Published in final edited form as:

J Asthma. 2019 July ; 56(7): 719–730. doi:10.1080/02770903.2018.1490753.

Validation of Fitness Tracker for Sleep Measures in Women with Asthma

Jessica Castner^{1,2,3}, Manoj J. Mammen^{4,5}, Carla Jungquist⁶, Olivia Licata⁷, John J. Pender⁸, Gregory E. Wilding⁹, and Sanjay Sethi⁵

¹The Rockefeller Heilbrunn Family Center for Nursing Research, New York, New York, USA

²University at Buffalo, Buffalo, New York, USA

³Castner Incorporated, Grand Island, New York, USA

⁴Department of Biomedical Informatics, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, New York, USA

⁵Department of Medicine, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, New York USA

⁶University at Buffalo School of Nursing, Buffalo, New York, USA

⁷University at Buffalo School of Engineering and Applied Sciences, Buffalo, New York, USA

⁸University at Buffalo School of Nursing, Buffalo, New York, USA

⁹Biostatistics, School of Public Health and Health Professions, University at Buffalo, Buffalo, New York, USA

Abstract

Objective: Night-time waking with asthma symptoms is key to assessment and therapy decisions, with no gold standard objective measure. The study aims were to 1) determine the feasibility, 2) explore equivalence, and 3) test concordance of a consumer-based accelerometer with standard actigraphy for measurement of sleep patterns in women with asthma as an adjunct to self-report.

Corresponding author: Jessica Castner, PhD, RN, FAEN, jpcastner15@gmail.com or jcastner@castnerincorp.com telephone: 716-417-7360; fax: 716-774-1050 ; 3480 Stony Point Road, Grand Island, NY 14072.

Detailed author contribution: J. Castner led the study conception and design of the work, collection, interpretation, and analysis of the data, and drafted the manuscript. M. Mammen, C. Jungquist and S. Sethi contributed to the conception of the study as content experts. G. Wilding and M. Mammen contributed to the analysis plan. O. Licata and J. Pender contributed to participant recruitment and retention, data collection, and processing. All authors revised the paper critically for intellectual content, substantive revisions, and approved the final version to be published.

Declaration of Interest: Authors CJ, MJM, JP, OL, GW, & SS declared that no competing interests exist. JC's research has been funded by grants from NSF (#1737617, #1645090), The Ohio State University, and had been committed support from the Environmental Health Study for Western New York (also documented as The Tonawanda Health Study: An Epidemiologic Study of Health Effects and Coke Oven Emissions from Tonawanda Coke). Dr. Castner has received support from the Emergency Nursing Association/Journal of Emergency Nursing for teaching, speaking honoraria, and editor/editorial board travel. Dr. Castner has had a research/data analysis consulting relationship with the Harvard T. H. Chan School of Public Health, American Lung Association, and hospitals in the region of study, with additional funding/consulting disclosures which are unlikely to be perceived as a conflict of interest available on request. The authors alone are responsible for the content and writing of the paper.

Methods: Panel study design of women with poorly controlled asthma from a university-affiliated primary care clinic system was used. We assessed sensitivity and specificity, equivalence and concordance of sleep time, sleep efficiency, and wake counts between the consumer-based accelerometer Fitbit Charge™ and Actigraph wGT3X+. We linked data between devices for comparison both automatically by 24-hour period and manually by sleep segment.

Results: Analysis included 424 938 minutes, 738 nights, and 833 unique sleep segments from 47 women. The fitness tracker demonstrated 97% sensitivity and 40% specificity to identify sleep. Between device equivalence for total sleep time (15 and 42-minute threshold) was demonstrated by sleep segment. Concordance improved for wake counts and sleep efficiency when adjusting for a linear trend.

Conclusions: There were important differences in total sleep time, efficiency, and wake count measures when comparing individual sleep segments versus 24-hour measures of sleep. Fitbit overestimates sleep efficiency and underestimates wake counts in this population compared to actigraphy. Low levels of systematic bias indicate the potential for raw measurements from the devices to achieve equivalence and concordance with additional processing, algorithm modification, and modeling. Fitness trackers offer an accessible and inexpensive method to quantify sleep patterns in the home environment as an adjunct to subjective reports, and require further informatics development.

Keywords

asthma; actigraphy; fitness tracker; sleep disruption; women

Background

Poorly controlled asthma has an estimated \$80 billion per year impact on the US economy, highlighting the public health need for increased monitoring and preventative intervention. (1) Night-time waking with asthma symptoms is a key indicator of disease control and severity, and provide an essential cue for self-management, behavior modification, healthcare utilization, and therapy adjustments. (2) Of the 8% of the population who currently have asthma, an estimated 24% had difficulty staying asleep in the prior month because of asthma. (3) Asthmatic airway obstruction often worsens at night, due to a complex interplay of factors such as cough suppression, hormonal changes, increased inflammatory mediator release, and airway resistance. (4) The prevalence and severity of adult asthma disproportionately affects women compared to men, with higher death rates, (5) emergency (6) and inpatient care utilization. (7) There is a paucity of research to identify clinically relevant differences in objectively measured sleep disruption in those with asthma, as exists with sleep disorders or self-reported asthma-specific awakenings. (9–11) While standard care is based on the self-report of asthma-related awakenings, (2) the ability to objectively quantify sleep patterns is promising as a real-time, adjunct, surveillance method for assessing disease control and severity in the future. (9)

