ORIGINAL RESEARCH

A Feasibility Study: Testing Whether a Sleep Application Providing Objective Sleep Data to Physicians Improves Patient–Physician Communication Regarding Sleep Experiences, Habits, and Behaviors

Sana Durrani · Sha Cao · Na Bo · Jennifer K. Pai · Jarod Baker ·

Lori Rawlings · Zaina P. Qureshi · Ninotchka L. Sigua ·

Shalini Manchanda · Babar Khan

ABSTRACT

Introduction: Sleep tracker data have not been utilized routinely in sleep-related disorders and their management. Sleep-related disorders are common in primary care practice and incorporating sleep tracker data may help in improving patient care. We conducted a pilot study to assess the feasibility of a sleep program using the Fitbit Charge 2™ device and SleepLife® application. The main aim of the study was to examine whether a program using a

S. Durrani · J. Baker · L. Rawlings · B. Khan (✉)
Regenstrief Institute, 1101 West 10th Street,
Indianapolis, IN 46202, USA
e-mail: bakhan@iu.edu

N. L. Sigua · S. Manchanda · B. Khan
Division of Pulmonary, Critical Care, Sleep and
Occupational Medicine, Department of Medicine,
Indiana University School of Medicine, 535 Barnhill
Drive, Indianapolis, IN 46202, USA

J. K. Pai
Merck & Co., Inc., Boston, MA, USA

Z. P. Qureshi
Merck & Co., Inc., Kenilworth, NJ, USA

S. Cao · N. Bo
Indiana University, 1050 Wishard Blvd.,
Indianapolis, IN 46202, USA

commercially available wearable sleep tracker device providing objective sleep data would improve communication in primary care settings between patients and their providers. Secondary aims included whether patient satisfaction with care would improve as result of the program.

Methods: A prospective, randomized, parallel group, observational pilot study was conducted in 20 primary care clinics in Indianapolis, IN from June 2018 to February 2019. Inclusion criteria included patients over the age of 18, have a diagnosis of insomnia identified by electronic medical record and/or a validated questionnaire, and were on a prescription sleep aid. The study was not specific to any sleep aid prescription, branded or generic, and was not designed to evaluate a drug or drug class. Each primary care clinic was randomized to either the SleepLife® intervention or the control arm. All patients were provided with a Fitbit Charge 2™ device. Only patients in the intervention arm were educated on how to use the SleepLife® application. Physicians in the intervention arm were set up with the SleepLife® portal on their computers.

Results: Forty-nine physicians and 75 patients were enrolled in the study. Patients had a mean age of 57 (SD 12.8) years and 61% were female. Mean age of physicians was 47 (SD 10.6) years. Patients showed high rates of involvement in the program with 83% completing all survey

questions. Physician survey completion rate was 55%. Only one physician logged into the SleepLife portal to check their patients' sleep status. At the end of the 6-week intervention, patients' composite general satisfaction scores with sleep health management decreased significantly in the intervention arm when compared to controls ($p = 0.03$). Patients' satisfaction with communication also decreased significantly in the intervention group ($p = 0.01$). The sleep outcomes, which were calculated on the basis of study questionnaire answers, improved significantly in the intervention group as compared to the control group ($p = 0.04$). Physician communication satisfaction scores remained unchanged ($p = 0.12$).

Conclusions: SleepLife® and its related physician portal can facilitate physician–patient communication, and it captures patient sleep outcomes including behaviors and habits. Patients were highly engaged with the program, while physicians did not demonstrate engagement. The study design and questionnaires do not specifically address the reasons behind the decreased patient satisfaction with care and communication, but it was perceived to be a result of physician non-responsiveness. Sleep quality scores on the other hand showed an improvement among SleepLife® users, suggesting that patients may have implemented good sleep practices on their own. Given that it was a feasibility study, and the sample size was small, we were not able to make major inferences regarding the difference between sleep disorder types. Additionally, we excluded patients with a history of alcohol use, substance abuse, or depression because of concerns that they may affect sleep independently. To promote the growth of technology in primary care, further research incorporating results from this study and physician engagement techniques should be included.

Keywords: Sleep; Tracker; Primary care; Technology; Communication; Smartphone app

Key Summary Points

Sleep-related disorders are common in primary care practice and incorporating sleep tracker data may help in improving patient care.

Our pilot study assessed the feasibility of using a Fitbit Charge 2 device and SleepLife application to improve communication in primary care settings between patients and their providers.

This prospective, randomized, parallel group pilot study was conducted in 20 primary care clinics, amongst primary care providers and their patients with insomnia diagnoses, on prescription sleep aid medications.

Only one physician logged into the SleepLife portal. The lack of physician engagement was a significant limitation of the study.

At the end of the 6-week intervention, patients' composite general satisfaction scores with sleep health management decreased significantly in the intervention arm when compared to controls. Their satisfaction with communication also decreased significantly.

Physician engagement can be improved by integrating sleep software into the electronic medical record (EMR), providing specific therapy/management suggestions based on sleep disturbances patients may discuss with their physicians about and by providing physicians extra education opportunities regarding sleep.

A study in the future utilizing the learning points and improving on the limitations from this pilot can lead to a big step in improving sleep communication between primary care physicians and their patients and integrating health technology into primary care for all age groups.

INTRODUCTION

On the basis of a National US survey in 2012, 69% of adults track at least one health indicator using either a tracking device or some other means [1]. The main health indicators tracked were diet, weight, and exercise. Compared to the aforementioned health indicators, studies evaluating sleep indicators through the use of trackers are few. Although one study did show that physicians do not consider sleep data tracking for patients to be useful, there is general lack of availability of reliable data discussing the use of sleep trackers in informing health decisions [2]. On the basis of our literature review, none of the available studies have looked at an intervention designed to utilize tracker-based data to improve physician–patient communication regarding sleep [3].

Commercially available fitness and sleep trackers are prevalent and consumer use is growing rapidly. Primary care physicians seldom use subjective patient-generated sleep data such as sleep diaries and have been slow to integrate objective sleep data collected from commercial sleep trackers. Education about sleep and sleep disorders has been inadequate, leading to a low rate of recognizing these disorders and discomfort with discussing them [4, 5].

The National Sleep Foundation (NSF) has led recent efforts to establish normative data (i.e., appropriate ranges) for sleep duration and sleep quality. NSF has also partnered with the Consumer Technology Association to establish a workgroup involving over 40 sleep tracking technology companies, which is working to standardize sleep tracking data collection and reporting. Additionally, NSF has developed a novel and innovative tool, the SleepLife® application and portal that translates data retrieved from commercially available sleep trackers into a personal sleep tracking record. Communication about sleep between patients and physicians is low at about 31–37% based on the Sleep and Healthy Activity Diet Environment and Socialization (SHADES) study [6]. This information presented the timely opportunity to test a new paradigm for patient and physician communication using objective patient sleep data obtained using the Fitbit Charge 2™ device and SleepLife® application.

