

Viewpoint

# The Importance of Data Quality Control in Using Fitbit Device Data From the *All of Us* Research Program

Lauren Lederer<sup>1</sup>, BS; Amanda Breton<sup>2</sup>, MS; Hayoung Jeong<sup>1</sup>, MS; Hiral Master<sup>3</sup>, PhD; Ali R Roghanizad<sup>1</sup>, PhD; Jessilyn Dunn<sup>1</sup>, PhD

<sup>1</sup>Department of Biomedical Engineering, Duke University, Durham, NC, United States

<sup>2</sup>Department of Electrical and Computer Engineering, Duke University, Durham, NC, United States

<sup>3</sup>Vanderbilt Institute for Clinical and Translational Research, Vanderbilt University Medical Center, Nashville, TN, United States

**Corresponding Author:**

Jessilyn Dunn, PhD

Department of Biomedical Engineering

Duke University

1427 FCIEMAS, Box 90281

Durham, NC, 27708

United States

Phone: 1 919-660-5131

Email: [jessilyn.dunn@duke.edu](mailto:jessilyn.dunn@duke.edu)

## Abstract

Wearable digital health technologies (DHTs) have become increasingly popular in recent years, enabling more capabilities to assess behaviors and physiology in free-living conditions. The *All of Us* Research Program (AoURP), a National Institutes of Health initiative that collects health-related information from participants in the United States, has expanded its data collection to include DHT data from Fitbit devices. This offers researchers an unprecedented opportunity to examine a large cohort of DHT data alongside biospecimens and electronic health records. However, there are existing challenges and sources of error that need to be considered before using Fitbit device data from the AoURP. In this viewpoint, we examine the reliability of and potential error sources associated with the Fitbit device data available through the AoURP Researcher Workbench and outline actionable strategies to mitigate data missingness and noise. We begin by discussing sources of noise, including (1) inherent measurement inaccuracies, (2) skin tone-related challenges, and (3) movement and motion artifacts, and proceed to discuss potential sources of data missingness in Fitbit device data. We then outline methods to mitigate such missingness and noise in the data. We end by considering how future enhancements to the AoURP's Fitbit device data collection methods and the inclusion of new Fitbit data types would impact the usability of the data. Although the reliability considerations and suggested literature are tailored toward Fitbit device data in the AoURP, the considerations and recommendations are broadly applicable to data from wearable DHTs in free-living conditions.

*JMIR Mhealth Uhealth* 2023;11:e45103; doi: [10.2196/45103](https://doi.org/10.2196/45103)

**Keywords:** wearable device; Fitbit; All of Us; data quality; noise; missingness; biometric monitoring

## Introduction

Wearable digital health technologies (DHTs) have become increasingly popular in recent years, especially as DHTs offer better user experiences, more capabilities, and greater functionality to assess behaviors and physiology in free-living conditions. The *All of Us* Research Program (AoURP) is an initiative that is seeking to collect health-related information, including DHT data, from a diverse cohort of over 1 million participants in the United States. In the AoURP, DHT data are collected alongside electronic health records, biospecimens, surveys, and standardized physical measurements. The goal

is to make these data accessible to both researchers and participants to advance precision diagnosis, prevention, and treatment [1].

In 2019, the AoURP expanded its data collection with the Fitbit Bring-Your-Own-Device (BYOD) project. This expansion has allowed participants to share their historical and ongoing Fitbit account data through the *All of Us* participant portal [2]. The AoURP's efforts to include Fitbit device data have continued to expand with the WEAR study, which gives eligible participants a no-cost Fitbit Charge 4 or Fitbit Versa 3 device [3]. Data from the BYOD project are available through the AoURP Researcher Workbench,

which offers data access and analysis tools for DHT data, electronic health record data, biospecimens, surveys, and physical measurements [4].

As of 2023, over 15,000 *All of Us* participants have shared Fitbit device data [5]. The Researcher Workbench provides access to these data, including Fitbit-defined heart rate by zones that are based on percentages of estimated maximum heart rate, minute-level heart rate, daily activity summaries, minute-level intraday steps, daily sleep summaries, and sleep levels [6].

Digital biomarkers derived from DHTs can potentially be used to improve clinical diagnostics, predict disease status, and support personalized clinical decision-making [7]. With the increasing use of DHTs like Fitbit devices in the AoURP and other research and clinical settings, it is important that those working with these data consider the inherent limitations of Fitbit devices, given their underlying technology. This will enable improved data processing and fit-for-purpose implementations of Fitbit devices in research and clinical settings. Researchers might ask the following: “How reliable is the data from these devices? What are the sources of noise, error, and bias that should be accounted for when using this data? How can these be accounted for?”

In this viewpoint paper, we examine the reliability of Fitbit device data in the context of the AoURP’s BYOD program. We focus on Fitbit devices, given their wide market share [8,9], the ongoing collection of data from Fitbit users in the AoURP (eg, BYOD program) [2,10], and these data’s availability to registered Researcher Workbench users [5]. Data are currently not available from the AoURP regarding the Fitbit device models included in their data set. For this reason, the data cleaning strategies we present are device model agnostic. This paper focuses specifically on data considerations around physical activity (steps and movement intensities) and heart rate measurements generated by Fitbit devices on a daily and per-minute basis. Given that Fitbit device sleep data are derived from the same underlying sensors that determine heart rate and motion intensity metrics, the same fundamental considerations surrounding inherent measurement reliability should be considered when working with Fitbit device sleep metrics.

## Sources of Noise

Measurement error can affect the reliability of Fitbit device data in both laboratory conditions and free-living conditions. In this section, we discuss the most commonly recognized sources of error that may be observed in Fitbit device data collected through the AoURP.

### *Inherent Measurement Inaccuracies*

Fitbit devices include a 3-axis accelerometer and photoplethysmography (PPG) sensor, with more recent device models including additional sensors, such as an altimeter, a gyroscope, a skin temperature sensor, and multipurpose electrical sensors [11]. It may be helpful for researchers to consider the data supply chain as they work with DHT data [12]; the data

that researchers generally have access to are processed, and the firmware that performs such data processing is regularly updated [13]. Aside from the data supply chain, there are also inherent limitations of the accelerometer and PPG sensors themselves that should be accounted for in data analysis.

All Fitbit models with Fitbit LLC’s patented PurePulse technology (eg, Fitbit Charge, Fitbit Charge 2, Fitbit Charge 3, Fitbit Alta, Fitbit Versa, Fitbit Blaze, and Fitbit Ionic) use the same PPG hardware and software for heart rate estimation [14]. PPG sensors, which optically measure light absorption under the skin, may be affected by user motion and activity intensity, skin tone, and the wavelength of light used by the sensor [15,16]. When compared to gold-standard electrocardiography, Fitbit devices tend to underestimate heart rate [14,15,17]. Further, Fitbit device heart rate measurements have higher reliability under stationary conditions [14,18].