There is no gold standard for an objective adjunct to measure asthma-specific waking. While asthma-specific waking is not an indication for polysomnography (PSG), PSG is the gold standard to measure sleep states and stages related to sleep-disorders and sleep

disordered breathing. PSG is impractical for long-term home use, and related measures of respiratory effort and desaturation have not been linked to the clinical presentation of asthma exacerbation.(11) Actigraphy uses an accelerometer to measure movement, worn on a limb or the hip. The Actigraph GT3X+ (Actigraph Corporation, Pensacola, Florida) is one of the most commonly used actigraphy devices in clinical and epidemiology research.(12, 13) The device has demonstrated validity, reliability,(12) and the ability to measure sleep patterns that predict health-related outcomes in several populations, including healthy adults,(12) diabetic adults,(14) cancer survivors,(15) community-dwelling older adults,(16) and children.(17) Actigraphy devices have been used to measure sleep patterns in the adult and child asthma population,(10) demonstrating greater accuracy in healthy controls, and the ability to distinguish worsened sleep patterns in those with asthma.(18, 19) Standard actigraphy devices are 5-10 times the cost of consumer-based fitness trackers.(20) While there are several generations of consumer fitness tracker products, Fitbit® fitness trackers demonstrate consistent performance for health-related measures.(12, 21–24) When PSG, standard actigraphy, and the Fitbit were compared as tools to measure sleep duration and efficiency, the Fitbit demonstrated inter-device reliability (.97-.99), was slightly more sensitive in measuring the presence of sleep than standard actigraphy (91-98% sensitivity), but less specific to measuring short-duration awakening.(20, 24) Fitbit has been validated with actigraphy in short-term studies in several populations,(12, 20, 22, 24) but there is a gap in evidence for use in women with asthma. Thus, because there is no indication to utilize PSG to measure asthma-specific wakening, and previous study has established comparisons among PSG, Fitbit, and actigraphy devices in several populations, we utilized actigraphy as the ground-truth objective adjunct for our study.

Objective

We explored feasibility, equivalence and concordance of the consumer-based accelerometer, Fitbit Charge™ (Fitbit, Inc.®, Pensacola, Florida) with standard actigraphy, Actigraph wGT3X+ (Actigraph Corp., Boston, Massachusetts) for the objective measurement of sleep patterns in women with poorly controlled asthma. Specifically, we tested total sleep time, sleep efficiency, and wake count measures.

Materials and Methods

Subjects and Setting

Fifty-one women with poorly controlled asthma provided written consent to be included in this study, which utilized an observational, repeated measures panel design from February to September of 2016. An *a priori* sample size of 50 was determined to enable adequate power for medium effect sizes in the repeated measures analyses or to adequately correct for missing data. Institutional Review Board approval was obtained. Participants were recruited from an Erie County, New York, university-affiliated primary care clinic system. A recruitment dataset was created from the electronic medical record repository, composed of women, ages 18-64 years with an asthma international diagnoses code (ICD-9) of 493.xx on the active problem list who had attended the clinic system for care within three years preceding the beginning of the study. Those with a current diagnosis of obstructive sleep apnea in the active problem list of the medical record repository or who lived along the less

populated perimeter or outside the county (for scheduling and budget reasons) were excluded. During telephone screening, potential participants were included if they had an Asthma Control Test (ACT) score ≥ 21 , (25) and had access to a mobile device or home computer capable of downloading Fitbit Connect software. Potential participants were excluded if they reported they were pregnant, actively trying to become pregnant or taking fertility medication, usually worked night shift, had dependents or pets who disrupted middle of the night sleep more than 2 times per week, at risk for sleep apnea using the Berlin screen (2+/3 categories with a positive score), (26) a diagnosis of neuromuscular disease (such as MS), or a current diagnosis of cancer of the lung, bronchus, or oropharynx which may affect the ability to complete spirometry. Those who had a self-reported history of sleep apnea, such as those who were no longer undergoing active therapy after weight loss, but did not have an active diagnosis on their primary care clinical record, were included. Data were not used from two participants who did not wear both devices at the same time and two who did not participate in recording sleep data. Additional information on self-report and medical record abstraction records can be found in the supplemental material.

Sleep Pattern Measures

Nightly measures of total sleep time, sleep efficiency, and wake counts (bouts, or number of awakenings after sleep onset) were collected using the Fitbit Charge™ and the Actigraph wGT3+. The devices were worn on the non-dominant wrist, (27) with the Fitbit distal to the Actigraph. Participants were scheduled to wear the devices together for approximately 2 weeks, extended depending on availability to schedule visits or difficulty with the devices. Based on previous study, the Fitbit was left in the default, rather than sensitive mode. (24) The Fitbit Charge™ proprietary algorithm (Firmware version 110 and Fitbit Connect software version 2.0.1.6782) processed the sensor data to automate the sleep variables. Fitbit data was synced by the participant at least twice a week to Fitabase (Small Steps Labs LLC, San Diego, California) third party integration cloud. Participants were instructed not to update the application for Fitbit in her smartphone or computer during the study period. The frequency of syncing was checked by the research staff weekly, with participant reminders as needed. The Actigraph device was configured to 100Hz, sleep enabled 60-second epoch, and processed with ActiLife (version 16.3.1) software.

