Mobile Phone Interventions for Sleep Disorders and Sleep Quality: Systematic Review

Jong Cheol Shin¹, MS; Julia Kim², RD, MPH; Diana Grigsby-Toussaint¹,², MPH, PhD

¹Department of Kinesiology and Community Health, University of Illinois-Urbana Champaign, Champaign, IL, United States
²Division of Nutritional Sciences, University of Illinois-Urbana Champaign, Urbana, IL, United States

Corresponding Author:
Diana Grigsby-Toussaint, MPH, PhD
Department of Kinesiology and Community Health
University of Illinois-Urbana Champaign
1206 S Fourth Street
Champaign, IL, 61820
United States
Phone: 1 2173339207
Fax: 1 2173332765
Email: dgrigs1@illinois.edu

Abstract

Background: Although mobile health technologies have been developed for interventions to improve sleep disorders and sleep quality, evidence of their effectiveness remains limited.

Objective: A systematic literature review was performed to determine the effectiveness of mobile technology interventions for improving sleep disorders and sleep quality.

Methods: Four electronic databases (EBSCOhost, PubMed/Medline, Scopus, and Web of Science) were searched for articles on mobile technology and sleep interventions published between January 1983 and December 2016. Studies were eligible for inclusion if they met the following criteria: (1) written in English, (2) adequate details on study design, (3) focus on sleep intervention research, (4) sleep index measurement outcome provided, and (5) publication in peer-reviewed journals.

Results: An initial sample of 2679 English-language papers were retrieved from five electronic databases. After screening and review, 16 eligible studies were evaluated to examine the impact of mobile phone interventions on sleep disorders and sleep quality. These included one case study, three pre-post studies, and 12 randomized controlled trials. The studies were categorized as (1) conventional mobile phone support and (2) utilizing mobile phone apps. Based on the results of sleep outcome measurements, 88% (14/16) studies showed that mobile phone interventions have the capability to attenuate sleep disorders and to enhance sleep quality, regardless of intervention type. In addition, mobile phone intervention methods (either alternatively or as an auxiliary) provide better sleep solutions in comparison with other recognized treatments (eg, cognitive behavioral therapy for insomnia).

Conclusions: We found evidence to support the use of mobile phone interventions to address sleep disorders and to improve sleep quality. Our findings suggest that mobile phone technologies can be effective for future sleep intervention research.

Introduction

Sleep disorders are an important public health problem that affects approximately 50 to 70 million people in the United States [1]. Sleep disorders are defined as having abnormal sleep behaviors, including insomnia, sleep apnea, restless leg syndrome, and narcolepsy [2]. Those who have chronic sleep disorders are at a greater risk for obesity, diabetes, hypertension, cardiovascular disease, stroke, and depression [3-6]. In addition to sleep disorders, insufficient sleep and irregular sleep patterns are also risk factors for obesity [7], impaired cardiovascular function [8,9], and diabetes [10]. According to the 2011-2014 report from the National Center for Health Statistics [11], 31.7% of US adults do not meet the National Sleep Foundation’s recommendation for at least 7 hours of sleep.
per night. Moreover, the 2011 Sleep in America poll reported that 63% of Americans did not meet the recommendations for sleep time during weekdays [12]. Given that sleep disorders and insufficient sleep increase the risk for chronic diseases, sleep interventions that improve sleep quality are increasingly becoming more important.

In 2011, approximately 39% of Americans, 72% of whom were adolescents, used mobile phones immediately before sleeping [12]. With the increasing use of mobile technology, mobile health (mHealth) is increasingly being used as a practical intervention tool in medicine and public health [13-19]. The term mHealth refers to the provision of health care services and delivery of personal health information using mobile technology such as mobile phones [20]. The portability of mobile devices used in mHealth addresses issues related to accessibility. This allows health care practitioners to provide services to patients regardless of geographical location [15,21]. Specifically, mHealth technology has additional unique functions that are not typically found in traditional care, such as text messaging-tailored medical advice or individualized phone alarms to encourage specific health behaviors (eg, exercising or taking medications) [22]. Regarding sleep, accelerometers in mobile phones can be used to measure and to evaluate sleep patterns [23], and voice recordings can detect abnormal sleep behavior such as snoring and sleep talking [24].

The potential and importance of mHealth technology in health care and health interventions are evident through more than 100,000 health apps in the app store, and a US $26 billion estimated mHealth market size in 2017 [25]. The 2016 mHealth App Developer Economics report [26] indicated that 53% of new apps were designed to improve or address various health conditions. In sum, mHealth is expected to play a significant role in health care due to its ability to be easily integrated into health care services and intervention studies.

Research Objective
Although new mHealth technologies are assumed to be able to improve the quality and quantity of sleep, limited examinations of behavioral sleep interventions using mobile phones exist [13,22]. For example, previous studies [27-31] have focused on examining the technological aspects of mobile phones only (ie, calibration and validation studies, device feasibility). Moreover, previous studies examining the use of the mobile phone’s effectiveness for addressing sleep quality or quantity have only focused on one specific sleep disorder, such as sleep disturbance [32] or obstructive sleep apnea [33,34]. Therefore, the primary purpose of this review was to determine the effectiveness for sleep disorder improvement. We hypothesized that interventions with mobile technology have a positive impact on improving sleep quality and various sleep disorders (eg, sleep apnea, snoring, or insomnia).

Methods

Study Selection Criteria
This study included articles if they met the following criteria:

1. Study design: randomized controlled trials (RCTs), pre-post studies, and case-control studies;
2. Sleep intervention study: using mHealth technology (eg, a mobile device or app);
3. Intervention outcome: sleep outcome measurement (eg, Insomnia Severity Index [ISI], Pittsburgh Sleep Quality Index [PSQI], or Epworth Sleepiness Scale [ESS]);
4. Language: written in English; and
5. Article type: peer-reviewed publications.

Search Strategy
Articles published between January 1983 (the year of the first handheld and commercial cellular phone from Motorola [35]) and December 2016 were searched from four electronic databases, including EBSCOhost: Academic Search Complete (ASC) & Cumulative Index to Nursing and Allied Health Literature (CINAHL), PubMed/Medline, Scopus, and Web of Science. The journal search process occurred from February 18 to 19, 2017. To identify components of three research topics, namely, “sleep,” “mHealth,” and “study design,” the specified search keywords were used as a search algorithm (see Multimedia Appendix 1).

After searching the electronic databases, one author (JCS) selected articles that were published in peer-reviewed journals and excluded books, case reports, conference proceedings, product reviews, newspapers, patents, serials, and theses. Secondly, the same author removed duplicate articles from the combined search results and screened articles based on the title and abstract. In addition, JCS hand-searched for relevant articles in the JMIR search engine on February 28, 2017. Once all full-text articles were found, two researchers (JCS and JK) reviewed each article using the eligibility criteria and eliminated unrelated articles. The entire procedure followed the guidelines for Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) [36].

Data Extraction
The following information was extracted from each article included in the systematic review: first author’s name, publication year, the country where the study was conducted, sample size, sample age, sample characteristics, study period, retention rate, study design, the technology used in the intervention, and measurement of sleep outcome. The retention rate was calculated using the number of participants who completed all assessments for the intervention study divided by the number of individuals who were originally recruited. Intervention technology was identified using two approaches: (1) intervention using mobile phone app and (2) supplementary mobile phone usage with traditional intervention. Additional information was extracted by using the empirical approach of statistical analysis such as effect size, standard error, and standard deviation for the difference before and after interventions.

