messaging program for smoking cessation.* American journal of preventive medicine, 2014. 47(3): p. 242-250.
241. Fjeldsoe, B.S., A.L. Marshall, and Y.D. Miller, *Behavior change interventions delivered by mobile telephone short-message service.* American journal of preventive medicine, 2009. 36(2): p. 165-173.
242. Park, L.G., J. Howie-Esquivel, and K. Dracup, *A quantitative systematic review of the efficacy of mobile phone interventions to improve medication adherence.* Journal of advanced nursing, 2014. 70(9): p. 1932-1953.

243. Hors-Fraile, S., et al., *Analyzing recommender systems for health promotion using a multidisciplinary taxonomy: A scoping review.* International journal of medical informatics, 2018. 114: p. 143-155.
244. Fogel, A.L. and J.C. Kvedar, *Artificial intelligence powers digital medicine.* NPJ digital medicine, 2018. 1(1): p. 5.
245. Vayena, E., A. Blasimme, and I.G. Cohen, *Machine learning in medicine: Addressing ethical challenges.* PLoS medicine, 2018. 15(11): p. e1002689.
246. Rigby, M.J., *Ethical dimensions of using artificial intelligence in health care.* AMA Journal of Ethics, 2019. 21(2): p. 121-124.
247. Price, I. and W. Nicholson, *Artificial Intelligence in Health Care: Applications and Legal Implications.* 2017.
248. Bamidis, P.D., et al. *Multimodal e-health services for smoking cessation and public health: the smokefreebrain project approach*. in *MEDINFO 2017: Precision Healthcare Through Informatics: Proceedings of the 16th World Congress on Medical and Health Informatics*. 2018. IOS Press.
249. Hors-Fraile, S., et al., *A recommender system to quit smoking with mobile motivational messages: Study protocol for a randomized controlled trial.* Trials, 2018. 19(1): p. 618.
250. Stanczyk, N.E., et al., *An economic evaluation of a video-and text-based computer-tailored intervention for smoking cessation: a cost-effectiveness and cost-utility analysis of a randomized controlled trial.* PloS one, 2014. 9(10): p. e110117.
251. de Ruijter, D., et al., *The effectiveness of a computer-tailored e-learning program for practice nurses to improve their adherence to smoking cessation counseling guidelines: randomized controlled trial.* Journal of medical Internet research, 2018. 20(5): p. e193.
252. Svensson, M., et al. *A recipe based on-line food store*. in *Proceedings of the 5th international conference on Intelligent user interfaces*. 2000. ACM.
253. Aberg, J. *Dealing with Malnutrition: A Meal Planning System for Elderly*. in *AAAI spring symposium: argumentation for consumers of healthcare*. 2006.
254. Tran, T.N.T., et al., *An overview of recommender systems in the healthy food domain.* Journal of Intelligent Information Systems, 2018. 50(3): p. 501-526.
255. Torkamaan, H. and J. Ziegler. *Multi-Criteria Rating-Based Preference Elicitation in Health Recommender Systems*. in *HealthRecSys@ RecSys*. 2018.
256. Torkamaan, H. and J. Ziegler. *Rating-based Preference Elicitation for Recommendation of Stress Intervention*. in *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*. 2019. ACM.
257. Tondello, G.F., R. Orji, and L.E. Nacke. *Recommender systems for personalized gamification*. in *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*. 2017. ACM.
258. Ghorai, K., et al. *An mhealth recommender for smoking cessation using case based reasoning*. in *2013 46th Hawaii International Conference on System Sciences*. 2013. IEEE.

259. Whittaker, R., et al., *A theory-based video messaging mobile phone intervention for smoking cessation: randomized controlled trial.* Journal of medical Internet research, 2011. 13(1): p. e10.
260. Shi, X., et al., *Long-term performance of collaborative filtering based recommenders in temporally evolving systems.* Neurocomputing, 2017. 267: p. 635-643.
261. Abroms, L.C., et al., *Text2Quit: results from a pilot test of a personalized, interactive mobile health smoking cessation program.* Journal of health communication, 2012. 17(sup1): p. 44-53.
262. Skov-Ettrup, L., et al., *Comparing tailored and untailored text messages for smoking cessation: a randomized controlled trial among adolescent and young adult smokers.* Health education research, 2014. 29(2): p. 195-205.
263. Strecher, V.J., et al., *Web-based smoking-cessation programs: results of a randomized trial.* American journal of preventive medicine, 2008. 34(5): p. 373-381.
264. Schwarzer, R., *Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors.* Applied psychology, 2008. 57(1): p. 1-29.
265. Haug, S., et al., *Efficacy of a text message-based smoking cessation intervention for young people: a cluster randomized controlled trial.* Journal of medical Internet research, 2013. **15**(8): p. e171.
266. Prochaska, J.O. and W.F. Velicer, *The transtheoretical model of health behavior change.* American journal of health promotion, 1997. 12(1): p. 38-48.
267. Noar, S.M., et al., *Tailored health communication to change lifestyle behaviors.* American Journal of Lifestyle Medicine, 2011. 5(2): p. 112-122.
268. Hawkins, R.P., et al., *Understanding tailoring in communicating about health.* 2008. 23(3): p. 454-466.
269. Department of Mental Health and Substance Dependence, W., *Encouraging People to Stop Smoking*. 2001.
270. Kasten, S., S. Eggers, and H. de Vries, *The importance of self-efficacy and action planning for smoking cessation.* European Health Psychologist, 2015. 17(S): p. 694.
271. de Vries, H., S.M. Eggers, and C. Bolman, *The role of action planning and plan enactment for smoking cessation.* BMC Public Health, 2013. 13(1): p. 393.
272. Dijkstra, A. and H.D. Vries, *Self-efficacy expectations with regard to different tasks in smoking cessation.* Psychology and Health, 2000. 15(4): p. 501-511.
273. Vries, H.D. and A.N. Mudde, *Predicting stage transitions for smoking cessation applying the attitude-social influence-efficacy model.* Psychology and Health, 1998. 13(2): p. 369-385.
274. De Vries, H., et al., *The Impact of Social Influences in the Context of Attitude, Self-Efficacy, Intention, and Previous Behavior as Predictors of Smoking Onset1.* Journal of applied social psychology, 1995. 25(3): p. 237-257.

275. Shahab, L. and A. McEwen, *Online support for smoking cessation: a systematic review of the literature.* Addiction, 2009. 104(11): p. 1792-1804.
276. Noar, S.M., C.N. Benac, and M.S. Harris, *Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions.* Psychological bulletin, 2007. 133(4): p. 673.
277. Macleod, M., et al., *The MUSiC performance measurement method.* Behaviour & Information Technology, 1997. 16(4-5): p. 279-293.
278. Abraham, C. and M. Kools, *Writing health communication: An evidence-based guide*. 2011: Sage.
279. Herlocker, J.L., J.A. Konstan, and J. Riedl. *Explaining collaborative filtering recommendations*. in *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. 2000. ACM.
280. Maltz, D. and K. Ehrlich. *Pointing the way: active collaborative filtering*. in *CHI*. 1995.
281. Park, S.-T., et al. *Naïve filterbots for robust cold-start recommendations*. in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2006. ACM.
282. Schein, A.I., et al. *Methods and metrics for cold-start recommendations*. in *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*. 2002. ACM.
283. Cacioppo, J.T., R.E.J.B. Petty, and A.S. Psychology, *Effects of message repetition on argument processing, recall, and persuasion.* 1989. **10**(1): p. 3-12.
284. Abroms, L.C., et al., *Developing and pretesting a text messaging program for health behavior change: recommended steps.* JMIR mHealth and uHealth, 2015. 3(4): p. e107.
285. Ahmad, W., *An empirical investigation of the association between creative advertising and advertising effectiveness in Pakistan*. 2010, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology (SZABIST ….
286. Claypool, H.M., et al., *The effects of personal relevance and repetition on persuasive processing.* Social Cognition, 2004. 22(3): p. 310-335.
287. Kim, J., D. Lee, and K.-Y. Chung, *Item recommendation based on context-aware model for personalized u-healthcare service.* Multimedia Tools and Applications, 2014. 71(2): p. 855-872.
288. Lin, Y., et al. *Motivate: Towards context-aware recommendation mobile system for healthy living*. in *2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*. 2011. IEEE.
289. López-Nores, M., et al., *Context-Aware Recommender Systems Influenced by the Users' Health-Related Data*, in *User Modeling and Adaptation for Daily Routines*. 2013, Springer. p. 153-173.
290. Choi, J., et al., *The influence of national culture on the attitude towards mobile recommender systems.* Technological Forecasting and Social Change, 2014. 86: p. 65-79.