We designed this pilot study with the primary objective to examine whether an application utilizing commercially available sleep tracker (Fitbit) data with continuous flow of sleep information to physicians and patients can be used to improve physician–patient communication regarding sleep and sleep-related behaviors and habits. The other objectives included improvement in physician and patient satisfaction with sleep counseling and effects on total sleep time. The program included a Health Insurance Portability and Accountability Act (HIPAA)-compliant, carefully controlled physician and patient interface (SleepLife®) and physician access to patient objective data under daily living environments collected via commercially available wearables.

METHODS

Design

This prospective, randomized, parallel group observational pilot study was conducted in Indianapolis, Indiana. Patients and physicians were enrolled from primary care clinics in the greater Indianapolis area. The study protocols were approved by the Indiana University Institutional Review Board (IRB), IRB Protocol Number 1802077293 and Merck & Co. DRC. The IRB (Protocol Number: 1802077293) provided approval coverage for the 20 primary care clinics where the study was conducted. This study was performed in accordance with the Helsinki Declaration of 164, and its later amendments. All subjects provided informed consent to participate in the study in addition to consent for publication if any identifying information is included in the manuscript.

Participants The DataCore team at the Regenstrief Institute used the Regenstrief Medical Record System (RMRS) to identify patients and physicians at primary care clinics in Indianapolis, Indiana, USA. RMRS is the primary instrument for processing data and monitoring patient and physician activity for the hospital.

The RMRS is a modular system, composed of Registration and Scheduling, Laboratory, Pharmacy, and Database modules; and maintains several other databases including vital signs, results of laboratory tests and diagnostic tests, discharge summaries, and inpatient and outpatient charges. The sample patient population at primary care clinics is representative of the population of a mid-sized city in America. The primary care clinics are dispersed throughout the greater Indianapolis area and represent a diverse group of subjects.

Inclusion Criteria

1. Age 18 years or older
2. Have a diagnosis of insomnia identified by electronic medical record (EMR) and/or a validated questionnaire
3. Taking one or more of the following prescription medications for insomnia: zolpidem (Ambien, Ambien CR, Edluar, Intermezzo, Zolpimist), suvorexant (Belsomra), butabarbital (Butisol), quazepam (Doral), estazolam (Prosom), flurazepam (Dalmane), triazolam (Halcion), tasimelteon (Heltioz), eszopiclone (Lunesta), temazepam (Restoril), ramelteon (Rozerem), secobarbital (Seconal), doxepin (Silenor), zaleplon (Sonata), and trazodone (Oleptro, Desyrel)
4. English speaking
5. Ability to consent in person
6. Have access to a telephone with smartphone capabilities (iOS/Android)
7. Have visited their primary care physician in the last 6 months or have an upcoming visit within 2 weeks
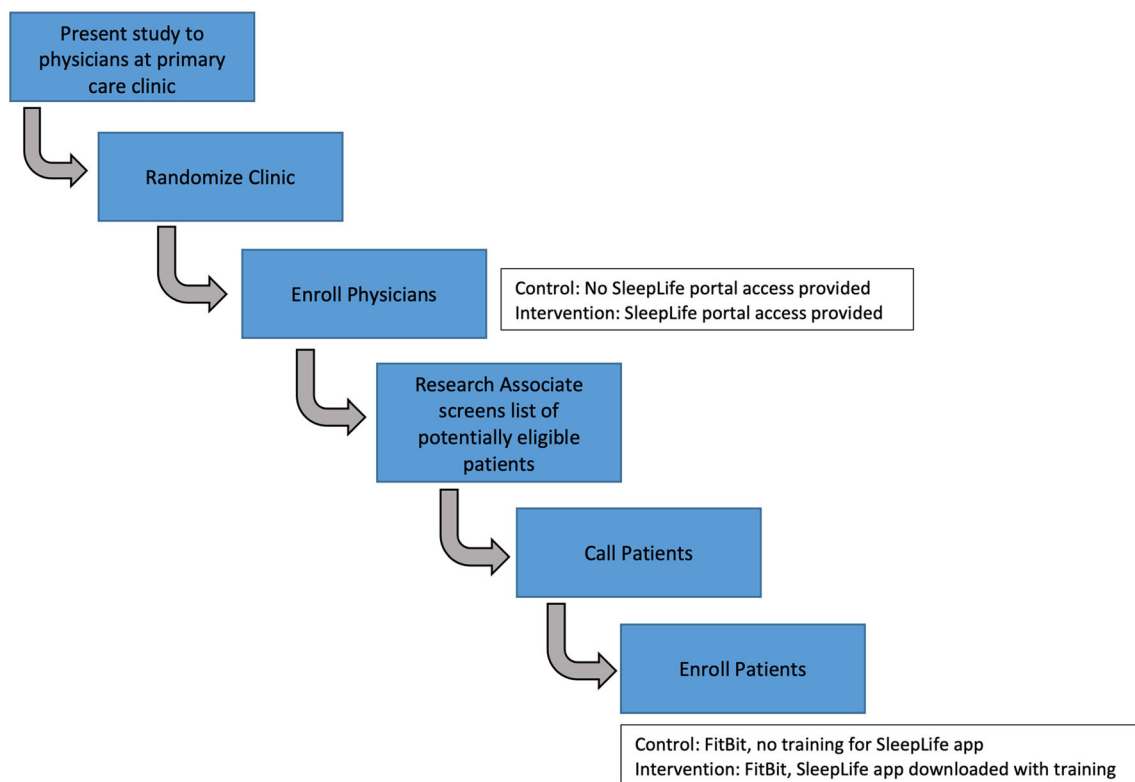

Fig. 1 Study flow

Exclusion Criteria

1. Have ischemic or cerebrovascular disease affecting collection of study outcomes (via ICD codes I6*, 43*)
2. History of dementia (via ICD codes F0*, 290*)
3. History of bipolar/schizophrenia/depression (via ICD codes F2*, F31*, F32*, F33*, 296*, 295*, 311*)
4. History of alcohol or substance abuse (via ICD codes F1*, 304*, 303*)
5. Be incarcerated or reside in long-term acute care/ skilled nursing facility
6. Pregnancy
7. Be unable to complete study questionnaires because of hearing loss or visual impairment

These criteria were chosen because they could either affect a person's sleep quality and habits or affect their ability to participate in this particular study. Each participant had to be taking a sleep aide to ensure that the patient's sleep problems were persistent enough to require a sleep aide.