Statistical Analysis
The effect size and standard deviation for each study’s sleep outcome measurement were obtained from data extraction. With each effect size, there was a difference of calculation for the total effect size based on the type of study. For example, the effect size of pre-post test was only considered the difference between pretest scores and posttest scores. On the other hand,
the effect size of RCTs was calculated using the difference between the intervention and the control groups’ effect sizes. Based on the total effect size, sample size, and standard error, two-sample t tests were performed to determine the mobile phone intervention effectiveness.

**Study Quality Assessment**

The study quality assessment tool was derived from Zhu and An [37], and then revised after discussion and peer review by the researchers for the purpose of this study. The quality of research was assessed using the following criteria: (1) the research question and objectives were stated clearly, (2) mobile health and/or mHealth was defined, (3) a control group was included, (4) participants were randomly recruited from a well-defined population, (5) sample size was more than 30, (6) attrition was analyzed and determined not to significantly differ by respondents’ baseline characteristics between control and experiment groups (<20%), (7) baseline characteristics between control and intervention groups were similar, (8) the intervention period was at least 4 weeks, (9) the sleep disorder measurement tools were shown to be reliable and valid in previous studies, and (10) demographic information was available to control for potential confounders in future analysis. A sample size of 30 was chosen due to a guarantee of a normal sample distribution based on the central limit theorem. The total study quality score was determined by summing items 1 to 10, and categories were created to define a general quality of research as poor (score 0-3); moderate (score 4-6), and high (score 7-10).

**Results**

**Study Selection**

The results of the literature search are summarized in Figure 1. A total of 2674 articles were initially selected from four electronic databases (EBSCOhost: CINAHL&ASC, PubMed/Medline, Scopus, and Web of Science). After excluding 440 duplicate articles, 896 articles were excluded from the first screening of peer-reviewed journal articles and 1339 articles were left. After the first screening, five articles were included from hand searching, and 1304 were excluded based on the title (n=1180) and abstract (n=124). All 39 remaining articles were then downloaded as PDF files for full-text peer review and discussion [13-16,21,22,38-70]. We excluded 23 articles after the full-text screening for the following reasons: (1) a device or program feasibility study (n=2) [66,69], (2) duplicated article (n=1) [21], (3) not an intervention study (n=5) [39,43,47,49,53], (4) not a mobile phone intervention (n=7) [14,42,48,56,58,63,68], (5) no sleep outcome (n=3) [16,44,65], and (6) a protocol or trial (n=5) [40,45,50-52]. The entire selection procedure was independently performed by two researchers (JCS and JK), and then the research team resolved differences by discussion. From individual selections, the interrater reliability analysis using the kappa statistic was performed with peer reviewers to determine consistency among raters. The interrater reliability was found to be Cohen kappa=.80, representing substantial agreement [20]. Therefore, a total of 16 articles that satisfied the inclusion criteria were used for the final analysis.

**Study Characteristics**

Table 1 provides summary information of individual studies. A total of 16 studies were included in the analysis: one case study [46], three pre-post studies [13,15,41], and 11 RCTs [21,22,38,54,55,59,61,62,64,67,70]. Of the 12 RCTs, three studies [62,64,67] used mobile phone apps as an intervention tool; one study [21] used text messages as part of the intervention. RCTs ranged in total sample size from 30 to 502 and had a larger sample size than pre-post studies, which ranged from three [13] to 12 [41]. The study period for RCTs was from 4 weeks to 6 months, whereas for pre-post studies it was between 2 and 5 weeks. For the study region, eight studies were performed in the United States [13,15,21,41,55,61,62,70], two studies in Canada [54,67], two studies in Asia (Hong Kong [57] and Taiwan [46]), and four studies in Europe (Finland [38], France [64], Netherlands [21], and Sweden [59]). Most of the studies were published after 2012, except one study that was published in 2006 [67]. Three studies [15,46,64] focused on an elderly population, whereas three studies [21,67,70] focused on a young adult population.
Figure 1. Study selection procedure according to the PRISMA guidelines.
Table 1. Basic characteristics of the studies of mobile phone interventions on sleep disorders (N=16).

<table>
<thead>
<tr>
<th>Author(s), Year</th>
<th>Study region</th>
<th>Sample characteristics (a)</th>
<th>Sample size</th>
<th>Mean age</th>
<th>Study period</th>
<th>Intervention methods (b)</th>
<th>Study design (c)</th>
<th>Sleep outcome measurement (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anttalainen et al, 2014 [38]</td>
<td>Finland</td>
<td>OSA patients</td>
<td>111</td>
<td>55.27</td>
<td>3 m</td>
<td>Phone + CPAP</td>
<td>RCT</td>
<td>AHI, ESS</td>
</tr>
<tr>
<td>Bauer et al, 2012 [41]</td>
<td>United States</td>
<td>Metropolitan area adults</td>
<td>12</td>
<td>32</td>
<td>4 w</td>
<td>App: “ShutEye” for Android</td>
<td>Pre-post</td>
<td>ESS</td>
</tr>
<tr>
<td>Chen et al, 2015 [46]</td>
<td>Taiwan</td>
<td>Elderly female</td>
<td>1</td>
<td>64</td>
<td>5 w</td>
<td>App: “Win-Win A Sleep”</td>
<td>Case study</td>
<td>SSR</td>
</tr>
<tr>
<td>Fox et al, 2012 [54]</td>
<td>Canada</td>
<td>OSA patients</td>
<td>75</td>
<td>53.54</td>
<td>3 m</td>
<td>Phone + CPAP</td>
<td>RCT</td>
<td>AHI, ESS</td>
</tr>
<tr>
<td>Freeman et al, 2015 [55]</td>
<td>United States</td>
<td>Breast cancer survivors</td>
<td>102</td>
<td>55.44</td>
<td>3 m</td>
<td>Phone-supported teleconference</td>
<td>RCT</td>
<td>PSQI</td>
</tr>
<tr>
<td>Ho et al, 2014 [57]</td>
<td>Hong Kong</td>
<td>Insomnia patients</td>
<td>149</td>
<td>38.5</td>
<td>12 w</td>
<td>Phone + CBT-I</td>
<td>RCT</td>
<td>PSQI, ISI, DBAS, SE, SQ, SOL, WASO, TST</td>
</tr>
<tr>
<td>Jernelov et al, 2012 [59]</td>
<td>Sweden</td>
<td>Adults with insomnia</td>
<td>133</td>
<td>47.9</td>
<td>6 w</td>
<td>Phone + bibliotherapy</td>
<td>RCT</td>
<td>ISI, DBAS, SE, SQ, SOL, WASO, TST, BTS, SRBQ</td>
</tr>
<tr>
<td>Lichstein et al, 2013 [15]</td>
<td>United States</td>
<td>Rural area adults with insomnia</td>
<td>5</td>
<td>65.8</td>
<td>5 w</td>
<td>Apps: “Skype” + CBT</td>
<td>Pre-post</td>
<td>ISI, HRDS, NWAK, SOL, SQ, WASO</td>
</tr>
<tr>
<td>Kauffman 2016 [61]</td>
<td>United States</td>
<td>Menopausal status with insomnia</td>
<td>106</td>
<td>54.85</td>
<td>24 w</td>
<td>Phone + CBT-I</td>
<td>RCT</td>
<td>ISI, PSQI, SE, SOL, TST, WASO</td>
</tr>
<tr>
<td>Mendelson et al, 2014 [64]</td>
<td>France</td>
<td>OSA patients</td>
<td>107</td>
<td>63</td>
<td>4 w</td>
<td>Researcher-built app + CPAP</td>
<td>RCT</td>
<td>ESS</td>
</tr>
<tr>
<td>Stremler et al, 2006 [67]</td>
<td>Canada</td>
<td>First-time mothers</td>
<td>30</td>
<td>31.85</td>
<td>6 w</td>
<td>Phone + sleep education</td>
<td>RCT</td>
<td>GSDS</td>
</tr>
<tr>
<td>Vuletic et al., 2016 [70]</td>
<td>United States</td>
<td>Soldiers With MTBI</td>
<td>356</td>
<td>29.35</td>
<td>6 m</td>
<td>Telephone-based problem-solving treatment</td>
<td>RCT</td>
<td>PSQI, SE, SOL, SQ</td>
</tr>
</tbody>
</table>