Fitbit devices use the 3-axis accelerometer to determine step count and categorize physical activity intensity (sedentary, light, moderate, vigorous, or moderate to vigorous) [19-21]. Comparisons between Fitbit device step counts using direct observation and gold-standard accelerometers, such as the ActiGraph GT3X+, demonstrate mixed reliability, depending on the type and speed of movement and the on-body placement of the Fitbit. During normal walking for example, the torso placement of the Fitbit device has resulted in the greatest accuracy, while ankle and wrist placement have been the most accurate in slow-walking and jogging, respectively [20]. Similar findings from other studies demonstrate that step count and physical activity intensity accuracy are affected by device placement and movement type [22-26].

### *Skin Tone*

Skin tone may be another inherent source of error for DHTs that rely on optical measurements (eg, PPG or pulse oximetry) [27,28]. Both melanin and skin with tattoos absorb more green light, that is, wavelengths of around 530 nm, which are the LED wavelengths commonly used in PPG sensors [29]. There have been mixed findings in this area; a study by Shcherbina et al [30] on older generations (2014-2016) of consumer smartwatches found that darker skin tones positively correlated with increased heart rate error, whereas our study, in which we used more recent devices (2014-2018), did not find a relationship between heart rate measurement accuracy and skin tone across a subset of consumer smartwatches [18]. Clinical-grade pulse oximeters that rely on red and infrared optical measurement technology may also be affected by skin tone [31].

Fitbit devices may use a combination of both green and red wavelengths to estimate heart rate [15,32]. Although green wavelengths can enable more accurate heart rate measurements during movement when compared to red wavelengths, green wavelengths are more readily absorbed by melanin before reaching the photodetector [29]. Additional research is needed on whether and how the accuracy of optical-based DHT measurements, such as heart rate and saturation of peripheral oxygen (SpO<sub>2</sub>), is affected by skin tone.

The data collected by the AoURP currently do not include data on skin color or the presence of tattoos under Fitbit devices; therefore, it is not possible to directly account for skin tone or wrist tattoos as a potential source of error in AoURP heart rate data. As a result, researchers working with *All of Us* data may need to take extra care when interpreting or translating results that may be influenced by skin tone or the presence of wrist tattoos, particularly when working with heart rate data.

### ***Movement and Motion Artifacts***

Motion artifacts can also be a source of error for heart rate, step count, and physical activity intensity data. Unexpected noise with random amplitudes and frequencies can be seen in raw sensor data and can cause the algorithms of Fitbit devices to falsely detect movement or a heart beat [33]. For example, the reliability of Fitbit Flex's step count and moderate to vigorous physical activity data was found to be dependent on the activity type (walking, stair stepping, jogging, and incline walking) [24], and step count error was shown to be higher during activity than during rest [20]. Therefore, it is likely that step count reliability varies, particularly during normal household activities, which may be logged by Fitbit devices as exercise movements.

Wearable device heart rate measurements are the most accurate under circumstances of rest, followed by physical activity and then rhythmic activity, such as walking or jogging. Our previous work demonstrated decreased reliability during rhythmic activities, such as walking or typing. This was likely due to Fitbit devices mistaking the periodic signal, which was being produced by the repetitive movements, for the cardiovascular cycle. Although walking resulted in heart rate measurements that were higher than the true heart rate, typing resulted in heart rate measurements that were lower than the true heart rate [18]. A study by Benedetto et al [15] assessed Fitbit Charge 2 heart rate accuracy during stationary biking and found that the device underestimated heart rate when compared to electrocardiography.

Some possible reasons for heart rate measurement error during motion include the device's sampling and interpolation methods, unstable device positioning, and variation in the pressure applied to the skin by the sensor [15,18,34]. Researchers should be aware of the impact of physical activity type (ie, motion intensity and periodicity) on Fitbit device heart rate measurement error.

The body positioning and fit of DHTs can also be sources of motion artifacts. For example, wrist-based Fitbit devices can misclassify nonambulatory arm movements as total body motion, which may result in the overestimation of physical activity and motion intensity [35]. This misclassification of physical activity may be worse if the device is not worn correctly. To address this challenge, the Fitbit device user manuals provide instructions for specific placement on the wrist to enable the acquisition of more reliable data [36].

## **Sources of Data Missingness in Fitbit Device Data**

It can often be challenging to determine the minimum amount of data necessary to achieve a particular analysis goal when using DHT data. A systematic review by Chan et al [37] pointed to a common definition for a "valid day" of wearable data—at least 10 intermittent hours of data present within 1 day—and a "valid week" of data—at least 3 valid days during the week. It is important for researchers to note that most Fitbit devices need to be charged at least once per week for 1 to 2 hours at the time of writing, and the need to remove the device from the wrist to charge it results in at least some data missingness. More than the minimal necessary data missingness can occur in the event that the wearer forgets to put the device back on their wrist after charging it [35]. Such nonwear is an example of structured missingness, where a contiguous block of missing observations occurs when the device is not being worn. At scale, there may be observable nonwear patterns, such as times when people commonly remove their devices (eg, during sleep) [38].

In addition to nonwear, improper device wear can also result in data missingness. Improper wear, such as insufficient tightening of the wrist strap, can lead to the sensor orientation being askew or a loss of sensor-to-skin contact, which is required for high-fidelity optical measurements, such as PPG-based heart rate measurements [39]. Moreover, observations can be impacted by large motion artifacts, and such observations (eg, high accelerometry values) may be removed by the device firmware. This leads to missing values in the final data set. To explore the extent of such data removal, our team recently compared data missingness in optical heart rate and SpO<sub>2</sub> observations across multiple wearables [18,40]. We found that, for heart rate measurements, the Fitbit Charge 2 had the highest amount of missing data during both rest (18.7%) and physical activity (10.4%) when compared to other consumer-grade wearables [18]. Data missingness due to improper wear or firmware attempts to account for motion artifacts may be seen in the data set as random missingness, lacking structure and predictability. However, some wearers may be more prone to improper wear or high-intensity activity, which can lead to higher amounts of missing data in individual data sets. Other factors that can affect the presence or absence of Fitbit device data include the frequency of syncing the device with the smartphone app and poor device connectivity [41,42].

There is a taxonomy of mechanisms for missing data, including (1) data missing completely at random (MCAR), where missingness is unrelated to observed characteristics; (2) data missing at random (MAR), where missingness is related to observed characteristics; and (3) data missing not at random (MNAR), where missingness is related to unobserved characteristics. Different methods are required to best account for these three missingness mechanisms during data preprocessing; thus, identifying the type of missingness is an important step in DHT data analysis. In the case of Fitbit device data, observations that are MCAR may be due

to nonsystematic device malfunctions, nonsystematic errors in data transfer, or sporadic improper device wear. Observations that are MAR may be the result of a particular device model missing a type of measurement capability (eg, a device that is known to not report heart rate measurements under high-intensity activity). An example of MNAR missingness might be nonwear during a bout of illness or due to a user having a poorly fitting device. In a free-living study, such as the AoURP, all three missingness mechanisms are likely to be present in the Fitbit device data and should be identified and appropriately addressed when possible by, for example, making assumptions about the reasons for the data missingness upon analyzing missingness patterns [43].