Data Analysis

Data were analyzed using STATA (Version 14.2, StataCorp, College Station, Texas). Differences between participants and those who were not recruited from the recruitment dataset (electronic medical record repository, see Figure 1) were assessed using Mann-Whitney U or χ^2 statistics, and correlations using Pearson product moment. The Actigraph and Fitbit raw sleep datasets were processed and merged automatically by 24-hour period and manually by matching sleep segments (see supplemental material for protocol). Data from each device distinguished disrupted sleep measures (wake bouts or restlessness within the total sleep duration) from separated segments of sleep identified with unique time in bed and sleep onset. Segmented sleep was defined as nights in which sleep occurred in two or more sleep segments, separated by enough wrist movement in which one of the device algorithms' indicated a time period out of bed, rather than a full night's sleep within a single time period. **Feasibility** was assessed by ascertaining the start and end dates in which the

devices were worn consecutively, and calculating the percent of study nights with data from both devices. Actigraphy was automatically scored each 60-second epoch as a binary asleep or awake state by the software. This version of fitness tracker results indicated asleep, restless, or awake. For the purposes of the epoch-by-epoch analysis, Fitbit restless codes were included in a binary “awake” score. Sensitivity, specificity, positive and negative predictive values were calculated from the 60-second epoch-by-epoch values of the two devices. Using the actigraphy results as the ground truth, a true positive was scored when both devices indicated sleep, true negative when both devices indicated wake, a false negative when actigraphy indicated sleep and Fitbit wake, and false positive when actigraphy indicated wake and Fitbit sleep. Three sensitivity analyses were conducted on these results by 1) alternatively coding the Fitbit restless category into the binary sleep state 2) using datasets with values excluded as implausible or unlikely after visual inspection of actigram, and 3) replacing missing Fitbit values with the “wake” code.

Logarithmic transformations were utilized for severe skew of the Fitbit wake counts.

Equivalence testing for each device was explored (two one-sided tests of equivalence) using the *a priori* equivalence levels that were 10% of the actigraphy mean score for each device (45 minutes total sleep time, 9% sleep efficiency, and 16 wake counts). **Concordance** was plotted using the Bland-Altman technique, including standard differences and regression lines between the devices.(28) Sensitivity analyses were conducted for datasets after additional data cleaning steps (e.g. deleting implausible values).

Results

Participants

Data was available from 47 participants for concordance testing (see Figure 1). Table 1 lists the participant’s demographic and clinical characteristics. Differences between those who consented to participate in the study and the remainder of the electronic medical record repository database can be found in the supplemental material.

Descriptive Statistics for Sleep

The box plots in Figure 2 depict the distribution of the variables total sleep time, sleep efficiency, and wake counts, comparing the manual and automated linked datasets. Table 2 summarizes the average nightly measures by each device with each night in which the participant wore both devices as the unit of analysis, and within each participant as the unit of analysis.

Feasibility

Combined device data was available on 75.5% of the possible consecutive nights of wear (738/978), with 41 (4.2%) nights of missing data from both devices in the same night, 78 (8.0%) missing nights from actigraphy alone, and 121 (12.4%) missing nights from the Fitbit fitness tracker alone. Reasons for missing data included device removal, damage, loss, or dead battery, participant skipping a night of sleep, participant activity and restlessness preventing sleep detection. Participants averaged 16.43 nights (SD=5.91) and a range of 8-37 nights per participant. 833 matched sleep segments were analyzed using manually

processed data. The distribution of the number of sleep segments each night for the fitness tracker data was 55.2% (n=536) for 1 segment, 22.3% (n=216) for 2 segments, 4.6% (n=46) for 3 sleep segments, and 2.2% (n=21) with 4-6 sleep segments.

The Fitbit fitness tracker demonstrated 97% sensitivity and 40% specificity to detect sleep in a minute-by-minute comparison to actigraphy, using data on a total of 424 938 minutes of sleep. Sensitivity analyses of the same tests with data cleaned, coded, and merged using different techniques, are presented in Supplemental Table 1 and 2. Overall, the results were consistent when using a dataset that was further cleaned and refined after removing questionable or implausible values from visual inspection of the actigram. Scoring the fitness tracker “restless” as sleep, rather than wake, increased the sensitivity (.97 to .99), but decreased the specificity (.40 to .10) of the fitness tracker compared to actigraphy. Finally, replacing fitness tracker missing data as “wake” increased the specificity (.93-.95) while sacrificing sensitivity (.55-.57) of the fitness tracker’s ability to detect sleep compared to actigraphy.

Equivalence

The mean difference for the manual and automatically merged data, respectively, are as follows: total sleep time −5.37 minutes (90% CI −13.17, 2.42) and 26.97 minutes (90% CI 13.52, 40.43), sleep efficiency 5.57% (90% CI 5.22, 5.91) and 5.29% (4.95, 5.63) and wake counts −11.81 (90% CI −12.25, −11.38) and −11.97 (−12.41, −11.53). The devices were equivalent at the *a priori* 42-minute threshold for total sleep time, as well as an exploratory 15-minute threshold, in the manually processed data by sleep segment. There were statistically relevant differences for the remaining measures (Supplemental Table 3).

