Framework to Evaluate Emerging Systems Designed to Health Field

A. Laflor-Hernandez[1], M. Vazquez-Briseno[1], J.I. Nieto-Hipolito[1], R. Conte[2], A. Garcia-Berumen[3], J.D. Sanchez - Lopez, E. Gutierrez[1]

[1]Universidad Autonoma de Baja California, FIAD, Ensenada, B.C., Mexico

[2]Centro de Investigacion cientifica y de estudios superiores de Ensenada, Ensenada, B.C., Mexico

[3]Instituto Tecnologico de Sonora, Ciudad Obregon, Sonora, Mexico

arturo.laflor@uabc.edu.mx, mabel.vazquez@uabc.edu.mx, jnieto@uabc.edu.mx, egutierrez@uabc.edu.mx, agberumen@itson.edu.mx, conte@cicese.mx

Abstract

In recent years, several information and communication technology systems have emerged as tools to improve sleep quality. Research reveals that poor sleep quality may produce irritability and deficits in performance, concentration, and learning ability in the short term, and is associated with chronic disease in the long term. ICT proposals range from the old Polysomnography (PSG) to innovative systems, such as wearable devices, smartphone applications, and suites of sensors embedded in the users' environment. Since these technological developments concern a health issue, they have raised important questions regarding their reliability and the level of rigor of the evaluations to which they are submitted. We found that some of the emerging systems that we studied, do not meet the requirements that health science demands to be accepted as clinical tools. The rationale behind this apparent weakness is explained with arguments from the field of evaluations for health interventions and evaluation of technological developments. We propose a framework to evaluate this kind of systems through appropriate scientific methods that provide valuable information to the research. These methods must be performed while designs mature and the feasibility of rigorous evaluations became appropriate.

Keywords: Sleep Quality, Technology Evaluation, Mobile Health, Framework, Emerging Mobile Applications.

Introduction

Many information and communication technology (ICT) systems have been proposed to help people to improve their sleep quality. Reasons have to do with the increasing knowledge about the impact of sleep on quality of life [1]. In the short term, poor sleep quality can cause irritability, as well as deficits in performance, concentration, learning ability, and decision-making processes [2, 3]. Poor sleep quality in the long term has been closely related to chronic diseases, such as diabetes [4], cardiovascular disease [5] depression [6], Alzheimer disease, and other mental health disorders [7].The most relevant concepts in this field of research and the relationships among them were synthesized in a mental map (See Fig. 1) explained in Subsection 2. Literature in this work covers technology for sleep from the old PSG, through actigraphy, to innovative technology that had been emerged as tools to monitor sleep, diagnose sleep disorders and persuade people to change habits to improve their own sleep quality. Emerging technologies include mobile applications involving environmental sensors, wearable monitoring devices complemented with mobile applications, desktop software and sensor networks, among others. To date, only two of the above technological developments have been approved by The American Academy of Sleep Medicine (AASM) and The American Sleep Disorders Association (ASDA) as reliable tools in clinical use. For many years, the PSG has been considered the gold standard test in the diagnosis of sleep disorders, and recently, actigraphy was approved as a reliable tool to assess sleep disorders [8]. Remaining devices and applications have no clinical recognition.

After a review on smartphone applications designed to attend sleep issues, [9] concludes that except for simple questionnaires, no existing sleep-related application available for smartphones is based on scientific evidence.

In Section 5, we discuss the rationale behind the Behars' conclusion. Additionally we explain our perspective of the problem based in works published in the field of assessment of technology and assessment of health interventions. Being consistent with the arguments in the Section 5 we propose in Section 4, a conciliatory answer to respond the rhetoric affirmation of Behar. We conclude the Section 4 with a framework to evaluate the early stages of design of technology addressed to sleep health issues. Finally, in Section 3 we show the results of evaluate eighteen systems in this field of research through the criteria of our own framework.