## Mitigating Missingness and Noise in DHT Data

### *Accounting for Data Missingness*

Avoiding data missingness is best done at the data collection stage. Prospective bring-your-own-device studies may increase wear time and improve device fit by incorporating reminders for users to wear the device, adding nonwear alerts for users or the study team, and educating users on fit and charging [44]. It should be noted that AoURP Fitbit device data collection is purely observational and, at this time, does not involve providing any alerts or interventions to improve Fitbit device wear habits.

Some missingness in the data is inevitable. Accounting for data missingness begins by thoroughly identifying the reason for missingness and deciding upon the most appropriate strategy for mitigation. For AoURP researchers using Fitbit device data collected in free-living conditions, it is not always possible to distinguish MAR, MCAR, and MNAR missingness; therefore, they may need to make assumptions

to decide how to proceed with mitigation [43]. In the context of Fitbit device data, we can distill a few practical solutions for addressing wear-related structured missingness [38].

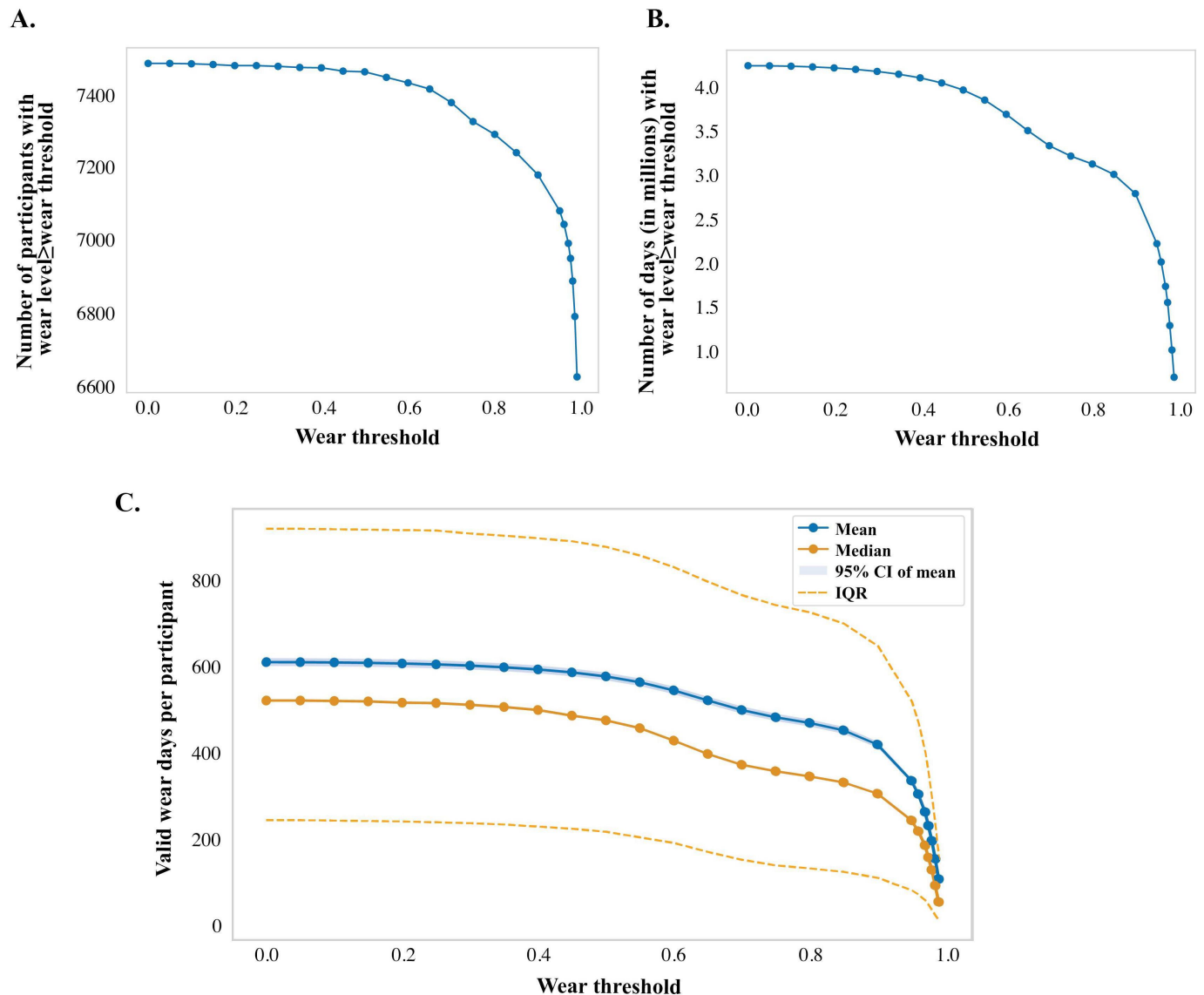
The first step is to decide upon a definition for “wear” in the context of the data and the question at hand. Depending on what analysis or question is of interest, the definition of “wear” can drastically change. For example, an analysis involving sleep quality and staging may require the definition of “wear” to include common sleep hours (eg, 8 PM to 8 AM) or a minimum monitored sleep duration time, such as that reported by Fitbit devices, but other analyses may only require a daily wear level, which gives an idea of a participant’s activity and physiological state without requirements regarding wear during specific activities, such as sleep and exercise, as described in the *Sources of Data Missingness in Fitbit Device Data* section [37].

One way to calculate a daily wear level is to leverage minute-level heart rate data, as most consumer devices only collect these data when they are worn on the wrist. This would, for example, help to avoid including step count data that may have been collected when a device is in a purse. By dividing the total count of minute-level heart rate observations collected within a single day by the total number of minutes when such data were possible to collect (1440 min in 1 d), we can derive a reasonable estimate of the proportion of the day when the device was on the wrist.

Once established, the daily wear level threshold can be used for filtering out nonwear data and participants; the optimal threshold for filtering should be selected carefully to avoid unnecessary data loss (Figure 1A, Figure 1B). The optimal threshold is where data loss is minimized and there is adequate statistical power to draw conclusions from the analysis (Figure 1C).



**Figure 1.** Graphs A and B show the total counts of participants and days, respectively, that meet each wear threshold, with an optimal threshold of <0.4. Graph C shows the mean and median total days (per participant) with a wear level greater than or equal to each wear threshold, demonstrating an optimal threshold of between 0.3 and 0.4. The 95% CI (indicated in blue) was calculated as follows:  $\mu \pm \sigma/\sqrt{n}$ . The IQR (indicated in orange dashed lines) indicates the first and third quartile values. The controlled tier AoURP version 7 data set (C2022Q4R9) was used to generate Figure 1. AoURP: *All of Us* Research Program.