Study Procedures

Each primary care clinic was randomized to either the SleepLife® intervention or control arm. The study flow is depicted in Fig. 1. The Regenstrief Data Core team used the Regenstrief Medical Record System (RMRS) to extract data on how many potential patients existed at each Indiana University Health Primary Care clinic in the greater Indianapolis area. The ICD-9 codes of 327.*, 780.5*, 347.* and ICD-10 code of G47* were used to screen for insomnia and other related sleep disorders. The research team including both the principal investigator and research staff, visited clinics to present the study protocol at individual clinics prior to randomization. Enrollment for physicians was completed at the Primary Care clinic. The details of the study were discussed with each physician and a signed copy of their informed consent was obtained before any study-related procedures. Once a physician was enrolled, his or her potential patients as originally identified by the

RMRS were screened by the research staff to ensure they met all inclusion and no exclusion criteria. Patients with a diagnosis of insomnia ascribed by a health care provider were approached for enrollment. If a patient was on a prescription sleep aid and did not have a diagnosis of insomnia, they were contacted by research staff and were screened through the insomnia severity index questionnaire to be eligible for enrollment. Patients were enrolled either at their primary care clinic or at a public place that they found comfortable and convenient, like a coffee shop or a library.

Intervention

Intervention components for this study included a Fitbit Charge 2™ as the sleep tracker and the smartphone application SleepLife®, developed by the National Sleep Foundation, as well as a website portal for physicians.

Our consent forms outlined the objectives of the study, noting that the primary objective was to improve physician–patient communication regarding sleep. For physicians, a short talk was given at each study site presenting the study to the clinic faculty, allowing for questions and concerns to be addressed. We then went to each physician individually, consented for the study, and if applicable (i.e., for the study site in the intervention group) downloaded the SleepLife portal onto their work computers, provided instructions on how to use it as well as a handout outlining these instructions. After completing the informed consent, physicians completed the demographic questionnaire and physician satisfaction questionnaire, which referred to their satisfaction with their communication about sleep with their patients. The SleepLife® web-portal was the primary method of intervention for the physicians. The SleepLife® web-portal was designed to connect to the SleepLife® application on the smartphones of patients. The SleepLife® web-portal imported and organized the patients' sleep data from the SleepLife® application on the patients' smartphone into an easy to view format for physicians and allows physicians the ability to send messages to their patients through the portal.

After enrollment, physicians in the intervention arm were provided with access and training to use the SleepLife® portal. A shortcut link to this portal was saved on their desktop in the clinic, or was bookmarked on their browser, according to their preference. In the control arm, physicians did not receive access to the SleepLife® portal. The physicians in the intervention arm received sleep data on their enrolled patients on a weekly basis, as compared to the physicians in the control group who did not receive any sleep data from their patients. On the basis of the sleep data, physicians in the intervention arm could provide, via the web portal, weekly feedback to patients on their sleep. Intervention physicians were also provided reminders about checking the portal.

Each physician was enrolled in the study for the entire duration of the study to accommodate for differential rates of patients' enrollment. Patients were enrolled in a staggered manner from various clinics instead of focusing on one physician or clinic at a time as a result of the variable response times from clinic staff, patients, and physicians.

All enrolled patients, after providing informed consent, completed a demographic questionnaire, a medication form, Katz and Lawton scales, patient satisfaction questionnaire, and a sleep health survey [6–10]. Fitbit Charge 2™ devices were provided to all patients at the time of enrollment. All patients downloaded the Fitbit application.

Patients enrolled in the intervention arm were provided with access to the SleepLife® application and research staff helped them download it onto their smartphone at the time of enrollment. They received detailed information about how to use the application including a one-page instructional document designed by the National Sleep Foundation. The SleepLife® application synced with the Fitbit application to obtain and organize the patients' sleep data. Through the SleepLife® application, patients could view their sleep data, input the time they got into bed, when they took their sleep medication, and send messages to their primary care provider. The SleepLife® application was not loaded onto the control group's smartphones.

The intervention period for patients was 6 weeks, beginning from the date of enrollment. Through this time, data points and questionnaires were collected via email from patients in the control arm, and through the SleepLife® application for patients in the intervention arm. The sleep health survey was collected from patients at the time of enrollment and then again at the end of 6 weeks in the study. The patient satisfaction questionnaire was collected at the time of enrollment, and afterwards at 2, 4, and 6 weeks from the date of enrollment.

In the event that a patient did not log any data via the tracker or app for 3 days in a row, research staff contacted the patient to determine if any adjustments or corrections were required to the tracker or the application.

Measures

Main outcomes collected were physician–patient communication, patient and physician satisfaction, sleep outcomes of total sleep time and satisfaction with sleep. We used a combination of different questionnaires and forms to obtain our outcomes and covariate information: a demographic questionnaire and a physician satisfaction questionnaire for physicians and a demographic questionnaire, Katz Scale, Lawton scale, medications form, patient satisfaction survey, and sleep health survey for patients [6–10]. The following questionnaires were used either as an entire collection instrument itself, or adapted to be a part of collecting study-specific outcomes:

Katz Scale Instrument to assess functional status, based on the patient's ability to perform activities of daily living independently. It includes items like bathing, dressing, toileting, transferring, and feeding.

Lawton Scale Instrument to assess patient's ability to perform activities like communication via telephone, laundry and finance management independently, helping assess their functional status.

Charlson Comorbidity Index Scale predicting the 1-year mortality for a patient on the basis of their comorbid conditions.

Patient Satisfaction Questionnaire A scale validated to assess patient satisfaction.

Campbell Assessment Validated tool to assess both doctor and patient perceptions of the communication.

Physician demographic details included gender, age, preferred method of contact, medical specialty, location of practice, number of patients seen per day, time allocated to each patient, and a question about whether they were satisfied with the amount of time they were allocated per patient. Physicians also filled out the physician satisfaction questionnaire at enrollment, and again at the end of the study.

