

Original Paper

# Determining a Likely Mechanism of Missingness in Repeated Measures Sleep Data From Wearable Fitness Trackers: Longitudinal Analysis

Kate Mobley<sup>1</sup>, BS, MS, PhD; Kevin B Gittner<sup>1,2</sup>, PhD; Karen Nielsen<sup>3</sup>, PhD; Lauren M Matheny<sup>1</sup>, MPH, PhD; Gita Taasoobshirazi<sup>1</sup>, PhD; Charles Natuhamya<sup>4</sup>, MS; Monica H Swahn<sup>5</sup>, PhD

<sup>1</sup>School of Data Science and Analytics, College of Computing and Software Engineering, Kennesaw State University, Marietta, GA, United States

<sup>2</sup>Department of Health Promotion and Physical Education, Wellstar College of Health and Human Services, Kennesaw State University, Kennesaw, GA, United States

<sup>3</sup>Department of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, GA, United States

<sup>4</sup>Uganda Youth Development Link, Kampala, Uganda

<sup>5</sup>School of Public Health, Virginia Commonwealth University, Richmond, VA, United States

## Corresponding Author:

Kate Mobley, BS, MS, PhD  
School of Data Science and Analytics  
College of Computing and Software Engineering, Kennesaw State University  
680 Arntson Drive Atrium J333, MD 9104  
Marietta, GA 30060  
United States  
Phone: 1-470-578-2865  
Email: [mobley.kate@gmail.com](mailto:mobley.kate@gmail.com)

## Abstract

**Background:** Wearable fitness trackers have become a valuable tool in public health research due to their ability to collect large-scale, individual-level data at a low cost. Because of this, wearables have the potential to mitigate the urgent need for health research in low-resource settings. However, their use has been largely limited to high-income settings, and a major challenge remains: high rates of missing data. This problem may be exacerbated in low-resource environments where logistical and operational barriers further complicate data collection. Wearable sleep data collected in an urban informal settlement in Uganda during The Onward Project on Well-being and Adversity project revealed substantial challenges related to missingness in low-resource research settings.

**Objective:** The study aimed to characterize the patterns and frequency of missing data and to demonstrate the use of analytical methods to determine the underlying missingness mechanism.

**Methods:** For this study, 300 women in slum communities in Kampala were equipped with Garmin smartwatches that collected data over 5 days. The nature of missingness was assessed through four methods: pattern analysis, the Little test, a random forest classification model, and a logistic regression classification model. Pattern analysis is an established method of identifying missing data patterns, while the Little test is used specifically to identify missing completely at random data. Random forest and logistic regression models are more recent methods proposed to identify the mechanism of missingness; both were used to examine agreement between the 2 methods and to ensure a thorough examination of missingness patterns.

**Results:** Approximately 30% of nighttime data were missing. Three patterns were identified that occurred in over 10% of participants' data: no data missing, fifth night of data missing, and all nights of data missing. Pattern analysis and the Little test ( $P < .001$ ) indicated that the data were not missing completely at random. Both the random forest (area under the curve=0.7) and logistic regression models suggested that the data were missing at random (MAR).

**Conclusions:** Evidence of missingness in this dataset was consistent with MAR. Potential causes of missingness include device removal, battery failure, and technical malfunctions. These findings have important implications for both wearable device users and future research. Understanding the mechanisms behind missingness can inform strategies to improve data quality, particularly in low-resource settings, and allow researchers to manage missingness using appropriate strategies designed to handle MAR data without introducing bias into the results.

**Keywords:** global health; wearable electronic devices; sleep; missing data; informal settlements

## Introduction

### *Wearable Fitness Trackers in Health Research*

Originally intended for personal health and fitness tracking, wearable fitness trackers have gained traction in research due to their ability to collect large-scale, real-world data from participants in natural settings [1]. According to Fitabase, a platform monitoring research applications of Fitbit devices, over 1300 studies incorporated Fitbit devices between 2012 and 2025 [2] and that number increases when factoring in the studies that have used other fitness trackers. Wearable fitness trackers provide real-time data collection, including physical activity, heart rate metrics, biometric and health indicators, location, behavior, and sleep. These data are critical for identifying risk factors for various health conditions, tracking health outcomes over time, evaluating interventions, and advancing personalized medicine. Wearable device data, when limitations are managed, have the potential to advance understanding of how lifestyle behaviors interact with physical and mental health across diverse populations, supporting more effective, data-driven public health strategies and clinical practices. Despite their growing adoption in research, these technologies remain disproportionately used in high-income countries and among wealthier populations [1, 3,4].

### *Research in Urban Slum Communities*

Wearable devices hold significant promise for research, particularly in low-resource settings such as urban slums [4,5]. These settings, defined by constraints across finances, health care, infrastructure, knowledge, research capacity, social resources, environmental factors, human resources, and cultural practices [6], pose unique challenges for data collection. Furthermore, over 1 billion people, or 24% of the world's urban population, currently live in slums, with this figure rising to 56% in sub-Saharan Africa [7]. Wearables offer a practical solution to the need for health research in these settings with large populations and disproportionately high risks of physical and mental health conditions [8-10]. They are cost-effective, scalable, minimally intrusive, and capable of gathering individual-level data for multiple days without reliance on external power or additional technology. These advantages make them especially valuable for research in resource-limited environments [4,5]. Understanding missingness in sleep data from wearable fitness trackers can provide important foundational information for this research.

Understanding the physical and mental health challenges in urban slums and other low-resource settings is a critical public health priority, and sleep serves as a key indicator of overall well-being. Sleep problems, including poor sleep quality and short sleep duration, are common among residents of slum communities and have been linked to mental illness,

but this topic is still understudied, especially so in low- and middle-income countries [4]. Wearable technology presents a transformative opportunity for such studies, offering cost-effective, minimally intrusive methods to collect longitudinal health data in challenging urban environments. For instance, while poor sleep is a documented issue in slums, existing research has relied primarily on subjective and unreliable self-reports rather than objective sensor-based measurements [4].

However, while missing data is an issue present in all research involving wearable fitness trackers, the difficulties associated with doing research in low-resource settings like urban slums can contribute to higher rates of missing data [11]. Residents of urban slums often have unreliable electricity access, are unfamiliar with wearable fitness tracker technologies, and face logistical barriers to implementing these technologies [12,13]. Additionally, many wearable fitness trackers are reliant on connection to a smartphone or the internet, and rates of internet access and smartphone usage are low in urban slums [4,14]. Because of these challenges associated with low-resource settings like urban slums, rates of missing data can be higher when wearable fitness trackers are used in these settings. Though wearable studies in slums remain limited and are often clinic-based, there is potential for scalable, real-world data collection using wearable devices [15-17]. To realize this potential, understanding critical challenges like missing data must become a priority.

### *Missing Data*

Missing data remains a pervasive challenge in human research, arising from diverse sources. Participant-related factors, such as missed study visits or attrition, and researcher limitations in follow-up capacity are well-documented contributors [18,19]. The growing integration of digital devices and sensors in research has introduced additional technical sources of missingness, including equipment malfunctions, battery failures, and connectivity issues [20]. Longitudinal studies, which require sustained data collection across multiple timepoints, face compounded risks of missing data, making this issue particularly acute in such designs [21].