Materials and Methods

We explore four databases to find literature on technology monitoring sleep, sleep quality and sleep hygiene. We found 1185 papers (151 from ACM, 69 from IEEE, 125 from SpringerLink, and 840 from ScienceDirect). After reading the title, the abstract and consulting the index citation level in Scopus Database, a total of 60 papers were included. In the second selection, full texts were analyzed. Articles testing electronic devices and those focused on medical field rather than to computer science field, were excluded. After applying these filters, 18 papers were selected to make the analysis. Fig. 1 shows a mental map representing the relationship between the most relevant concepts of technological systems designed to address sleep problems. It was found that the systems are mainly aimed at addressing three aspects related to sleep: a) Sleep patterns, b) sleep disorders, and c) sleep hygiene habits (SHH). These aspects are approached from three different perspectives but often are used together to achieve the desired objectives: a) Perspective of monitoring, b) Perspective of diagnosis and c) Perspective of persuasion.

Monitoring systems are aimed to observe sleep patterns, such as positions or complex movements [10], and they typically record data that physicians can see in real time or as required. Some monitoring systems can be used to determine good sleep health or emit warnings to a specialist as a signal of that further studies are needed. Diagnostic systems provide results that physicians can use when a sleep disorder is being addressed. These results offer desirable benefits, especially the possibility of reducing hospital-related costs [11]. For instance, Sleep Apnea Monitor helps to determine if a patient has sleep apnea, before a physician orders advanced and expensive sleep tests [12].Persuasive systems are intended to take the role of a coach based on proven psychological theories. For instance [13] uses Cognitive Behavioral Therapy (CBT) arguing that many of the problems that are related to beliefs, attitudes and emotions, can be treated by means of therapies, avoiding the necessity of pharmacological treatments. Insomnia is one of sleep disorders that on repeated occasions has its' origin in emotional factors or lifestyle habits that can be addressed and corrected through appropriate therapy. Persuasive systems involve strategies such as serious games, social activities with group challenges, and timely reminders to motivate people to improve sleep habits [1]. The purpose of changing habits in adults is questionable [14], however, researchers such as [15] and [16] have founded that providing users information about their sleep behavior could motivate them to engage in healthier sleep habits and could effectively promote changes in user activity.

The evaluation of the systems determines the maturity of the system. There are systems clinically approved and those that are in early stages of design so called emerging systems. Two aspects are considered in the evaluation process: the efficiency of the system to achieve its proposed goal, and the features of design that users qualify as desirables or undesirables. From the perspective of the users, the most desirable feature in a system designed to help people to sleep better is the unobtrusiveness.

Results and Discussion

The health area is the most rigorous area to evaluate protocols, procedures, human interventions and so forth. Therefore, all technological developments designed with the intention of participating in clinical solutions, must be evaluated with the same level of rigor.

Table I was generated after analyzing three frameworks to classify methodologies for evidence in the health field [17,18,19]. The table includes two types of evaluations that produce valuable information in two areas, efficiency and appropriateness/user perception.

Table I categorizes Randomized Controlled Trial (RCT) as good (Level II), placing it under evidence through systematic reviews which are categorized as excellent due to the guarantee of reliable generalization of outcomes. The other methodologies (Level II-V) are categorized as fair or poor.

When literature on a subject of research is not available or is not sufficient to performs a systematic review, RCT is the most reliable method to obtain scientific evidence. RCTs have the highest level of internal validity [18]. The avoidance of confusion variables, ensures that the variables included in the experimentation, explain a high percent of the phenomenon studied. RCT includes rigorous procedures to guarantee the confidence of its' outcomes. Two relevant components of an RCT are the intervention time and the strategy to calculate the sample size. The compute of these important components is based on these three metrics: the significance level of the test (α), the power of the test ($1-\beta$), and, an effective size that guarantees the practical convenience of implementing a new treatment (Δ).Apart from PSG and actigraphy the emerging technology have no credentials to contribute as part of clinical treatments or protocols to diagnosis. In the words of Behar, after a review on smartphone applications designed to address sleep issues: "*With the exception of simple questionnaires, no existing sleep related application available for smartphones is based on scientific evidence*" [9]. It means that no one of the analyzed smartphone applications have been evaluated through RCT or higher methodologies in rigor. In this context, the phrase applies not only to smartphone applications, but also to the diverse systems analyzed in this work. The question in this scenario is: Why have emerging sleep technologies not been evaluated with the rigor that scientific health evidence requires them to be approved as clinical tools?