Concordance

The results of concordance testing are illustrated in Figures 3–5 with Bland-Altman plots for standard differences and adjusted for linear trend for total sleep time, sleep efficiency, and number of awakenings. Supplemental Figure 1 is also included with the plots of the wake counts before logarithmic transformation for skew. Table 3 summarizes raw score correlations and the concordance results with the number of nights outside of the 95% limits of agreement between the devices, with and without adjusting for linear trend. Correlations of the raw scores were low (.09-.52, $p < .05$).

Discussion

Nearly three fourths of the adult United States population routinely tracks and monitors at least one health related measure for themselves or their loved ones.(17) Given the ubiquitous use of fitness trackers, the purpose of this research was to assess feasibility,(29–31) to explore equivalence, and test concordance of a fitness tracker and actigraphy device to measure total sleep time, sleep efficiency, and wake counts in women with poorly controlled asthma. The strengths of this current study design were to 1) add important evidence in the feasibility and accuracy of using inexpensive fitness trackers to measure sleep patterns longitudinally in the home environment, 2) study a sample representative of diverse and

vulnerable populations, and 3) lengthen the observation period of assessing abnormal sleep patterns in those with asthma from previous study.(9)

Only 55% of the nights of sleep in our study were recorded as a single, uninterrupted sleep period. Our data demonstrated a substantial proportion of nights with segmented sleep (See Supplemental Figure 2 for an example actigram). This evidence opens a potential new line of future inquiry to ascertain if segmented sleep should be considered normal or pathological in our population.(32) Due to the large proportion of nights with segmented sleep, we innovated a method to compare the two devices by linking the data in two different ways, 1) automatically by 24-hour noon-to-noon period and 2) manually by individual sleep segments. Manually processing the sleep segments was labor intensive in our study, and not feasible for clinical practice translation without additional software and algorithm updates to the fitness tracker data. However, the findings in our study from the different methods of linking these data demonstrate the need for additional clinical and research inquiry into the measurement of potentially segmented sleep habits of working-age women with asthma.

Our feasibility analysis demonstrated that roughly 75% of the consecutive study nights resulted in usable data, which is an improvement from a previous study.(33) Pragmatic feasibility challenges including the battery life of the devices, potential user forgetfulness to wear the device every night, broken or uncomfortable wristbands (particularly on nights with warmer temperatures), lost devices, or unwillingness to wear the device on certain days of the week or during social activities. While we did not formally collect data on ongoing use of fitness trackers, several participants spontaneously notified study staff they were purchasing a fitness tracker for personal use, which corroborates evidence from other studies on the perceived value of sharing fitness tracker data with healthcare providers.(34) Similar to our experience in educating our participants to utilize the technology, studies in other populations have found that participants may struggle with the technology during initial learning and set-up, but were able to wear and care for the devices, and are satisfied with the comfort and design features.(33–35)

With 97% sensitivity and 40% specificity, the Fitbit fitness tracker's ability to detect sleep also demonstrated the likelihood to overestimate sleep. Further study is needed to ascertain 24-hour epoch-by-epoch sensitivity and specificity. The version of fitness tracker data we used does not enable 24-hours of epoch-by-epoch data downloads. The comparisons in our present study included the 60-second epochs between sleep onset and final waking determined by the fitness tracker algorithm. Our results corroborated findings from 21 patients with a major depressive disorder, demonstrating 98% sensitivity and 35% specificity compared to PSG in a one-night sleep lab study.(24) However, this previous study found no differences between the fitness tracker and actigraphy, as we have demonstrated. The sensitivity analysis in our study (supplemental material) demonstrated that coding missing daytime data as "wake" for the fitness tracker resulted in a substantial increase in the specificity to measure sleep (95%), with an expected drop in sensitivity (55–56%). The equivalence and concordance results of our study suggest the devices have substantially different algorithms to process and produce the sleep metrics for clinical and research use. Overall, the Fitbit demonstrated borderline accuracy, compared to the Actigraph, in this population. The more disrupted the sleep (with >5 actigraphy-measured awakenings per

night), the less accuracy has been demonstrated in total sleep time measures overall (33) and in those with asthma.(16)

Overall, our results corroborate similar evidence of overestimating sleep efficiency and underestimating waking measures noted in other studies of healthy adults (12, 20) and those with major depressive disorders (24) when we linked the two device's data automatically with a programming code. Actigraphy and fitness trackers consistently overestimate sleep time and efficiency by miscategorizing motionless wakefulness as sleep when compared to PSG.(20) However, the inconsistent findings in our study between merging the data manually by sleep segments, and by automated methods for the 24-hour period, highlight a potential area for improved understanding of the validity of fitness trackers in those with more segmented sleep patterns. We found that manually processing the data into sleep segments resulted in the Fitbit underestimating total sleep time compared to actigraphy, which was novel and opposite of previous study focusing on calendar night measures.(12, 23, 24) An alternative and plausible explanation is the additional cleaning and matching of the datasets falsely enhanced the equivalence by disregarding unmatched measurements.