(a) CUD: cannabis use disorders; MTBI: mild traumatic brain injury; OSA: obstructive sleep apnea.
(b) CBT-I: Cognitive Behavioral Therapy for Insomnia; CPAP: continuous positive airway pressure.
(c) RCT: randomized controlled trial.
(d) AHI: Apnea-Hypopnea Index; BTS: bed time stress; DBAS: Dysfunctional Beliefs and Attitudes about Sleep scale; ESS: Epworth Sleepiness Scale; GSDS: General Sleep Disturbance Scale; HRSD: Hamilton Rating Scale for Depression with sleep; ISI: Insomnia Severity Index; PSQI: Pittsburgh Sleep Quality Index; NWAK: number of awakenings; SE: sleep efficiency; SOL: sleep-onset latency; SQ: sleep quality; SQ_PSQI: extracted sleep quality score based on PSQI; SRBQ: Sleep-Related Behavior Questionnaire; SSR: Sleep Satisfaction Rate; TST: total sleep time; WASO: wakefulness after sleep onset.
Table 2. Summary of sleep measurement tools and study design for mobile phone sleep intervention included in the systematic review (N=16).

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Case study &amp; pre-post test (n=4)</th>
<th>RCT (n=12)</th>
<th>Standard treatment: CBT-I &amp; CPAP (n=5)(^a)</th>
<th>Other recognized treatment (n=7)</th>
<th>Waitlist (n=4)(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apnea-Hypopnea Index (AHI)</td>
<td>Score (total apneas event/TST)</td>
<td>0</td>
<td>2 [38,54]</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bed time stress (BTS)</td>
<td>Scored on scale 0-5</td>
<td>0</td>
<td>0</td>
<td>1 [59]</td>
<td>1 [59]</td>
</tr>
<tr>
<td>Dysfunctional Beliefs and Attitudes about Sleep Scale (DBAS)</td>
<td>30 items with scale 1-10 (total:300)</td>
<td>0</td>
<td>1 [57]</td>
<td>1 [59]</td>
<td>2 [57,59]</td>
</tr>
<tr>
<td>Epworth Sleepiness Scale (ESS)</td>
<td>8 items with scale 0-3 (total: 24)</td>
<td>1 [41]</td>
<td>2 [38,54]</td>
<td>1 [64]</td>
<td>0</td>
</tr>
<tr>
<td>General Sleep Disturbance Scale (GSDS)</td>
<td>21 items with scale 0-7 (total:147)</td>
<td>0</td>
<td>0</td>
<td>1 [67]</td>
<td>0</td>
</tr>
<tr>
<td>Hamilton Rating Scale for Depression with sleep (HRSD)</td>
<td>Total 21 items:(^c); 10 items with scale 0-4, 2 items with scale 0-3, and 10 items with scale 0-2 (total: 66)</td>
<td>1 [15]</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Insomnia Severity Index (ISI)</td>
<td>7 items with scale 0-4 (total: 28)</td>
<td>1 [15]</td>
<td>3 [15,57,61]</td>
<td>2 [59,62]</td>
<td>2 [57,59]</td>
</tr>
<tr>
<td>Number of awakenings (NWAK)</td>
<td>Frequency of awakening</td>
<td>1 [15]</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pittsburgh Sleep Quality Index (PSQI)</td>
<td>7 components calculated from 9 questions with scale 0-3 (total: 21)</td>
<td>1 [13]</td>
<td>2 [57,61]</td>
<td>2 [55,70]</td>
<td>3 [22,55,57]</td>
</tr>
<tr>
<td>Sleep efficiency (SE)(^f)</td>
<td>Percentage (TST/total time in bed)</td>
<td>0</td>
<td>2 [57,61]</td>
<td>2 [59,70]</td>
<td>2 [57,59]</td>
</tr>
<tr>
<td>Sleep quality (SQ)(^f)</td>
<td>Scored on scale 1-5</td>
<td>1 [15]</td>
<td>1 [57]</td>
<td>2 [59,70]</td>
<td>2 [57,59]</td>
</tr>
<tr>
<td>Extracted sleep quality score based on PSQI (SQ_PSQI)</td>
<td>8 items with scale 1-4 (total: 32)</td>
<td>0</td>
<td>0</td>
<td>1 [21]</td>
<td>1 [21]</td>
</tr>
<tr>
<td>Sleep-Related Behavior Questionnaire (SRBQ)</td>
<td>32 items with scale 1-5 (total: 160)</td>
<td>0</td>
<td>0</td>
<td>1 [59]</td>
<td>1 [59]</td>
</tr>
<tr>
<td>Sleep Satisfaction Rate (SSR)(^f)</td>
<td>Scored on scale 0-3</td>
<td>1 [46]</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total sleep time (TST)</td>
<td>Hours</td>
<td>0</td>
<td>2 [57,61]</td>
<td>2 [21,59]</td>
<td>3 [21,57,59]</td>
</tr>
</tbody>
</table>

\(^a\)CBT-I: Cognitive Behavioral Therapy for Insomnia; CPAP: continuous positive airway pressure.

\(^b\)Three articles have two comparison groups (waitlist vs other).

\(^c\)Higher scores indicate lesser severity.

Table 2 shows the number of studies based on study design and sleep measurement tools. Several questionnaires were used to evaluate various aspects of sleep quality and quantity.