Table I. Evidence level of methodologies for assessments in the heath field. [17,18,19]

Efficiency		**Perception of the user**	
Level	**Methodology**	**Level**	**Methodology**
I	Systematic Reviews (SR)	I	Systematic Reviews (SR)
II	RCT Observational studies (OS)	II	Randomized Controlled Trials (RCT)
III	Non-Randomized Control Trials (NRCT) Before and after studies (BAS)	III	Cross sectional surveys Focus groups (FG) Phenomenological study
IV	Case of Study Correlational study (CS) Single qualitative study (SQS)	IV	Expert Opinions (EO) Other qualitative designs.
V	Expert Opinion (EO)		

In the following Section, we approach the answer of this question from a conciliatory perspective. While it is true that systems do not apply to be approved as tools of clinical intervention, it is also true, that they can be evaluated using scientific methods. The obtained results contribute to the advancement of future technological developments and research. Knowing the goals that the systems pursue and understanding the strategies of approaching the sleep issue, it is possible to have a better intuition about the rationale behind the methodology that research used to evaluate each system.

A. Evaluating Emerging Technology for Sleep Health

In synthesis, we have found that emerging applications lack scientific evidence to demonstrate effectiveness and reliability as a clinical tool. The affirmation seems too hard. However, if we analyze in detail, it refers to the fact that the developments do not evaluate the proposals through the rigorous methodologies that are required in health science, and not necessarily that the scientific methodology was ignored to evaluate the system. Our response to the question is based on arguments from researchers in HCI and health science, due to these two areas are closely related with the subject that we are analyzing. Klansja et al. [14] argues that is not feasible to evaluate emerging technology through RCT studies in the early stages of design. This kind of evaluation would have a high cost in resources while the outcomes would be not comparable in quality. Instead of preparing a rigorous study, researchers in the early stages would prepare evaluations that provide information in two directions: 1) The efficiency that developments have regarding the goals proposed in the current stage of design; 2) The perception that users have of the efficiency of use, ease of use, intrusiveness and comfort. These types of evidence in exploratory studies could be obtained through quasi-experiments, correlational studies, case studies and qualitative evaluations, among others. The outcomes through these types of evaluations, do not provide evidence of causality, but establish relations between variables and identify patterns [18].In concordance with the arguments above, Evans et.al. [17] introduce two important aspects in evaluating an intervention in addition to efficiency, appropriateness and feasibility. In its' framework, as well as efficiency, these concepts can be evaluated through RCTs, but not in the early stages of design, which is the case of emerging technologies design to address sleep health. Methodologies to evaluate these metrics include correlational studies, focus groups, before and after studies, phenomenological studies, expert opinion, among others. It is relevant to note that the questions proposed by Evans et al. to evaluate appropriateness and feasibility are qualitative. Specifically, in the case of appropriateness the questions are: What is the experience of the user? What health issues are important to the user? Does the user feel the outcomes as beneficial? These questions can be answered with responses based on the users' perception. These questions are equivalent to those that HCI researchers use to evaluate applications implemented in various domains. These ideas and those proposed by Klansja et. al. to evaluate developments in the HCI area, can be easily adopted by technology applications in the sleep research domain. RCTs are not able to respond to all type of questions. There is a lot of valuable information that can be obtained through alternative scientific methods of evaluation. Such is the case of quasi-experiments, quantitative and qualitative studies that help to respond to important questions, maintaining the quality of research [18]. The key is to identify the type of question that is posed, since there exist appropriate methods to diverse questions [19]. By selecting the appropriate methodology to the type of question, evaluations will obtain valuable outcomes that provide reliable evidence for proposed purposes.In various scenarios, it is pertinent to obtain valuable information through scientific evaluations with less rigor than those required by RCT [14,19,18,17]. In emerging areas as sleep technology for health, it is recommendable to perform these types of evaluations instead of RCTs, since the study field is in the early stages of development. These types of evaluations provide more valuable information than RCT by its' nature. The outcomes will be used to understand the impact that systems produce in people closely related with the issue [12]. Outcomes in these stages do not provide evidence for clinical use, however, they contribute by providing guidelines for new design stages where more rigorous evaluations will be required. If the project is continued, the time will come when evaluations to provide clinical evidence will be appropriate. Collaterally, the scientific community interested in the same subject, will design from a more robust platform [14]. We conclude this section with the Table II. This table is a proposal of a framework to evaluate emerging systems designed to attend sleep health issues, according with the arguments exposed in Section 4. Researchers in technology for sleep health, must be careful when describing the goals that they want to reach in their current stage of work. The goal and relevant questions define the appropriate evaluation for the research. It avoids the thought of this emerging technology ignoring scientific evidence, but demonstrates scientific evidence depending on its design stage.In the next section, we use six metrics to observe eighteen systems oriented to contributing to technological research for addressing and preventing sleep disorders. Though this evaluation, it will be possible to observe the methodology used to evaluate them. It will reveal a fairer perspective of strengths and weaknesses of the evidence that authors claim with their proposals.