Missing data in wearable device studies typically stems from two primary sources: device-related factors and wearer-related behaviors [22,23]. Technical issues such as battery failure, sensor displacement, software malfunctions, connectivity problems, or synchronization errors frequently contribute to device-related missingness [22-24]. Wearer-related missingness most commonly results from nonadherence, including device removal or improper wear [23]. There has been some study of missing data in wearable data, but very little in low-resource settings where wearables have only been used in a small number of studies outside of clinical settings, where conditions are much more controlled. Researchers in low-resource settings have theorized that a higher prevalence

of manual labor in these settings results in exposure to extreme conditions, contributing to both device- and wearer-related challenges, which may contribute to the deterioration of wearable devices and increased missingness [5]. Wearables also differ widely in the types of data they collect, and only a few studies have considered missingness in wearable sleep data specifically [23,25]. Despite the growing use of wearable devices to collect sleep data, limited research has examined patterns and mechanisms of missingness, particularly in low-resource settings, where contextual challenges may further exacerbate missing data [11].

The prevalence of missing data significantly impacts data quality and poses substantial analytical challenges [26]. Effective handling of missing data through methods like imputation becomes crucial for ensuring accurate analyses [24]. A comprehensive review of wearable studies in biomedical research confirms that data missingness remains among the most significant obstacles researchers face [26, 27]. Few researchers have used wearables in low-resource settings, but among those who have, numerous studies have found data missingness to be a substantial issue, underscoring the need to understand the mechanisms of missingness and to find robust solutions to this persistent issue [5,28].

Missing data can be categorized into three primary mechanisms: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). When data are MCAR, the probability of missingness is entirely independent of both observed and unobserved data [29,30]. In contrast, MAR occurs when missingness depends only on observed variables, not the missing values themselves [30]. The most challenging scenario, MNAR, arises when missingness is related to the unobserved data [18, 31]. Understanding these common classifications of missing data is necessary, as these distinctions depend on whether missingness is related to observed or unobserved information and has important implications for analysis and interpretation [29-31].

### **Analytical Methods for Determining the Mechanism of Missingness**

Determining the exact mechanism of missingness is challenging, but several methods can provide evidence for specific patterns. Visualization techniques offer a preliminary approach to identifying missingness patterns through graphical exploration [32]. For formal testing, the Little test, a well-established method, assesses whether data are MCAR [33]. To explore whether data are MAR, two analytical approaches are commonly used: random forest and logistic regression. Both methods model relationships between observed variables and a binary indicator of missingness. If significant associations exist between observed variables and missingness, this supports the MAR mechanism [34,35]. However, in longitudinal studies, where observations within participants are correlated, these methods require modifications to account for dependencies. In addition to contributing to the body of knowledge on missingness in wearable sleep data, this study demonstrates the use of these analytical methods on wearable sleep data, and on repeated

measures data, where adjustments are necessary to preserve the dependencies between observations from the same person.

### **Wearable Fitness Trackers in Sleep Research**

Since around 2017, consumer wearable fitness trackers have measured time spent in each sleep stage using a combination of motion detection, heart rate variability, and respiratory rate monitoring [36]. While questions remain about the absolute accuracy of sleep-stage classification, the devices' ability to detect sleep-wake patterns has significantly improved [37, 38]. Importantly, wearable fitness trackers provide more objective sleep data than self-reports, as individuals often cannot accurately recall specific sleep metrics like sleep onset time or nighttime awakenings [36]. Although wearable sleep tracking technology still has limitations, it represents a substantial improvement over subjective reporting for research purposes, with accuracy continuing to advance through technological innovations [37-39].

Garmin devices like the ones used in this study have been assessed for the validity of their measurements. Existing studies indicate that they provide highly valid step count measurements, though their accuracy for distance, energy expenditure, and heart rate data remains more variable [40]. Garmin devices demonstrate reasonable agreement with polysomnography for measuring sleep duration and staging [41], and they more accurately identify deep sleep stages than some other wearable fitness trackers [42]. Manufacturers release new models frequently, and different models from the same brand can vary. Therefore, the comprehensive evaluation of wearable devices across all models and brands, including validity, reliability, proprietary algorithms, cost differences, and feature variability, remains challenging. Overall, however, modern wearable technologies that combine heart rate and motion tracking demonstrate comparable effectiveness for sleep monitoring, despite some variation in specific metrics across models and brands.

### **Wearable Sleep Data Missingness**

While the problem of missingness in wearables data, including sleep data, has been well documented, only 2 studies were identified that considered how to address this problem in sleep data [23,25]. However, these studies differed substantially from this work. First, both studies used sleep and wakefulness data collected every few seconds for up to 50 days, so that missingness was determined by measuring the length of time during the data collection period that no data was collected (a method typically referred to as measuring data completeness). Additionally, one of these studies used data from individuals with very irregular sleep patterns so that the results may be used by those studying sleep disorders, while the other removed participants who had removed the device at any point, a common cause of missing data [23,25]. Neither study used data from low-resource settings where missingness is likely to be even higher than in data from high-income populations.

## Study Purpose

The limited research on missing data in wearable sleep data that has been conducted so far uses sleep data that is in a significantly different format from that used in this study. To the authors' knowledge, this is the first study to examine missingness in wearable sleep data aggregated by night within a low-resource setting. This innovative approach to analyzing missingness in wearable sleep data is closely aligned with the typical use and analysis of this type of data.

Wearable fitness trackers offer a promising tool for collecting sleep data in diverse environments, which can allow for further study of sleep and the impact of poor sleep on health. However, missing data remain a significant limitation, and the unique logistical challenges of conducting research in low-resource settings, where sleep and health research is urgently needed, may exacerbate this issue. Addressing the patterns of missingness is essential not only for improving data quality but also for advancing research on sleep's role in health in all settings. This study has two primary objectives: (1) to characterize the patterns and frequency of missing data and (2) to determine the underlying missingness mechanism using analytical techniques appropriate for longitudinal repeated measures data. Enders [43] comprehensive work on missing data argues that MAR is the most reasonable assumption for the mechanism of missing data in most cases of research, and this assumption should be adopted unless there is evidence to the contrary. Therefore, this study will focus on evidence that the missingness is MCAR or MAR.

## Methods

### The TOPOWA Project

The data for this study were collected through The Onward Project on Wellbeing and Adversity (TOPOWA), a mixed methods investigation examining the social determinants of mental health among women in Kampala, Uganda's urban slums. Details of the study have been reported previously [4,44]. Participant recruitment was conducted across 3 slum communities (Banda, Bwaise, and Makindye) in collaboration with the local community leaders and Uganda Youth Development Link, a nongovernmental organization serving youth in Kampala's informal settlements.

TOPOWA project used multiple research components, including focus group discussions, Photovoice methodology, and a longitudinal cohort study which includes the collection of sleep data via wearable fitness trackers to study the prevalence of sleep disturbance and its association with mental health in this population.

### Ethical Considerations

All study procedures were designed in compliance with the Declaration of Helsinki and received ethical approval from both the Makerere University School of Social Sciences Research Ethics Committee (MAKSSREC MUSSS-2022-172) and the Uganda National Council of Science and Technology (SS1560ES).

Uganda Youth Development Link staff at each study site described the TOPOWA study including risks, benefits, and expectations regarding follow-up assessments to interested participants. Informed consent in writing was obtained from those who agreed to participate in the study. Participants were compensated in the local currency for each component of the study that they participated in, including wearing and returning the wearable device and completing the survey, as well as being provided with food and transportation on the days that they traveled to the research hub to complete assessments. All research assistants involved in participant enrollment and data collection were trained in participant privacy, confidentiality, and safety measures. Data were deidentified before being uploaded to REDCap (Research Electronic Data Capture) for storage and shared with members of the research team for analysis.