The Fitbit sleep efficiency scores should be interpreted as an overestimation, or different measure, from traditional actigraphy. The operational definition of sleep efficiency continues to be a matter of debate, particularly for those with insomnia or segmented sleep who do not spend the entire night in bed as observed in our study.(37) While sleep efficiency is traditionally calculated as the total sleep time divided by the minutes in bed, this often inaccurately includes middle-night, out of bed activity between sleep segments as time in which sleep was attempted in some algorithms.(37) More novel approaches that affect equivalence, and may be incorporated into the proprietary fitness tracker algorithm, include sleep latency, total sleep time, wake after sleep onset, and time attempting to sleep after final awakening by segment in the efficiency calculations.(37) We observed an additional average 5% overestimation for the Fitbit on a 0-100% scale in addition to previous 8-9% sleep efficiency overestimation of actigraphy compared to PSG. (12, 20) Because the variability in sleep efficiency measured by actigraphy was greater than the fitness tracker (a wider distribution of scores), the devices were not statistically equivalent at the 9% threshold. The devices demonstrated concordance (or lack of a systematic bias) for sleep efficiency that improved when adjusting for a linear trend in the data. These trends should be considered in future computer programming development and informatics solutions for real-time clinical and research data use. The fitness tracker website (<https://community.fitbit.com/>) currently includes specific instructions to register for and obtain an application program interface (API) for ongoing software programming and application development for clinical and research applications. Our finding of concordance, without equivalence, also supports that it is plausible that the Fitbit measures sleep efficiency using a different calculation. Standardized methods of merging or aggregating sleep segments into a weighted daily measure for both devices is needed for ongoing Fitbit use in the clinical or research measurement of sleep efficiency used as an adjunct to self reports, to diagnose or treat patients.

Wake counts were concordant, but poorly correlated and not equivalent between the devices, with actigraphy estimating an average of 12-14 more wakening per night than the Fitbit. Clinically, an average of more than 5 actigraphy-recorded awakenings per night requires further investigation to assess disrupted sleep.(36) We did not use the Fitbit's sensitive setting in this study, based on poor performance in previous study of other populations.(24) Mathematically transforming the wake counts reduced systematic bias, which should also be considered in informatics and algorithm development for future use. Similarly, a previous study on the use of regression modeling meant to enhance the accuracy and further process of fitness tracker data for research and clinical use has been demonstrated for step counts and energy expenditure, and our study adds to the evidence on the need to create these adjustments for wake counts as well.(22) In the asthma population, Fitbit's underestimation of wake counts may still theoretically be of research and therapeutic value. While the actigraphy device that is most sensitive to detecting disturbed sleep is clinically relevant for sleep disorder detection, it is unclear if the decreased sensitivity will enhance asthma-specific wakening measurement. Since we anticipate night-time wakening with asthma symptoms will involve greater levels of movement and action (e.g. taking rescue inhaler, repositioning for maximal airflow), further research is needed in this population to ascertain the level of accuracy in recording general, as opposed to asthma-specific, sleep disruption. In addition, the Fitbit continues to demonstrate potential relevance in self-management due to its accessibility for individuals to track personalized patterns of sleep disturbance.

In addition to several indications for future research and informatics development, our findings have preliminary relevance for clinical practice. Disrupted sleep is a key indicator of poorly controlled asthma and asthma related quality of life. There is currently no gold standard by which to establish an objective adjunct to subjective self-report of asthma specific wakening. This study provides the preliminary evidence on an objective assessment of sleep patterns in women with asthma using two different potentially clinically relevant devices. A prior investigation indicates that patients at highest risk for poor health outcomes are the least likely to use wearable health trackers or communicate the data to their clinicians. (35) In consort with software development collaborators, our results are relevant to designing self-management and clinical dashboard displays. Fitness trackers, with an adequate user clinician interpretation of the overestimation of sleep efficiency and sleep time as presented in this and other studies, may be used as an adjunct assessment and as possibly intervention. We were able to reach a high proportion of participants belonging to high risk for poorly controlled asthma and overall vulnerable populations with a racially diverse sample, those with higher BMI's, and covered by Medicaid insurance. Therefore, studies focusing on asthma as a health outcome have a potential to extend the understanding, relevance, and reach of mobile health (mHealth) applications using fitness trackers to diverse populations with higher health risks for poorly controlled asthma. Future study is warranted to compare fitness trackers and actigraphy in specific subgroups of women with asthma, such as those with co-morbid obesity, sleep apnea, COPD, or by smoking status.

The findings from this study should be interpreted with several limitations. Actigraphy and fitness trackers measure limb movement as a proxy for sleep, and do not directly measure sleep states. Skin and ambient temperature may affect readings.(20) The labor intensive process to manually match the nightly sleep data, with several hours per participant, is not

currently replicable in the clinical setting, but demonstrates the need for improved algorithms that address sleep segments, rather than self-reported sleep measures. The act of wearing the device, with access to the data, may have changed the participants' behaviors from their natural patterns. Participants were instructed not to install updates in this study. However, the proprietary algorithm can be updated and patched without researcher knowledge at any time. Access to the raw data and proprietary algorithms are not only a research limitation, but also a broader clinical limitation of the fitness tracker's applications and use in practice. An automated log of updated software and installation date will enhance the potential use of the device as a research tool. Newer generations of the device and software are currently on the market.