Specifically, the Pittsburgh Sleep Quality Index (PSQI) [13,21,22,55,57,61,70] had a list of items to evaluate sleep quality and it was the most frequently used measure in our review. In addition, three studies used a sleep quality [15,57,59] measure that consisted of a simple question about the participant’s perception about sleep quality using a five-point scale. Moreover, five studies [15,21,57,59,61] used measures of sleep quantity: (1) sleep-onset latency (n=4; the length of time to transition from full wakefulness to sleep completely) [15,57,59,61], (2) number of awakenings (n=1, the frequency of awakening during sleep) [15], (3) wake after sleep onset (n=4; the amount of time [eg, minutes] one wakes up during sleep) [15,57,59,61], and (4) total sleep time (n=4; total length of sleep time) [21,57,59,61]. Total sleep time is also used to calculate sleep efficiency [57,59,61]; that is, the percentage of total sleep time per total bed time. For the measurement of behavioral aspects of sleep, some researchers used the Epworth Sleepiness Scale (ESS) (n=4) [38,41,54,64] to capture participant sleepiness. Also two additional studies measured other specific sleep-related behaviors, such as the Sleep-Related Behavior Questionnaire (SRBQ) [59], the Dysfunctional Beliefs and Attitudes about Sleep Scale (DBAS) [57,59], and Sleep Satisfaction Rate (SSR) [46]. Finally, some sleep measurement tools focused on specific sleep disorders such as the severity of insomnia. The Apnea-Hypopnea Index (AHI) [38,54] is used to evaluate sleep apnea, and the General Sleep Disturbance Scale (GSDS) [67] is meant to examine factors that interrupt sleep. Two studies [15,59] used tools to examine the emotional aspects of sleep, such as bed time stress (BTS) levels [59], and the Hamilton Rating Scale for Depression with sleep (HRSD) [15]. All sleep outcome measurements were conducted by subjective measurement (ie, self-report questionnaire).
### Table 3. Summary statistics: quality assessment of each journal and the effects of intervention through \( t \) test by intervention type and sleep outcome measurement.

<table>
<thead>
<tr>
<th>Author(s), year</th>
<th>Quality score</th>
<th>Study design: intervention type</th>
<th>Sleep measurement</th>
<th>Effect size, mean (SD)</th>
<th>95% CI</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anttalainen et al, 2014 [38]</td>
<td>7</td>
<td>RCT: mobile + CPAP vs standard CPAP</td>
<td>AHI</td>
<td>-1.9 (2.9)</td>
<td>-2.4, -1.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ESS</td>
<td>0.0 (1.7)</td>
<td>-0.3, 0.3</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>Babson et al, 2015 [13]</td>
<td>3</td>
<td>Pre-post</td>
<td>PSQI</td>
<td>-1.5 (2.4)</td>
<td>-6.2, 3.2</td>
<td>NA</td>
</tr>
<tr>
<td>Bauer et al, 2012 [41]</td>
<td>5</td>
<td>Pre-post</td>
<td>ESS</td>
<td>-1.7 (1.3)</td>
<td>-2.4, -0.9</td>
<td>.001</td>
</tr>
<tr>
<td>Chen et al, 2015 [46]</td>
<td>2</td>
<td>Case study</td>
<td>SSR</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Filion et al, 2015 [21]</td>
<td>8</td>
<td>RCT: smoking prevention text vs sleep or activity</td>
<td>SQ_PSQI</td>
<td>-0.2 (1.7)</td>
<td>-0.5, 0.1</td>
<td>.14</td>
</tr>
<tr>
<td>Fox et al, 2012 [54]</td>
<td>8</td>
<td>RCT: mobile + CPAP vs standard CPAP</td>
<td>Promotion text</td>
<td>TST (weekday)</td>
<td>0.6 (0.5)</td>
<td>0.5, 0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TST (weekend)</td>
<td>0.6 (0.7)</td>
<td>0.5, 0.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Freeman et al, 2015 [55]</td>
<td>9</td>
<td>RCT: mobile vs waitlist</td>
<td>PSQI</td>
<td>-2.6 (1.5)</td>
<td>-2.9, -2.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Ho et al, 2014 [57]</td>
<td>8</td>
<td>RCT: mobile + CBT-I vs waitlist</td>
<td>DBAS</td>
<td>-23.5 (15.1)</td>
<td>-25.5, -21.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ISI</td>
<td>-3.1 (1.8)</td>
<td>-3.3, -2.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PSQI</td>
<td>-2.3 (1.5)</td>
<td>-2.5, -2.1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SE</td>
<td>4.5 (4.8)</td>
<td>3.8, 5.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SOL</td>
<td>-7.4 (12.9)</td>
<td>-9.2, -5.6</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SQ</td>
<td>-0.2 (0.2)</td>
<td>-0.2, -0.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TST</td>
<td>0.0 (8.8)</td>
<td>1.2, 1.2</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WASO</td>
<td>-10.8 (14.8)</td>
<td>-12.8, -8.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Jernelov et al, 2012 [59]</td>
<td>9</td>
<td>RCT: mobile + bibliotherapy vs waitlist</td>
<td>BTS</td>
<td>-0.9 (0.4)</td>
<td>-1.0, -0.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DBAS</td>
<td>-56.4 (9.8)</td>
<td>-58.4, -54.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ISI</td>
<td>-8.8 (1.6)</td>
<td>-9.1, -8.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SE</td>
<td>15.3 (7.1)</td>
<td>13.8, 16.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SOL</td>
<td>-30.7 (20.4)</td>
<td>-35.0, -26.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SQ</td>
<td>0.9 (0.2)</td>
<td>0.9, 0.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SRBQ</td>
<td>-21.9 (4.8)</td>
<td>-22.9, -20.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TST</td>
<td>0.4 (0.5)</td>
<td>0.3, 0.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WASO</td>
<td>-30.1 (18.9)</td>
<td>-34.1, -26.1</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

- **DBAS**: Daily Behavioral Adjustment Scale
- **ISI**: Insomnia Severity Index
- **PSQI**: Pittsburgh Sleep Quality Index
- **SE**: Sleep Efficiency
- **SOL**: Sleep Onset latency
- **SQ**: Sleep Quality
- **SQ_PSQI**: Sleep Quality and Pittsburgh Sleep Quality Index
- **TST**: Total Sleep Time
- **WASO**: Wake After Sleep Onset
Table 3 provides a summary of the effectiveness of each sleep intervention. It is categorized by sleep outcome measurement, and the statistics can be comparable if there are more than two sleep measures. Specifically, half of studies [13,22,41,46,62,64,67,70] used a single sleep outcome measurement, while others had at least two or more measurements. Of nine studies [13,15,38,46,54,57,61,62,64] of standard treatment, five studies compared a mobile phone intervention and standard treatment intervention: (1) continuous positive airway pressure (CPAP) [38,54,64] and (2) cognitive behavioral therapy for insomnia (CBT-I) [57,61]. Five studies [13,22,41,46,61] focused on a mobile phone app. Two of these studies [13,61] used a “CBT-I coach” app that was developed based on CBT-I, which is considered one of the traditional standard treatments for insomnia (Figure 2).