### Data Collection and Processing

The TOPOWA project enrolled 300 women from Kampala's urban slums, equipping each participant with a Garmin vivoactive 3 device to wear continuously for 5 days and nights. These devices were specifically selected for their field suitability: they operate independently without smartphone connectivity, feature a 5-day battery life, and automatically collect sleep data each night. During data collection, 299 devices were successfully returned, with only 1 unit lost. The vivoactive 3 devices record heart rate, heart rate variability, and movement data, which Garmin's proprietary algorithms analyze to classify sleep into 4 stages (rapid eye movement [REM], deep, light, and awake) plus an "unmeasurable" category for periods that cannot be classified [45]. For each sleep session, the devices generate detailed outputs including the duration (in seconds) spent in each sleep stage and precise timestamps marking the beginning and end of the primary sleep period. Notably, when multiple sleep periods occur within a day, the system prioritizes and reports only the longest continuous sleep bout, overriding shorter naps. This design ensures researchers obtain standardized, comparable metrics for a person's main sleep period, which will be discussed as nightly sleep for simplicity, while maintaining the practicality required for field studies in low-resource settings.

During the compilation of baseline sleep data, specific criteria were established for identifying missing data. These criteria focused on total sleep duration, treating nights with no sleep information as missing while retaining nights with vague or low-quality data. Notably, the absence of data for a particular sleep stage on nights when other stages were recorded was not classified as missing data. This pattern could indicate either that the participant genuinely did not enter that sleep stage or that the device failed to detect it—a distinction that cannot be determined from the available data. Consequently, 0 values for specific stages on otherwise recorded nights were treated as valid observations rather than missing data. No nights with data had a 0 recorded for light sleep, but there were nights that contained 0 for the deep sleep, REM sleep, and awake categories. Less than 2.5% of nights with data had a 0 recorded for multiple

sleep stages. Similarly, “unmeasurable” sleep seconds were considered (periods the device could not classify into any stage) as legitimate data points and included in total sleep duration calculations, rather than being treated as missing values. Zeros were common in the “unmeasurable” category, indicating that there were no seconds of sleep that the device was unable to classify into one of the other sleep stage categories during those nights. This approach ensured that the analysis reflected the complete sleep architecture as captured by the device while maintaining methodological consistency.

The presence or absence of heart rate and activity data was also recorded. If heart rate data were present for the prior day, then a 1 was recorded, and if heart rate data were not present for the day before the night in question, then a 0 was recorded. If at least 1 activity was recorded during the 5 days that the participant wore the device, then a 1 was recorded for activity data, and if no activity was recorded over the 5 days, then a 0 was recorded. The presence or absence of heart rate data was used in the statistical analysis to determine the likely mechanism of missingness; however, since activity data was only collected during prolonged periods of physical activity and was therefore much less frequent, they were not included in the statistical models. While the Garmin sleep stage classification algorithm uses concurrently collected heart rate data, the missing data definition used in this study allowed for inclusion of nights with unclassified stages—therefore, the presence of heart rate data was not perfectly correlated with missingness.

## Statistical Analysis

The analysis of missing data followed a systematic approach. First, a pattern analysis was conducted using the *nanjar* (version 1.1.0; Nicholas Tierney and Dianne Cook) package in R to visually examine the structure of missingness in the dataset [46]. Next, the Little test, a likelihood ratio-based test that compares the expected means for each variable across different missingness patterns [33], was applied to evaluate whether the data were MCAR. For the Little test, the data were converted to wide format, and the 5 sleep category variables (deep, light, REM, awake, and unmeasurable) were included as well as the variables for marital status, household size, education level, and employment status. This test was conducted first because the MCAR criteria are more restrictive than the MAR criteria.

Following this initial assessment, 2 distinct classification methods, a random forest model and a generalized estimating equations (GEE) model with a logit link function, were used to explore whether the missingness pattern was consistent with MAR. For both models, a binary dependent variable indicating whether data were present or missing for each night was created. Both models incorporated six predictor variables: (1) night sequence number (1-5), (2) household size, (3) education level, (4) employment status (binary), (5) marital status (binary), and (6) presence of heart rate data for the day before the night of sleep (binary). These covariates were selected based on established relationships between socioeconomic factors and sleep behaviors in prior research [47,48]. The results from both models were then used to

strengthen the conclusions about the nature of the missing data mechanism.

The dataset included participants with complete missingness across all 5 nights of sleep monitoring, represented by NA (not available) values for all sleep stages. While standard practice might suggest listwise deletion for such cases, these participants were retained in this analysis for three key reasons: (1) they provided complete data for all independent variables (household size, marital status, education level, employment status, and heart rate data availability), (2) this study specifically aimed to develop alternatives to conventional listwise deletion approaches, and (3) missingness, even complete missingness, can be informative when evaluating the mechanism and patterns behind the missingness.

While the random forest classification and GEE modeling with logit link functions serve similar analytical purposes for MAR assessment, their concurrent application provides methodological robustness, given that (1) each approach has distinct strengths in handling different data characteristics and (2) their use for MAR determination remains relatively novel in the literature, with foundational publications emerging only in the early 2000s [34,35,49]. This dual-method approach strengthened the conclusions by assessing consistency across complementary analytical frameworks.

## Random Forest Classification Model

A random forest classification model was used using the *randomForest* R package (version 4.7.1.1; Andy Liaw and Matthew Wiener) to assess whether the data were consistent with MAR, as this approach can effectively capture complex, nonlinear relationships among observed variables [35]. Model performance was evaluated using the area under the receiver operating characteristic curve. An area under the curve (AUC) value significantly greater than 0.5 provided evidence for MAR, as this indicates that the missingness pattern was systematically related to the observed variables in the dataset [35]. Low AUC values (less than 0.5) indicate that the model performs no better than random guessing and would be expected if the data were MCAR. AUC values above 0.5 can occur if the data are MNAR but are more likely to occur if the data are MAR [35].

A critical consideration in the analysis was that standard random forest classification assumes independence between observations, an assumption violated in longitudinal data where multiple measurements come from the same individual. To address this dependency, a cluster-adjusted bootstrap sampling approach [50] was implemented, which resampled participants (rather than individual nights) to preserve within-subject correlations. While the use of a random forest model to assess evidence of MAR is not novel, modifying that model to account for dependencies between observations in repeated measures data is not a standard use of the method. For this reason, a second method of modeling patterns between observed variables was used.

## ***GEE Model With Logit Link Function***

The second approach implemented for exploring a MAR mechanism was a logistic regression model. This method tests whether observed variables can systematically predict missingness patterns—a key criterion for MAR. While several logistic regression-based approaches exist for MAR assessment (including Ridout's method for dropout analysis in longitudinal studies [34] and Fairclough's method for missingness prediction), Fairclough's framework [34] was adapted for this analysis since its purpose was closely aligned with the goal of this study. Independent variables were entered sequentially into the model to assess their association with missingness. A variable's contribution was evaluated based on the regression coefficient estimates, with statistical significance determined using robust standard errors derived from the GEE, to characterize predictors of missingness. Five total models were created with an additional variable added in each model. This iterative approach using forward selection allowed for assessing how different participant characteristics and study factors influenced the probability of data being missing for a given night.

While traditional logistic regression assumes independence between observations, this assumption is violated in repeated measures designs where observations from the same participant are inherently correlated [51]. To address this dependency, a GEE model with a logit link function was used using the *geepack* package in R (version 1.3.12; Søren Højsgaard, Ulrich Halekoh, Jun Yan, and Claus Thorn Ekstrøm), which extends generalized linear models by incorporating specified covariance structures [52]. Working correlation structures were assessed for fit, including the unstructured, exchangeable, and autoregressive (1). The quasi-likelihood under the independence model information criterion was used to assess which working correlation structure was most suitable [52]. For this analysis, an exchangeable working correlation structure was selected, which assumes that the correlations between observations from the same participant are equally correlated [52]. This selection is appropriate because the data were collected over just 5 consecutive days, so significant correlation decay is not expected in such a short time period. The GEE framework allowed for properly accounting for within-participant correlations while examining how various factors predicted missingness patterns.