After removing nonwear data (ie, contiguous blocks of missing observations), it is important to identify other sources of data missingness and determine whether mitigation is best done by using imputation or by using the complete case method [43]. Sometimes, the decision can be made based on the extent of the missingness relative to the overall data volume needed for analysis, and at times, it may be deemed that the original analysis cannot be performed as planned due to insufficient data. As an example, a recent study on COVID-19 detection via Fitbit device data calculated the mean over 5-minute intervals of heart rate data; subsequently, any missing data over full 5-minute intervals were imputed by using the median heart rate value from a previously defined 14-day window [45]. Although it is common in wearable data analysis to use the mean and median values of heart rate data for imputation, new imputation methods for biomedical wearable data have also been developed to incorporate machine learning for improved imputation accuracy [46,47].

Although imputation can be beneficial because it preserves data, it should be noted that imputation is not always a good idea, particularly in cases with substantial missingness for which typical values cannot be established.

### Accounting for Noise

Many methods exist to reduce noise in raw signal data (ie, sample-level or high-frequency signal data), particularly when the source of the noise is well characterized. Unfortunately, consumer devices typically do not give access to such high-frequency data, but their firmware and adaptive data collection methods are thought to include steps that account for skin tone-related errors and motion artifacts. Unfortunately, the public has no way of assessing how well these methods perform. Based on this, it is difficult for researchers to directly address errors resulting from skin tone or motion artifact errors. However, some general data cleaning strategies exist that may help to mitigate noise in the data,

regardless of the source of the noise (Table 1). In this section, we discuss how to leverage changes from an individual's baseline, filtering during repeated activities, and  $z$  score

normalization to improve the signal to noise ratio (SNR). We recommend using these techniques in combination with one another to best mitigate noise.

**Table 1.** Examples of noise mitigation methods for wearable data.

	Baseline comparisons	Sampling during periods of similar activity	$z$ score normalization
Description	<ul style="list-style-type: none"> <li>Calculate the median value during a defined baseline period.</li> <li>Calculate the <math>\Delta</math> from the baseline for all other data.</li> </ul>	<ul style="list-style-type: none"> <li>Establish specific wear times and use the "Activity Type" metric to filter an individual's Fitbit device data during similar time periods each day for comparable activity types.</li> <li>Conduct analysis using these segmented data sets.</li> </ul>	<ul style="list-style-type: none"> <li>Subtract the mean from each observation and divide by the SD.</li> </ul>
Applicability	<ul style="list-style-type: none"> <li>Mitigate consistent measurement error (bias).</li> </ul>	<ul style="list-style-type: none"> <li>Mitigate noise that is exacerbated under specific conditions.</li> </ul>	<ul style="list-style-type: none"> <li>Mitigate short periods of noise.</li> </ul>
Benefits	<ul style="list-style-type: none"> <li>Provides a "usual" picture of an individual [48].</li> </ul>	<ul style="list-style-type: none"> <li>Assists in isolating confounding effects that may arise in different activity types and heart rate zones.</li> <li>Standardizes the comparison of data across different individuals.</li> </ul>	<ul style="list-style-type: none"> <li>Allows direct comparison of 2 observations originating from different segments of temporal data [49].</li> </ul>
Limitations	<ul style="list-style-type: none"> <li>Need ample data to establish a baseline [48].</li> <li>Baselines can change over time.</li> </ul>	<ul style="list-style-type: none"> <li>Recommended for large sets of longitudinal data to make accurate comparisons.</li> </ul>	<ul style="list-style-type: none"> <li>Abstraction of units and range may make it difficult to interpret data.</li> </ul>

On an individual level, the comparison of observations to a reliable baseline can be helpful for determining changes in biosignals over time while reducing the influence of both skin tone and motion artifacts. Reliable baselines can be established by first summarizing an individual's measurements during periods of sleep or inactivity or before a perturbation. The determination of which time period to use to establish a baseline is study dependent. Depending on the timescale of the analysis, it is also useful to consider a sliding window approach, wherein new baselines are established during predefined time periods to account for baseline changes over time. The median value of the biosignal serves as a useful baseline value because it is less susceptible to noise and outliers compared to other statistical summary metrics and provides a way to amplify the SNR during the next steps of the analysis. The establishment of and comparisons to reliable baselines have been performed in multiple studies [50-53]. One limitation of this approach is that substantial monitoring time may be needed to establish a reliable baseline for an individual due to inherent biological and behavioral variability and the effects of external factors that may be difficult to control for (eg, seasonality, circadian rhythms, weekdays vs weekends, etc) [48]. It should also be noted that comparisons to a reliable baseline would not improve the SNR in scenarios where there is a compound effect of the source of noise and the conditions of measurement [54]. For example, skin tone may only increase measurement error for certain heart rate zones (eg, high heart rate) or under circumstances of high motion. In such cases, removing data collected under certain conditions that exacerbate measurement error may be the most appropriate approach.

Another way to handle the challenge of confounding sources of measurement error is to only compare segments of data that are measured under the same conditions (eg, similar movement types and heart rate zones) [15,18,20,24,55]. This technique allows researchers to further isolate confounding

sources of measurement error that may be exacerbated during different activities. First, one must define specific wear times of interest, such as wear during specific times of the day, which helps account for circadian variability. Second, when available, researchers should use the "Activity Type" provided by Fitbit devices to segment heart rate data into comparable sections. Researchers can also leverage this activity information to anticipate activities for which heart rate data may be less accurate.

In circumstances where there are short periods of incorrectly reported heart rates or step counts (eg, during high-intensity motion), simple normalization methods, such as  $z$  score normalization and minimum-maximum normalization, are the most useful [56,57]. Minimum-maximum normalization is useful when extreme outliers are not present in the data, especially when the data have a fixed possible range.  $z$  score normalization is particularly useful because it centers and scales the Fitbit device data to a mean of 0 and an SD of 1.  $z$  score normalization helps to reduce the comparatively higher impact of outliers within shorter data segments because it leverages information from longer segments (ie, the mean and SD) for normalization. Once normalized, the data can be compared across wearable data types and participants.

## Future Directions

As the AoURP continues efforts to provide wearable data to researchers and expand the scope of the Fitbit device data made available on the Researcher Workbench, there are several future directions to be considered. Although the ideas presented herein are tailored to Fitbit device data originating from the bring-your-own-device facet of AoURP, Fitbit device data are now actively being collected from other studies, and these data may one day be integrated into the Researcher Workbench [3]. Each additional study may have

unique characteristics, including the target population, which may play a role in the overall quality of the data. For example, whether a Fitbit device was provided to participants or whether they were using an existing device they purchased may be a factor in a participant's comfort level with using the device properly and regularly. Understanding the nuances and potential variations in data quality arising from different study protocols and data sources within the AoURP ecosystem necessitates further research. Investigating how specific study designs, participant demographics, and data collection protocols within the AoURP may influence the overall quality of the collected data will be crucial for researchers seeking to derive meaningful insights and improve the designs of future studies that implement DHTs.