Patient demographics included age, race, gender, and years of education. Height, weight, body mass index, heart rate, blood pressure, and Charlson Comorbidity Index were also included [11, 12]. Activities and instrumental activities of daily living (ADL/IADLs) were obtained through Katz and Lawton scales, and the medication form was filled out through information from the patients themselves and their medical records [6, 7]. The medication information was updated at the end of the study again, via the patient's medical chart, to note if there had been any changes to the patient's medication regime. Patients also filled out the sleep health survey and patient satisfaction questionnaire at enrollment and then again periodically throughout the study. The sleep health survey was filled out at the end of 6 weeks, and the patient satisfaction questionnaire was filled out at 2 weeks, 4 weeks, and then at the end of 6 weeks. We collected data about patients' objective sleep times automatically throughout the study duration through the Fitbit and SleepLife® applications. Questions from the "National Sleep Foundation: Sleep Health Index", Patient Satisfaction Questionnaire (PSQ), and the Campbell Assessment were adapted for the study purposes [8–10]. They were used to create the sleep health survey, patient satisfaction questionnaire, and the physician satisfaction questionnaire. Scores on the individual questions were tabulated to generate overall general satisfaction (GS) and communication (C) scores. The adapted questions, mapped to their outcomes, are included

as Online Appendix A. Brief descriptions of these have been given earlier.

Our direct tool of measurement was the Fitbit. Fitbits were chosen because a systematic review showed that sleep-staging Fitbit models compared to a polysomnogram were accurate in determining total sleep time and wake after sleep onset even though Fitbit data tend to underestimate sleep-onset latency [13]. It was also one of the two commercial sleep trackers that had been validated by independent testing [14, 15]. In particular, Fitbit Charge 2™ in a validation study had shown 0.96 sensitivity in detecting sleep and a 0.61 specificity in detecting wake [16]. Fitbit models have their inherent limitations and tend to overestimate total sleep time and sleep efficiency, and underestimate wakefulness after sleep onset (WASO) [13]. Sleep stage transition is difficult to assess using the Fitbit technology [2]. Fitbit sleep data has less agreement with polysomnography findings in patients with insomnia compared to those who are good sleepers [3]. However a systematic review article concluded that they have been found to be appropriate for patients to judge their own sleep hygiene and data, and for physicians to be able to gain a superficial perspective on how a patient is doing in terms of their sleep habits [5]. We utilized the Fitbit Charge 2 as a device to measure patient's sleep for this study.

Statistical Analysis

Physician Data
Physician Satisfaction/Communication Survey For each individual at each week, composite scores for physician general satisfaction (GS) and communication (C) were calculated as means of the scores for the relevant questions being answered. Questions 6, 7, 10, 11, and 13–16 were used to calculate GS; and questions 1–5, 8, 9, and 12 for the C score (Online Appendix A). Each question had a scale of 1–5. No reverse coding was performed here. Higher scores meant worse GS and C.

Demographic Variables Medical specialty was reclassified as "Family medicine" and "Internal

medicine". Number of years in practice was reclassified as < 1, 1–5, 6–10, and > 10. Descriptive statistics were summarized. This included count and frequency for categorical variables, and mean, standard deviation (SD), median, and quantiles for continuous variables.

Statistical Tests For demographic variables, two-sample *t* tests were performed to compare the difference of means of continuous demographic variables between intervention and control groups; chi-squared tests were performed to examine the association between categorical demographic variables and study groups.

For individual questions and composite scores, the difference between the scores was calculated by subtracting the baseline scores from the scores obtained at week 6. The raw means (with standard deviation) of each difference in score at week 6 were calculated for intervention and control group, respectively. The significance of the difference in the difference of scores previously calculated was evaluated using linear regression models, adjusting for age, gender, and race. The *p* values less than 0.05 were considered as statistically significant. A positive difference in scores meant worse GS and C.

Sample size and power calculations were completed using PASS. Statistical significance threshold is set as 0.05; intra-cluster correlation is set at 0.3, and effect size 0.5. For five clusters in each arm, with 20 patients in each cluster, we have statistical power of 0.85.

Patient Data
Patient Satisfaction Survey Patient satisfaction surveys were obtained at baseline, week 2, week 4, and week 6 for each patient. Participants who chose "don't know" or "refuse to answer" were coded as missing. Questions 1–4 and 10–12 were reversely coded weekly to reduce bias. For each individual at each week, composite scores for "general satisfaction (GS)", "communication (C)", and "time with doctors (TD)" were calculated as means of the scores for the relevant questions being answered respectively. Questions 1, 3, 6–8, 10, 12, and 13 were used to calculate GS; 2, 4, 5, 9, and 11 for C; and

8 and 12 for TD (Online Appendix A). Each question has a scale of 1–5. Here, when calculating a composite score, patients that answered fewer than two questions in a category (GS, C, or TD) were coded as missing. After reverse coding, higher scores meant a higher level of GS, better C, and longer TD.

Patient Sleep Survey Patient sleep surveys were obtained at baseline and week 6 for each patient. Participants who chose "don't know" or "refuse to answer" were coded as missing. Question 1 was rescaled as excellent = 100, very good = 75, good = 50, only fair = 25, and poor = 0. Question 2 was inversely coded. Questions 2–7 were rescaled as 7d = 100, 6d = 86, 5d = 71, 4d = 57, 3d = 43, 2d = 29, 1d = 14, 0d = 0. Thus, each question had a scale of 0 to 100. The mean of questions 1–6 was calculated for each survey the patients answered and provided the sleep quality index (SQ) score. The mean of questions 1–7 was calculated for each survey the patient answered and provided the sleep outcome (SL) score. Higher scores meant worse sleep quality and sleep outcome.

Demographic Variables Race was reclassified as Black/African American, White/Caucasian, and other. Education was reclassified as high school, GED or vocational school, 13 years—bachelor's degree, and higher degree. Age was calculated from each patient's birth date, with reference to 30 November 2018. "Not available" in heart rate was coded as missing. Descriptive statistics including count and frequency for categorical variables, and mean, standard deviation (SD), median, and quantiles for continuous variables were summarized.

Statistical Tests For demographic variables, two-sample *t* tests were performed to compare the difference between the means of the continuous demographic variables of the intervention and control groups. Chi-squared tests were performed to examine the association between categorical demographic variables and the study groups.

For individual questions and composite scores, the difference between the scores was obtained by subtracting the baseline scores from

week 2, 4, and 6, respectively. The raw means (with standard deviation) of each difference in scores at weeks 2, 4, and 6 were calculated for the intervention and control group separately. The significance of the difference in the difference of scores previously calculated was evaluated using generalized estimating equation (GEE) models, adjusting for age, gender, and race. The GEE model could account for potential correlation of patients seeing the same physician, and physician is hence used as the clustering variable. A p value less than 0.05 was considered as statistically significant. A positive difference in scores meant a higher level of GS, better C, and longer TD.