291. Babitsch, B., et al., *Doctor's perception of doctor-patient relationships in emergency departments: What roles do gender and ethnicity play?* BMC health services research, 2008. 8(1): p. 82.
292. Organization, W.H. *Global health risks: mortality and burden of disease attributable to slect major risks.* 2009.
293. Health, U.D.o. and H. Services, *The health consequences of smoking: a report of the Surgeon General*. Vol. 62. 2004: US Department of Health and Human Services, Centers for Disease Control and ….
294. (CDC), C.f.D.C.a.P. *Smoking and Tobacco Use.* 2017.
295. (WHO), W.H.O. *Tobacco*. 2017.
296. Messer, K., et al., *Smoking cessation rates in the United States: a comparison of young adult and older smokers.* American journal of public health, 2008. 98(2): p. 317-322.
297. Grimshaw, G. and A. Stanton, *Tobacco cessation interventions for young people.* Cochrane database of systematic reviews, 2006(4).
298. Silagy, C., et al., *Nicotine replacement therapy for smoking cessation.* Cochrane Database Syst Rev, 2004. 3(3).
299. Stead, L.F., et al., *Combined pharmacotherapy and behavioural interventions for smoking cessation.* Cochrane database of systematic reviews, 2016(3).
300. Rice, V.H. and L.F. Stead, *Nursing interventions for smoking cessation.* Cochrane database of systematic reviews, 2008(1).
301. Organization, W.H. *Addiction to Nicotine*. 2010.
302. Hughes, J.R., *Effects of abstinence from tobacco: etiology, animal models, epidemiology, and significance: a subjective review.* Nicotine & Tobacco Research, 2007. 9(3): p. 329-339.
303. Hughes, J.R., *Effects of abstinence from tobacco: valid symptoms and time course.* Nicotine & Tobacco Research, 2007. 9(3): p. 315-327.
304. Kassel, J.D., L.R. Stroud, and C.A. Paronis, *Smoking, stress, and negative affect: correlation, causation, and context across stages of smoking.* Psychological bulletin, 2003. 129(2): p. 270.
305. Dawkins, L., et al., *First-versus second-generation electronic cigarettes: predictors of choice and effects on urge to smoke and withdrawal symptoms.* Addiction, 2015. 110(4): p. 669-677.
306. Bancej, C., et al., *Smoking cessation attempts among adolescent smokers: a systematic review of prevalence studies.* Tobacco control, 2007. 16(6): p. e8-e8.
307. Gilbert, H., et al., *Effectiveness of personalised risk information and taster sessions to increase the uptake of smoking cessation services (Start2quit): a randomised controlled trial.* The Lancet, 2017. 389(10071): p. 823-833.
308. Kale, D., H.M. Gilbert, and S. Sutton, *Are predictors of making a quit attempt the same as predictors of 3-month abstinence from smoking? Findings from a sample of smokers*

recruited for a study of computer-tailored smoking cessation advice in primary care. Addiction, 2015. 110(10): p. 1653-1664.

309. Lotrean, L.M., et al., *Process evaluation of the first computer tailored program for smoking cessation among Romanian smokers.* Asian Pac J Cancer Prev, 2014. 15(20): p. 8809-8814.
310. Buller, D.B., et al., *Randomized trial of a smartphone mobile application compared to text messaging to support smoking cessation.* Telemedicine and e-Health, 2014. 20(3): p. 206-214.
311. Naughton, F., et al., *Randomized controlled trial to assess the short-term effectiveness of tailored web-and text-based facilitation of smoking cessation in primary care (i Q uit in P ractice).* Addiction, 2014. 109(7): p. 1184-1193.
312. Smokefreebrain.eu. *SmokeFreeBrain*. 2017 12 Oct 2017]; Available from: http://smokefreebrain.eu/.
313. De Vries, H., *An integrated approach for understanding health behavior; the I-change model as an example.* Psychol Behav Sci Int J, 2017. 2(2): p. 555-585.
314. De Vries, H., et al., *The European Smoking prevention Framework Approach (EFSA): an example of integral prevention.* Health education research, 2003. 18(5): p. 611-626.
315. de Vries, H., M. Dijkstra, and P. Kuhlman, *Self-efficacy: the third factor besides attitude and subjective norm as a predictor of behavioural intentions.* Health education research, 1988. 3(3): p. 273-282.
316. de Ruijter, D., et al., *Web-based computer-tailoring for practice nurses aimed to improve smoking cessation guideline adherence: a study protocol for a randomized controlled effectiveness trial.* Contemporary clinical trials, 2016. 48: p. 125-132.
317. Stearns, M., et al., *Towards evaluating and enhancing the reach of online health forums for smoking cessation.* Network Modeling Analysis in Health Informatics and Bioinformatics, 2014. 3(1): p. 69.
318. Salumedia. *Quit&Return - Android Apps on Google Play*. 2017 12 Oct 2017]; Available from: https://play.google.com/store/apps/details?id=com.salumedia.quitandreturn.
319. Gabarron, E., et al., *Impact of Facebook ads for sexual health promotion via an educational web app: a case study*, in *Healthcare Policy and Reform: Concepts, Methodologies, Tools, and Applications*. 2019, IGI Global. p. 990-1003.
320. Realsun. 2017 13 Oct 2017]; Available from: http://www.realsun.com.tw/.
321. CKAN. *CKAN*. 1 Oct 2017]; Available from: https://ckan.org.
322. Corporation, M. *The Mirth open Source Portal | Mirth Corporation*. 2017 12 Oct 2017]; Available from: https://nextgen.com.
323. Google.com. *Google Fit*. 2017 12 Oct 2017]; Available from: https://www.google.com/fit/.
324. Marlin, B.M., et al. *Towards collaborative filtering recommender systems for tailored health communications*. in *AMIA annual symposium proceedings*. 2013. American Medical Informatics Association.

325. Eur-lex.europa.eu. *EUR-Lex - 32002L0058 - EN*. 2017 1 Oct 2017]; Available from: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:en:HTML.
326. Eur-lex.europa.eu. *EUR-Lex - 31995L0046 - EN*. 2017 1 Oct 2017]; Available from: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML.
327. Brug, J., et al., *The impact of computer-tailored feedback and iterative feedback on fat, fruit, and vegetable intake.* Health Education & Behavior, 1998. 25(4): p. 517-531.
328. Enwald, H.P.K. and M.-L.A. Huotari, *Preventing the obesity epidemic by second generation tailored health communication: an interdisciplinary review.* Journal of medical Internet research, 2010. 12(2): p. e24.
329. Ghalibaf, A.K., et al., *Comprehensive overview of computer-based health information tailoring: a systematic scoping review.* BMJ open, 2019. 9(1): p. e021022.
330. Kreuter, M., et al., *Tailored health messages: Customizing communication with computer technology.* Mahwah, New Jersey: Lewrence Erlbaum, 1999.
331. Glasgow, R., et al., *Engagement in a diabetes self-management website: usage patterns and generalizability of program use.* Journal of medical Internet research, 2011. 13(1): p. e9.
332. Schneider, F., et al., *The influence of user characteristics and a periodic email prompt on exposure to an internet-delivered computer-tailored lifestyle program.* Journal of medical Internet research, 2012. 14(2): p. e40.
333. Bennett, G.G. and R.E. Glasgow, *The delivery of public health interventions via the Internet: actualizing their potential.* Annual review of public health, 2009. 30: p. 273-292.
334. Kreps, G.L. and L. Neuhauser, *New directions in eHealth communication: opportunities and challenges.* Patient education and counseling, 2010. 78(3): p. 329-336.
335. Kaptein, M. and D. Eckles. *Selecting effective means to any end: Futures and ethics of persuasion profiling*. in *International conference on persuasive technology*. 2010. Springer.
336. Lederman, R., et al., *Tailoring patient information to encourage patient engagement.* International Journal of Networking and Virtual Organisations, 2016. 16(2): p. 191-203.
337. Spittaels, H., et al., *Effectiveness of an online computer-tailored physical activity intervention in a real-life setting.* Health education research, 2007. 22(3): p. 385-396.
338. Morrison, L., et al., *Optimizing engagement with I nternet-based health behaviour change interventions: Comparison of self-assessment with and without tailored feedback using a mixed methods approach.* British journal of health psychology, 2014. 19(4): p. 839-855.
339. Elahi, M., et al. *Interaction Design in a Mobile Food Recommender System*. in *IntRS@RecSys*. 2015.

R

340. Musto, C., et al. *Towards a knowledge-aware food recommender system exploiting holistic user models*. in *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*. 2020.
341. Gómez-del-Río, N., et al., *Health promotion for childhood obesity: An approach based on self-tracking of data.* Sensors, 2020. 20(13): p. 3778.
342. Elsweiler, D., et al. *Bringing the" healthy" into Food Recommenders*. in *DMRS*. 2015.
343. Zenun Franco, R. *Online recommender system for personalized nutrition advice*. in *Proceedings of the Eleventh ACM Conference on Recommender Systems*. 2017.
344. Ribeiro, D., et al. *SousChef: Mobile Meal Recommender System for Older Adults*. in *ICT4AgeingWell*. 2017.
345. Hors-Fraile, S., et al., *Opening the Black Box: Explaining the Process of Basing a Health Recommender System on the I-Change Behavioral Change Model.* IEEE Access, 2019. 7: p. 176525-176540.
346. Pu, P., L. Chen, and R. Hu, *Evaluating recommender systems from the user's perspective: survey of the state of the art.* User Modeling and User-Adapted Interaction, 2012. 22(4-5): p. 317-355.
347. Gunawardana, A. and G. Shani, *Evaluating recommender systems*, in *Recommender systems handbook*. 2015, Springer. p. 265-308.
348. Said, A., et al. *Recommender Systems Evaluation: A 3D Benchmark*. in *RUE@ RecSys*. 2012.
349. Herlocker, J.L., et al., *Evaluating collaborative filtering recommender systems.* ACM Transactions on Information Systems (TOIS), 2004. 22(1): p. 5-53.
350. Çano, E. and M. Morisio, *Hybrid recommender systems: A systematic literature review.* Intelligent Data Analysis, 2017. 21(6): p. 1487-1524.
351. Heatherton, T.F., et al., *The Fagerström test for nicotine dependence: a revision of the Fagerstrom Tolerance Questionnaire.* British journal of addiction, 1991. 86(9): p. 1119-1127.
352. Richmond, R.L., L.A. Kehoe, and I.W. Webster, *Multivariate models for predicting abstention following intervention to stop smoking by general practitioners.* Addiction, 1993. 88(8): p. 1127-1135.
353. Berkup, S.B., *Working with generations X and Y in generation Z period: Management of different generations in business life.* Mediterranean Journal of Social Sciences, 2014. 5(19): p. 218.
354. Gupta, S.K., *Intention-to-treat concept: a review.* Perspectives in clinical research, 2011. 2(3): p. 109.
355. Cacheda, F., et al., *Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems.* ACM Transactions on the Web (TWEB), 2011. 5(1): p. 1-33.