While the GEE model with a logit link function and the random forest model perform similarly in that they both model relationships between observed variables, neither is a standard method of assessing the mechanism of missingness

that has been well documented in previous publications, as is the case for the Little test. Literature was identified that supported the validity of these methods, but no prior studies using these methods were identified. For these reasons, both modeling methods were performed so that their results could be compared.

## **Results**

### ***Participant Characteristics***

All women who participated in the TOPOWA project were aged between 18 and 24 years old. All participants had completed some primary school and could read and write, but only 8% (24/300) had completed secondary school. Over half, 62% (186/300), had given birth and 82% (152/186) of those who had given birth lived with her child(ren). Only 15% (45/300) of the women lived in a single generation household, while the remaining 85% (255/300) lived in households with 2 or more generations. Women who had a job represented 32% (97/300) while the remaining 68% (203/300) of the women were unemployed.

### ***Missing Data in the TOPOWA Baseline Sleep Data***

Participants wore Garmin devices continuously for 5 days and nights, yielding sleep data from 299 individuals. However, 37 (12%) devices recorded no sleep data across all 5 nights (complete missingness). The dataset exhibited a notable missing data pattern: nights were either fully observed (all sleep stages recorded) or fully missing (no data for any sleep stage). Consequently, missingness proportions were identical across all sleep metrics (deep, light, REM, awake, and unmeasurable seconds), with an overall missing data rate of 30% (455 nights missing out of 1495 total nights). The rate of missingness was similar across the 3 study sites (Banda, Bwaise, and Makindye) with Banda having a slightly higher rate of missingness than Bwaise and Makindye (Table 1). The rate of missingness in the sleep data remained consistent as household size increased but increased slightly with each additional child a woman had, so that participants with no children had the lowest rate of missingness (155/570 nights of data, 27%) and participants with 3 children had the highest rate of missingness (18/25 nights of data, 72%) (Table 1). The rate of missingness in the sleep data was approximately equal across changes in income, employment status, education level, and marital status. Notably, 58% (173/299) of participants experienced at least 1 night of missing data, demonstrating the pervasiveness of this issue.

**Table 1.** Proportion of missing data by study site, night of data collection, and number of children.

Rate and number of nights of data missing	Nights of data (n=495) missing from Banda participants, n/N (%)	Nights of data (n=500) missing from Bwaise participants, n/N (%)	Nights of data (n=500) missing from Makindye participants, n/N (%)	Total nights of data (n=1495) missing from all participants, n/N (%)
Night of data collection				
1	22/99 (22)	25/100 (25)	20/100 (20)	67/299 (22)
2	32/99 (32)	26/100 (26)	21/100 (21)	79/299 (26)
3	32/99 (32)	25/100 (25)	24/100 (24)	81/299 (27)
4	35/99 (35)	21/100 (21)	32/100 (32)	88/299 (29)
5	49/99 (49)	44/100 (44)	47/100 (47)	140/299 (47)
Number of children				
0	34/155 (22)	65/155 (42)	56/155 (36)	155/570 (27)
1	89/188 (47)	45/188 (24)	54/188 (29)	188/625 (30)
2	35/94 (37)	26/94 (28)	33/94 (35)	94/275 (34)
3	11/18 (61)	7/18 (39)	N/A <sup>a</sup>	18/25 (72)

<sup>a</sup>N/A: not available.

The 37 cases of complete missingness for all 5 nights of sleep data can be categorized into 3 distinct scenarios with varying probable causes, detailed in [Table 2](#).

**Table 2.** Categorization of the 37 cases of complete sleep data missingness.

Status of sleep and heart rate data	Number of cases
No sleep or heart rate data collected	1 device
Partial data collection—either heart rate or activity (but not both) data collected for at least 1 day	8 devices
Selective sleep data missingness—devices collected both heart rate and activity data for at least 1 day, but no sleep data	28 devices

## Data Exploration

The distributions of time spent in each sleep stage and total time asleep were examined to determine if the distributions differed based on the day of the data collection (1-5). For each sleep stage, as well as total time asleep, the distribution was the same both when the 5 nights of data were considered together and when the data were grouped by night. There were no changes in the distribution of sleep times over the course of the 5 days of data collection.

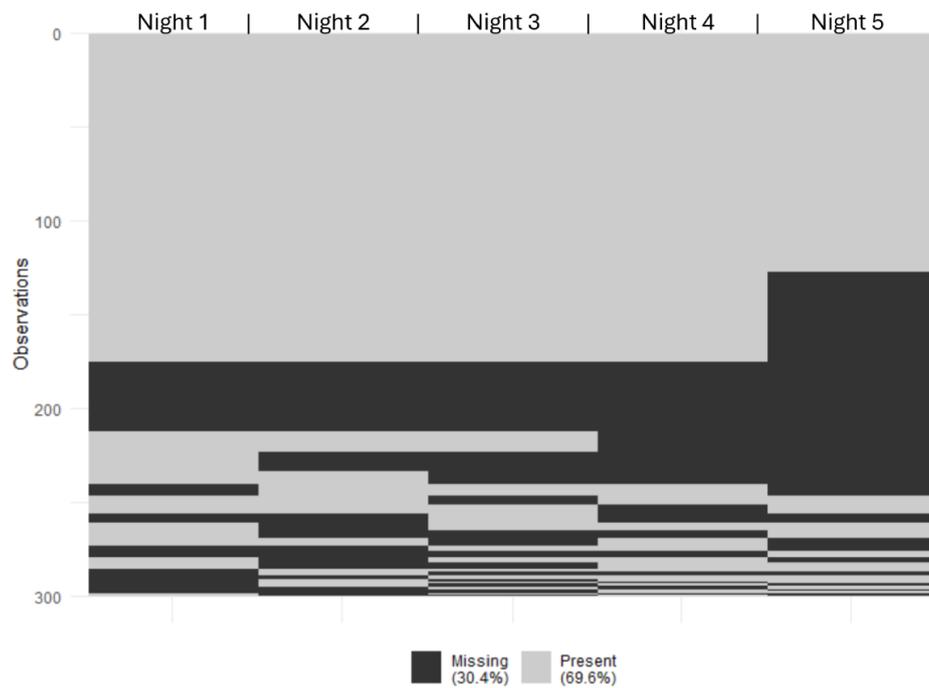
There were some outliers in the total sleep durations. For 1 participant, less than 1 hour of sleep was recorded in 1 night, and for another participant, just over 12.5 hours of sleep was recorded on 1 night. Because the outlier values

present in the data are not outside the realm of possibility, these observations were retained when assessing the missing data mechanism.

## Missingness Patterns

The first night of data collection has the least amount of missing data, with 22% (66/300) of nights missing data. The rate of missing data increases for each subsequent night, and the fifth and final night of data collection has 47% (141/300) of data missing ([Table 1](#)). There were 29 missingness patterns identified in the data ([Figure 1](#)), with no nights missing being the most common (126 cases), followed by missing only the last night of data (48 cases) and then missing all nights of data (37 cases).

**Figure 1.** Patterns of missingness in the sleep data. Missingness patterns are ordered from most common at the top (no missing nights of data) to the least common (bottom).