The sample size and power calculation was conducted prior to participant enrollment. We were planning to recruit 20 patients in each of the five clusters. However, the actual enrollment of the patients and physicians was less than what we expected. While we hoped to recruit the same number of patients and physicians in the intervention and control groups, we believe that our eventual numbers of 28 (control) vs 21 (intervention) physicians and 41 (control) vs 34 (intervention) patients are not highly unbalanced. Hence, we did not consider adjustment of this issue. There are demographic differences between the control and intervention groups. Given this, we have adjusted for age, gender, and race in the GEE-based analysis, for both patient and physician satisfaction data.
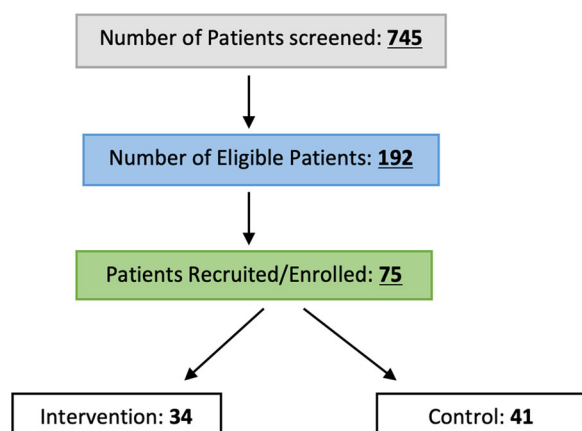
RESULTS

Study Results

Physicians and patients were recruited between June 2018 and February 2019 at Indiana University Health primary care clinics in the greater Indianapolis area. Out of the 20 enrolled primary care clinics, 12 were randomized to the control arm and eight to the intervention arm. Forty-nine physicians were enrolled with 28 physicians in the control arm and 21 in the intervention arm. We enrolled 75 patients, 41 in the control arm and 34 in the intervention arm.

Figure 2 provides the number of patients that participated and their intervention allocation status. Two patients were lost to follow-up and one patient withdrew from the study.

Physician demographics are reported in Table 1 for both the intervention and control arms. The mean age of the physicians was 42 years in the control arm compared to 53 years in the intervention arm. Patient demographic characteristics are displayed in Table 2 for both the intervention and control arms. A higher number of patients in the control group had bachelors or a higher degree [intervention, $n = 23$ (67.6%); controls, $n = 37$ (90.25%)].

The survey completion rate amongst physicians was 55%. Only one physician logged onto the SleepLife portal during the course of the study to check their patients' sleep status. Physician communication scores did not change significantly between the intervention and control group (Table 2). Physician general satisfaction (GS) scores were also not different between the groups at the end of the study (Table 3). Patients showed high rates of involvement in the program with 83% completing all survey questions. At the end of the 6 weeks, patients' composite general satisfaction scores in the intervention group decreased (i.e., worsening satisfaction with care) significantly as compared to the change in the control group ($p = 0.03$) (Table 4). Patients' perception of physician–patient communication also decreased in the intervention arm as compared

Fig. 2 Patient study sample and allocation

Number of Patients screened: **745**

↓

Number of Eligible Patients: **192**

↓

Patients Recruited/Enrolled: **75**

↓

Intervention: **34**   Control: **41**

Table 1 Physician demographics: comparison between intervention and control groups

Variable	Intervention (N = 21)	Control (N = 28)	Total (N = 49)	p value
Age in years mean (SD)	52.65 (10.15)	42.46 (8.99)	46.68 (10.65)	0.002
Gender, n (%)				0.026
Female	4 (19.05%)	14 (50.00%)	18 (36.73%)	
Male	17 (80.95%)	14 (50.00%)	31 (63.27%)	
Medical specialty, n (%)				0.017
Family medicine	9 (42.86%)	20 (76.92%)	29 (61.70%)	
Internal medicine	12 (57.14%)	6 (23.08%)	18 (38.30%)	
Number of years in practice, n (%)				0.341
< 1	0 (0.00%)	1 (4.00%)	1 (2.38%)	
> 10	13 (76.47%)	14 (56.00%)	27 (64.29%)	
1–5	1 (5.88%)	6 (24.00%)	7 (16.67%)	
6–10	3 (17.65%)	4 (16.00%)	7 (16.67%)	
Number of patients seen per day, n (%)				0.255
11–15	2 (11.76%)	6 (23.08%)	8 (18.60%)	
16–20	8 (47.06%)	16 (61.54%)	24 (55.81%)	
21–25	5 (29.41%)	3 (11.54%)	8 (18.60%)	
26–30	1 (5.88%)	0 (0.00%)	1 (2.33%)	
6–10	1 (5.88%)	0 (0.00%)	1 (2.33%)	
Over 30	0 (0.00%)	1 (3.85%)	1 (2.33%)	
Time allotted per patient in mins, n (%)				0.856
0–20 min	10 (58.82%)	14 (56.00%)	24 (57.14%)	
21–40 min	7 (41.18%)	11 (44.00%)	18 (42.86%)	
I am satisfied by the time allotted to see each patient, n (%)				0.158
Yes	12 (70.59%)	22 (88.00%)	34 (80.95%)	
No	5 (29.41%)	3 (12.00%)	8 (19.05%)	

to the change in the control arm, at the end of 6 weeks. This was indicated by the lower "communication" scores for the intervention versus control group at the end of the 6 weeks ($p = 0.01$). The sleep outcomes on the sleep health questionnaire improved significantly in the intervention group as compared to the control group ($p = 0.04$) (Table 4).