356. Candillier, L., et al., *State-of-the-art recommender systems*, in *Collaborative and Social Information Retrieval and Access: Techniques for Improved User Modeling*. 2009, IGI Global. p. 1-22.
357. Desrosiers, C. and G. Karypis, *A comprehensive survey of neighborhood-based recommendation methods*, in *Recommender systems handbook*. 2011, Springer. p. 107-144.
358. Ziegler, C.-N., et al. *Improving recommendation lists through topic diversification*. in *Proceedings of the 14th international conference on World Wide Web*. 2005.
359. Schafer, J.B., et al., *Collaborative filtering recommender systems*, in *The adaptive web*. 2007, Springer. p. 291-324.
360. Serrano-Guerrero, J., et al., *A google wave-based fuzzy recommender system to disseminate information in University Digital Libraries 2.0.* Information Sciences, 2011. 181(9): p. 1503-1516.
361. Porcel, C., et al., *A hybrid recommender system for the selective dissemination of research resources in a technology transfer office.* Information Sciences, 2012. 184(1): p. 1-19.
362. Torous, J., et al., *Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements.* Evidence-based mental health, 2018. 21(3): p. 116-119.
363. Sandborg, J., et al., *Participants' Engagement and Satisfaction With a Smartphone App Intended to Support Healthy Weight Gain, Diet, and Physical Activity During Pregnancy: Qualitative Study Within the HealthyMoms Trial.* JMIR mHealth and uHealth, 2021. 9(3): p. e26159.
364. Dovaliene, A., Z. Piligrimiene, and A. Masiulyte, *Factors influencing customer engagement in mobile applications.* Engineering Economics, 2016. 27(2): p. 205-212.
365. Casetta, B., et al., *Association between cigarette smoking prevalence and income level: a systematic review and meta-analysis.* Nicotine & Tobacco Research, 2017. 19(12): p. 1401-1407.
366. Hosseinpoor, A.R., et al., *Social determinants of smoking in low-and middle-income countries: results from the World Health Survey.* PloS one, 2011. 6(5): p. e20331.
367. Arnold, C.L., et al., *Smoking status, reading level, and knowledge of tobacco effects among low-income pregnant women.* Preventive medicine, 2001. 32(4): p. 313-320.
368. Fakir, S.E., et al., *Knowledge and perceptions of smoking according to income level in Morocco.* American Journal of Health Promotion, 2011. 25(6): p. 387-391.
369. Kaye, L., et al., *Impact of Socioeconomic Status on the Relationship Between Smartphone App Engagement and Adherence Among Patients with Asthma*, in *B102. CARING FOR LUNG DISEASES IN UNDERSERVED AND VULNERABLE POPULATIONS*. 2020, American Thoracic Society. p. A4281-A4281.
370. Crafoord, M.-T., et al., *Engagement in an Interactive App for Symptom Self-Management during Treatment in Patients With Breast or Prostate Cancer: Mixed Methods Study.* Journal of Medical Internet Research, 2020. 22(8): p. e17058.

371. Garnett, C., et al., *Predictors of engagement, response to follow up, and extent of alcohol reduction in users of a smartphone app (Drink Less): secondary analysis of a factorial randomized controlled trial.* JMIR mHealth and uHealth, 2018. 6(12): p. e11175.
372. Zheng, Y., B. Mobasher, and R. Burke. *Similarity-based context-aware recommendation*. in *International Conference on Web Information Systems Engineering*. 2015. Springer.
373. Guo, G., J. Zhang, and N. Yorke-Smith. *A novel Bayesian similarity measure for recommender systems*. in *IJCAI*. 2013.
374. Sarwar, B., et al. *Analysis of recommendation algorithms for e-commerce*. in *Proceedings of the 2nd ACM conference on Electronic commerce*. 2000.
375. Nelson, D.B., et al., *Why assigning ongoing tobacco use is not necessarily a conservative approach to handling missing tobacco cessation outcomes.* Nicotine and Tobacco Research, 2009. 11(1): p. 77-83.
376. Blankers, M., et al., *The missing= smoking assumption: a fallacy in internet-based smoking cessation trials?* Nicotine & Tobacco Research, 2016. 18(1): p. 25-33.
377. Westmaas, J.L., et al., *Randomised controlled trial of stand-alone tailored emails for smoking cessation.* Tobacco Control, 2018. 27(2): p. 136-146.
378. Regmi, K., et al., *Effectiveness of mobile apps for smoking cessation: a review.* Tobacco Prevention & Cessation, 2017. 3.
379. Bricker, J.B., et al., *Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy.* Drug and alcohol dependence, 2014. 143: p. 87-94.
380. Ubhi, H.K., et al., *A mobile app to aid smoking cessation: preliminary evaluation of SmokeFree28.* Journal of medical Internet research, 2015. 17(1): p. e17.
381. Iacoviello, B.M., et al., *Clickotine, a personalized smartphone app for smoking cessation: initial evaluation.* JMIR mHealth and uHealth, 2017. 5(4): p. e56.
382. Meyer, C., et al., *Proactive interventions for smoking cessation in general medical practice: a quasi-randomized controlled trial to examine the efficacy of computer-tailored letters and physician-delivered brief advice.* Addiction, 2008. 103(2): p. 294-304.
383. Karumur, R.P., T.T. Nguyen, and J.A. Konstan, *Personality, user preferences and behavior in recommender systems.* Information Systems Frontiers, 2018. 20(6): p. 1241-1265.
384. Willemse, B.J.P.C., M.C. Kaptein, and F. Hasaart, *Developing Effective Methods for Electronic Health Personalization: Protocol for Health Telescope, a Prospective Interventional Study.* JMIR research protocols, 2020. 9(7): p. e16471.
385. Kennedy-Martin, T., et al., *A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results.* Trials, 2015. 16(1): p. 495.
386. Sahoo, A.K., et al., *DeepReco: deep learning based health recommender system using collaborative filtering*. Computation, 2019. 7(2): p. 25.

387. Rivero-Rodriguez, A., et al. *A health information recommender system: Enriching YouTube health videos with Medline Plus information by the use of SnomedCT terms*. in *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*. 2013. IEEE.
388. Bocanegra, C.L.S., et al., *HealthRecSys: A semantic content-based recommender system to complement health videos.* BMC medical informatics and decision making, 2017. 17(1): p. 1-10.
389. Zeng, E.Y., et al., *Get with the program: adherence to a smartphone app for smoking cessation.* Addictive behaviors, 2016. 63: p. 120-124.
390. Hutton, H.E., et al., *A systematic review of randomized controlled trials: Web-based interventions for smoking cessation among adolescents, college students, and adults.* Nicotine & Tobacco Research, 2011. 13(4): p. 227-238.
391. Benowitz, N., et al., *SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation.* Nicotine Tob Res, 2002. 4(2): p. 149-159.
392. Stanczyk, N., et al., *Effectiveness of video-versus text-based computer-tailored smoking cessation interventions among smokers after one year.* Preventive medicine, 2016. 82: p. 42-50.
393. Patrick, D.L., et al., *The validity of self-reported smoking: a review and meta-analysis.* American journal of public health, 1994. 84(7): p. 1086-1093.
394. Knijnenburg, B.P., et al., *Explaining the user experience of recommender systems.* User Modeling and User-Adapted Interaction, 2012. 22(4): p. 441-504.
395. Kaptein, M. and D. Eckles, *Heterogeneity in the effects of online persuasion.* Journal of Interactive Marketing, 2012. 26(3): p. 176-188.
396. Kaptein, M., et al., *Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles.* International Journal of Human-Computer Studies, 2015. 77: p. 38-51.
397. Smith, A.L., S. Chapman, and S.M. Dunlop, *What do we know about unassisted smoking cessation in Australia? A systematic review, 2005–2012.* Tobacco control, 2015. 24(1): p. 18-27.
398. Moore, D., et al., *Effectiveness and safety of nicotine replacement therapy assisted reduction to stop smoking: systematic review and meta-analysis.* Bmj, 2009. 338.
399. Shiffman, S., et al., *Real-world efficacy of prescription and over-the-counter nicotine replacement therapy.* Addiction, 2002. 97(5): p. 505-516.
400. Davis, R., et al., *Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review.* Health psychology review, 2015. 9(3): p. 323-344.
401. Lustria, M.L.A., et al., *A meta-analysis of web-delivered tailored health behavior change interventions.* Journal of health communication, 2013. 18(9): p. 1039-1069.