### The Little Test

When running the Little test, the null hypothesis is that the data are MCAR while the alternative hypothesis is that the data are not MCAR and are therefore either MAR or MNAR. The  $\chi^2$  statistic was 969 and the degrees of freedom was 587, which resulted in a *P* value for the Little test of *P*<.001. Therefore, the null hypothesis is rejected. This suggests that the data were not MCAR.

distinguish missingness, which is consistent with data MAR. The sensitivity, a measure of how well true positives (ie, missing data) are predicted, was 0.30 and the specificity, a measure of how well true negatives (ie, data present) are identified, was 0.94. The overall accuracy for the model was 0.75. Table 3 shows the confusion matrix for the random forest classification model.

### Random Forest Classification Model

The AUC value of 0.70 for the random forest classification model indicated that the model was generally able to

**Table 3.** Random forest classification model confusion matrix.

Predicted values	True values	
	0 (data present)	1 (data missing)
0 (data present)	978	319
1 (data missing)	62	136

### GEE Model With Logit Link Function

The results of the GEE model with a logit link function also yielded results that were consistent with a MAR assumption. The independent variables were added to the model in an iterative process to determine the extent to which each of the variables was able to predict missingness. The final

model (Table 4) contained only 2 predictor variables that were statistically significant, number of night in the 5-night sequence (1-5) and whether the participant had a job. The model had an accuracy of 0.70. The sensitivity was 0.12, while the specificity was 0.96. The confusion matrix for the model is displayed in Table 5.

**Table 4.** Results of generalized estimating equations model with logit link function.

	Parameter estimate	SE	Odds ratio (95% CI)	<i>P</i> value
Intercept	-1.8546	0.1850	0.156 (0.109-0.225)	<.001
Number of night in the 5-night sequence	0.2554	0.0357	1.291 (1.204-1.384)	<.001
Job	0.6671	0.2016	1.948 (1.312-2.893)	.001

**Table 5.** Confusion matrix for the generalized estimating equations model with logit link function.

Predicted values	True values	
	0 (data present)	1 (data missing)
0 (data present)	996	402
1 (data missing)	44	53

## Comparing the Results of the 2 Classification Models

The random forest model and the logistic regression model agreed in 90% (1346) of cases. Table 6 contains the confusion

matrix showing how often the 2 models agreed and how often they disagreed in their classifications.

**Table 6.** Confusion matrix comparing the random forest and logistic regression models.

Logistic regression predictions	Random forest predictions	
	0 (data present)	1 (data missing)
0 (data present)	1281	117
1 (data missing)	32	65

## Discussion

### Principal Results

Our comprehensive analysis provides evidence that the missing sleep data in the TOPOWA study follows a MAR pattern, thereby permitting the use of imputation and likelihood-based methods that would be inappropriate under an MNAR mechanism. Furthermore, Enders [43] recommends adopting MAR as the default assumption in most realistic research scenarios. The findings in this study indicate that missingness was systematically related to observable variables rather than unobserved sleep characteristics [21,53]. Both the random forest model and the GEE model with a logit link function had high specificity values, indicating that the models identified true negatives well (present data values), but the values for sensitivity were low, indicating that the models performed worse at identifying true positives (missing data values). The classification models more accurately identified when data values were present than when data values were missing. However, previous researchers who have used this method to check for evidence that missing data are MAR have focused on the AUC value and did not report sensitivity and specificity [35].

While additional studies using varying types of data, such as supplementary self-reported sleep data, are necessary to further validate the methodology used in this study, the results suggest that a GEE model with a logit link function and a random forest classification model can be used to identify evidence that missing data are MAR in repeated measures data when modifications are made to account for the dependencies between observations from the same person. This finding has the potential to improve researchers' handling of missing data in longitudinal data, common in health research, by providing a way to empirically determine the likely mechanism of missingness prior to analysis and decide how best to handle the missing data based on that mechanism of missingness.

### Possible Explanations for Missing Data

Several key observations support the conclusion that the missingness is MAR. First, the temporal pattern of missingness showed a clear progression, with later nights (particularly night 5) having higher rates of missingness than earlier nights. Second, known explanations for missingness—including device failure (n=1 likely case), nonadherence (n=3 self-reported cases), or improper device fit—would generally qualify as MAR mechanisms. While 37 participants had no sleep data for any of the 5 nights, the available evidence suggests these cases resulted from technical or behavioral factors rather than being related to the missing sleep values themselves.

Missing data in the baseline sleep dataset likely stemmed from multiple sources, each with distinct implications for analysis. Battery depletion emerged as a primary factor, with some devices failing to maintain charge through all 5 nights, according to reports from the research assistants. This explanation aligns with the observed increase in missingness across successive nights and qualifies as MAR, as battery failure was unrelated to participants' sleep patterns. The most common missingness pattern in the data was data missing on the fifth night only, which likely occurred because the battery was depleted before the fifth night. The third most common missingness pattern was data present for the first 3 nights and missing for the fourth and fifth nights, which also likely occurred due to a depleted battery. A subset of cases involved unpredictable device malfunctions (MCAR), where hardware failures caused complete data loss independent of user behavior or sleep characteristics. This was likely the case for the 1 participant for which no heart rate, activity, or sleep data was collected, and may have been the case for the 8 participants for whom either heart rate data or activity data was collected (but not both) and no sleep data was collected. Another possibility is that the missing data in the underlying raw data (heart rate variability and motion data) resulted in insufficient physiological data for sleep algorithm processing. This pattern highlights the technical and behavioral

challenges inherent in wearable research, where algorithm-derived outcomes (like sleep duration and stages) depend on sensor data.

Participant-related issues also contributed: some women reported removing devices due to social discomfort (eg, partner or child concerns) [14], while improper fit (eg, loose wrist placement) prevented sensor contact, mimicking nonwear. In the case of the 8 participants for whom no sleep data was collected but either heart rate or activity data was collected, and the 28 participants for which heart rate and activity data was collected but sleep data was not, intermittent device removal (ie, at night) may have been a factor contributing to sleep data missingness. The coexistence of these mechanisms, and the potential for both technical and behavioral reasons for missingness, highlights the complexity of missingness in wearable research and underscores the need for robust analytical approaches to mitigate bias.

While only 3 women reported removing the device for any period of time during the 5-day data collection period, the missingness patterns in the data from considerably more devices suggested that device removal was a possible factor. The second most common missingness pattern in the data was data missing for all 5 nights, which could indicate the participants removed the device. Notably, the discrepancy between self-reported and probable nonwear suggests possible social desirability bias, as participants received compensation for participation.

## Limitations

Several important limitations should be considered when interpreting this study's findings. The presence or absence of other physiological or behavioral metrics that wearable devices can capture (eg, activity levels, heart rate variability) was examined, but only the presence or absence of this data was tracked. Detailed information on heart rate or activity was not analyzed. Low-quality data may also mask data that should have been marked as missing. Some extreme sleep duration values (both small and large) were identified that could represent either genuine biological variation or measurement artifacts. Wearable devices may misclassify wakefulness as sleep during prolonged sedentary periods or conversely interpret restless sleep as wakefulness. While these devices generally provide more objective sleep measurements than self-reports, such classification errors underscore the importance of interpreting extreme values cautiously. Future studies could benefit from incorporating complementary data sources such as self-reported sleep to validate these outlier observations.

---

## Acknowledgments

The authors would like to thank their partners at Uganda Youth Development Link staff and research assistants who supported The Onward Project on Well-being and Adversity project.

---

## Funding

Research reported in this publication was funded by the National Institute of Mental Health, of the National Institutes of Health, under award number R01MH128930 to Swahn. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

While it is possible that the day of the week may impact missingness, this was not explored in this analysis for 2 reasons. The women in this study were issued Garmin devices for 5-day periods on a rolling basis, meaning that not all women wore the devices over the weekend. The second reason that considering weekday versus weekend missingness is not ideal using this data is that the lifestyle and employment status of these women are very different from those of middle- and working-class populations, as well as Western assumptions of a structured 5-day work week, who typically have consistent and differing routines on weekdays and weekends. Most of these women did not have jobs with regular Monday-Friday hours.