<table>
<thead>
<tr>
<th>Author(s), year</th>
<th>Quality score</th>
<th>Study design: intervention type&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Sleep measurement&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Effect size, mean (SD)</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Koffel et al, 2016 [62]</td>
<td>8</td>
<td>RCT: app use vs non app</td>
<td>ISI</td>
<td>−4.5 (1.5)</td>
<td>−4.8, −4.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SE</td>
<td>10.2 (7.0)</td>
<td>8.7, 11.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SOL</td>
<td>−14.8 (21.8)</td>
<td>−19.3, −10.3</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SQ</td>
<td>0.5 (0.2)</td>
<td>0.5, 0.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SRBQ</td>
<td>−15.3 (5.4)</td>
<td>−16.4, −14.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TST</td>
<td>0.0 (0.5)</td>
<td>−0.1, 0.1</td>
<td>.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WASO</td>
<td>−21.1 (19.4)</td>
<td>−25.1, −17.1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Lichstein et al, 2013 [15]</td>
<td>6</td>
<td>Pre-post</td>
<td>ISI</td>
<td>3.4 (4.6)</td>
<td>1.2, 5.5</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HRSD</td>
<td>−7.0 (3.6)</td>
<td>−10.2, −3.8</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ISI</td>
<td>−11.3 (3.3)</td>
<td>−14.2, −8.3</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NWAK</td>
<td>−1.7 (2.1)</td>
<td>−3.5, 0.1</td>
<td>.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SOL</td>
<td>−18.4 (20.9)</td>
<td>−36.7, −0.1</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SQ</td>
<td>0.5 (0.1)</td>
<td>0.4, 0.6</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WASO</td>
<td>−23.8 (9.2)</td>
<td>−31.9, −15.7</td>
<td>.004</td>
</tr>
<tr>
<td>Kauffman 2016 [61]</td>
<td>8</td>
<td>RCT: mobile + CBT-I vs menopause education</td>
<td>ISI</td>
<td>−4.0 (5.6)</td>
<td>−5.1, −2.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PSQI</td>
<td>−1.6 (3.0)</td>
<td>−2.2, −1.0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SE</td>
<td>3.2 (12.3)</td>
<td>0.8, 5.6</td>
<td>.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SOL</td>
<td>−11.9 (36.3)</td>
<td>−18.8, −5.0</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TST</td>
<td>0.2 (1.2)</td>
<td>0.0, 0.4</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WASO</td>
<td>−6.8 (47.3)</td>
<td>−15.8, 2.2</td>
<td>.14</td>
</tr>
<tr>
<td>Mendelson et al, 2014 [64]</td>
<td>8</td>
<td>RCT: app + CPAP vs standard CPAP</td>
<td>ESS</td>
<td>−0.2 (1.5)</td>
<td>−0.5, 0.1</td>
<td>.17</td>
</tr>
<tr>
<td>Stremler et al, 2006 [67]</td>
<td>8</td>
<td>RCT: mobile + education vs education</td>
<td>GSDS</td>
<td>−13.3 (8.2)</td>
<td>−16.2, −10.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>van Drongelen et al, 2014 [22]</td>
<td>9</td>
<td>RCT: app use vs non app</td>
<td>PSQI</td>
<td>−0.6 (1.3)</td>
<td>−0.7, −0.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Vuletic et al, 2016 [70]</td>
<td>8</td>
<td>RCT: mobile + education vs education</td>
<td>PSQI</td>
<td>−1.5 (1.5)</td>
<td>−1.7, −1.4</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

<sup>a</sup>CBT-I: cognitive behavioral therapy for insomnia; CPAP: continuous positive airway pressure.

<sup>b</sup>AHI: Apnea-Hypopnea Index; BTS: bed time stress; DBAS: Dysfunctional Beliefs and Attitudes about Sleep Scale; ESS: Epworth Sleepiness Scale; GSDS: General Sleep Disturbance Scale; HRSD: Hamilton Rating Scale for Depression with sleep; ISI: Insomnia Severity Index; PSQI: Pittsburgh Sleep Quality Index; NWAK: number of awakenings; SE: sleep efficiency; SOL: sleep-onset latency; SQ: sleep quality; SQ_PSQI: extracted sleep quality score based on PSQI; SRBQ: Sleep-Related Behavior Questionnaire; SSR: Sleep Satisfaction Rate; TST: total sleep time; WASO: wakefulness after initial sleep onset.
Figure 2. Sample size based on the intervention methods (N=16).

Table 4. Quality assessment of studies to determine the impact of sleep intervention with mobile technology on sleep disorders.

<table>
<thead>
<tr>
<th>Criteria item</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Research question and objective were stated clearly</td>
<td>0.94 (0.24)</td>
</tr>
<tr>
<td>2. Definition of telehealth and/or mHealth was stated</td>
<td>0.18 (0.39)</td>
</tr>
<tr>
<td>3. A control group was included</td>
<td>0.82 (0.39)</td>
</tr>
<tr>
<td>4. Participants were randomly recruited from well-defined population</td>
<td>0.82 (0.39)</td>
</tr>
<tr>
<td>5. Sample size was &gt;30</td>
<td>0.76 (0.44)</td>
</tr>
<tr>
<td>6. Attrition was analyzed and determined not to significantly differ by respondents’ baseline characteristics between control and experiment groups (&lt;20%)</td>
<td>0.47 (0.51)</td>
</tr>
<tr>
<td>7. Baseline characteristics between control and intervention groups were similar</td>
<td>0.76 (0.44)</td>
</tr>
<tr>
<td>8. The intervention period was at least 4 weeks</td>
<td>0.94 (0.24)</td>
</tr>
<tr>
<td>9. The sleep disorder measurement tools were shown to be reliable and valid in previously published studies</td>
<td>0.88 (0.33)</td>
</tr>
<tr>
<td>10. Demographic information is available to control potential confounders for future analysis</td>
<td>0.82 (0.39)</td>
</tr>
<tr>
<td>Total study quality score&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7.41 (0.44)</td>
</tr>
</tbody>
</table>

<sup>a</sup>By summing up items 1 to 10 (range 3-10).

Study Quality

Table 4 summarizes the results of the quality assessment of the studies included in this review. The quality score of the final articles ranged from 3 to 10 out of a possible score of 10, with an overall mean score of 7.41 (SD 0.44). The distribution of quality scores differed substantially across each criterion. Of 16 articles, 12 articles were considered high quality [21,22,38,54,55,59,61,62,64,67,70], two articles moderate quality [15,41], and two articles poor quality [13,46]. Most of the studies had clear research questions and objectives [13,15,21,22,38,41,46,54,55,57,59,61,62,64,67,70], and used valid sleep measurement tools [13,15,21,22,38,41,54,55,57,59,61,62,64,70] for outcome measures after 4 weeks [15,21,22,38,41,46,54,55,57,59,61,62,64,67,70]. All RCTs [21,22,38,54,55,57,59,61,62,64,67,70] satisfied items 3 (i.e., existence of control group) and 5 (i.e., sample size greater than 30). The mean intraclass correlation (ICC) between peer reviewers for quality assessment was ICC 0.987 (95% CI 0.964-0.996, P<.001).

Effectiveness of Mobile Phones Usage as a Sleep Intervention Tool

The mobile phone can be used as a tool to effectively deliver and enhance traditional behavioral interventions due to its...
portability. We found eight articles that described the advantages of using mobile phones to improve traditional intervention methods for sleep disorders [38,54,55,57,59,61,67,70].

**Mobile Phone Usage as Alternative Intervention Tools**

Mobile phones are a tool for intervention studies due to the portability for the participant. Within the results of our review, two studies focused on the effectiveness of mobile phones as intervention tools.