402. Fjeldsoe, B.S., Y.D. Miller, and A.L. Marshall, *MobileMums: a randomized controlled trial of an SMS-based physical activity intervention.* Annals of Behavioral Medicine, 2010. 39(2): p. 101-111.
403. Riley, W.T., et al., *Health behavior models in the age of mobile interventions: are our theories up to the task?* Translational behavioral medicine, 2011. 1(1): p. 53-71.
404. Dahlke, D.V., et al., *Apps seeking theories: results of a study on the use of health behavior change theories in cancer survivorship mobile apps.* JMIR mHealth and uHealth, 2015. 3(1): p. e31.
405. West, J.H., et al., *Health behavior theories in diet apps.* Journal of consumer health on the internet, 2013. 17(1): p. 10-24.
406. Cowan, L.T., et al., *Apps of steel: are exercise apps providing consumers with realistic expectations? A content analysis of exercise apps for presence of behavior change theory.* Health Education & Behavior, 2013. 40(2): p. 133-139.
407. Conroy, D.E., C.-H. Yang, and J.P. Maher, *Behavior change techniques in top-ranked mobile apps for physical activity.* American journal of preventive medicine, 2014. 46(6): p. 649-652.
408. Prochaska, J.O., C.A. Redding, and K.E. Evers, *The transtheoretical model and stages of change.* Health behavior: Theory, research, and practice, 2015. 97.
409. Champion, V.L. and C.S. Skinner, *The health belief model.* Health behavior and health education: Theory, research, and practice, 2008. 4: p. 45-65.
410. Bandura, A., *Health promotion from the perspective of social cognitive theory.* Psychology and health, 1998. 13(4): p. 623-649.
411. Ajzen, I., *The theory of planned behavior.* Organizational Behavior and Human Decision Processes, 1991. 50(2): p. 179-211.
412. Yürüten, O., *Recommender Systems for Healthy Behavior Change*. 2017, EPFL.
413. Bandura, A., *Social foundations of thought and action.* Englewood Cliffs, NJ, 1986. 1986: p. 23-28.
414. Nakamura, J., et al., *Handbook of positive psychology*. 2002, Oxford University Press New York.
415. Cremonesi, P., Y. Koren, and R. Turrin. *Performance of recommender algorithms on top-n recommendation tasks*. in *Proceedings of the fourth ACM conference on Recommender systems*. 2010.
416. Chang, S., *Leveraging Collective Intelligence in Recommender System.* 2016.
417. Glass, R.L. and I. Vessey, *Contemporary application-domain taxonomies.* IEEE Software, 1995. 12(4): p. 63-76.
418. Lousame, F.P. and E. Sánchez, *A taxonomy of collaborative-based recommender systems*, in *Web Personalization in Intelligent Environments*. 2009, Springer. p. 81-117.
419. Friedrich, G. and M. Zanker, *A taxonomy for generating explanations in recommender systems.* AI Magazine, 2011. 32(3): p. 90-98.

420. Papadimitriou, A., P. Symeonidis, and Y. Manolopoulos, *A generalized taxonomy of explanations styles for traditional and social recommender systems.* Data Mining and Knowledge Discovery, 2012. 24(3): p. 555-583.
421. Pu, P. and L. Chen, *Trust-inspiring explanation interfaces for recommender systems.* Knowledge-Based Systems, 2007. 20(6): p. 542-556.
422. McClure, J.B., A.L. Hartzler, and S.L. Catz, *Design considerations for smoking cessation apps: feedback from nicotine dependence treatment providers and smokers.* JMIR mHealth and uHealth, 2016. 4(1): p. e17.
423. Dijkstra, A. and H. De Vries, *The development of computer-generated tailored interventions.* Patient education and counseling, 1999. 36(2): p. 193-203.
424. Jawaheer, G., P. Weller, and P. Kostkova, *Modeling user preferences in recommender systems: A classification framework for explicit and implicit user feedback*. ACM Transactions on Interactive Intelligent Systems (TiiS), 2014. 4(2): p. 1-26.
425. Baskerville, N.B., L.L. Struik, and D. Dash, *Crush the crave: development and formative evaluation of a smartphone app for smoking cessation.* JMIR mHealth and uHealth, 2018. 6(3): p. e52.
426. Fogg, B.J. and J. Hreha. *Behavior wizard: a method for matching target behaviors with solutions*. in *International Conference on Persuasive Technology*. 2010. Springer.
427. Bindoff, I., et al., *Quittr: the design of a video game to support smoking cessation.* JMIR Serious Games, 2016. 4(2): p. e19.
428. Weiser, P., et al., *A taxonomy of motivational affordances for meaningful gamified and persuasive technologies.* 2015.
429. Mummah, S.A., et al., *IDEAS (Integrate, Design, Assess, and Share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior.* Journal of medical Internet research, 2016. 18(12): p. e317.
430. Quiñonez, S.G., et al., *mHealth or eHealth? Efficacy, use, and appreciation of a Web-based computer-tailored physical activity intervention for Dutch adults: a randomized controlled trial.* Journal of medical Internet research, 2016. 18(11): p. e278.
431. Burley, C., et al. *Increasing Engagement in eHealth Interventions Using Personalization and Implementation Intentions*. in *2020 Systems and Information Engineering Design Symposium (SIEDS)*. 2020. IEEE.
432. Heerwegh, D., *Effects of personal salutations in e-mail invitations to participate in a web survey.* Public Opinion Quarterly, 2005. 69(4): p. 588-598.
433. Mora, A., et al. *Effect of personalized gameful design on student engagement*. in *2018 IEEE Global Engineering Education Conference (EDUCON)*. 2018. IEEE.
434. Muñoz-Leiva, F., et al., *Improving the response rate and quality in Web-based surveys through the personalization and frequency of reminder mailings.* Quality & Quantity, 2010. 44(5): p. 1037-1052.

R

435. Graham, A.L., et al., *Optimizing Text Messages to Promote Engagement With Internet Smoking Cessation Treatment: Results From a Factorial Screening Experiment.* Journal of Medical Internet Research, 2020. 22(4): p. e17734.
436. McClure, J.B., et al., *The effect of program design on engagement with an internet-based smoking intervention: randomized factorial trial.* Journal of medical Internet research, 2013. 15(3): p. e69.
437. Heminger, C.L., et al., *Text2Quit: an analysis of participant engagement in the mobile smoking cessation program*. The American journal of drug and alcohol abuse, 2016. 42(4): p. 450-458.
438. Festinger, L., *A theory of cognitive dissonance*. Vol. 2. 1957: Stanford university press.
439. Greene, J. and J.H. Hibbard, *Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes.* Journal of general internal medicine, 2012. 27(5): p. 520-526.
440. Hibbard, J.H., et al., *Do increases in patient activation result in improved self-management behaviors?* Health services research, 2007. 42(4): p. 1443-1463.
441. Luxton, D.D., *An introduction to artificial intelligence in behavioral and mental health care*, in *Artificial intelligence in behavioral and mental health care*. 2016, Elsevier. p. 1-26.
442. Robila, M. and S.A. Robila, *Applications of artificial intelligence methodologies to behavioral and social sciences.* Journal of Child and Family Studies, 2019: p. 1-13.
443. Verification, S.S.o.B., *Biochemical verification of tobacco use and cessation.* Nicotine & Tobacco Research, 2002. 4(2): p. 149-159.
444. Toyama, M., et al., *Design and content validation of a set of SMS to promote seeking of specialized mental health care within the Allillanchu Project.* Global health, epidemiology and genomics, 2018. 3.
445. Alismail, S. and L. Olfman, *A tailored motivational messages library for a mobile health sleep behavior change support system to promote continuous positive airway pressure use among patients with obstructive sleep apnea: Development, content validation, and testing.* JMIR mHealth and uHealth, 2020. 8(8): p. e18793.
446. Kassam-Adams, N., et al., *A new method for assessing content validity in model-based creation and iteration of eHealth interventions.* Journal of medical Internet research, 2015. 17(4): p. e3811.
447. Rosenstock, I.M., *Historical origins of the health belief model.* Health education monographs, 1974. 2(4): p. 328-335.
448. Rogers, R.W., *A protection motivation theory of fear appeals and attitude change1.* The journal of psychology, 1975. 91(1): p. 93-114.
449. Ryan, R.M. and E.L. Deci, *Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.* American psychologist, 2000. 55(1): p. 68.

450. Bartholomew, L.K., G.S. Parcel, and G. Kok, *Intervention mapping: a process for developing theory and evidence-based health education programs.* Health education & behavior, 1998. 25(5): p. 545-563.
451. Kok, G., et al., *Intervention mapping: protocol for applying health psychology theory to prevention programmes.* Journal of health psychology, 2004. 9(1): p. 85-98.
452. Revathy, V. and S.P. Anitha, *Cold start problem in social recommender systems: State-of-the-art review*, in *Advances in Computer Communication and Computational Sciences*. 2019, Springer. p. 105-115.
453. Xu, J., et al. *Ice-breaking: Mitigating cold-start recommendation problem by rating comparison*. in *Twenty-Fourth International Joint Conference on Artificial Intelligence*. 2015.
454. Guo, G., J. Zhang, and D. Thalmann, *Merging trust in collaborative filtering to alleviate data sparsity and cold start.* Knowledge-Based Systems, 2014. 57: p. 57-68.
455. Ilarri, S., R. Trillo-Lado, and R. Hermoso. *Datasets for context-aware recommender systems: Current context and possible directions*. in *2018 IEEE 34th International Conference on Data Engineering Workshops (ICDEW)*. 2018. IEEE.
456. Pasinato, M., et al. *Generating synthetic data for context-aware recommender systems*. in *2013 BRICS Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence*. 2013. IEEE.
457. Wiegand, T., et al., *WHO and ITU establish benchmarking process for artificial intelligence in health.* The Lancet, 2019. 394(10192): p. 9-11.
458. Wilkinson, M.D., et al., *The FAIR Guiding Principles for scientific data management and stewardship.* Scientific data, 2016. 3(1): p. 1-9.
459. Holub, P., et al., *Enhancing reuse of data and biological material in medical research: from FAIR to FAIR-health.* Biopreservation and biobanking, 2018. 16(2): p. 97-105.
460. Meyffret, S., et al., *RED: a rich epinions dataset for recommender systems.* 2012.
461. McJones, P., *Eachmovie collaborative filtering data set.* DEC Systems Research Center, 1997. 249: p. 57.
462. Bennett, J. and S. Lanning. *The netflix prize*. in *Proceedings of KDD cup and workshop*. 2007. New York.
463. McFee, B., et al. *The million song dataset challenge*. in *Proceedings of the 21st International Conference on World Wide Web*. 2012.
464. Smyth, J.M. and A.A. Stone, *Ecological momentary assessment research in behavioral medicine.* Journal of Happiness studies, 2003. 4(1): p. 35-52.
465. Shiffman, S., A.A. Stone, and M.R. Hufford, *Ecological momentary assessment.* Annu. Rev. Clin. Psychol., 2008. 4: p. 1-32.
466. Bolman, C. and H. de Vries, *Psycho-social determinants and motivational phases in smoking behavior of cardiac inpatients.* Preventive Medicine, 1998. 27(5): p. 738-747.