Although Fitbit device model data are not currently available on the Researcher Workbench, it is worth considering how differing device models, software, and firmware may affect the data collected. At the time of writing, the underlying PurePulse PPG technology is the same across all Fitbit LLC heart rate tracking devices [14]. The largest differences in accelerometer-derived data have been observed between Fitbit LLC's early torso clip-on trackers and the newer wrist-based devices [20]. Additional research is needed to investigate whether there are any substantial differences in accelerometry performance across Fitbit LLC wrist-based models. With regard to data derived from heart rate and accelerometry, such as sleep tracking data, prior to Fitbit LLC's release of heart rate tracking devices in 2014 [58], Fitbit LLC's early accelerometry-only devices estimated sleep metrics based on movement alone. Only Fitbit devices with heart rate tracking will include sleep staging, wake heart rate, and sleep-time heart rate [59]. Identifying whether sleep

staging metrics are available for a particular individual may be a convenient way to identify the broad type of Fitbit device that was worn. The future incorporation of other contextual information, such as environmental factors, user behaviors, and device models, will enhance the ability to detect and mitigate noise, improve overall data quality, and provide a more comprehensive understanding of an individual's health.

## Conclusion

The development and validation of Fitbit device-derived digital biomarkers offer the potential for remote and continuous measurement of physiological data. Such digital biomarkers can help inform medical decisions and predict disease states [7]. The wide adoption of DHTs by both consumers and programs like the AoURP make DHTs a great source of data for researchers. Researchers can use various analytical, statistical, and machine learning approaches to further develop DHT data into digital biomarkers [60-63]. Like with any technology, there are inherent limitations and sources of error that stakeholders (eg, researchers using DHT data in their analyses) should be aware of. We encourage the *All of Us* community to use data processing techniques that address noise and missingness to reduce problems downstream in the data analysis. Although we focused on heart rate and motion data in this work, the error mitigation methods described are applicable to other forms of wearable data, including sleep data. For example, changes in total sleep time and sleep stages can be compared against baselines over time. Researchers should consider their study goals and expected outcomes when determining which data cleaning strategies are the most salient to their goals.

## Acknowledgments

This work was funded by the Data and Research Center (5U2COD023196-05). The *All of Us* Research Program would not be possible without the partnership of its participants. The *All of Us* Research Program is supported by the National Institutes of Health, Office of the Director: regional medical centers (1 OT2 OD026549; 1 OT2 OD026554; 1 OT2 OD026557; 1 OT2 OD026556; 1 OT2 OD026550; 1 OT2 OD 026552; 1 OT2 OD026553; 1 OT2 OD026548; 1 OT2 OD026551; 1 OT2 OD026555; IAA: AOD21037, AOD22003, AOD16037, AOD21041), federally qualified health centers (HHSN 263201600085U), Data and Research Center (5 U2C OD023196), Biobank (1 U24 OD023121), The Participant Center (U24 OD023176), Participant Technology Systems Center (1 U24 OD023163), communications and engagement partners (3 OT2 OD023205; 3 OT2 OD023206), and community partners (1 OT2 OD025277; 3 OT2 OD025315; 1 OT2 OD025337; 1 OT2 OD025276).

## Data Availability

To ensure the privacy of participants, the *All of Us* Research Program data used for this study are available to approved researchers following registration, completion of ethics training, and attestation of a data use agreement through the *All of Us* Research Workbench platform.

## Conflicts of Interest

None declared.

## References

1. All of Us Research Program Investigators, Denny JC, Rutter JL, et al. The "All of Us" Research Program. *N Engl J Med*. 2019 Aug 15;381(7):668-676. [doi: [10.1056/NEJMs1809937](https://doi.org/10.1056/NEJMs1809937)] [Medline: [31412182](https://pubmed.ncbi.nlm.nih.gov/31412182/)]
2. All of Us Research Program expands data collection efforts with Fitbit. All of Us Research Program. Jan 16, 2019. URL: <https://allofus.nih.gov/news-events/announcements/all-us-research-program-expands-data-collection-efforts-fitbit> [Accessed 2023-07-09]