Table 2 Patient demographics: comparison between intervention and control groups

Variable	Intervention (N = 34)	Control (N = 41)	Total (N = 75)	p value
Age, years, mean (SD)	58.68 (9.77)	55.15 (14.90)	56.75 (12.87)	0.240
Gender, n (%)				0.066
Female	17 (50.00%)	29 (70.73%)	46 (61.33%)	
Male	17 (50.00%)	12 (29.27%)	29 (38.67%)	
Race, n (%)				0.232
White/Caucasian	26 (76.47%)	36 (87.80%)	62 (82.67%)	
Black/African American	7 (20.59%)	3 (7.32%)	10 (13.33%)	
Other	1 (2.94%)	2 (4.88%)	3 (4.00%)	
Education, years, n (%)				0.011
High school	10 (29.41%)	2 (4.88%)	12 (16.00%)	
GED or vocational school	1 (2.94%)	2 (4.88%)	3 (4.00%)	
13 years—bachelor's degree	20 (58.82%)	25 (60.98%)	45 (60.00%)	
Higher degree	3 (8.82%)	12 (29.27%)	15 (20.00%)	
Height, cm, mean (SD)	170.74 (11.50)	168.58 (7.99)	169.56 (9.73)	0.341
Weight at most recent clinic visit, kg, mean (SD)	85.64 (19.51)	89.55 (26.65)	87.78 (23.61)	0.479
BMI mean (SD)	29.33 (5.97)	31.58 (9.74)	30.56 (8.27)	0.243
Heart rate (HR), beats/min, mean (SD)	70.99 (17.32)	73.74 (13.78)	72.67 (15.17)	0.492
Diastolic blood pressure, mmHg, mean (SD)	77.12 (9.69)	75.20 (8.08)	76.07 (8.84)	0.352
Charlson comorbidity index, mean (SD)	0.62 (1.16)	0.59 (0.89)	0.60 (1.01)	0.892
Katz scale score, mean (SD)	6.00 (0.00)	6.00 (0.00)	6.00 (0.00)	0.275
Lawton scale score, mean (SD)	8.00 (0.00)	8.00 (0.00)	8.00 (0.00)	0.275
Insurance, n (%)				0.269
Yes	33 (97.06%)	41 (100.00%)	74 (98.67%)	
No	1 (2.94%)	0 (0.00%)	1 (1.33%)	
Prior experience with a sleep tracker, n (%)				0.788
Yes	25 (73.53%)	29 (70.73%)	54 (72.00%)	
No	9 (26.47%)	12 (29.27%)	21 (28.00%)	
Medications at baseline n (%)				0.182
Antidiabetic	2 (5.88%)	1 (2.44%)	3 (4.00%)	
Antihypertensive	5 (14.71%)	1 (2.44%)	6 (8.00%)	
Sedating	15 (44.12%)	19 (46.34%)	34 (45.33%)	
Other	12 (35.29%)	20 (48.78%)	32 (42.67%)	

Table 3 Comparison between general satisfaction and communication scores between the physicians in the intervention and control groups

Score	Intervention	Control	p value
General satisfaction (GS)	− 0.28 (0.80)	0.11 (0.87)	0.123
Communication (C)	0.09 (0.84)	0.19 (0.76)	0.328

Data is shown in the format of mean (SD)

The second column shows the mean and standard deviation of the score difference between week 6 and baseline (week 6 minus baseline) in the intervention group, and the third column shows the difference in the control group. The fourth column shows the GEE test significance of the difference of the difference values

Table 4 Comparison between the patients' satisfaction and sleep composite scores between intervention and control groups

Score	Intervention	Control	p value
Patient satisfaction survey			
General satisfaction (GS)			
Week 2	− 0.5 (0.80)	0.05 (0.51)	< 0.001
Week 4	− 0.47 (0.8)	0 (0.42)	< 0.001
Week 6	− 0.32 (0.62)	− 0.09 (0.49)	0.039
Communication			
Week 2	− 0.76 (0.84)	0.01 (0.58)	< 0.001
Week 4	− 0.72 (0.79)	0.07 (0.44)	< 0.001
Week 6	− 0.49 (0.85)	− 0.03 (0.49)	0.011
Time with doctor (TD)			
Week 2	− 0.64 (1.39)	− 0.04 (0.58)	0.014
Week 4	− 0.76 (1.22)	− 0.04 (0.64)	< 0.001
Week 6	− 0.32 (0.82)	− 0.01 (0.69)	0.112
Patient sleep survey			
Sleep outcome (SL)	− 7.23 (19.74)	1.5 (15.62)	0.048
Sleep quality (SQ)	− 7.27 (22.26)	2.67 (15.37)	0.033

Data is shown in the form of mean (SD)

The second column shows the mean and standard deviation of the score difference between weeks 2/4/6 and baseline (week 2/4/6 minus baseline) in the intervention group, and the third column shows the difference in the control group. The fourth column shows the GEE test significance of the difference of the difference values

A decrease in the patients' composite general satisfaction scores signifies worsening satisfaction with care. A lower communication score signifies a decreased patient perception of physician–patient communication. Higher scores on patient sleep survey meant worse sleep quality and sleep outcome

Fitbit Results for Sleep Time

The SleepLife® application collected sleep-related measures directly through the Fitbit device. Fitbit data contained the measurements for seven variables, which were available for 43 patients over the duration of the study. These seven variables included sleep duration, total time the patient spent in bed, latency to sleep onset, number of times they woke up after falling asleep, time they spent awake after falling asleep, percentage of time they spent awake in bed, and sleep efficiency (Online Appendix B Figs. B.1–B.7). These data demonstrate the feasibility of capturing sleep data through our program using the SleepLife® application and Fitbit.

Safety

No adverse events occurred during the course of the study.

DISCUSSION

Physician–patient communication and patient satisfaction regarding sleep and sleep-related behaviors and habits can be collected through the SleepLife® application. Patients demonstrated very high engagement levels with the program, while physicians had very poor engagement. Patients in the intervention group showed worsening satisfaction with their sleep care and physician–patient communication, possibly highlighting a perception gap with the expectation that they would be receiving enhanced attention from their primary care provider regarding their sleep complaints. Sleep health outcomes improved in both groups but there was a significant increase in the intervention group. This discrepancy could be a result of patients paying more attention to their sleep habits and taking steps that may improve their sleep regardless of their physicians' input. However, since both groups received a Fitbit and were able to view their own sleep health data, this difference could reflect the impact that the thought of more engagement, time and focus spent on sleep health could affect their actual sleep. The use of a symptom diary to self-monitor and provide data to both patients and physicians has been well studied and shown to be beneficial and have a positive impact on outcomes when both sets of participants are engaged in the activity [17]. Sleep diaries have been prescribed by physicians for a long time to help patients self-monitor and identify sleep patterns. Given the popularity of sleep trackers, as mentioned earlier, many have started monitoring their sleep similarly but using technology. If physicians were able to directly view their patients' sleep outcomes, this could help to appropriately prescribe medications for patients depending on which sleep behavior they are having trouble with and to help them with better sleep hygiene and non-pharmacological measures. They may be able to have objective outcomes to work with once patients implement these measures that can provide positive feedback.

As stated earlier, physicians were not engaged in the project and did not close the feedback loop as we had expected. This could be a possible reason for the significant decrease in general patient satisfaction and physician–patient communication scores in the intervention arm compared to the control arm. Patients likely expressed less satisfaction with their doctors when data pertaining to their sleep was sent from their wearable devices to their primary care physicians, but no action was taken. This observation could be due to multiple reasons. Firstly, it may appear as if the physicians are not spending adequate time on communication regarding sleep. During the process of enrollment, a couple of patients commented that they felt as though their sleep disturbances were not a priority for their physicians and were glad to hear about the study because it felt like it was finally given significance. Many patients expressed frustration over their sleeping problems. Secondly, the patients themselves may also simply be more aware of their sleeping habits because of using the Fitbit devices, the SleepLife® application, and being a part of the study itself. Additionally, an absence of strong bidirectional communication around the patient's disordered sleep habits could foster

feelings of decreased satisfaction with respect to sleep care. There was no post-participation debriefing regarding the lack of communication between physicians and patients that was supervised by the research team, although this would be useful for patients in future studies.