467. Nierkens, V., K. Stronks, and H. de Vries, *Attitudes, social influences and self-efficacy expectations across different motivational stages among immigrant smokers: Replication of the pattern*. Preventive medicine, 2006. 43(4): p. 306-311.
468. Osada, K., et al., *Comparative Usability Analysis of Selected Data Entry Methods for Web Systems.* Cybernetics and Systems, 2020. 51(2): p. 192-213.
469. Wentzel, J., et al., *Mixing online and face-to-face therapy: how to benefit from blended care in mental health care.* JMIR mental health, 2016. 3(1): p. e9.
470. Bossen, D., et al., *A blended intervention for patients with knee and hip osteoarthritis in the physical therapy practice: development and a pilot study.* JMIR research protocols, 2016. 5(1): p. e32.
471. Metz, M., et al., *Effectiveness of a multi-facetted blended eHealth intervention during intake supporting patients and clinicians in Shared Decision Making: A cluster randomised controlled trial in a specialist mental health outpatient setting.* PloS one, 2018. 13(6): p. e0199795.
472. Talboom-Kamp, E.P., et al., *From chronic disease management to person-centered eHealth; a review on the necessity for blended care.* Clinical eHealth, 2018. 1(1): p. 3-7.
473. Mizgajski, J. and M. Morzy, *Affective recommender systems in online news industry: how emotions influence reading choices.* User Modeling and User-Adapted Interaction, 2019. 29(2): p. 345-379.
474. Abroms, L., P. Padmanabhan, and W. Evans, *Mobile phones for health communication to promote behavior change.* eHealth applications: Promising strategies for behavior change, 2012: p. 147-166.
475. Jódar-Sánchez, F., et al., *Using the social-local-mobile app for smoking cessation in the SmokeFreeBrain project: protocol for a randomized controlled trial.* JMIR research protocols, 2018. 7(12): p. e12464.
476. O'brien, H.L. and E.G. Toms, *Examining the generalizability of the User Engagement Scale (UES) in exploratory search.* Information Processing & Management, 2013. 49(5): p. 1092-1107.
477. Hussain, J., et al. *Adaptive user interface and user experience based authoring tool for recommendation systems*. in *International Conference on Ubiquitous Computing and Ambient Intelligence*. 2014. Springer.
478. Sezgin, E., et al., *A scoping review of patient-facing, behavioral health interventions with voice assistant technology targeting self-management and healthy lifestyle behaviors.* 2019.
479. Savage, R.L., A. Zekarias, and P. Caduff-Janosa, *Varenicline and abnormal sleep related events.* Sleep, 2015. 38(5): p. 833-837.
480. Ebbert, J.O., et al., *Varenicline and bupropion sustained-release combination therapy for smoking cessation.* Nicotine & Tobacco Research, 2009. 11(3): p. 234-239.
481. Filozof, C., M. Fernandez Pinilla, and A. Fernández-Cruz, *Smoking cessation and weight gain.* Obesity reviews, 2004. 5(2): p. 95-103.

482. Froom, P., S. Melamed, and J. Benbassat, *Smoking cessation and weight gain.* Journal of Family Practice, 1998. 46: p. 460-464.
483. Sazonov, E., P. Lopez-Meyer, and S. Tiffany, *A wearable sensor system for monitoring cigarette smoking.* Journal of studies on alcohol and drugs, 2013. 74(6): p. 956-964.
484. Echebarría, I.T.U., et al. *Monitoring smoking behaviour using a wearable acoustic sensor*. in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2017. IEEE.
485. Parate, A., et al. *Risq: Recognizing smoking gestures with inertial sensors on a wristband*. in *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*. 2014.
486. Senyurek, V.Y., et al., *A CNN-LSTM neural network for recognition of puffing in smoking episodes using wearable sensors.* Biomedical Engineering Letters, 2020: p. 1-9.
487. Patil, Y., S. Tiffany, and E. Sazonov. *Understanding smoking behavior using wearable sensors: Relative importance of various sensor modalities*. in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2014. IEEE.
488. Oard, D.W. and J. Kim. *Implicit feedback for recommender systems*. in *Proceedings of the AAAI workshop on recommender systems*. 1998. WoUongong.
489. Lian, J., et al. *xdeepfm: Combining explicit and implicit feature interactions for recommender systems*. in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2018.
490. Wray, T.B., et al., *Exploring the Use of Smartphone Geofencing to Study Characteristics of Alcohol Drinking Locations in High-Risk Gay and Bisexual Men.* Alcoholism: clinical and experimental research, 2019. 43(5): p. 900-906.
491. Hébert, E.T., et al., *An ecological momentary intervention for smoking cessation: the associations of just-in-time, tailored messages with lapse risk factors.* Addictive behaviors, 2018. 78: p. 30-35.
492. Heron, K.E. and J.M. Smyth, *Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments.* British journal of health psychology, 2010. 15(1): p. 1-39.
493. Businelle, M.S., et al., *An ecological momentary intervention for smoking cessation: evaluation of feasibility and effectiveness.* Journal of Medical Internet Research, 2016. 18(12): p. e321.
494. Velupillai, S., et al., *Recent advances in clinical natural language processing in support of semantic analysis.* Yearbook of medical informatics, 2015. 24(01): p. 183-193.
495. Stephens, T.N., et al., *Feasibility of pediatric obesity and prediabetes treatment support through Tess, the AI behavioral coaching chatbot.* Translational behavioral medicine, 2019. 9(3): p. 440-447.
496. Gabrielli, S., et al., *A Chatbot-Based Coaching Intervention for Adolescents to Promote Life Skills: Pilot Study.* JMIR Human Factors, 2020. 7(1): p. e16762.

497. Piao, M., et al., *Use of the Healthy Lifestyle Coaching Chatbot App to Promote Stair-Climbing Habits Among Office Workers: Exploratory Randomized Controlled Trial.* JMIR mHealth and uHealth, 2020. 8(5): p. e15085.
498. Fadhil, A., G. Schiavo, and Y. Wang, *CoachAI: A Conversational Agent Assisted Health Coaching Platform.* arXiv preprint arXiv:1904.11961, 2019.
499. Hauser-Ulrich, S., et al., *A smartphone-based health care chatbot to promote self-management of chronic pain (SELMA): pilot randomized controlled trial.* JMIR mHealth and uHealth, 2020. 8(4): p. e15806.
500. Kamita, T., et al., *A chatbot system for mental healthcare based on SAT counseling method.* Mobile Information Systems, 2019.
501. Abd-alrazaq, A.A., et al., *An overview of the features of chatbots in mental health: A scoping review.* International Journal of Medical Informatics, 2019. 132: p. 103978.
502. Bendig, E., et al., *The next generation: chatbots in clinical psychology and psychotherapy to foster mental health–a scoping review.* Verhaltenstherapie, 2019: p. 1-13.
503. Provoost, S., et al., *Validating automated sentiment analysis of online cognitive behavioral therapy patient texts: An exploratory study.* Frontiers in psychology, 2019. 10: p. 1065.
504. Griol, D., J.M. Molina, and A. Sanchis, *A multimodal conversational coach for active ageing based on sentient computing and m-health.* Expert Systems, 2020. 37(2): p. e12454.
505. Fadhil, A., et al. *The effect of emojis when interacting with conversational interface assisted health coaching system*. in *Proceedings of the 12th EAI international conference on pervasive computing technologies for healthcare*. 2018.
506. Burke, R., et al. *Detecting profile injection attacks in collaborative recommender systems*. in *The 8th IEEE International Conference on E-Commerce Technology and The 3rd IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services (CEC/EEE'06)*. 2006. IEEE.
507. Williams, C.A., B. Mobasher, and R. Burke, *Defending recommender systems: detection of profile injection attacks.* Service Oriented Computing and Applications, 2007. **1**(3): p. 157-170.
508. Mobasher, B., R. Burke, and J.J. Sandvig. *Model-based collaborative filtering as a defense against profile injection attacks*. in *AAAI*. 2006.
509. Pitsilis, G. and S.J. Knapskog, *Social Trust as a solution to address sparsity-inherent problems of Recommender systems.* arXiv preprint arXiv:1208.1004, 2012.
510. Massa, P. and P. Avesani. *Trust-aware collaborative filtering for recommender systems*. in *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*. 2004. Springer.
511. Massa, P. and P. Avesani. *Trust-aware recommender systems*. in *Proceedings of the 2007 ACM conference on Recommender systems*. 2007.