Stremler et al [67] was the first published research article that used a mobile phone as a behavioral-educational strategy tool for a sleep intervention. The focus of the study was to evaluate the feasibility, acceptability, and effects of a mobile phone intervention on sleep in the early postpartum period. During the 6-week intervention period, the intervention group was provided a well-established booklet, in-depth counseling, and a nurse to call for sleep advice, whereas the control group was only provided a one-page pamphlet and brief meeting session. Because the primary target subjects were first-time mothers and their children, researchers used GSDS to evaluate the sleep disturbances in employed women. As a result, the sleep intervention group had lower GSDS scores as much as 13.3 points compared to the control group (t_{28} = –8.84, P < .001).

Similarly, Vuletic et al [70] had two intervention groups: (1) a group with 12 biweekly “telecounseling” and (2) an “education-only” group with an educational brochure. The research team developed the telephone-based problem-solving treatment, which was a sleep intervention that related to sleep quality to improve post-deployment soldier’s traumatic brain injury. Since the researchers used various sleep outcome measurements such as SOL, total sleep time, sleep disturbance, and daytime dysfunction that were part of PSQI, only PSQI was included for data analysis. After 6 months of intervention, the intervention group with telecounseling had significantly lower scores than the education-only intervention group (effect size = –1.54, SE = 0.19; t_{274} = –16.62, P < .001).

**Mobile Phone Used as Auxiliary Equipment**

Both CPAP and CBT-I are considered standard treatments for obstructive sleep apnea (OSA) [38,54] and sleep disturbances [13,57,62]. Three studies [38,54,61] applied mobile phones as supplementary tools in addition to these standard treatments.

Fox and colleagues [54] conducted a RCT to determine the effectiveness and adherence of telemonitoring and CPAP combined. To demonstrate the effectiveness of the intervention, AH1 and ESS were used to measure sleep outcomes. Researchers found that participants in the intervention group had an improvement in adherence to CPAP, but no significant differences were found for AH1 and ESS between the control and intervention groups. Based on t test results, in addition, the effect size of AH1 showed significantly decreased (t_{57} = –3.82, P < .001).

Aanttalainen et al [38] had a similar research method as Fox and colleagues’ [54], such as intervention duration, study design, intervention methods, presence of a control group, and sleep outcome measurement. Aanttalainen et al did not find a statistically significant difference in ESS between groups, whereas participants in the intervention group had a significantly higher score on the AH1 than standard in-person treatment (t_{109} = 6.91, P < .001). Because ESS between standard CPAP and mobile-supported CPAP were not significantly different, the effects of mobile-supported CPAP did not differ from standard CPAP.

Kaufman [61] conducted a RCT to determine the effectiveness of telephone-based CBT-I to determine the effectiveness of telephone-based CBT-I compared to menopause education control. The target population was menopausal middle-aged women with moderate insomnia (ISI>15). Based on the t test results, there was a significantly different effect size for the index type sleep measurements between the telephone-based CBT-I and menopause education control (ISI: t_{104} = 7.33, P < .001; PSQI: t_{104} = –5.54, P < .007), whereas sleep time measures were not significant such as sleep onset latency, wake after sleep onset, total sleep time, and sleep efficiency.

**Multidimensional Approach to Mobile Phone Usage**

Three studies [55,57,59] examined the effectiveness of mobile phone usage as an alternative intervention tool and as an auxiliary equipment.

Jernelov et al [59] evaluated the effectiveness of mobile phone support with CBT bibliotherapy (CBT). To examine the effectiveness of mobile phone support, a total of 133 participants were randomly allocated to three comparable groups (44 for CBT bibliotherapy with mobile support; 45 for only CBT bibliotherapy; 44 for the control group). After the 6-week intervention, a total of 116 participants completed the follow-up assessment (40 for CBT bibliotherapy with mobile support; 37 for CBT bibliotherapy; 39 for the control group), with an attrition rate of 12.78%. Participants in the mobile-supported CBT bibliotherapy group had significant improvements in all sleep measurements compared to the control group (BTS: t_{96} = –20.80, P < .001; DBAS: t_{96} = –53.96, P < .001; ISI: t_{96} = –50.64, P < .001; SE: t_{95} = –20.18, P < .001; SOL: t_{95} = –14.06, P < .001; SQ: t_{95} = –39.59, P < .001; SRQ: t_{95} = –42.81, P < .001; TST: t_{95} = –6.77, P < .001; WASO: t_{95} = –14.83, P < .001), and even CBT bibliotherapy only (BTS: t_{57} = –21.95, P < .001; DBAS: t_{57} = –31.74, P < .001; ISI: t_{57} = –28.88, P < .001; TST: t_{57} = –13.78, P < .001; SOL: t_{57} = –6.39, P < .001; SQ: t_{57} = –24.41, P < .001; SRQ: t_{57} = –26.77, P < .001; WASO: t_{57} = –10.26, P < .001), except TST (t_{57} = 0.685, P = .84).

Although Ho et al [57] had a similar study design and approach to Jernelov and colleagues [59], Ho et al had a different intervention tool: the comparison between the self-help CBT-I only. Of Jernelov and colleagues [59], Ho et al had a different study design and approach to Jernelov and colleagues [59], Ho et al had a different intervention tool: the comparison between the self-help CBT-I only. Of Jernelov and colleagues [59], Ho et al had a different study design and approach to Jernelov and colleagues [59], Ho et al had a different intervention tool: the comparison between the self-help CBT-I only.
evaluation. Therefore, the statistics in Table 3 were based on the follow-up session at 4 weeks to compare whether both interventions significantly improved the sleep measurement score compared to the initial assessment. From the t test, several sleep outcome measures indicated that mobile phone-supported intervention had a statistically significant improvement compared to the waitlist control group (PSQI: $t_{206} = –7.54, P < .001$; ISI: $t_{206} = 24.37, P < .001$; database: $t_{206} = –22.48, P < .001$; SOL: $t_{206} = 8.27, P < .001$; WASO: $t_{206} = 10.56, P < .001$; SE: $t_{206} = 13.57, P < .001$; and SQ: $t_{206} = 14.14, P < .001$) and also the standard treatment intervention (PSQI: $t_{205} = –7.63, P < .001$; ISI: $t_{205} = 8.64, P < .001$; SOL: $t_{205} = 7.94, P < .001$; WASO: $t_{205} = 7.53, P < .001$; SE: $t_{205} = 3.62, P < .001$; and SQ: $t_{205} = 7.85, P < .001$).

Filion et al [21] established the “Envision the Rhythms of Life” program, which is a group intervention to increase the quality of life for breast cancer survivors, and also to evaluate its advantages on aspects of quality of life. A total of 118 participants were randomly assigned to three groups (48 participants for the in-person delivery; 23 participants for the mobile phone-supported teleconferencing; 47 participants for the waitlists) at the beginning of the study, and 102 participants remained (40 participants for the in-person delivery; 19 participants for the mobile phone-supported teleconferencing; 43 participants for the waitlists) at the end of the study. According to the t test, participants in the mobile phone-supported group had a significant increase in PSQI scores among breast cancer survivors compared to both control groups (standard treatment: mean = –0.68, SD 1.39; $t_{270} = 4.11, P < .001$; waitlist: mean = –2.55, SD 1.53; $t_{270} = 13.97, P < .001$). In addition, authors performed the linear multilevel modeling analysis, and found that PSQI was only considerable for group effect ($P < .001$), not for time ($P = .35$) or group × time ($P = .30$).