Research has shown that primary care physicians and advanced practice registered nurses perceive that wearable devices are helpful in managing patients' health and that patients are interested in these technologies. However, the studies do not find wearable devices useful for tracking patients' sleep, compared to patients' physical activity and diet [2]. This was attributed to most likely be due to sleep being a secondary tracking feature on most wearable devices. Although our study did not focus on this in particular, it may have played a role in the low rates of physician engagement. The SHADES study showed that patients or physicians were not discussing sleep more than a third of the time over the course of a patient's care [6]. This is consistent with our study which showed generally low levels of general satisfaction and communication in the patient satisfaction questionnaire which were about their communication with their primary care physicians about sleep. The work in the SHADES study also indicated that patients that were depressed were more likely to raise their concerns about sleep with their physicians [6]. This was something that we could not assess for in our study given that we excluded patients with the diagnosis of depression, but would be interesting to discuss further; it is possible that the presence of a mental health condition may prompt the idea of disordered sleep to come up more in conversation, and improving discussions about sleep would require us to think about why this would happen.

One of the major barriers in the diagnosis of sleep disorders is the fact that patients tend not to volunteer information pertaining to their sleep to their primary care physicians [18]. Physician education in sleep and sleep disorders has been studied as being inadequate as well [4]. This has led to a low rate of recognizing, and consequently managing, sleep disorders in primary care at both academic and community centers [5]. Primary care doctors are often uncertain regarding their diagnosis, or do not assess for insomnia at all either as a result of feeling uncomfortable with the diagnosis, treatment, or because of time constraints [18]. Our study did not adequately assess for either of these barriers. It highlights the need for increased physician awareness and engagement to address the needs of patients suffering from insomnia. Only one physician actually accessed their patients' sleep data via the SleepLife portal. This was a major limitation of our study, indicating that the current SleepLife portal and application may not be a feasible mode of delivery/presentation of objective patient sleep data to physicians. We performed a small focus group with four physicians from one study site after the study had been completed to evaluate reasons why physician engagement was low in the study. The reasons for non-engagement included lack of time, not having the motivation or not remembering to log into a patient portal when that patient was not scheduled for a visit, and not feeling comfortable interpreting and acting upon the data received. The lack of physician engagement in our study meant that patients and physicians were not communicating about sleep, leading to results not accurately reflecting satisfaction with that communication. In fact, patients may have been more frustrated having been told that they would be able to discuss their sleep concerns with their physicians and then not getting a response from them. If the study had gone on for more than 6 weeks and would have included an in person visit with the physician where the patient presented their Fitbit results to the physician and they discussed them together, this may have meant significantly improved communication. In short, wearable devices allow patients to view their own sleep data, communicate with their physicians, and may instill a greater sense of independence. However, in the absence of reciprocity from their physicians, patient satisfaction is adversely affected. This supports the theory that collaborative physician–patient communication, in which the patients share responsibility and control over their health along with their physicians, could result in better care and outcomes.

Physicians' perception of physician–patient communication and physician satisfaction scores did not demonstrate significant differences between the intervention and control arms. These results are not robust as there was minimal to no physician engagement during the course of the study with only one physician logging into the SleepLife® portal. One of the major discussion points with physicians as we were ending the study was about how to integrate the SleepLife® portal into their workflow. Physicians reported that if they had the time, or possibly an additional incentive, they would be more open to review patient data via the SleepLife® portal. Of note, studies have attempted to formulate care coordination plans that eased time constraints and engaged physicians in patient communication more regularly [19–22]. This needs to be explored further in a study with a larger sample size of physicians, a more robust system of obtaining physician outcome assessments or questionnaires, and utilizing behavioral nudges incorporating the principles of behavioral economics to improve engagement and collaboration.

Our pilot study provides insight into what can be taken forward to design a future study. One of the strengths was the randomized design which minimized bias and possible confounding factors to a minimum in both groups. The SleepLife® portal and application were easy to navigate and had a well-designed user interface. Troubleshooting of the SleepLife® application was performed quickly and efficiently. We received this feedback from physicians themselves and by patients who called in for assistance with the application. Patient enrollment was streamlined as the study moved forward. Phone calls were found to be an acceptable and effective method of gaining assent from patients. Utilization of the SleepLife® application to collect questionnaire data was efficient and led to over 80% of patients completing all questionnaires. We were able to enroll several patients over the age of 60 years, allowing us to view and learn from their interaction with a wearable device and technology generally.

Limitations

There were several limitations to this study. The major one was the lack of physician engagement. They did not check the SleepLife® portal and sending them reminders was both not acceptable to them and did not work. It had a significant impact on our results. Physicians would prefer to have sleep data brought to them by a patient at their individual clinic visit, or have an incentive, financial or otherwise, to look at the portal.

One way to improve physician engagement is to create a tool that automatically sends a physician their patient's sleep data via EMR message every couple of weeks in an easy-to-read format. This kind of software would eliminate the need for physicians to access a separate website, remember to log in to the website, and thus save time. It would integrate with the patient's EMR and allow physicians to access their information in a timely manner without them having to learn new software or check a different interface themselves. We could also add a page on specific therapy/management suggestions based on what difficulties patients may face with sleep, e.g., trouble initiating sleep, staying asleep, etc., on sleep hygiene, and on medications that may be appropriate for particular age groups. As part of a study, another way to do so would be to provide an incentive for physicians like CME credits or meal gift cards with every message they send to their patients regarding sleep. When applying the study results to the real world, as stated earlier, a tool that is integrated with the EMR and provides the physicians with data instead of them having to seek it out is the best way to improve their engagement. Other ways include providing physicians in training, and already trained physicians with easy access to resources on sleep education, allowing them to be more comfortable with diagnosing, discussing, and managing sleep disorders.