Table II. Proposed framework to evaluate emerging systems designed to health fied.

Methodology		
Level	**Efficiency**	**User´s perception**
High	Non-Randomized Control Trials (NRCT) Before and after studies (BAS)	Cross sectional surveys Focus groups (FG) Phenomenological study
Medium	Case of study Correlational study (CS) Single qualitative study (SQS)	Expert (EO) Opinions Other qualitative designs
Low	Expert Opinion (EO)	
Poor	Poor Methodology/Does not report efficiency evaluation	Does not report a user perception evaluation

B. Evaluation of Systems Designed for Sleep Health

Table III shows the evaluation performed on the systems based on the following arguments: 1) it is possible to make some tests in short term and with small sample sizes to obtain results that contribute to knowing whether technology is achieving aims that researchers are supposed to be doing; 2) it is feasible to compare techniques, algorithms and implementations to validate their efficacy and test intervention strategies; and 3) the most important issue in short evaluation tests, are the conclusions obtained from users about their experiences using the systems [34, 35, 19, 18, 17,14].The last two columns in Table 3 evaluate the systems based on Table I. The penultimate column, shows the level of evaluation to obtain evidence in the efficiency achieved by the system according with its' proposed goals. This column is closely related with the column two. However, a (yes) answer in column two, does not mean that systems are ready to be implemented in a real scenario, but demonstrate that the goals proposed for the current stage of development are being achieved. This is a valid position over all in the early stages of design, where the outcomes do not provide conclusive, but suggestive results [14]. The last column, shows the level of evaluation to obtain evidence in terms of users' perception. These features allow to evaluate the systems in early stages of development and the evaluation can be performed independently of the nature of system and the technological design that they have. Other columns in the table explains its' content by itself.As Table 3 shows, most of these systems are very far from meeting the RTC requirements or equivalent. Only [30] has a period intervention comparable to the shortest sample sizes and interventions for RCTs analyzed by [35] in a review of interventions in the health field. On the other hand, no evaluations in this list have a sample size near to the samples size of RCT. However, all authors of evaluated systems argue that they have achieved their proposed goals based on their metrics and evaluations performed. Based on Table 2 that excludes RCTs or higher evaluations, but considering those methodological evaluations appropriated to the first stages of design, 7/18 systems were evaluated at High level of evidence, 9/18 systems were evaluated at Medium level of evidence and 2/18 were evaluated at Low level. Systems were evaluated through a scientific methodology that does not meet the requirements for clinical evidence, however, the outcomes are useful in several ways.