Lastly, questions regarding the validity and reliability of consumer-grade wearable devices remain and should be further researched, especially because validity can vary between different device models from the same brand using similar technology. As health research using wearable devices continues and expands, the study of the validity of these devices should continue as well, especially as new models and technology that have not been previously studied are released. Additionally, self-reported sleep data has long been used for sleep research and can be a valuable tool when used as complementary data for validating sleep data from wearable technology.

## Conclusion

This work demonstrates the process researchers can undertake when assessing missing data assumptions prior to the analysis of wearable data. Taken together, these findings suggest that a MAR mechanism can be assumed for sleep data in the TOPOWA study, with missingness driven primarily by observable technical and behavioral factors rather than unobserved sleep characteristics. This distinction is critical for selecting appropriate imputation strategies and ensuring valid inferences in future analyses. As wearable methods expand in global health research, especially in urban informal settlements, addressing the practical and ethical complexities of data missingness will be essential for advancing equity, rigor, and reproducibility. In addition to considering data collected in a low-resource setting, where research faces additional challenges, this study contributes meaningful information on missingness in wearable sleep data aggregated by night, closely aligning it with the common use and analysis of this type of data.

---

### Data Availability

The deidentified participant-level sleep data may be obtained by reasonable request to the principal investigator or corresponding author.

---

### Authors' Contributions

Conceptualization: KM (lead), KBG (supporting), KN (supporting), LMM (supporting)

Data curation: MHS, KM, KN

Formal analysis: KM

Funding acquisition: MHS

Investigation: MHS, KM, KN, KBG

Methodology: KM (lead), KBG (supporting), KN (supporting), LMM (supporting), GT (supporting)

Project administration: MHS (lead), KM, KN, CN, KBG

Resources: MHS

Supervision: KBG, MHS, KN, LMM, GT

Writing – original draft: KM

Writing – review & editing: KM, KBG, KN, LMM, GT, CN, MHS

---

### Conflicts of Interest

None declared.

---

### References

1. Huhn S, Axt M, Gunga HC, et al. The impact of wearable technologies in health research: scoping review. *JMIR Mhealth Uhealth*. Jan 25, 2022;10(1):e34384. [doi: [10.2196/34384](https://doi.org/10.2196/34384)] [Medline: [35076409](https://pubmed.ncbi.nlm.nih.gov/35076409/)]
2. Fitbit research library. Fitabase. URL: <https://www.fitabase.com/research-library/> [Accessed 2025-06-02]
3. Holko M, Litwin TR, Munoz F, et al. Wearable fitness tracker use in federally qualified health center patients: strategies to improve the health of all of us using digital health devices. *NPJ Digit Med*. Apr 25, 2022;5(1):53. [doi: [10.1038/s41746-022-00593-x](https://doi.org/10.1038/s41746-022-00593-x)] [Medline: [35469045](https://pubmed.ncbi.nlm.nih.gov/35469045/)]
4. Nielsen K, Mobley K, Culbreth R, Palmier J, Nabulya A, Swahn MH. Feasibility and acceptability of wearable devices and daily diaries to assess sleep and other health indicators among young women in the slums of Kampala, Uganda. *Digit Health*. 2024;10:20552076241288754. [doi: [10.1177/20552076241288754](https://doi.org/10.1177/20552076241288754)] [Medline: [39430703](https://pubmed.ncbi.nlm.nih.gov/39430703/)]
5. Huhn S, Matzke I, Koch M, et al. Using wearable devices to generate real-world, individual-level data in rural, low-resource contexts in Burkina Faso, Africa: a case study. *Front Public Health*. 2022;10:972177. [doi: [10.3389/fpubh.2022.972177](https://doi.org/10.3389/fpubh.2022.972177)] [Medline: [36249225](https://pubmed.ncbi.nlm.nih.gov/36249225/)]
6. van Zyl C, Badenhorst M, Hanekom S, Heine M. Unravelling “low-resource settings”: a systematic scoping review with qualitative content analysis. *BMJ Glob Health*. Jun 2021;6(6):e005190. [doi: [10.1136/bmjgh-2021-005190](https://doi.org/10.1136/bmjgh-2021-005190)] [Medline: [34083239](https://pubmed.ncbi.nlm.nih.gov/34083239/)]
7. Make cities and human settlements inclusive, safe, resilient and sustainable. United Nations Statistics Division. URL: <https://unstats.un.org/sdgs/report/2021/goal-11/> [Accessed 2025-06-02]
8. Sclar ED, Garau P, Carolini G. The 21st century health challenge of slums and cities. *Lancet*. 2005;365(9462):901-903. [doi: [10.1016/S0140-6736\(05\)71049-7](https://doi.org/10.1016/S0140-6736(05)71049-7)] [Medline: [15752535](https://pubmed.ncbi.nlm.nih.gov/15752535/)]
9. Nejad FN, Ghamari MR, Kamal SHM, Tabatabaee SS, Ganjali R. The most important social determinants of slum dwellers' health: a scoping review. *J Prev Med Public Health*. Jul 2021;54(4):265-274. [doi: [10.3961/jpmph.21.073](https://doi.org/10.3961/jpmph.21.073)] [Medline: [34370940](https://pubmed.ncbi.nlm.nih.gov/34370940/)]
10. Ompad DC, Galea S, Caiaffa WT, Vlahov D. Social determinants of the health of urban populations: methodologic considerations. *J Urban Health*. May 2007;84(3 Suppl):i42-53. [doi: [10.1007/s11524-007-9168-4](https://doi.org/10.1007/s11524-007-9168-4)] [Medline: [17458704](https://pubmed.ncbi.nlm.nih.gov/17458704/)]
11. Ustinov Y, Lichstein KL. Actigraphy reliability with normal sleepers. *Behav Sleep Med*. 2013;11(5):313-320. [doi: [10.1080/15402002.2012.688779](https://doi.org/10.1080/15402002.2012.688779)] [Medline: [23268697](https://pubmed.ncbi.nlm.nih.gov/23268697/)]
12. Gupta N, Crowley P, Holtermann A, Straker L, Stamatakis E, Ding D. Are we ready for wearable-based global physical activity surveillance? *Br J Sports Med*. Mar 21, 2024;58(7):356-358. [doi: [10.1136/bjsports-2023-106780](https://doi.org/10.1136/bjsports-2023-106780)] [Medline: [38336382](https://pubmed.ncbi.nlm.nih.gov/38336382/)]
13. Liverani M, Ir P, Perel P, Khan M, Balabanova D, Wiseman V. Assessing the potential of wearable health monitors for health system strengthening in low- and middle-income countries: a prospective study of technology adoption in Cambodia. *Health Policy Plan*. Sep 13, 2022;37(8):943-951. [doi: [10.1093/heapol/czac019](https://doi.org/10.1093/heapol/czac019)] [Medline: [35262172](https://pubmed.ncbi.nlm.nih.gov/35262172/)]
14. Swahn MH, Gittner KB, Lyons MJ, et al. Advancing mHealth research in low-resource settings: young women's insights and implementation challenges with wearable smartwatch devices in Uganda. *Sensors (Basel)*. Aug 29, 2024;24(17):5591. [doi: [10.3390/s24175591](https://doi.org/10.3390/s24175591)] [Medline: [39275502](https://pubmed.ncbi.nlm.nih.gov/39275502/)]