Conclusion

Fitness trackers offer an easily available and inexpensive method to quantify sleep patterns in the home environment as an adjunct to the self-report of asthma symptom sleep disruption. With no current clinical gold standard to objectively measure asthma-specific sleep disruption, this study provides preliminary evidence on longitudinal wrist-worn device assessment of sleep patterns in the home environment. Although fitness trackers have a tendency to overestimate total sleep time and efficiency, they have demonstrated preliminary validity with actigraphy in several populations and the need for further algorithm and informatics development.(20) We demonstrated feasibility and equivalence for total sleep time at the 15-minute threshold for sleep segments, with a lack of systematic bias for wake counts and sleep efficiency. There were important differences in the devices when comparing by specific sleep segment, rather than by 24-hour period.

Fitness trackers continue to hold promise for use in research and clinical care, with ongoing research needed to apply adjustments in data aggregation, mathematical modeling, comparison to PSG and actigraphy, and interpretation relative to the overestimation of total sleep time and efficiency and underestimation of wake counts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

The authors gratefully acknowledge the contributions of Jillian Barry, RN, Camille Applegate, Albert Titus, PhD, and Peter Elkin, MD.

Funding: Research co-sponsored by the Rockefeller University Heilbrunn Family Center for Research Nursing through the generosity of the Heilbrunn Family, University at Buffalo HomeBASE Center, and the Patricia H. Garman Behavioral Health through the generosity of the Garman Family. Drs. Wilding and Sethi were supported by NIH 1UL1TR001412-01 Buffalo Clinical and Translational Research Center, and Dr. Caster was committed to receive salary support by the University at Buffalo School of Nursing and additional support from NIH 1UL1TR001412-01 Buffalo Clinical and Translational Research Center.

References

1. Nurmagambetov T, Kuwahara R, Garbe P. The Economic Burden of Asthma in the United States, 2008-2013. *Annals of the American Thoracic Society*. 2018;15(3):348–56. [PubMed: 29323930]

2. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention 2017 [Available from: Available from: www.ginasthma.org/2017-gina-report-global-strategy-for-asthma-management-and-prevention/].
3. Centers for Disease Control and Prevention. Asthma Call-Back Survey United States Codebook Report, 2012 2012 [Available from: http://www.cdc.gov/brfss/acbs/2012/pdf/ACBS_2012_ADULT_LLCP_CODEBOOK.pdf].
4. Khan WH, Mohsenin V, D'Ambrosio CM. Sleep in Asthma. *Clin Chest Med*. 2014;35(3):483–93. [PubMed: 25156764]
5. Moorman JE. National surveillance for asthma--United States, 1980–2004: Department of Health and Human Services, Centers for Disease Control and Prevention Atlanta, GA; 2007.
6. Szyszkowicz M, Kousha T. Emergency department visits for asthma in relation to the Air Quality Health Index: A case-crossover study in Windsor, Canada. *Canadian journal of public health= Revue canadienne de sante publique*. 2013;105(5):e336–e41.
7. Lin RY, Lee GB. The gender disparity in adult asthma hospitalizations dynamically relates to age. *Journal of Asthma*. 2008;45(10):931–5. [PubMed: 19085585]
8. Lin RY, Lee GB. The gender disparity in adult asthma hospitalizations dynamically relates to age. *J Asthma*. 2008;45(10):931–5. [PubMed: 19085585]
9. Krouse HJ, Yarandi H, McIntosh J, Cowen C, Selim V. Assessing sleep quality and daytime wakefulness in asthma using wrist actigraphy. *J Asthma*. 2008;45(5):389–95. [PubMed: 18569232]
10. Koinis-Mitchell D, Kopel SJ, Seifer R, LeBourgeois M, McQuaid EL, Esteban CA, et al. Asthma-related lung function, sleep quality, and sleep duration in urban children. *Sleep health*. 2017;3(3):148–56. [PubMed: 28526251]
11. Morgenthaler T, Alessi C, Friedman L, Owens J, Kapur V, Boehlecke B, et al. Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: an update for 2007. *Sleep*. 2007;30(4):519–29. [PubMed: 17520797]
12. Rosenberger ME, Buman MP, Haskell WL, McConnell MV, Carstensen LL. Twenty-four Hours of Sleep, Sedentary Behavior, and Physical Activity with Nine Wearable Devices. *Med Sci Sports Exerc*. 2016;48(3):457–65. [PubMed: 26484953]
13. Centers for Disease Control and Prevention. National health and Nutrition Survey (NHANES) physical activity monitor (PAM) procedures manual 2011 [updated July 12, 2017. Available from: https://www.cdc.gov/nchs/data/nhanes/nhanes_11_12/Physical_Activity_Monitor_Manual.pdf].
14. Johnson ST, Thiel D, Al Sayah F, Mundt C, Qiu W, Buman MP, et al. Objectively measured sleep and health-related quality of life in older adults with type 2 diabetes: a cross-sectional study from the Alberta's Caring for Diabetes Study. *Sleep health*. 2017;3(2):102–6. [PubMed: 28346155]
15. Vallance JK, Buman MP, Lynch BM, Boyle T. Reallocating time to sleep, sedentary, and active behaviours in non-Hodgkin lymphoma survivors: associations with patient-reported outcomes. *Ann Hematol*. 2017;96(5):749–55. [PubMed: 28197722]
16. Kim M, Yoshida H, Sasai H, Kojima N, Kim H. Association between objectively measured sleep quality and physical function among community-dwelling oldest old Japanese: A cross-sectional study. *Geriatr Gerontol Int*. 2015;15(8):1040–8. [PubMed: 25312049]
17. Gomes TN, dos Santos FK, Santos D, Pereira S, Chaves R, Katzmarzyk PT, et al. Correlates of sedentary time in children: a multilevel modelling approach. *BMC Public Health*. 2014;14:890. [PubMed: 25174736]
18. Babcock J, Krouse HJ. Evaluating the sleep/wake cycle in persons with asthma: Three case scenarios. *J Am Acad Nurse Pract* 2010;22(5):270–7. [PubMed: 20500742]
19. Ward TM, Lentz M, Kieckhefer GM, Landis CA. Polysomnography and actigraphy concordance in juvenile idiopathic arthritis, asthma and healthy children. *J Sleep Res*. 2012;21(1):113–21. [PubMed: 21592248]
20. 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. [PubMed: 26684758]
21. Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, et al. Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort. *J Pers Med*. 2017;7(2).