Table III. Evaluation of methodology used to assess systems designed to attend sleep health.

- The information generated could be used as part of the platform to build new stages.
- Refinements will be performed in the new stage of design to achieve more efficiency in the goals proposed.

System	Reach its own goals	Sample	Weeks	LEEP	Proof	Efficiency	User's Perception
[20]	*yes*	12	4	ED	DI	Medium	High
[21]	*yes*	15	4	EP	DS	Medium	Low
[22]	*yes*	5	6	ED	DS	Medium	High
[23]	*yes*	8	1	ED	DS	Medium	High
[24]	*yes*	18	—	ED	tt	Low	High
[25]	*yes*	7		ED	DS	High	Low
[26]	*yes*	4	2	ED	DS, DI	Medium	High
[27]	*yes*	8	$\frac{1}{2}$ - 26	ED	tt	High	High
[16]	*yes*	27	3	ED	tt	High	Medium
[28]	*yes*	26	1	ED	PC, tt	High	High
[29]	*yes*	–	–	EB	–	Medium	Low
[3]	*yes*	1	$\frac{1}{2}$	EP	ML, BS	Medium	Poor
[30]	*yes*	6	12	ED	DS	High	Low
[31]	*yes*	7	—	EP	ML	High	Low
[32]	*yes*	10	1	ED	ML, tt	High	Low
[10]	*yes*	–	–	EB	–	Medium	Low
[15]	*yes*	8	6	ED	AV, PC	Medium	High
[33]	*yes*	–	–	EB	–	Low	Low

LEEP: Level of description of the evaluation procedure; **ED:** Explained in Detail; **EP:** Explained Partially; **EB:** Explained Briefly; **ML:** Machine Learning; **tt:** Student Test; **AV:** ANOVA Test; **PC:** Pearson Correlation; **DS:** Descriptive Statistics; **DI:** Deductive/Inductive

- Findings can be made that will improve the design of the technological developments for its' next version.
- The improvements will be traduced in satisfaction of the users and other interested parties, which is very useful in the scope of services and businesses. For example, because of these evaluations and the reports proportioned by works which include qualitative studies, we identified that the researchers found a high level of acceptance in those systems that were unnoticed. This is a valuable finding that helps researchers in sleep technology to design systems with an emphasis on the concept of unobtrusiveness.
- Although systems were evaluated through a scientific methodology in low levels of the evidence framework, some weaknesses were found, and it is important to highlight them. From eighteen evaluated systems, a 47% do not include a qualitative study. From all systems with a qualitative study, only three ([22, 26, 20]) reported a methodology plan based on considerations and recommendations from the literature. Furthermore, it is important to observe that no article explains the reason of the sample size, neither why or how the duration time for the evaluations was chosen.