512. Akter, S., J. D'Ambra, and P. Ray, *Trustworthiness in mHealth information services: an assessment of a hierarchical model with mediating and moderating effects using partial least squares (PLS).* Journal of the American Society for Information Science and Technology, 2011. 62(1): p. 100-116.
513. Papagelis, M., et al. *Incremental collaborative filtering for highly-scalable recommendation algorithms*. in *International Symposium on Methodologies for Intelligent Systems*. 2005. Springer.
514. Smith, B. and G. Linden, *Two decades of recommender systems at Amazon. com.* Ieee internet computing, 2017. 21(3): p. 12-18.
515. Sittig, D.F., E. Belmont, and H. Singh. *Improving the safety of health information technology requires shared responsibility: It is time we all step up*. in *Healthcare*. 2018. Elsevier.
516. Nebeker, C., J. Torous, and R.J.B. Ellis, *Building the case for actionable ethics in digital health research supported by artificial intelligence.* BMC medicine, 2019. 17(1): p. 137.
517. George, C., D. Whitehouse, and P. Duquenoy, *eHealth: legal, ethical and governance challenges*. 2012: Springer Science & Business Media.
518. Kargl, F., et al. *Security, privacy and legal issues in pervasive ehealth monitoring systems*. in *2008 7th International Conference on Mobile Business*. 2008. IEEE.
519. Sahama, T., L. Simpson, and B. Lane. *Security and Privacy in eHealth: Is it possible?* in *2013 IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013)*. 2013. IEEE.

R

Summary

Smoking has several harmful effects on our health and affects our organs, leading to the incidence of many life-threatening diseases. Furthermore, it is one of the most preventable causes of death. Despite its detrimental effect on our health, quitting smoking is challenging due to the tobacco addictive chemicals and humans' psychological dependency on it. Nonetheless, there are different approaches to support people willing to stop smoking. One method is eHealth computer tailoring, which helps personalize feedback given to smokers based on psychological models of behavioral change based on pre-defined if-then-else rules. These methods showed to generate positive results in terms of high abstinence rates and cost-effectiveness. However, new innovative solutions are available to improve the eHealth methods for smoking cessation further. One of those methods is related to recommender systems technology. Recommender systems are AI algorithms that can select the most relevant item (such as a piece of text, book, movie, or product) from a set of items for each user. Depending on the type of recommender system, relevance is determined considering different methods and variables. A commonly used method for calculating relevance is the "collective intelligence" approach. This approach uses algorithms to generate a user profile for each user (e.g., using demographic variables) and calculate how relevant a specific item is based on the given relevance of that item for users with similar user profiles. These systems can learn from user feedback over time in that the users rate the relevance of the recommended items, which helps train the system for making future recommendations. For decades, the scientific community has explored the relevance of these systems in other fields such as leisure (movie recommendations on Netflix) and e-commerce (product recommendations on Amazon). Due to their potential and proven effectiveness in other fields but limited application in the healthcare sector, which began only a few years ago, studying how these systems can be applied for smoking cessation is crucial.

In Chapter 2 of this dissertation, we have conducted a scoping review to assess the existing knowledge and research gaps using recommender systems in healthcare, also known as health recommender systems (HRSs). We assessed their technical and healthcare aspects through this review. Based on its results, we then generated a new taxonomy for these types of systems. Next, we provided a detailed description of a health recommender system (HRS) design process with collective intelligence grounded in behavioral science for smoking cessation using the I-Change model as an example. In Chapter 3, we explained all the steps and the system design, including algorithm components, messages creation, and user interface design, to help interested stakeholders better understand such systems, which would provide inspiration and a basis for future studies. Furthermore, we performed an assessment study to test the created HRS using collective intelligence in a real-world setting with a follow-up period of six months. The control condition was a simpler version of the created HRS in this assessment, except for the collective intelligence component. In Chapter 4, we reported the protocol of this study and analyzed the actual results regarding the appreciation, engagement, dropouts, and smoking abstinence generated by the system (Chapter 5).

Chapter 1 provides a general introduction to the problem associated with smoking cessation. First, it introduces different existing support approaches, focusing on the ones related to behavioral change and their application in computer-tailored interventions. Then, it presents the recommender system technology and its different types that exist as an option for facilitating computer-tailored interventions. Further, it highlights the appreciation and engagement metrics, which are the factors that complement abstinence for intervention success.

Chapter 2 contains a scoping review that provides an analysis of the state-of-the-art HRS, identifying the research gaps and the elements that should be improved when applying this technology to the healthcare sector. From this study, we identified that the collaborative filtering technique was the most-used information filtering method. However, it was also observed that there is a lack of applying behavioral change theories and factors in HRS studies. Furthermore, these studies neither implemented the principles of tailoring nor assessed their (cost)-effectiveness. Therefore, a taxonomy was proposed to facilitate consistent classification and better comprehension of these systems. This taxonomy included the domain of the study (e.g., the type of population, country, therapeutic area), the methodology and procedures of the study (the duration, number of users, outcomes), health behavior change factors (e.g., self-efficacy, social influence, attitudes), and the technical aspects required to understand the algorithm (e.g., recommendation technology, profile generation techniques).

Chapter 3 provides a multidisciplinary and comprehensive description of the design process of an HRS for supporting smoking cessation that uses collective intelligence in combination with the I-Change behavioral change model. This detailed description contributed to help reveal the process of how an HRS can be built to support behavioral change interventions. This process had not been disclosed in detail before, and this lack of transparency can act as a barrier for behavioral change researchers in using HSR technology. The new system was built based on a previous HRS that utilized a mobile app to support smokers trying to stay abstinent by sending them motivational messages. First, we identified the areas that needed improvements based on the app's usage data. Then, we implemented relevant changes to our new system design (e.g., increasing the granularity of the possible user feedback from three options to five options). Our final mobile app was supposed to be more streamlined and usable than the first version. The generated HRS was a hybrid algorithm with a knowledge-based step and a collaborative-filtering step in cascade. It used 58 variables to compute the similarity formula for choosing recommendations; from the total, 47 were related to the determinants of the

I-Change model. Altogether, 331 motivational messages were created, and ten different health communication methods were considered for their design.

Chapter 4 explains the protocol to be followed to assess the system created in Chapter 3. This protocol included the description of a clinical pilot and a public pilot. We used the latter one to analyze the HRS in this dissertation.

Chapter 5 presents, discusses and reflects on the results obtained from the public pilot. The public pilot was a double-blinded experiment. Those smokers who can read English or Mandarin and download a mobile app from the Internet were eligible to participate. After creating their account and answering questions relevant to their user profile (e.g., name, age, gender, level of addiction, and motivation to quit), they can set a quitting day to start receiving personalized motivational text messages via the mobile app. Smokers were randomly allocated to the group where such messages were generated by the new HRS, which was described in Chapter 3, or to the group associated with a simpler version of the algorithm, without collective intelligence (using only the knowledge-based step), selected and sent these messages. A total of 371 participants were eligible to be part of the study analysis. Smokers were followed up for six months, starting from their quitting day, and were asked weekly about their smoking abstinence through a voluntary question in the app.

Moreover, we measured their message appreciation and engagement. The attributes (factors) considered as possible indicators of differences in the study outcomes included the motivation to quit, nicotine dependence, age, gender, and completion of the extended user profile questionnaire. They were studied as potential covariates in the statistical analysis. No statistically significant differences were found neither for the analysis on available data of the 7D-PP abstinence averaging the abstinence reports across the study nor for the penalized imputation analysis of both the 7D-PP abstinence averaging the abstinence reports across the study and the 7D-PP considering only the last available abstinence report. However, the analysis on available data for the 7D-PP considering only the last available abstinence report showed lower abstinence rates in the HRS using collective intelligence. Also, the results showed that the HRS using collective intelligence did not have statistically significant differences for message appreciation, number of rated messages, and number of quitting attempts. However, the collective intelligence algorithm performed worse regarding the number of abstinence reports and active days. The sub-group analysis showed that the completion of the extended user profile did significantly impact the engagement of the participants reducing the number of dropouts in both groups and increasing the number of quitting attempts in participants who received messages selected with the collective intelligence.

Finally, Chapter 6 provides a general discussion of the main findings and conclusions of all the studies presented in this dissertation (from chapters 2–5). It also contains the main methodological considerations for this dissertation, such as the strengths and limitations, risks, reflections for practice, and the impact of this thesis on the scientific community. In conclusion, the studies presented in this dissertation showed that although HRSs are gaining traction in the healthcare sector, they are still novel, with underreported details and suboptimal application, as they do not take advantage of the behavioral change theories. However, we have shown that they can be used as an alternative approach to traditional tailoring for behavioral interventions by embedding behavioral science in the design of these emergent systems. We compared the HRSs with and without collective intelligence technology for a trial for smoking cessation, measuring their performance in real-life conditions. The results showed that despite showing some positive results in terms of engagement – number of quitting attempts - when completing the extended user profile, the HRS using collective intelligence did not manage to improve smoking behavior, appreciation, and engagement compared to the other HRS. In addition, some of the engagement and abstinence metrics led to worse results. Furthermore, although we achieved better smoking cessation outcomes than quitting cold turkey or with brief clinician advice, our HRS did not improve the abstinence rates achieved by other approaches in smoking cessation, such as traditional computer tailoring. Further, it is still unclear why the theoretical potential of collective intelligence did not provide the expected benefits in our study. Therefore, future research is needed to find out how HRS-based interventions, using or not using the collective intelligence technology, can be improved to achieve better outcomes in terms of behavioral change.

Acknowledgements (Agradecimientos)

Acknowledgements

I would not have been able to do this thesis without many people's direct and indirect support and contribution. I want to thank all of them, of whom I would like to mention some in particular.