3. Through ‘All of Us’ program, Scripps Research launches wearable technology study to accelerate precision medicine. Scripps Research. Feb 24, 2021. URL: <https://www.scripps.edu/news-and-events/press-room/2021/20210224-aou-fitbit-study.html> [Accessed 2023-07-09]
4. Researcher workbench. All of Us Research Hub. URL: <https://www.researchallofus.org/data-tools/workbench/> [Accessed 2023-07-09]
5. Master H, Kouame A, Hollis H, Marginean K, Rodriguez K. 2022Q4R9 v7 data characterization report. All of Us Research Program. 2023. URL: <https://support.researchallofus.org/hc/en-us/articles/14558858196628-2022Q4R9-v7-Data-Characterization-Report> [Accessed 2023-08-12]
6. Fitbit data. All of Us Research Hub. 2023. URL: <https://databrowser.researchallofus.org/fitbit> [Accessed 2023-07-09]
7. Motahari-Nezhad H, Fgaier M, Abid MM, Péntek M, Gulácsi L, Zrubka Z. Digital biomarker-based studies scoping review of systematic reviews. *JMIR Mhealth Uhealth*. 2022 Oct 24;10(10):e35722. [doi: [10.2196/35722](https://doi.org/10.2196/35722)] [Medline: [36279171](https://pubmed.ncbi.nlm.nih.gov/36279171/)]
8. Curry D. Fitbit revenue and usage statistics (2023). *Business of Apps*. 2023. URL: <https://www.businessofapps.com/data/fitbit-statistics/> [Accessed 2023-07-09]
9. Wearables market share companies 2023. Statista. 2023. URL: <https://www.statista.com/statistics/435944/quarterly-wearables-shipments-worldwide-market-share-by-vendor/> [Accessed 2023-07-09]
10. Holko M, Litwin TR, Munoz F, Theisz KI, Salgin L, Jenks NP, 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*. 2022 Apr 25;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/)]
11. General info and specifications. Fitbit Sense User Manual. 2023. URL: <https://help.fitbit.com/manuals/sense/Content/manuals/html/General%20Info%20and%20Specifications.htm> [Accessed 2023-07-09]
12. Goldsack JC, Coravos A, Bakker JP, Bent B, Dowling AV, Fitzer-Attas C, et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for biometric monitoring technologies (BioMeTs). *NPJ Digit Med*. 2020 Apr 14;3:55. [doi: [10.1038/s41746-020-0260-4](https://doi.org/10.1038/s41746-020-0260-4)] [Medline: [32337371](https://pubmed.ncbi.nlm.nih.gov/32337371/)]
13. Woolley SI, Collins T, Mitchell J, Fredericks D. Investigation of wearable health tracker version updates. *BMJ Health Care Inform*. 2019 Oct;26(1):e100083. [doi: [10.1136/bmjhci-2019-100083](https://doi.org/10.1136/bmjhci-2019-100083)] [Medline: [31597642](https://pubmed.ncbi.nlm.nih.gov/31597642/)]
14. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR. Accuracy of PurePulse photoplethysmography technology of Fitbit Charge 2 for assessment of heart rate during sleep. *Chronobiol Int*. 2019 Jul;36(7):927-933. [doi: [10.1080/07420528.2019.1596947](https://doi.org/10.1080/07420528.2019.1596947)] [Medline: [30990098](https://pubmed.ncbi.nlm.nih.gov/30990098/)]
15. Benedetto S, Caldato C, Bazzan E, Greenwood DC, Pensabene V, Actis P. Assessment of the Fitbit Charge 2 for monitoring heart rate. *PLoS One*. 2018 Feb 28;13(2):e0192691. [doi: [10.1371/journal.pone.0192691](https://doi.org/10.1371/journal.pone.0192691)] [Medline: [29489850](https://pubmed.ncbi.nlm.nih.gov/29489850/)]
16. Pollreis D, TaheriNejad N. Detection and removal of motion artifacts in PPG signals. *Mobile Networks and Applications*. 2022;27:728-738. [doi: [10.1007/s11036-019-01323-6](https://doi.org/10.1007/s11036-019-01323-6)]
17. Gorny AW, Liew SJ, Tan CS, Müller-Riemenschneider F. Fitbit Charge HR wireless heart rate monitor: validation study conducted under free-living conditions. *JMIR Mhealth Uhealth*. 2017 Oct 20;5(10):e157. [doi: [10.2196/mhealth.8233](https://doi.org/10.2196/mhealth.8233)] [Medline: [29055881](https://pubmed.ncbi.nlm.nih.gov/29055881/)]
18. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. *NPJ Digit Med*. 2020 Feb 10;3(1):18. [doi: [10.1038/s41746-020-0226-6](https://doi.org/10.1038/s41746-020-0226-6)] [Medline: [32047863](https://pubmed.ncbi.nlm.nih.gov/32047863/)]
19. How does my Fitbit device calculate my daily activity? Fitbit Help. URL: [https://help.fitbit.com/articles/en\\_US/Help\\_article/1141.htm](https://help.fitbit.com/articles/en_US/Help_article/1141.htm) [Accessed 2023-07-09]
20. Feehan LM, Goldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. *JMIR Mhealth Uhealth*. 2018 Aug 9;6(8):e10527. [doi: [10.2196/10527](https://doi.org/10.2196/10527)] [Medline: [30093371](https://pubmed.ncbi.nlm.nih.gov/30093371/)]
21. 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 Dec 18;12(1):159. [doi: [10.1186/s12966-015-0314-1](https://doi.org/10.1186/s12966-015-0314-1)] [Medline: [26684758](https://pubmed.ncbi.nlm.nih.gov/26684758/)]
22. Chow JJ, Thom JM, Wewege MA, Ward RE, Parmenter BJ. Accuracy of step count measured by physical activity monitors: the effect of gait speed and anatomical placement site. *Gait Posture*. 2017 Sep;57:199-203. [doi: [10.1016/j.gaitpost.2017.06.012](https://doi.org/10.1016/j.gaitpost.2017.06.012)] [Medline: [28666177](https://pubmed.ncbi.nlm.nih.gov/28666177/)]
23. Jung HC, Kang M, Lee NH, Jeon S, Lee S. Impact of placement of Fitbit HR under laboratory and free-living conditions. *Sustainability*. 2020;12(16):6306. [doi: [10.3390/su12166306](https://doi.org/10.3390/su12166306)]
24. Sushames A, Edwards A, Thompson F, McDermott R, Gebel K. Validity and reliability of Fitbit Flex for step count, moderate to vigorous physical activity and activity energy expenditure. *PLoS One*. 2016 Sep 2;11(9):e0161224. [doi: [10.1371/journal.pone.0161224](https://doi.org/10.1371/journal.pone.0161224)] [Medline: [27589592](https://pubmed.ncbi.nlm.nih.gov/27589592/)]