**Text Message for Sleep Intervention**

Filion et al [21] used a text message-based intervention for young adult smokers to promote better sleep and physical activity habits. A total of 164 baseline participants were assigned to two groups: (1) sleep/physical activity group (n = 63) and (2) smoking cessation group (n = 101). A total of 129 participants completed the study (smoking cessation group: n = 77; sleep group: n = 44). Participants in the sleep/physical activity group received sleep- and activity-related messages for 6 weeks, whereas the smoking cessation group received quitting smoking-related messages. Although the difference of the PSQI score between the smoking group (mean = –1.47, SD 1.77) and sleep group (mean = –1.7, SD 1.48) was not significant ($P = .60$), the amount of sleep was significantly different ($P = .03$). Specifically, the amount of sleep (hours) for the sleep/physical activity group increased (mean 0.52, SD 0.44), whereas sleep among the smoking cessation group decreased (mean = –0.03, SD 0.46).

**Effectiveness of Mobile Phone Apps as a Sleep Intervention Tool**

As a measurement tool for sleep disorder intervention, mobile phone apps are able to perform diverse functions (ie, tracking [13,41,46,62], sleep advice for behavioral change [13,15,41,46], and optimized alarms [13,22,62]). In our review, seven studies evaluated the effectiveness of these app-based interventions [13,15,22,41,46,62,64].

**Mobile Phone Apps as Auxiliary Equipment**

Lichstein et al [15] used the Skype app for teleconferencing in addition to CBT as a sleep intervention. The purpose of the study was to examine the feasibility and effectiveness of interventions to address challenges faced by rural older adults with comorbid conditions, namely insomnia and depression. The researchers demonstrated the limitation of mobile phone apps for older population (age: mean 65.8, SD 10.4 years). From the t test, participants in the Skype group had significant improved sleep quality (sleep quality: $t_{4} = 9.7, P < .001$), less insomnia (ISI: $t_{4} = –7.54, P < .002$), less wakefulness after sleep onset ($t_{4} = –4.33, P = .01$). Despite the improvement in sleep quality, this study had a higher attrition rate; only five of 18 baseline participants completed the entire study (Multimedia Appendix 2).

Mendelson et al [64] evaluated whether a combination of the mobile phone app using telemedicine and CPAP had better performance compared to standard care (CPAP only). After 4 weeks of intervention, a total 82 of 107 baseline participants (76.6%) completed the entire test procedure, and 54 participants adhered to CPAP for more than 4 hours. Although the ESS was statistically significantly different between pretest and posttest (combined intervention: mean = –2.3, SD 4.0, $P < .05$; standard care: mean = –2.1, SD 4.1, $P < .001$), the effectiveness of treatment did not differ in each intervention.

**Mobile Phone Apps Developed by Researchers to Improve Sleep Behavior**

Bauer et al [41] were the first to assess the use of mobile phone apps for sleep intervention programs. The research team developed the “Shuteye” app in 2011, which offered real-time recommendations to promote awareness related to healthy sleep behaviors. They also highlighted the following as being important for effective sleep apps: design of the app, adherence of mobile phone, appearance, awareness of app usage, and learnability (ie, user-friendly interface). For the intervention study, ESS was measured for 12 participants before and after the intervention. We calculated the changes in participant scores pre- and postintervention based on the raw data provided by the research team. Based on the data in the articles, we performed the paired t test, and ESS decreased after intervention (mean = –1.67, SD 1.32; $t_{11} = –4.36, P = .001$), although this was not found to be statistically significant, most likely due to the small sample size.

Van Drongelen et al [22] conducted a RCT addressing sleep using a mobile phone app with international aircraft pilots. The “More Energy” app was invented to reduce pilots’ fatigue and also to improve their healthy behavior and sleep. During the 6 months of intervention, a total of 390 participants (77.7% of 502 baseline participants: 191 in the intervention group, 199 in waitlists) completed the study. The intervention program showed statistically significant fatigue improvement, and also showed a positive effect on sleep. We calculated the difference between the intervention group (score decreased by mean 0.2, SD 1.2)
and waitlists (score increased by mean 0.37, SD 1.32); the PSQI score significantly improved (beta=-0.59, P=.001).

Mobile Phone Apps With CBT-I

Three studies [13,46,62] used mobile phone apps to reform CBT-I, which is one of the traditional standard treatments for improving sleep behavior. Babson et al [13] used a mobile phone app version of CBT-I to determine the feasibility and efficacy for veterans using cannabis to treat sleep disorders. A total of four participants were included in the study (two in the intervention group; two in the control group). According to the t tests result, the mobile phone intervention did not significantly affect the PSQI score (t=−0.87, P=.48). However, this study was considered a pre-post study because all participants in the control group dropped out during the test procedure. In addition to the loss of a follow-up evaluation in the control group, several limitations with the study, such as small sample size and the rarity of the target population (ie, veterans using cannabis treatment for sleep disorders), caused the authors’ skeptical conclusion of the ineffectiveness of mHealth intervention.

Chen et al [46] used the “Win-Win a Sleep” app, which was developed from CBT-I, and revealed effectiveness and feasibility of sleep interventions. In this case study, a participant who was female aged 63 years suffered from insomnia and used the app during a 5-week test period. Because the participant took hypnotic medications, researchers also considered the impact of decreasing the amount of drugs used to treat sleep disorders. The participant had better sleep satisfaction, from no preference (mean 1.5, SD 0.7) to good (mean 2.0, SD 0.0), based on the result of a Likert scale change (score range between 0=“very bad” and 3=“very good”).

Koffel et al [62] recruited from the general population to conduct a RCT to determine the feasibility and acceptability of CBT-I coach apps. A total of 18 participants evaluated their ISI scores at each therapy session (five times), and also reported the adherence rate. We did not perform a t test because there was no distinguishable statistic of ISI included in the article. However, ISI scores tended to decrease based on figures that were analyzed by hierarchical linear models. There was a significant difference (t15=3.04, P=.008) in ISI scores between baseline and after the third session, even though the standard CBT-I group had lower ISI scores than the app user group.

Lack of Data for Performing Meta-Analysis

Although meta-analysis is a great way to examine the effectiveness of interventions with a specific parameter, we did not have enough articles to conduct a meta-analysis. Sleep measurements were not included in all articles selected for the systematic literature review. Because more than four articles were required to perform a meta-analysis, according to Table 2, there were not enough articles when stratified by sleep measurement and study design. Also, a sample size of N=2 or N=3 is insufficient to perform the publication bias test (Begg rank correlation test or Egger regression test) and to determine the level of heterogeneity.

Discussion

Summary of Evidence

The purpose of our systematic review was to investigate whether mobile phones are a feasible and usable tool to improve sleep disorders and sleep quality in intervention studies. We also assumed that auxiliary mobile phone use helps to enhance the performance of existing intervention methods.

This review presented the articles based on two intervention methods. The first method was utilizing mobile phone apps and the second was conventional mobile phone support through telephone calls or text messages. Although many mobile phone apps were used independently, the conventional mobile phone supporting methods were (1) to use the mobile phone itself as only an intervention tool (eg teleconferencing or telecounseling), or (2) to combine with another treatment (eg, CBT-I, CBT-B, and CPAP) to enhance the effectiveness of sleep interventions.