First of all, I would like to thank my thesis supervisory team from the University of Seville and Maastricht University. Starting with the University of Seville representatives, Antón Civit-Balcells trusted me to lead the European SmokeFreeBrain project tasks assigned to the University of Seville. He gave me the freedom to manage the project budget and the actual project tasks execution approach. Without such space, I would not have grown my coordination and interpersonal skills as I have done, which I consider crucial in any researcher. He also facilitated that I could develop myself as a more business-oriented researcher with a multidisciplinary perspective. This thesis is the result of it.

Also, from the University of Seville, my supervisor, business partner, and friend Luis Fernández-Luque taught me most of what I know about eHealth. His daily guidance in research methods, extensive knowledge and contacts in the eHealth domain, and entrepreneurial vision greatly impacted me. I thoroughly enjoyed his advice and felt he was there for me regardless of the distance and time difference that separated us. Furthermore, his leadership style made me push my limits and achieve goals that I would have never imagined.

From Maastricht University, Francine Schneider was my reliable source of good pieces of advice, inspiration, positiveness, and practical research guidance. She was my go-to person when everything seemed to lead to dead-ends in my Ph.D. journey. I admire her capability of quickly understanding the technology behind my research, helping me identify its strengths and limitations, and adding the needed behavioral science perspective that I miss from my original background as an engineer. I do not know how I would have continued my research without her excellent disposition and mentoring under challenging moments. Thus, I want to express my deepest gratitude to her.

The other cornerstone of my support from Maastricht University was my Ph.D. promotor, Hein de Vries. Hein was the senior strategist and mentor who helped me in my journey to writing this thesis. His previous work was part of the foundations of this thesis, and I am honored he accepted me as a Ph.D. student, although my background was in software engineering, not health sciences. As a result, Hein had to be patient with me, bear with the limitations of the project I was using to get data and encourage me to find alternative solutions. In addition, he provided many rounds of critical feedback in the scientific publications we co-authored. I appreciate every single one of those comments because they were the difficult-to-swallow pills I needed to advance as a researcher. I cannot conclude this acknowledgment to Hein without mentioning that, on top of his superb leadership in the academic field, he showed extraordinary empathy for me as a person as a whole, and not just as a Ph.D. student, for which I will always be thankful.

I also want to stress that my four supervisors provided me with the comprehension, flexibility, and support I needed to complete this thesis while working in a company full-time. As a result, neither could I devote all my time to this thesis nor even have a consistent minimum weekly dedication. My job made me advance in this thesis slowly when I had heavy workloads in my company. I want to express my gratitude for their patience with this situation.

I acknowledge this thesis would not have been possible without the European Project SmokeFreeBrain. I want to thank all the people who ideated the project, the European Commission who accepted and funded it, and especially the coordinator of the project Panagiotis Bamidis from the Aristotle University of Thessaloniki. Special thanks to José Luis Sevillano, who was also part of the project from the University of Seville, and who kick-started my passion for recommender systems. As part of the development team in the SmokeFreeBrain project, who implemented the software used in the research of this thesis, I want to thank Francisco Luna-Perejón from the University of Seville. He was my sidekick in the development of the software solution. I want to state his devotion, can-do attitude, and extraordinary skills coding the mobile app and algorithm following my specifications with tight deadlines. Also, V ictor Alonso from Salumedia Labs helped me automate the extraction from the database and clean the collected data. Without his ‘magic’ scripts, the data analysis conducted in this thesis would just not have been possible.

I would also like to extend my special thanks to the International Center for Health Information Technology (ICHIT) team from the Taipei Medical University in Taiwan: Shwetambara Malwade, Shabbir Syed-Abdul, and Yu-Chuan Jack Li. They hosted me there for two months to deploy the software and launch the study. In addition, they contributed to the subsequent research and dissemination of the results. The logistic, economic, and scientific support provided by ICHIT played an important role in the completion thesis, and I am very thankful for their contributions.

The core chapters of this thesis would not have been possible without the work and inputs of my other co-authors Octavio Rivera, Chien-Tien Su, Wei-Li Jeng, Dimitris Spachos, Claudio Amaya, and Kiko Núñez Benjumea. I am really grateful to all of you for the time you spent and your good disposition to collaborate and create impactful research. Also, I want to express my most sincere thanks to Math Candel, who provided magnificent and diligent support for the statistical analysis of the results in the last sprint of this journey.

Another group of people to whom I want to thank as part of my research training experience is the team from the Virgen del Rocío University Hospital, represented by Francisco Ortega Ruiz, Laura Carrasco-Hernández, Marco Mesa González, Jesús Moreno-Conde, Francisco Jódar-Sanchez, and Carlos Luis Parra-Calderón. Their work in the SoLoMo study within the SmokeFreeBrain project made a solid foundation for this thesis.

I will also be eternally grateful to María Soriano and Zafira Castaño for their commitment to IMFAHE and with Carlos Castro-González, my mentor at IMFAHE, for awarding me with a fellowship to do a research stay at the Department of Clinical Informatics in Harvard Medical School Teaching Hospital for three months. There, I met Yuri Quintana and Charles Safran, to whom I want to express my gratitude. They were exceptional professionals and taught me new perspectives, methodologies, and approaches to better comprehend the e-health field in the USA. Also, I want to thank the ENJECT COST Action project, led by the University College Dublin, as an economic supporter of my research stays and training. The workshops it organized, and the research network it created were an excellent complement to my Ph.D. training.

Last but not least, I want to thank my close friends and relatives. Despite not directly contributing to this thesis, they emotionally supported me in different ways through these exciting but challenging years. Among all of them, my uncle Pedro Fraile Balbín was one of the first people to encourage me to undertake a Ph.D. Furthermore, he was always supportive and followed up on my progress with great interest. Finally, my warmest thanks go to my father, Santiago Hors Mocholi, and my mother, Rosario Fraile Balbín. They always were there for me, giving me their invaluable care, love, and support these years. In addition, they taught me the values and principles of effort, perseverance, and honesty, which I did my best to apply throughout the execution of this thesis.

Agradecimientos

No habría podido realizar esta tesis sin el apoyo y la contribución directa e indirecta de muchas personas. Quiero dar las gracias a todas ellas, de las que me gustaría mencionar a algunas en particular.

En primer lugar, quiero dar las gracias a mi equipo de directores de tesis de la Universidad de Sevilla y de la Universidad de Maastricht. Empezando por los representantes de la Universidad de Sevilla, Antón Civit-Balcells confió en mí para dirigir las tareas del proyecto europeo SmokeFreeBrain asignadas a la Universidad de Sevilla. Me dio libertad para gestionar el presupuesto del proyecto y el enfoque de ejecución de las tareas del mismo. Sin ese espacio, no habría crecido como lo he hecho mis habilidades de coordinación e interpersonales, que considero cruciales en cualquier investigador. También facilitó que pudiera desarrollarme como un investigador más orientado a la empresa y con una perspectiva multidisciplinar. Esta tesis es el resultado de ello.

También, desde la Universidad de Sevilla, mi supervisor, socio y amigo Luis Fernández-Luque me enseñó la mayor parte de lo que sé sobre eSalud. Su orientación diaria en métodos de investigación, sus amplios conocimientos y contactos en el ámbito de la eSalud y su visión empresarial me impactaron enormemente. Disfruté mucho de sus consejos y sentí que estaba a mi lado a pesar de la distancia y la diferencia horaria que nos separaba. Además, su estilo de liderazgo me hizo superar mis límites y alcanzar metas que nunca hubiera imaginado.

De la Universidad de Maastricht, Francine Schneider fue mi fuente fiable de buenos consejos, inspiración, positivismo y orientación práctica para la investigación. Fue la persona a la que recurría cuando todo parecía llevarme a un callejón sin salida en mi doctorado. Admiro su capacidad para entender rápidamente la tecnología que hay detrás de mi investigación, ayudándome a identificar sus puntos fuertes y sus limitaciones, y añadiendo la necesaria perspectiva de las ciencias del comportamiento que echaba de menos por mi formación original como ingeniero. No sé cómo habría continuado mi investigación sin su excelente disposición y tutoría en los momentos difíciles. Por ello, quiero expresarle mi más profunda gratitud.

La otra piedra angular de mi apoyo de la Universidad de Maastricht fue mi promotor del doctorado, Hein de Vries. Hein fue el estratega y mentor principal que me ayudó en mi camino para escribir esta tesis. Su trabajo anterior formó parte de los cimientos de esta tesis, y me siento honrado de que me aceptara como estudiante de doctorado, aunque mi formación fuera en ingeniería de software, no en ciencias de la salud. Como resultado, Hein tuvo que ser paciente conmigo, soportar las limitaciones del proyecto que estaba utilizando para obtener datos y animarme a encontrar soluciones alternativas. Además, aportó muchas rondas de comentarios críticos en las publicaciones científicas de las que fuimos coautores. Agradezco cada uno de esos comentarios porque fueron las píldoras difíciles de tragar que necesitaba para avanzar como investigador. No puedo concluir este reconocimiento a Hein sin mencionar que, además de su magnífico liderazgo en el ámbito académico, mostró una extraordinaria empatía por mí como persona en su conjunto, y no sólo como estudiante de doctorado, lo que siempre le agradeceré.

También quiero destacar que mis cuatro supervisores me proporcionaron la comprensión, la flexibilidad y el apoyo que necesitaba para completar esta tesis mientras trabajaba en una empresa a tiempo completo. Como resultado, no pude dedicar todo mi tiempo a esta tesis ni siquiera tener una dedicación mínima semanal constante. Mi trabajo me hizo avanzar en esta tesis lentamente cuando tenía mucha carga de trabajo en mi empresa. Quiero expresar mi agradecimiento por su paciencia con esta situación.