25. Modave F, Guo Y, Bian J, Gurka MJ, Parish A, Smith MD, et al. Mobile device accuracy for step counting across age groups. *JMIR Mhealth Uhealth*. 2017 Jun 28;5(6):e88. [doi: [10.2196/mhealth.7870](https://doi.org/10.2196/mhealth.7870)] [Medline: [28659255](https://pubmed.ncbi.nlm.nih.gov/28659255/)]
26. Tedesco S, Sica M, Ancillao A, Timmons S, Barton J, O'Flynn B. Validity evaluation of the Fitbit Charge2 and the Garmin vivosmart HR+ in free-living environments in an older adult cohort. *JMIR Mhealth Uhealth*. 2019 Jun 19;7(6):e13084. [doi: [10.2196/13084](https://doi.org/10.2196/13084)] [Medline: [31219048](https://pubmed.ncbi.nlm.nih.gov/31219048/)]
27. Colvonen PJ. Response to: investigating sources of inaccuracy in wearable optical heart rate sensors. *NPJ Digit Med*. 2021 Feb 26;4(1):38. [doi: [10.1038/s41746-021-00408-5](https://doi.org/10.1038/s41746-021-00408-5)] [Medline: [33637822](https://pubmed.ncbi.nlm.nih.gov/33637822/)]
28. Bent B, Enache OM, Goldstein B, Kibbe W, Dunn JP. Reply: matters arising 'Investigating sources of inaccuracy in wearable optical heart rate sensors'. *NPJ Digit Med*. 2021 Feb 26;4(1):39. [doi: [10.1038/s41746-021-00409-4](https://doi.org/10.1038/s41746-021-00409-4)] [Medline: [33637842](https://pubmed.ncbi.nlm.nih.gov/33637842/)]
29. Lee J, Matsumura K, Yamakoshi K, Rolfe P, Tanaka S, Yamakoshi T. Comparison between red, green and blue light reflection photoplethysmography for heart rate monitoring during motion. *Annu Int Conf IEEE Eng Med Biol Soc*. 2013;2013:1724-1727. [doi: [10.1109/EMBC.2013.6609852](https://doi.org/10.1109/EMBC.2013.6609852)] [Medline: [24110039](https://pubmed.ncbi.nlm.nih.gov/24110039/)]
30. 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 May 24;7(2):3. [doi: [10.3390/jpm7020003](https://doi.org/10.3390/jpm7020003)] [Medline: [28538708](https://pubmed.ncbi.nlm.nih.gov/28538708/)]
31. Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial bias in pulse oximetry measurement. *N Engl J Med*. 2020 Dec 17;383(25):2477-2478. [doi: [10.1056/NEJMc2029240](https://doi.org/10.1056/NEJMc2029240)] [Medline: [33326721](https://pubmed.ncbi.nlm.nih.gov/33326721/)]
32. US patent for multi-channel photoplethysmography sensor patent (patent # 11,633,117). *Justia Patents*. Oct 3, 2019. URL: <https://patents.justia.com/patent/11633117> [Accessed 2023-07-09]
33. Li M, Kim YT. Design of a wireless sensor system with the algorithms of heart rate and agility index for athlete evaluation. *Sensors (Basel)*. 2017 Oct 17;17(10):2373. [doi: [10.3390/s17102373](https://doi.org/10.3390/s17102373)] [Medline: [29039763](https://pubmed.ncbi.nlm.nih.gov/29039763/)]
34. Weiler DT, Villajuan SO, Edkins L, Cleary S, Saleem JJ. Wearable heart rate monitor technology accuracy in research: a comparative study between PPG and ECG technology. *Proc Hum Factors Ergon Soc Annu Meet*. 2017;61(1):1292-1296. [doi: [10.1177/1541931213601804](https://doi.org/10.1177/1541931213601804)]
35. Wright SP, Brown TSH, Collier SR, Sandberg K. How consumer physical activity monitors could transform human physiology research. *Am J Physiol Regul Integr Comp Physiol*. 2017 Mar 1;312(3):R358-R367. [doi: [10.1152/ajpregu.00349.2016](https://doi.org/10.1152/ajpregu.00349.2016)] [Medline: [28052867](https://pubmed.ncbi.nlm.nih.gov/28052867/)]
36. Düking P, Fuss FK, Holmberg HC, Sperlich B. Recommendations for assessment of the reliability, sensitivity, and validity of data provided by wearable sensors designed for monitoring physical activity. *JMIR Mhealth Uhealth*. 2018 Apr 30;6(4):e102. [doi: [10.2196/mhealth.9341](https://doi.org/10.2196/mhealth.9341)] [Medline: [29712629](https://pubmed.ncbi.nlm.nih.gov/29712629/)]
37. Chan A, Chan D, Lee H, Ng CC, Yeo AHL. Reporting adherence, validity and physical activity measures of wearable activity trackers in medical research: a systematic review. *Int J Med Inform*. 2022 Apr;160:104696. [doi: [10.1016/j.ijmedinf.2022.104696](https://doi.org/10.1016/j.ijmedinf.2022.104696)] [Medline: [35121356](https://pubmed.ncbi.nlm.nih.gov/35121356/)]
38. Mitra R, McGough SF, Chakraborti T, Holmes C, Copping R, Hagenbuch N, et al. Learning from data with structured missingness. *Nat Mach Intell*. 2023;5(1):13-23. [doi: [10.1038/s42256-022-00596-z](https://doi.org/10.1038/s42256-022-00596-z)]
39. Yamamoto S, editor. *Human Interface and the Management of Information: Information and Interaction Design*, 15th International Conference, HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part I. Springer; 2013. [doi: [10.1007/978-3-642-39209-2](https://doi.org/10.1007/978-3-642-39209-2)]
40. Jiang Y, Spies C, Magin J, Bhosai SJ, Snyder L, Dunn J. Investigating the accuracy of blood oxygen saturation measurements in common consumer smartwatches. *PLOS Digit Health*. 2023 Jul 12;2(7):e0000296. [doi: [10.1371/journal.pdig.0000296](https://doi.org/10.1371/journal.pdig.0000296)] [Medline: [37437005](https://pubmed.ncbi.nlm.nih.gov/37437005/)]
41. Constantinou V, Felber AE, Chan JL. Applicability of consumer activity monitor data in marathon events: an exploratory study. *J Med Eng Technol*. 2017 Oct;41(7):534-540. [doi: [10.1080/03091902.2017.1366560](https://doi.org/10.1080/03091902.2017.1366560)] [Medline: [28954563](https://pubmed.ncbi.nlm.nih.gov/28954563/)]
42. Cleland I, Donnelly MP, Nugent CD, Hallberg J, Espinilla M, Garcia-Constantino M. Collection of a diverse, realistic and annotated dataset for wearable activity recognition. 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops. Institute of Electrical and Electronics Engineers; 2018;555-560. [doi: [10.1109/PERCOMW.2018.8480322](https://doi.org/10.1109/PERCOMW.2018.8480322)]
43. Di J, Demanuele C, Kettermann A, Karahanoglu FI, Cappelleri JC, Potter A, et al. Considerations to address missing data when deriving clinical trial endpoints from digital health technologies. *Contemp Clin Trials*. 2022 Feb;113:106661. [doi: [10.1016/j.cct.2021.106661](https://doi.org/10.1016/j.cct.2021.106661)] [Medline: [34954098](https://pubmed.ncbi.nlm.nih.gov/34954098/)]
44. Demanuele C, Lokker C, Jhaveri K, Georgiev P, Sezgin E, Geoghegan C, et al. Considerations for conducting bring your own "device" (BYOD) clinical studies. *Digit Biomark*. 2022 Jul 4;6(2):47-60. [doi: [10.1159/000525080](https://doi.org/10.1159/000525080)] [Medline: [35949223](https://pubmed.ncbi.nlm.nih.gov/35949223/)]