22. Andalibi V, Honko H, Christophe F, Viik J. Data correction for seven activity trackers based on regression models. *Conf Proc IEEE Eng Med Biol Soc* 2015;2015:1592–5.
23. Williams S, Han M, Ye X, Zhang H, Meyring-Wosten A, Bonner M, et al. Physical Activity and Sleep Patterns in Hemodialysis Patients in a Suburban Environment. *Blood Purif.* 2017;43(1–3): 235–43. [PubMed: 28114147]
24. Cook JD, Prairie ML, Plante DT. Utility of the Fitbit Flex to evaluate sleep in major depressive disorder: A comparison against polysomnography and wrist-worn actigraphy. *J Affect Disord.* 2017;217:299–305. [PubMed: 28448949]
25. Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, et al. Development of the asthma control test: a survey for assessing asthma control. *J Allergy Clin Immunol.* 2004;113(1): 59–65. [PubMed: 14713908]
26. Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. *Ann Intern Med.* 1999;131(7):485–91. [PubMed: 10507956]
27. Zinkhan M, Berger K, Hense S, Nagel M, Obst A, Koch B, et al. Agreement of different methods for assessing sleep characteristics: a comparison of two actigraphs, wrist and hip placement, and self-report with polysomnography. *Sleep Med.* 2014;15(9):1107–14. [PubMed: 25018025]
28. Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;327(8476):307–10.
29. Castner J, Sullivan SS, Titus AH, Klingman KJ. Strengthening the Role of Nurses in Medical Device Development. *Journal of professional nursing : official journal of the American Association of Colleges of Nursing.* 2016;32(4):300–5. [PubMed: 27424930]
30. Castner J, Suffoletto H. Emergency department crowding and time at the bedside: A wearable technology feasibility study. *J Emerg Nurs.* Forthcoming, Accepted March 6, 2018.
31. Castner J, Gehrke GE, Shapiro N, Dannemiller KC. Community interest and feasibility of using a novel smartphone-based formaldehyde exposure detection technology. *Public health nursing* (Boston, Mass). 2018.
32. Samson DR, Yetish GM, Crittenden AN, Mabulla IA, Mabulla AZP, Nunn CL. What is segmented sleep? Actigraphy field validation for daytime sleep and nighttime wake. *Sleep health.* 2016;2(4): 341–7. [PubMed: 29073393]
33. Kim J Analysis of health consumers' behavior using self-tracker for activity, sleep, and diet. *Telemed J E Health.* 2014;20(6):552–8. [PubMed: 24745608]
34. Rosenberg D, Kadokura EA, Bouldin ED, Miyawaki CE, Higano CS, Hartzler AL. Acceptability of Fitbit for physical activity tracking within clinical care among men with prostate cancer. *AMIA Annu Symp Proc.* 2016;2016:1050–9. [PubMed: 28269902]
35. Kaewkannate K, Kim S. A comparison of wearable fitness devices. *BMC Public Health.* 2016;16:433. [PubMed: 27220855]
36. Martin JL, Hakim AD. Wrist Actigraphy. *Chest.* 2011;139(6):1514–27. [PubMed: 21652563]
37. Reed DL, Sacco WP. Measuring Sleep Efficiency: What Should the Denominator Be? *J Clin Sleep Med.* 2016;12(2):263–6. [PubMed: 26194727]
38. Pevnick JM, Fuller G, Duncan R, Spiegel BM. A Large-Scale Initiative Inviting Patients to Share Personal Fitness Tracker Data with Their Providers: Initial Results. *PLoS One.* 2016;11(11):e0165908. [PubMed: 27846287]

*Reasons for non-recruitment on phone screen (n=60): No device/computer (n=6), ACT score (n=20), pregnant or trying to become (n=3), night shift (n=2), dependents/pets routinely disrupt middle of the night sleep (n=7), neuromuscular disease (n=1), Berlin Score for Sleep Apnea (n=18), not interested after phone eligibility (n=3).