Reconozco que esta tesis no habría sido posible sin el proyecto europeo SmokeFreeBrain. Quiero dar las gracias a todas las personas que idearon el proyecto, a la Comisión Europea que lo aceptó y financió, y especialmente al coordinador del proyecto Panagiotis Bamidis de la Universidad Aristóteles de Tesalónica. Un agradecimiento especial a José Luis Sevillano, que también formó parte del proyecto desde la Universidad de Sevilla, y que inició mi pasión por los sistemas de recomendación. Como parte del equipo de desarrollo del proyecto SmokeFreeBrain, que implementó el software utilizado en la investigación de esta tesis, quiero agradecer a Francisco Luna-Perejón de la Universidad de Sevilla. Él fue mi compañero en el desarrollo de la solución de software. Quiero dejar constancia de su dedicación, su actitud positiva y sus extraordinarias habilidades para codificar la aplicación móvil y el algoritmo siguiendo mis especificaciones con plazos muy ajustados. Además, Víctor Alonso de Salumedia Labs me ayudó a automatizar la extracción de la base de datos y a limpiar los datos recogidos. Sin sus códigos “mágicos”, el análisis de datos realizado en esta tesis no habría sido posible.

Además quiero destacar mi especial agradecimiento al equipo del Centro Internacional de Tecnología de la Información Sanitaria (ICHIT) de la Universidad Médica de Taipei, en Taiwán: Shwetambara Malwade, Shabbir Syed-Abdul y Yu-Chuan Jack Li. Me acogieron allí durante dos meses para desplegar el software y lanzar el estudio. Además, contribuyeron a la investigación posterior y a la difusión de los resultados. El apoyo logístico, económico y científico proporcionado por el ICHIT desempeñó un papel importante en la realización de la tesis, y estoy muy agradecido por sus contribuciones.

Los capítulos centrales de esta tesis no habrían sido posibles sin el trabajo y las aportaciones de mis otros coautores Octavio Rivera, Chien-Tien Su, Wei-Li Jeng, Dimitris Spachos, Claudio Amaya y Kiko Núñez Benjumea. Les agradezco mucho a todos el tiempo que han dedicado y su buena disposición para colaborar y crear una investigación de impacto. Asimismo, quiero expresar mi más sincero agradecimiento a Math Candel, quien brindó un magnífico y diligente apoyo para el análisis estadístico de los resultados en el último sprint de esta jornada.

Otro grupo de personas a las que quiero dar las gracias como parte de mi experiencia de formación en investigación es el equipo del Hospital Universitario Virgen del Rocío, representado por Francisco Ortega Ruiz, Laura Carrasco-Hernández, Marco Mesa González, Jesús Moreno-Conde, Francisco Jódar-Sánchez y Carlos Luis Parra-Calderón. Su trabajo en el estudio SoLoMo dentro del proyecto SmokeFreeBrain supuso una sólida base para esta tesis.

También estaré eternamente agradecido a María Soriano y Zafira Castaño por su compromiso con el programa IMFAHE y con Carlos Castro-González, mi mentor en IMFAHE, por concederme una beca para realizar una estancia de investigación en el Departamento de Informática Clínica del Harvard Medical School Teaching Hospital durante tres meses. Allí conocí a Yuri Quintana y Charles Safran, a quienes también quiero expresar mi gratitud. Fueron unos profesionales excepcionales y me enseñaron nuevas perspectivas, metodologías y enfoques para comprender mejor el campo de la salud digital en los Estados Unidos. También quiero dar las gracias al proyecto ENJECT COST Action, dirigido por el University College de Dublín, por haber apoyado económicamente mis estancias de investigación y mi formación. Los talleres que organizó y la red de investigación que creó fueron un excelente complemento para mi formación de doctorado.

Por último, pero no menos importante, quiero dar las gracias a mis amigos y familiares más cercanos. A pesar de no haber contribuido directamente a esta tesis, me apoyaron emocionalmente de diferentes maneras a lo largo de estos apasionantes pero difíciles años. Entre todos ellos, mi tío Pedro Fraile Balbín fue una de las primeras personas que me animó a realizar un doctorado. Además, siempre me apoyó y siguió mis progresos con gran interés. Por último, mi más sincero agradecimiento a mi padre, Santiago Hors Mocholi, y a mi madre, Rosario Fraile Balbín. Siempre estuvieron a mi lado, dándome su inestimable cuidado, cariño y apoyo durante estos años. Además, me enseñaron los valores y principios del esfuerzo, la perseverancia y la honestidad, que hice lo posible por aplicar a lo largo de la realización de esta tesis.

Curriculum Vitae

Curriculum Vitae

Santiago Hors-Fraile was born on November 19th, 1987, in Seville, Spain. After completing secondary school (Scientific-Technological, and Natural Sciences and Health Baccalaureate) in 2005 at the Institute of Secondary Education (ISE) San Antonio María Claret (Seville, Spain), he started a bachelor in Technical Engineering in Computer Systems at the School of Computer Science of the University of Seville in Spain (2005-2008). Then, he dedicated two additional years to doing an MSc in Computer Science at the University of Seville (2009-2010). In 2011, he studied another MSc in Computational & Software Techniques in Engineering, specializing in Software Engineering for Technical Computing at Cranfield University in the UK. In 2012, he started working at the eHealth company Salumedia Labs as a Digital Health Solutions Manager. During the following years, he complemented his work in the eHealth and mHealth fields with continuous training in areas such as entrepreneurship (University of Seville, 2013), gamification (Engagement Alliance, 2013), and healthcare digital marketing (ESIC, 2014). During his job in Salumedia Labs until 2015, he designed and coordinated digital health solutions for pharmaceutical companies, research institutes, and worldwide health organizations.

In 2015, he undertook a Medical and Industrial Applications Computer Technology Ph.D. at the University of Seville. In 2016, he took a course on Health communication and Health promotion at Maastricht University (Netherlands). He decided to add a behavioral science dimension to his Ph.D. in collaboration with the Department of Health Promotion of Maastricht University. The results of this double Ph.D. between the University of Seville and Maastricht University are presented in this dissertation. During the seven years of his Ph.D., he did research stays at University College Dublin (Ireland), Taipei Medical University (Taiwan), and Harvard Medical School (USA). In addition, he disseminated his work and research at different conferences around the world. He also run specialized seminars at both academic institutions such as IMF and Loyola University and non-academic entities such as the European Cancer Leagues association and the leading Middle-East start-up Droobi Health. Currently, he holds the role of Director of Product at Adhera Health, Inc., where he is applying the principles of using the combination of artificial intelligence and behavioral to the product of the company and validating them in different health contexts within international pilot studies.

Publication list

Publications in this thesis

Hors-Fraile, S., et al., *Analyzing recommender systems for health promotion using a multidisciplinary taxonomy: a scoping review.* International journal of medical informatics, 2018. 114: p. 143-155.

Hors-Fraile, S., et al., *A recommender system to quit smoking with mobile motivational messages: study protocol for a randomized controlled trial.* Trials, 2018. 19(1): p. 618.

Hors-Fraile, S., et al., *Opening the black box: Explaining the process of basing a health recommender system on the I-Change behavioral change model.* IEEE Access, 2019. 7: p. 176525-176540.

Hors-Fraile, S., et al., *Applying collective intelligence in health recommender systems for smoking cessation: a comparison trial.* (Submitted to UMUAU SI: Recommender Systems for Health and Wellbeing)

Publications in the same line of research

Hors-Fraile, S., et al., *Design of two combined health recommender systems for tailoring messages in a smoking cessation app.* arXiv preprint arXiv:1608.07192, 2016.

Schäfer, H., et al. *Towards health (aware) recommender systems*. in *Proceedings of the 2017 international conference on digital health*. 2017. ACM.

Jódar-Sánchez, F., et al., *Using the social-local-mobile app for smoking cessation in the SmokeFreeBrain project: protocol for a randomized controlled trial.* JMIR research protocols, 2018. 7(12): p. e12464.

Hors-Fraile, S., et al., *Tailoring motivational health messages for smoking cessation using an mHealth recommender system integrated with an electronic health record: a study protocol.* BMC public health, 2018. 18(1): p. 1-10.

Jódar-Sánchez, F., et al., *Social-Local-Mobile intervention for supporting smoking cessation in the SmokeFreeBrain project: a study protocol of a 12-month randomized open-label parallel-group trial.*

Fernandez-Luque, L., et al. *Implementation and feasibility study of a tailored health education bot in Telegram for mothers of children with obesity and overweight.* Qatar Foundation Annual Research Conference Proceedings Volume 2018 Issue 2 (Vol. 2018, No. 2, p. HBPD506). Hamad bin Khalifa University Press

Carrasco-Hernandez, L., et al., *A mobile health solution complementing psychopharmacology-supported smoking cessation: Randomized controlled trial.* JMIR mHealth and uHealth, 2020. 8(4): p. e17530.

Beristain Iraola, A., et al. *Virtual Coaching for Older Adults at Home using SMART Goal Supported Behavior Change*. in *Companion Publication of the 2020 International Conference on Multimodal Interaction*. 2020.

Cheung, K.L., S. Hors-Fraile, and H. de Vries, *How to use the Integrated-Change Model to design digital health programs*, in *Digital Health*. 2021, Elsevier. p. 143-157.

Cheung, K.L., S. Hors-Fraile, and H. de Vries, *Illustration of tailored digital health and potential new avenues*, in *Digital Health*. 2021, Elsevier. p. 159-169.