In our review, six studies were related to CBT-I [13,15,46,57,61,62] and three studies to CPAP [38,54,64]. This indicated that sleep intervention studies using mobile phones have been developed based on the standard treatment, which has already shown reliability [33,53,71-73].

Among the sleep outcome measurements, PSQI was the most frequently reported, and PSQI scores of all seven studies [13,21,22,55,57,61,70] were significantly decreased in the intervention group using mobile phones compared to standard treatment and a waitlist group. Participants in the intervention groups with mobile phones had a mean decrease in PSQI scores of 1.73 points compared to the waitlist group, and a mean decrease of 0.97 points compared to other treatment groups. Although the difference of PSQI was statistically significant, clinicians must consider the clinical significance of a one-point change in PSQI scores. According to the Buysse et al [74], a score of 5 of 21 is the cut-off point for poor sleep. Thus, the clinical significance of a one-point difference depends on the baseline score.

The ISI was the second most frequently used sleep outcome measurement tool, and all five articles using the measure [15,57,59,61,62] found a statistically significant decrease in ISI score in the mobile phone use intervention group compared to non-mobile phone use group. Participants in the intervention group with mobile phones had a mean decrease in ISI scores of 5.09 points compared to the waitlist groups [57,59,62], and a mean decrease of 3.2 points in ISI scores compared to other treatments [57,59,61]. Because a change of 7 points in the ISI is interpreted as a different level of insomnia, a 3- or 5-point decrease in ISI scores might be considered a critical improvement. Although ESS, WASO, SOL, and TST were included in at least four studies, there was at least more than one study that reported a lack of statistical significance of intervention effect [15,38,41,54,57,59,61]. Therefore, ISI and PSQI seem to be the most applicable measurement tools to investigate the effect of mobile phone intervention on sleep disorders.

In general, 87.5% (14/16) of the studies [15,21,22,38,41,54,55,57,59,61,62,64,67,70] reviewed supported...
the evidence of capability and efficacy of mobile phone usage interventions. When mobile phones were used as auxiliary equipment, the intervention that applied telephone calls or mobile phone apps clearly demonstrated equal or enhanced efficacy for sleep quality and index score compared to the traditional sleep intervention [15,38,54,55,57,59,61,64,67,70]. This difference was seen not only for the waitlist control group [54,55,57,59], but also for the telephone-supported intervention and traditional treatment [15,38,61,64,67,70]. The advantage of mobile phone intervention was also illustrated in apps using intervention [13,22,41,46,62] or text message interventions [21].

Limitations of Review

There are several limitations of the studies included in this review. First, there is no standardized study design, especially for the test period, procedure, and sleep intervention tools. For example, there is a limitation to use all mobile phone apps to compare each study because researchers used their own personally developed or nonpublicly available apps. As such, we were unable to make comparisons between interventions due to differences in app functions and interfaces. Due to the heterogeneity of the study design and sleep measurement tools, it was difficult to compare the effectiveness of the sleep app interventions. This will be an issue for reproducibility in further research. Second, although it is a common limitation for RCT study designs, the uniqueness of each study’s target population limits further analysis and replication. For instance, results from studies with a cannabis disorder use group [13] and a post-deployment soldier with mild traumatic brain injury [70] were difficult to extrapolate to the general population. Third, there are few interventions using mobile technology that have been published in peer-reviewed journals, so we were limited with respect to the number of articles we could include in the study. Although many t test for each sleep measurement showed significant effectiveness of sleep interventions, the small number of studies for subgrouping limits performing a meta-analysis. Fourth, there is the possibility of missed articles. Because the mHealth market is growing rapidly, our search possibly missed some articles that used new mobile phone intervention tools.

On the other hand, it also might be possible to miss some articles during the search stage due to the usage of different language or jargon unique to the evolving mHealth industry.

Conclusion

This study has several strengths. To the best of our knowledge, this study was the first to review the effectiveness of mobile phone interventions on sleep quality, quantity, and sleep disorders. By focusing exclusively on the mobile phone itself, it allows us to tailor future mHealth interventions for sleep. Also, our study examined various aspects of sleep measurement tools that account for sleep quality, sleep quantity, and many sleep disorders such as insomnia and sleep latency. Additionally, our study provided evidence of the potential of mobile-based interventions for improving sleep disorders. Along with current research that support the benefits of cost-efficiency of mHealth interventions, these findings provide an impetus for further research examining the empirical evidence of sleep interventions using mobile phones.

In conclusion, our systematic review supports the evidence that mobile technology-based interventions are an effective tool to improve symptoms of sleep disorders and quality of sleep than traditional intervention without mobile phone. Also, we suggest the following research design for future sleep intervention studies: (1) PSQI and ISI as sleep outcome measurements, (2) RCTs, (3) compare with standard treatment (ie, CPAP, CBT-I), and (4) compare to a waitlist control group. In addition to intervention methods, because mobile phone apps vary and many of these apps are not being studied, it is important to perform a content analysis on commercially available apps to determine common functionalities prior to undertaking interventions [75]. Our finding was not only applicable to those with sleep disorders who need clinical care, but also to medical professionals who are interested in ways to determine effective sleep interventions. Moving forward, app developers and sleep experts need to develop evidence-based guidelines with behavioral change components for sleep apps to maximize their efficacy and to take advantage of mobile phones to apply to existing standard treatments for sleep interventions.

Acknowledgments

The authors would like to thank Dr Ruopeng An for his advice regarding systematic reviews.

Authors’ Contributions

JCS contributed to the concept and design for the current study, performed data analyses, and drafted the manuscript. JK participated in conception and design, data analysis of review procedure, and provided critical feedback on the draft of the manuscript. DGT provided guidance and critical feedback on the concept of the study and drafts of the manuscript. All authors contributed critical revision of the manuscript for important intellectual content and read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Searching Keywords.

[PDF File (Adobe PDF File), 15KB - mhealth_v5j9e131_app1.pdf]
Multimedia Appendix 2

Quality Assessment Score.

References


Abbreviations

AHI: Apnea-Hypopnea Index  
BTS: bed time stress  
CINAHL: Cumulative Index to Nursing and Allied Health Literature  
CPAP: continuous positive airway pressure  
DBAS: Dysfunctional Beliefs and Attitudes about Sleep Scale  
ESS: Epworth Sleepiness Scale  
GDS: General Sleep Disturbance Scale  
HRSD: Hamilton Rating Scale for Depression with sleep  
ISI: Insomnia Severity Index  
OSA: obstructive sleep apnea  
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis  
PSQI: Pittsburgh Sleep Quality Index  
RCT: randomized controlled trial  
SE: sleep efficiency  
SOL: sleep-onset latency  
SQ: sleep quality  
SRBQ: Sleep-Related Behavior Questionnaire  
SSR: Sleep Satisfaction Rate  
TST: total sleep time  
WASO: wakefulness after initial sleep onset
Mobile Phone Interventions for Sleep Disorders and Sleep Quality: Systematic Review

Shin JC, Kim J, Grigsby-Toussaint D
URL: http://mhealth.jmir.org/2017/9/e131/
doi:10.2196/mhealth.7244
PMID:28882808

©Jong Cheol Shin, Julia Kim, Diana Grigsby-Toussaint. Originally published in JMIR Mhealth and Uhealth (http://mhealth.jmir.org), 07.09.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.