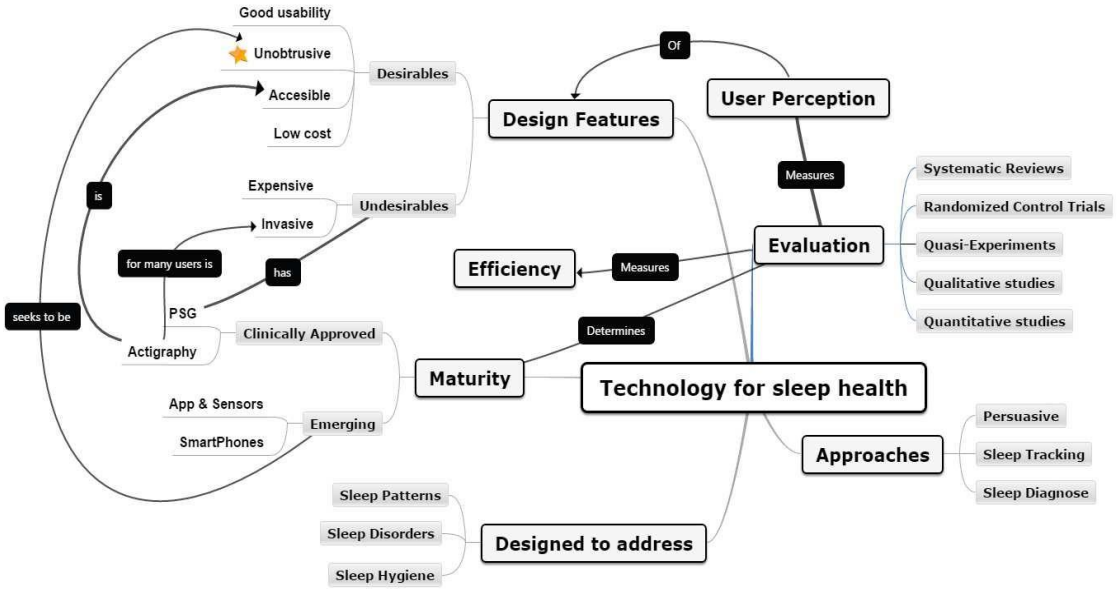

Fig. 1 Technologies for sleep health

Conclusions

Technology to help people to sleep better evolved from the PSG, to embedded systems in small devices. Many systems are embedded in smartphones that use small sensors to obtain raw data from their environment, while other systems combine a set of external sensors to collect data and software applications to show the users the information that was collected. Both types of systems process raw data to produce high level information, e.g. messages for physicians to know the patients' state of sleep health. The systems are classified in three application areas, monitoring sleep patterns, diagnosis of sleep disorders and persuasive systems to help people to sleep better. A valid concern exists among the community regarding the reliability of emerging designs. While the oldest PSG and recently the actigraphy have been approved for clinical use, no one of the reviewed systems have been approved, even when some of them are based on the principles of PSG or actigraphy. The basic reason is that emerging technology by their nature has not been evaluated through RCT or higher methodologies. Instead of RCTs, for new designs, researchers should choose assessments that evaluate systems in two directions: 1) Efficiency in meeting the proposed goals; 2) Level of acceptance among users, physicians and other stakeholders in the context which the systems are implemented. The outcomes provided by these types of evaluations will be valuable evidence to take decisions in the following stages of design in the project. On the other hand, information obtained will provide evidence to guide new designs in this scope. We analyzed eighteen systems, and we found that any system was evaluated through RCT methodology or higher. Instead, researchers use quasi-experiments or other methodologies in the same level of rigor or less. Even when these methodologies provide scientific evidence in the early stages of design, opportunities exist to improve the quality of evaluations. Based in the recommendation of the literature reviewed and analyzing the systems selected in this work, we found opportunities for improvement in three aspects principally:

- Include qualitative studies to obtain information of the users' perception.
- Describe clearly the procedure through which the evaluation was performed.
- Justify the method used to decide the intervention period and the sample size.

Acknowledgments

This work is supported by the CONACYT and by the MyDCI program of Autonomous University of Baja California. Additionally, we acknowledge the work of reviewers' team, especially, we thank the collaboration of Bradley Aitken as language reviewer.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