45. Liu S, Han J, Puyal EL, Kontaxis S, Sun S, Locatelli P, et al. Fitbeat: COVID-19 estimation based on wristband heart rate using a contrastive convolutional auto-encoder. *Pattern Recognit.* 2022 Mar;123:108403. [doi: [10.1016/j.patcog.2021.108403](https://doi.org/10.1016/j.patcog.2021.108403)] [Medline: [34720200](https://pubmed.ncbi.nlm.nih.gov/34720200/)]
46. Feng T, Narayanan S. Imputing missing data in large-scale multivariate biomedical wearable recordings using bidirectional recurrent neural networks with temporal activation regularization. *Annu Int Conf IEEE Eng Med Biol Soc.* 2019 Jul;2019:2529-2534. [doi: [10.1109/EMBC.2019.8856966](https://doi.org/10.1109/EMBC.2019.8856966)] [Medline: [31946412](https://pubmed.ncbi.nlm.nih.gov/31946412/)]
47. Lin S, Wu X, Martinez G, Chawla NV. Filling missing values on wearable-sensory time series data. *Proceedings of the 2020 SIAM International Conference on Data Mining (SDM)*. Society for Industrial and Applied Mathematics; 2020; 46-54. [doi: [10.1137/1.9781611976236](https://doi.org/10.1137/1.9781611976236)]
48. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized trial of a Fitbit-based physical activity intervention for women. *Am J Prev Med.* 2015 Sep;49(3):414-418. [doi: [10.1016/j.amepre.2015.01.020](https://doi.org/10.1016/j.amepre.2015.01.020)] [Medline: [26071863](https://pubmed.ncbi.nlm.nih.gov/26071863/)]
49. Standard score - understanding Z-scores and how to use them in calculations. Laerd Statistics. URL: <https://statistics.laerd.com/statistical-guides/standard-score.php> [Accessed 2023-07-09]
50. Shandhi MMH, Cho PJ, Roghanizad AR, Singh K, Wang W, Enache OM, et al. A method for intelligent allocation of diagnostic testing by leveraging data from commercial wearable devices: a case study on COVID-19. *NPJ Digit Med.* 2022 Sep 1;5(1):130. [doi: [10.1038/s41746-022-00672-z](https://doi.org/10.1038/s41746-022-00672-z)] [Medline: [36050372](https://pubmed.ncbi.nlm.nih.gov/36050372/)]
51. Dooley EE, Golaszewski NM, Bartholomew JB. Estimating accuracy at exercise intensities: a comparative study of self-monitoring heart rate and physical activity wearable devices. *JMIR Mhealth Uhealth.* 2017 Mar 16;5(3):e34. [doi: [10.2196/mhealth.7043](https://doi.org/10.2196/mhealth.7043)] [Medline: [28302596](https://pubmed.ncbi.nlm.nih.gov/28302596/)]
52. Viboud C, Santillana M. Fitbit-informed influenza forecasts. *Lancet Digit Health.* 2020 Feb;2(2):e54-e55. [doi: [10.1016/S2589-7500\(19\)30241-9](https://doi.org/10.1016/S2589-7500(19)30241-9)] [Medline: [33334559](https://pubmed.ncbi.nlm.nih.gov/33334559/)]
53. Master H, Annis J, Huang S, Beckman JA, Ratsimbazafy F, Marginean K, et al. Association of step counts over time with the risk of chronic disease in the All of Us Research Program. *Nat Med.* 2022 Nov;28(11):2301-2308. [doi: [10.1038/s41591-022-02012-w](https://doi.org/10.1038/s41591-022-02012-w)] [Medline: [36216933](https://pubmed.ncbi.nlm.nih.gov/36216933/)]
54. Louie A, Feiner JR, Bickler PE, Rhodes L, Bernstein M, Lucero J. Four types of pulse oximeters accurately detect hypoxia during low perfusion and motion. *Anesthesiology.* 2018 Mar;128(3):520-530. [doi: [10.1097/ALN.0000000000002002](https://doi.org/10.1097/ALN.0000000000002002)] [Medline: [29200008](https://pubmed.ncbi.nlm.nih.gov/29200008/)]
55. Stein PK, Kleiger RE. Insights from the study of heart rate variability. *Annu Rev Med.* 1999;50:249-261. [doi: [10.1146/annurev.med.50.1.249](https://doi.org/10.1146/annurev.med.50.1.249)] [Medline: [10073276](https://pubmed.ncbi.nlm.nih.gov/10073276/)]
56. Böttcher S, Vieluf S, Bruno E, Joseph B, Epitashvili N, Biondi A, et al. Data quality evaluation in wearable monitoring. *Sci Rep.* 2022 Dec 10;12(1):21412. [doi: [10.1038/s41598-022-25949-x](https://doi.org/10.1038/s41598-022-25949-x)] [Medline: [36496546](https://pubmed.ncbi.nlm.nih.gov/36496546/)]
57. Jain A, Nandakumar K, Ross A. Score normalization in multimodal biometric systems. *Pattern Recognit.* 2005 Dec;38(12):2270-2285. [doi: [10.1016/j.patcog.2005.01.012](https://doi.org/10.1016/j.patcog.2005.01.012)]
58. Burns M. Fitbit's latest activity trackers feature heart monitoring, smartwatch functions. *TechCrunch.* Oct 27, 2014. URL: <https://techcrunch.com/2014/10/27/fitbits-latest-activity-trackers-feature-heartheart-monitoring-smartwatch-functions/> [Accessed 2023-07-11]
59. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of wristband Fitbit models in assessing sleep systematic review and meta-analysis. *J Med Internet Res.* 2019 Nov 28;21(11):e16273. [doi: [10.2196/16273](https://doi.org/10.2196/16273)] [Medline: [31778122](https://pubmed.ncbi.nlm.nih.gov/31778122/)]
60. Husom EJ, Dautov R, Videsjorden AN, Gonidis F, Papatzelos S, Malamas N. Machine learning for fatigue detection using Fitbit fitness trackers. In: Capelli C, Verhagen E, Pazarat-Correia P, Vilas-Boas J, Cabri J, editors. *Proceedings of the 10th International Conference on Sport Sciences Research and Technology Support (icSPORTS 2022)*. SciTePress; 2022; 41-52. [doi: [10.5220/0011527500003321](https://doi.org/10.5220/0011527500003321)]
61. Fuller D, Anaraki JR, Simango B, Rayner M, Dorani F, Bozorgi A, et al. Predicting lying, sitting, walking and running using Apple Watch and Fitbit data. *BMJ Open Sport Exerc Med.* 2021 Apr 8;7(1):e001004. [doi: [10.1136/bmjsem-2020-001004](https://doi.org/10.1136/bmjsem-2020-001004)] [Medline: [33907628](https://pubmed.ncbi.nlm.nih.gov/33907628/)]
62. Gavhane A, Kokkula G, Pandya I, Devadkar K. Prediction of heart disease using machine learning. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA). Institute of Electrical and Electronics Engineers; 2018; 1275-1278. [doi: [10.1109/ICECA.2018.8474922](https://doi.org/10.1109/ICECA.2018.8474922)]
63. Faust L, Purta R, Hachen D, Striegel A, Poellabauer C, Lizardo O, et al. Exploring compliance: observations from a large scale Fitbit study. *SocialSens'17: Proceedings of the 2nd International Workshop on Social Sensing*. Association for Computing Machinery; 2017; 55-60. [doi: [10.1145/3055601.3055608](https://doi.org/10.1145/3055601.3055608)]

## Abbreviations

**AoURP:** *All of Us* Research Program

**BYOD:** Bring-Your-Own-Device

**DHT:** digital health technology

**MAR:** missing at random

**MCAR:** missing completely at random

**MNAR:** missing not at random

**PPG:** photoplethysmography

**SNR:** signal to noise ratio

**SpO<sub>2</sub>:** saturation of peripheral oxygen

*Edited by Lorraine Buis; peer-reviewed by Alessandra Angelucci, Hossein Motahari-Nezhad, Vincent Ukachukwu, Weizhuang Zhou; submitted 16.12.2022; final revised version received 14.08.2023; accepted 08.09.2023; published 03.11.2023*

*Please cite as:*

*Lederer L, Breton A, Jeong H, Master H, Roghanizad AR, Dunn J*

*The Importance of Data Quality Control in Using Fitbit Device Data From the All of Us Research Program*

*JMIR Mhealth Uhealth 2023;11:e45103*

*URL: <https://mhealth.jmir.org/2023/11/e45103>*

*doi: [10.2196/45103](https://doi.org/10.2196/45103)*

© Lauren Lederer, Amanda Breton, Hayoung Jeong, Hiral Master, Ali R Roghanizad, Jessilyn Dunn. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 03.11.2023. 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.