Fig. 1.
Study Flow Diagram

Fig. 2.

Box Plots for Total Sleep Time, Sleep Efficiency, Wake Counts by Device Type

Note: Panel A-C Automatically Merged Device Data by 24-hour period; Panel D-F Manually Linked Device Data by Sleep Segment

Fig. 3.

Bland Altman Plot with Concordance for Total Sleep Time

Note: Panel A-B Automatically Merged Device Data by 24-hour period; Panel C-D Manually Linked Device Data by Sleep Segment

Fig. 4.

Bland Altman Plot with Concordance for Sleep Efficiency

Note: Panel A-B Automatically Merged Device Data by 24-hour period; Panel C-D Manually Linked Device Data by Sleep Segment

Fig. 5.
Bland Altman Plot with Concordance for Wake Counts
Note: Panel A-B Automatically Merged Device Data by 24-hour period; Panel C-D
Manually Linked Device Data by Sleep Segment

Table 1.

Baseline Characteristics of Participants (N=47)

Variable	M (SD) or %	Range
Age, yr	40.6 (13.0)	19-64
Race, %		
Non-Hispanic Black	44.7	
Non-Hispanic White	36.2	
Hispanic/Latina White	4.3	
More than one race (Non-Hispanic)	10.6	
Hispanic/Latina More than one race	2.1	
Native Hawaiian/Pacific Islander	2.1	
Highest Education Completed, %		
Grades 9-11	8.5	
Grade 12 or GED	27.7	
1-3 years of college	36.2	
4+ years of college/college graduate	25.5	
Income, %		
<\$15,000	21.3	
\$15,000-25,000	40.4	
\$26,000-50,000	17.0	
\$51,000+	12.8	
Asthma Control Test	14.0 (3.38)	7-21
Well Controlled, %	4.3	
Not Well Controlled, %	25.5	
Very Poorly Controlled, %	70.2	
Quality of Life (asthma-specific)	3.8 (1.1)	1.8-6
Epworth Sleepiness Scale	9.6 (5.9)	1-24
Emotionality – General Fatigue	3.2 (.96)	1.5-5.0
History of GERD, %	31.9	
History of Obstructive Sleep Apnea, %	12.8	
Ever cigarette smoker, %	51.0	
Current cigarette smoker, %	34.0	
Smoking exposure inside home in last week, %	42.6	
Age of Asthma Onset, yrs	20.0 (15.3)	0-55
Body mass index	33.6 (9.8)	19-60
Normal, %	17.0	
Overweight, %	27.7	
Obese Class I, %	17.0	
Obese Class II, %	14.9	
Obese Class III, %	23.4	

Variable	M (SD) or %	Range
Polypharmacy Score	10.8 (6.5)	2-29
Medical Diagnoses (last 3 years)	17.4 (9.6)	2-43

Note: BMI=Body Mass Index; GERD=Gastro-Esophageal Reflux Disease

Table 2

Average Nightly Sleep Measures by Device, by Night/Sleep Segment and Participant as Unit of Analysis

	Manually Merged Data by Sleep Segment			Automated Merged Data by Day		
	Total Sleep Time, min	Sleep Efficiency	Wake Counts	Total Sleep Time, min	Sleep Efficiency	Wake Counts
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
By Night/Segment as Unit of Analysis						
	(N=833)			(N=861)		
Fitness Tracker	410.6 (134.3)	92.9 (3.4)	2.9 (3.3)	476.4 (195.0)	92.7 (3.3)	4.0 (4.4)
Actigraphy	419.8 (144.9)	87.3 (6.3)	14.9 (7.1)	469.7 (189.9)	87.7 (6.0)	15.8 (7.4)
By Participant as Unit of Analysis						
	(N=47)			(N=47)		
Fitness Tracker	412.4 (66.5)	92.9 (1.9)	2.9 (1.5)	486.1 (93.3)	92.7 (1.8)	3.9 (2.1)
Actigraphy	420.5 (74.0)	87.4 (3.1)	14.8 (4.2)	471.9 (106.2)	87.7 (3.1)	15.8 (4.3)

Table 3.

Concordance Testing: Percent of Observations Outside 95% Limits of Agreement

	Correlation	Unadjusted	Adjusted for Trend
Device Data Manually Merged			
Total Sleep Time	.52 ***	5.72%, N=47/822	9.61%, N=79/822
Sleep Efficiency	.36 ***	6.21%, N=51/821	5.12%, N=42/821
Wake Counts	.09 *	4.63%, N=38/821	2.68%, N=22/821
Device Data Automated Merge			
Total Sleep Time	.28 ***	7.73%, N=57/737	10.85%, N=80/737
Sleep Efficiency	.35 ***	5.54%, N=41/740	4.46%, N=33/740
Wake Counts	.11 **	4.85%, N=36/742	3.37%, N=25/742

Note:

^a Pearson Product Moment Correlation of Raw Scores*
p<.05**
p<.01***
p<.001