15. Garbern S, Mbanjumucyo G, Umuhoza C, et al. Feasibility and accuracy of a wearable biosensor device for vital sign monitoring in septic emergency department patients in Rwanda. *Prehosp Disaster med.* May 2019;34(s1):s85-s86. [doi: [10.1017/S1049023X1900178X](https://doi.org/10.1017/S1049023X1900178X)]
16. Ghomrawi HM, Many BT, Holl JL, et al. Clinicians' perspectives on wearable sensor technology as an alternative bedside monitoring tool in two West African countries. *Int J Med Inform.* Jul 2023;175:105046. [doi: [10.1016/j.ijmedinf.2023.105046](https://doi.org/10.1016/j.ijmedinf.2023.105046)] [Medline: [37148867](https://pubmed.ncbi.nlm.nih.gov/37148867/)]
17. Van HMT, Hao NV, Phan Nguyen Quoc K, et al. Vital sign monitoring using wearable devices in a Vietnamese intensive care unit. *BMJ Innov.* Mar 2021;7(Suppl 1):s7-s11. [doi: [10.1136/bmjinnov-2021-000707](https://doi.org/10.1136/bmjinnov-2021-000707)]
18. Cao Y, Allore H, Vander Wyk B, Gutman R. Review and evaluation of imputation methods for multivariate longitudinal data with mixed-type incomplete variables. *Stat Med.* Dec 30, 2022;41(30):5844-5876. [doi: [10.1002/sim.9592](https://doi.org/10.1002/sim.9592)] [Medline: [36220138](https://pubmed.ncbi.nlm.nih.gov/36220138/)]
19. Liu Y, De A. Multiple imputation by fully conditional specification for dealing with missing data in a large epidemiologic study. *Int J Stat Med Res.* 2015;4(3):287-295. [doi: [10.6000/1929-6029.2015.04.03.7](https://doi.org/10.6000/1929-6029.2015.04.03.7)] [Medline: [27429686](https://pubmed.ncbi.nlm.nih.gov/27429686/)]
20. Adhikari D, Jiang W, Zhan J, et al. A comprehensive survey on imputation of missing data in Internet of Things. *ACM Comput Surv.* Jul 31, 2023;55(7):1-38. [doi: [10.1145/3533381](https://doi.org/10.1145/3533381)]
21. Huque MH, Carlin JB, Simpson JA, Lee KJ. A comparison of multiple imputation methods for missing data in longitudinal studies. *BMC Med Res Methodol.* Dec 12, 2018;18(1):168. [doi: [10.1186/s12874-018-0615-6](https://doi.org/10.1186/s12874-018-0615-6)] [Medline: [30541455](https://pubmed.ncbi.nlm.nih.gov/30541455/)]
22. Braem CIR, Yavuz US, Hermens HJ, Veltink PH. Missing data statistics provide causal insights into data loss in diabetes health monitoring by wearable sensors. *Sensors (Basel).* Feb 27, 2024;24(5):1526. [doi: [10.3390/s24051526](https://doi.org/10.3390/s24051526)] [Medline: [38475061](https://pubmed.ncbi.nlm.nih.gov/38475061/)]
23. Weed L, Lok R, Chawra D, Zeitzer J. The impact of missing data and imputation methods on the analysis of 24-hour activity patterns. *Clocks & Sleep.* 2022;4(4):497-507. [doi: [10.3390/clockssleep4040039](https://doi.org/10.3390/clockssleep4040039)]
24. Feng T, Narayanan S. Imputing missing data in large-scale multivariate biomedical wearable recordings using bidirectional recurrent neural networks with temporal activation regularization. Presented at: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Jul 23-27, 2019; Berlin, Germany. URL: <https://ieeexplore.ieee.org/document/8856966> [Accessed 2026-03-09] [doi: [10.1109/EMBC.2019.8856966](https://doi.org/10.1109/EMBC.2019.8856966)]
25. Lee MP, Hoang K, Park S, et al. Imputing missing sleep data from wearables with neural networks in real-world settings. *Sleep.* Jan 11, 2024;47(1):zsad266. [doi: [10.1093/sleep/zsad266](https://doi.org/10.1093/sleep/zsad266)] [Medline: [37819273](https://pubmed.ncbi.nlm.nih.gov/37819273/)]
26. Cho S, Weng C, Kahn MG, Natarajan K. Identifying data quality dimensions for person-generated wearable device data: multi-method study. *JMIR Mhealth Uhealth.* Dec 23, 2021;9(12):e31618. [doi: [10.2196/31618](https://doi.org/10.2196/31618)] [Medline: [34941540](https://pubmed.ncbi.nlm.nih.gov/34941540/)]
27. Hicks JL, Althoff T, Sosic R, et al. Best practices for analyzing large-scale health data from wearables and smartphone apps. *NPJ Digit Med.* 2019;2:45. [doi: [10.1038/s41746-019-0121-1](https://doi.org/10.1038/s41746-019-0121-1)] [Medline: [31304391](https://pubmed.ncbi.nlm.nih.gov/31304391/)]
28. Matzke I, Huhn S, Koch M, et al. Assessment of heat exposure and health outcomes in rural populations of western Kenya by using wearable devices: observational case study. *JMIR Mhealth Uhealth.* Jul 4, 2024;12(1):e54669. [doi: [10.2196/54669](https://doi.org/10.2196/54669)] [Medline: [38963698](https://pubmed.ncbi.nlm.nih.gov/38963698/)]
29. McKnight PE, McKnight KM, Sidani S, Figueredo AJ. *Missing Data: A Gentle Introduction.* Guilford Press; 2007. URL: <https://www.guilford.com/books/Missing-Data/McKnight-McKnight-Sidani-Figueredo/9781593853938> [Accessed 2026-03-09] ISBN: 978-1593853938
30. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res.* Mar 2011;20(1):40-49. [doi: [10.1002/mpr.329](https://doi.org/10.1002/mpr.329)] [Medline: [21499542](https://pubmed.ncbi.nlm.nih.gov/21499542/)]
31. Madley-Dowd P, Hughes R, Tilling K, Heron J. The proportion of missing data should not be used to guide decisions on multiple imputation. *J Clin Epidemiol.* Jun 2019;110:63-73. [doi: [10.1016/j.jclinepi.2019.02.016](https://doi.org/10.1016/j.jclinepi.2019.02.016)] [Medline: [30878639](https://pubmed.ncbi.nlm.nih.gov/30878639/)]
32. Fernstad SJ. To identify what is not there: a definition of missingness patterns and evaluation of missing value visualization. *Inf Vis.* Apr 2019;18(2):230-250. [doi: [10.1177/1473871618785387](https://doi.org/10.1177/1473871618785387)]
33. Little RJA. A test of missing completely at random for multivariate data with missing values. *J Am Stat Assoc.* Dec 1988;83(404):1198-1202. [doi: [10.1080/01621459.1988.10478722](https://doi.org/10.1080/01621459.1988.10478722)]
34. Fielding S, Fayers PM, Ramsay CR. Investigating the missing data mechanism in quality of life outcomes: a comparison of approaches. *Health Qual Life Outcomes.* Jun 22, 2009;7:57. [doi: [10.1186/1477-7525-7-57](https://doi.org/10.1186/1477-7525-7-57)] [Medline: [19545408](https://pubmed.ncbi.nlm.nih.gov/19545408/)]
35. Sondhi A, Weberpals J, Yerram P, et al. A systematic approach towards missing lab data in electronic health records: a case study in non-small cell lung cancer and multiple myeloma. *CPT Pharmacom Syst Pharma.* Sep 2023;12(9):1201-1212. [doi: [10.1002/psp4.12998](https://doi.org/10.1002/psp4.12998)]