1. E. K. Choe, S. Consolvo, N. F. Watson, and J. A. Kientz, "Opportunities for computing technologies to support healthy sleep behaviors," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11. New York, NY, USA: ACM, 2011, pp. 3053–3062.
2. M. J. Deen, "Information and communications technologies for elderly ubiquitous healthcare in a smart home," *Personal and Ubiquitous Computing*, vol. 19, no. 3-4, pp. 573–599, 2015.
3. W.-H. Liao and J.-H. Kuo, "Sleep monitoring system in real bedroom environment using texture-based background modeling approaches," *Journal of Ambient Intelligence and Humanized Computing*, vol. 4, no. 1, pp. 57–66, 2013.
4. B.-Q. Zhu, X.-M. Li, D. Wang, and X.-F. Yu, "Sleep quality and its impact on glycaemic control in patients with type 2 diabetes mellitus," *International Journal of Nursing Sciences*, vol. 1, no. 3, pp. 260–265, 2014.
5. M. Sekine, T. Tatsuse, N. Cable, T. Chandola, and M. Marmot, "U-shaped associations between time in bed and the physical and mental functioning of japanese civil servants: The roles of work, family, behavioral and sleep quality characteristics," *Sleep Medicine*, vol. 15, no. 9, pp. 1122–1131, 2014.
6. K. Chang, S. Son, Y. Lee, J. Back, K. Lee, S. Lee, Y. Chung, K. Lim, J. Noh, H. Kim, S. Koh, H. Roh, M. Park, J. Kim, and C. Hong, "Perceived sleep quality is associated with depression in a korean elderly population," *Archives of gerontology and geriatrics*, vol. 59, no. 2, pp. 468–473, 2014.
7. S. Bolitho, S. Naismith, S. Rajaratnam, R. Grunstein, J. Hodges, Z. Terpening, N. Rogers, and S. Lewis, "Disturbances in melatonin secretion and circadian sleep-wake regulation in parkinson disease," *Sleep Medicine*, vol. 15, no. 3, pp. 342–347, 2014.
8. A. Sadeh, "The role and validity of actigraphy in sleep medicine: An update," *Sleep Medicine Reviews*, vol. 15, no. 4, pp. 259–267, 2011.
9. J. Behar, A. Roebuck, J. S. Domingos, E. Gederi, and G. D. Clifford, "A review of current sleep screening applications for smartphones," *Physiological Measurement*, vol. 34, no. 7, p. R29, 2013.
10. C. Occhiuzzi and G. Marrocco, "The rfid technology for neurosciences: Feasibility of limbs' monitoring in sleep diseases," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 1, pp. 37–43, 2010.
11. J. Behar, A. Roebuck, M. Shahid, J. Daly, A. Hallack, N. Palmius, J. Stradling, and G. Clifford, "Sleepap: An automated obstructive sleep apnoea screening application for smartphones," *Biomedical and Health Informatics, IEEE Journal of*, vol. 19, no. 1, pp. 325–331, Jan 2015.
12. L. Parra, S. Sendra, J. M. Jim´enez, and J. Lloret, "Multimedia sensors embedded in smartphones for ambient assisted living and e-health," *Multimedia Tools and Applications*, pp. 1–27, 2015, article in Press.
13. R. Beun, "Persuasive strategies in mobile insomnia therapy: alignment, adaptation, and motivational support," *Personal and Ubiquitous Computing*, vol. 17, no. 6, pp. 1187–1195, 2013.
14. P. Klasnja, S. Consolvo, and W. Pratt, "How to evaluate technologies for health behavior change in hci research," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11.

New York, NY, USA: ACM, 2011, pp. 3063–3072. [online]. Available: http://doi.acm.org/10.1145/1978942.1979396

15. A. Shirazi, J. Clawson, Y. Hassanpour, M. Tourian, A. Schmidt, E. Chi, M. Borazio, and K. Van Laerhoven, "Already up? using mobile phones to track & share sleep behavior," *International Journal of Human Computer Studies*, vol. 71, no. 9, pp. 878–888, 2013.
16. N. D. Lane, M. Lin, M. Mohammod, X. Yang, H. Lu, G. Cardone, S. Ali, A. Doryab, E. Berke, A. T. Campbell *et al.*, "Bewell: Sensing sleep, physical activities and social interactions to promote wellbeing," *Mobile Networks and Applications*, vol. 19, no. 3, pp. 345–359, 2014.
17. D. Evans, "Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions," *Journal of clinical nursing*, vol. 12, no. 1, pp. 77–84, 2003.
18. D. Pati, "A framework for evaluating evidence in evidence-based design," *HERD: Health Environments Research & Design Journal*, vol. 4, no. 3, pp. 50–71, 2011.