Physicians were also not enthused at having to complete physician satisfaction questionnaires. Many therefore may have been in a rush and not answered the questionnaires thoughtfully. Primary care physicians may benefit from additional sleep-based educational modules and

discussions around how to integrate technology into primary care and how helpful it can be. Another major limitation was that we did not include patients with mental health conditions or with a history of alcohol or substance abuse. This decreased our pool of patients greatly because of the high percentage of overlap between mental health conditions and sleep disorders. The study team felt that including mental health disorders may affect a patient's sleep independently and make it difficult to see a real difference due to the intervention itself. The small sample size of the study precluded us from making any meaningful inference on findings amongst the different sleep disorders as well. It was also very time intensive to enroll clinics and physicians. The response time for clinics was slow, leading to the study going on for longer than planned. Although we were able to enroll many patients over the age of 60, it was quite time intensive to enroll and teach them how to use the Fitbit and the SleepLife® application. Encouraging them to use the application and the device may be difficult.

CONCLUSION

The use of wearable devices to measure patients' sleep measures and to provide data to primary care physicians is an area that has the potential for a significant amount of research and discussion. Patient engagement levels were high; however, physician engagement was extremely poor. The absence of engagement of physicians in the program could explain the decrease in physician–patient communication and patient general satisfaction scores in the intervention arm. Future studies that consider and work around the limitations reported in our pilot project are required to further consumer health technology use in primary care, which could potentially lead to improved physician–patient communication and sleep-related health outcomes.

ACKNOWLEDGEMENTS

Shalini Manchanda and Babar Khan have nothing to disclose. Jennifer K. Pai and Zaina P. Qureshi are employees of Merck & Co. but have nothing additional to disclose.

Compliance with Ethic Guidelines. This prospective, randomized, parallel group observational pilot study was conducted in Indianapolis, Indiana. Patients and physicians were enrolled from primary care clinics in the greater Indianapolis area. The study protocols were approved by the Indiana University Institutional Review Board (IRB), IRB Protocol Number 1802077293 and Merck & Co. DRC. The IRB (Protocol Number 1802077293) provided approval coverage for the 20 primary care clinics where the study was conducted. This study was performed in accordance with the Helsinki Declaration of 1964, and its later amendments. All subjects provided informed consent to participate in the study in addition to consent for publication if any identifying information is included in the manuscript.

Data Availability. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

REFERENCES

1. Fox S, Duggan M. Tracking For Health. Pew Research Center's Internet and American Life Project; 2013. https://www.pewinternet.org/2013/01/28/tracking-for-health/. Accessed 6 May 2019.

2. Holtz B, Vasold K, Cotten S, Mackert M, Zhang M. Health care provider perceptions of consumer-grade devices and apps for tracking health: a pilot study. JMIR Mhealth Uhealth. 2019;7(1):9929. https://doi.org/10.2196/mhealth.9929.

3. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act. 2015;12:159. https://doi.org/10.1186/s12966-015-0314-1.

4. Rosen RC, Rosekind M, Rosevear C, Cole WE, Dement WC. Physician education in sleep and sleep disorders: a national survey of U.S. medical schools. Sleep. 1993;16:249–54.

5. Rosen RC, Zozula R, Jahn EG, Carson JL. Low rates of recognition of sleep disorders in primary care: comparison of a community-based versus clinical academic setting. Sleep Med. 2001;2(1):47–55.

6. Klingman KJ, Williams NJ, Perlis ML, Grandner MA. Doctor-patient sleep discussions for US adults: results from the SHADES study. Sleep Health. 2019;5(6):658–65. https://doi.org/10.1016/j.sleh.2019.07.004.

7. Katz S, Ford A, Moskowitz R. Studies of illness in the aged. JAMA. 1963;185(12):914. https://doi.org/10.1001/jama.1963.03060120024016.

8. Sicotte C, Tilquin C, D'Hoore W. Risk adjustment in outcome assessment: the Charlson comorbidity index. Methods Inf Med. 1993;32(05):382–7. https://doi.org/10.1055/s-0038-1634956.

9. Knutson K, Phelan J, Paskow M, et al. The National Sleep Foundation's Sleep Health Index. Sleep Health. 2017;3(4):234–40. https://doi.org/10.1016/j.sleh.2017.05.011.

10. Hays RD, Davies AR, Ware JE. Scoring the medical outcomes study patient satisfaction questionnaire: PSQ-III. MOS Memorandum. 1987. https://www.rand.org/content/dam/rand/www/external/health/surveys_tools/psq/psq3_scoring.

11. Herr KA, Garand L. Assessment and measurement of pain in older adults. Clin Geriatr Med. 2001;17(3):457–vi.

12. Charlson M, Sax F, Mackenzie C. Resuscitation: how do we decide? JAMA. 1986;255(10):1316.

https://doi.org/10.1001/jama.1986.03370100110027.

13. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of wristband fitbit models in assessing sleep: systematic review and meta-analysis. J Med Internet Res. 2019;21(11): e16273. https://doi.org/10.2196/16273.

14. de Zambotti M, Baker FC, Willoughby AR, et al. Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. Physiol Behav. 2016;158: 143–9. https://doi.org/10.1016/j.physbeh.2016.03.006.

15. de Zambotti M, Claudatos S, Inkelis S, Colrain IM, Baker FC. Evaluation of a consumer fitness-tracking device to assess sleep in adults. Chronobiol Int. 2015;32(7):1024–8. https://doi.org/10.3109/07420528.2015.1054395.

16. de Zambotti M, Goldstone A, Claudatos S, Colrain IM, Baker FC. A validation study of Fitbit Charge 2^{TM} compared with polysomnography in adults. Chronobiol Int. 2019;35(4):465–76. https://doi.org/10.1080/07420528.2017.1413578.

17. Ferrari R, Russel AS. Effect of a symptom diary on symptom frequency and intensity in healthy subjects. J Rheumatol. 2010;37(11):2387–9.

18. Ogeil RP, Chakraborty SP, Young AC, Lubman DI. Clinician and patient barriers to the recognition of insomnia in family practice: a narrative summary of reported literature analyzed using the theoretical domains framework. BMC Fam Pract. 2020;21(1):1. https://doi.org/10.1186/s12875-019-1070-0.

19. Bardach SH, Schoenberg NE. Primary care physicians' prevention counseling with patients with multiple morbidity. Qual Health Res. 2012;22(12): 1599–611. https://doi.org/10.1177/1049732312458183.

20. Jaen CR, Stange KC. Competing demands of primary care: a model for the delivery of clinical preventive services. J Fam Pract. 1994;38(2):166–71.

21. Engelman D, Benjamin E. Physician engagement: the "secret sauce" to success in bundled health care. Am J Med Qual. 2018;33(1):100–2.

22. Ha JF, Longnecker N. Doctor-patient communication: a review. Ochsner J. 2010;10(1):38–43.