36. 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*. Nov 28, 2019;21(11):e16273. [doi: [10.2196/16273](https://doi.org/10.2196/16273)] [Medline: [31778122](https://pubmed.ncbi.nlm.nih.gov/31778122/)]
37. Berryhill S, Morton CJ, Dean A, et al. Effect of wearables on sleep in healthy individuals: a randomized crossover trial and validation study. *J Clin Sleep Med*. May 15, 2020;16(5):775-783. [doi: [10.5664/jcsm.8356](https://doi.org/10.5664/jcsm.8356)] [Medline: [32043961](https://pubmed.ncbi.nlm.nih.gov/32043961/)]
38. Roberts DM, Schade MM, Mathew GM, Gartenberg D, Buxton OM. Detecting sleep using heart rate and motion data from multisensor consumer-grade wearables, relative to wrist actigraphy and polysomnography. *Sleep*. Jul 13, 2020;43(7):zsaa045. [doi: [10.1093/sleep/zsaa045](https://doi.org/10.1093/sleep/zsaa045)] [Medline: [32215550](https://pubmed.ncbi.nlm.nih.gov/32215550/)]
39. Sathyanarayana A, Joty S, Fernandez-Luque L, et al. Sleep quality prediction from wearable data using deep learning. *JMIR Mhealth Uhealth*. Nov 4, 2016;4(4):e125. [doi: [10.2196/mhealth.6562](https://doi.org/10.2196/mhealth.6562)] [Medline: [27815231](https://pubmed.ncbi.nlm.nih.gov/27815231/)]
40. Evenson KR, Spade CL. Review of validity and reliability of Garmin activity trackers. *J Meas Phys Behav*. Jun 2020;3(2):170-185. [doi: [10.1123/jmpb.2019-0035](https://doi.org/10.1123/jmpb.2019-0035)] [Medline: [32601613](https://pubmed.ncbi.nlm.nih.gov/32601613/)]
41. Schyvens AM, Van Oost NC, Aerts JM, et al. Accuracy of Fitbit Charge 4, Garmin Vivosmart 4, and WHOOP versus polysomnography: systematic review. *JMIR Mhealth Uhealth*. Mar 27, 2024;12:e52192. [doi: [10.2196/52192](https://doi.org/10.2196/52192)] [Medline: [38557808](https://pubmed.ncbi.nlm.nih.gov/38557808/)]
42. Grandner MA, Lujan MR, Ghani SB. Sleep-tracking technology in scientific research: looking to the future. *Sleep*. May 14, 2021;44(5):zsab071. [doi: [10.1093/sleep/zsab071](https://doi.org/10.1093/sleep/zsab071)] [Medline: [33782687](https://pubmed.ncbi.nlm.nih.gov/33782687/)]
43. Enders CK. *Applied Missing Data Analysis*. 2nd ed. Guilford Publications; 2022. URL: <https://www.guilford.com/books/Applied-Missing-Data-Analysis/Craig-Enders/9781462549863> [Accessed 2026-03-09] ISBN: 978-1462549863
44. Culbreth RE, Nielsen KE, Mobley K, Palmier J, Bukuluki P, Swahn MH. Life satisfaction factors, stress, and depressive symptoms among young women living in urban Kampala: findings from the TOPOWA Project pilot studies. *Int J Environ Res Public Health*. Feb 6, 2024;21(2):184. [doi: [10.3390/ijerph21020184](https://doi.org/10.3390/ijerph21020184)] [Medline: [38397675](https://pubmed.ncbi.nlm.nih.gov/38397675/)]
45. How Garmin watches track your sleep, calculate sleep score. Garmin. URL: <https://www.garmin.com/en-US/blog/fitness/how-garmin-watches-track-your-sleep-calculate-sleep-score/> [Accessed 2025-06-03]
46. Zhang Z. Missing data exploration: highlighting graphical presentation of missing pattern. *Ann Transl Med*. Dec 2015;3(22):356. [doi: [10.3978/j.issn.2305-5839.2015.12.28](https://doi.org/10.3978/j.issn.2305-5839.2015.12.28)] [Medline: [26807411](https://pubmed.ncbi.nlm.nih.gov/26807411/)]
47. Grandner MA, Jackson NJ, Izci-Balserak B, et al. Social and behavioral determinants of perceived insufficient sleep. *Front Neurol*. 2015;6:112. [doi: [10.3389/fneur.2015.00112](https://doi.org/10.3389/fneur.2015.00112)] [Medline: [26097464](https://pubmed.ncbi.nlm.nih.gov/26097464/)]
48. Mazloomi SN, Talebi S, Kazemi M, et al. Food insecurity is associated with the sleep quality and quantity in adults: a systematic review and meta-analysis. *Public Health Nutr*. Apr 2023;26(4):792-802. [doi: [10.1017/S1368980022002488](https://doi.org/10.1017/S1368980022002488)] [Medline: [36416108](https://pubmed.ncbi.nlm.nih.gov/36416108/)]
49. Fairclough DL. *Design and Analysis of Quality of Life Studies in Clinical Trials*. Chapman and Hall/CRC; 2002. URL: <https://www.taylorfrancis.com/books/mono/10.1201/9781420035438/design-analysis-quality-life-studies-clinical-trials-diane-fairclough> [Accessed 2026-03-09] ISBN: 9780429122743
50. Karpievitch YV, Hill EG, Leclerc AP, Dabney AR, Almeida JS. An introspective comparison of random forest-based classifiers for the analysis of cluster-correlated data by way of RF++. *PLoS One*. Sep 18, 2009;4(9):e7087. [doi: [10.1371/journal.pone.0007087](https://doi.org/10.1371/journal.pone.0007087)] [Medline: [19763254](https://pubmed.ncbi.nlm.nih.gov/19763254/)]
51. Hoffman L. *Longitudinal Analysis: Modeling within-Person Fluctuation and Change*. Routledge; 2015. [doi: [10.4324/9781315744094](https://doi.org/10.4324/9781315744094)] ISBN: 9781315744094
52. Ballinger GA. Using generalized estimating equations for longitudinal data analysis. *Organ Res Methods*. Apr 2004;7(2):127-150. [doi: [10.1177/1094428104263672](https://doi.org/10.1177/1094428104263672)]
53. Jadhav A, Pramod D, Ramanathan K. Comparison of performance of data imputation methods for numeric dataset. *Appl Artif Intell*. Aug 24, 2019;33(10):913-933. [doi: [10.1080/08839514.2019.1637138](https://doi.org/10.1080/08839514.2019.1637138)]

## Abbreviations

- AUC:** area under the curve
- GEE:** generalized estimating equations
- MAR:** missing at random
- MCAR:** missing completely at random
- MNAR:** missing not at random
- REDCap:** Research Electronic Data Capture
- REM:** rapid eye movement
- TOPOWA:** The Onward Project on Well-being and Adversity

*Edited by Georgian Badicu; peer-reviewed by Antonio Martinko, Matthew Louis Mauriello; submitted 22.Jul.2025; final revised version received 25.Jan.2026; accepted 09.Feb.2026; published 31.Mar.2026*

*Please cite as:*

*Mobley K, Gittner KB, Nielsen K, Matheny LM, Taasoobshirazi G, Natuhama C, Swahn MH*

*Determining a Likely Mechanism of Missingness in Repeated Measures Sleep Data From Wearable Fitness Trackers: Longitudinal Analysis*

*JMIR Mhealth Uhealth 2026;14:e81123*

URL: <https://mhealth.jmir.org/2026/1/e81123>

doi: [10.2196/81123](https://doi.org/10.2196/81123)

© Kate Mobley, Kevin B Gittner, Karen Nielsen, Lauren M Matheny, Gita Taasoobshirazi, Charles Natuhama, Monica H Swahn. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.