Demographic Imbalances Resulting from Bring-Your-Own-Device Study Design


Submitted to: JMIR mHealth and uHealth
on: April 23, 2021

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressively prohibit redistribution of this draft paper other than for review purposes.
## Table of Contents

Original Manuscript........................................................................................................... 4  
Supplementary Files............................................................................................................ 16  
Figures ............................................................................................................................................. 17  
  Figure 1 ........................................................................................................................................ 18  
  Figure 2 ........................................................................................................................................ 19  
  Figure 3 ........................................................................................................................................ 20  
Multimedia Appendixes.......................................................................................................... 21  
  Multimedia Appendix 1 .............................................................................................................. 22  
  Multimedia Appendix 2 .............................................................................................................. 22  
  Multimedia Appendix 3 .............................................................................................................. 22  
  Multimedia Appendix 4 .............................................................................................................. 22
Demographic Imbalances Resulting from Bring-Your-Own-Device Study Design

Peter Jaeho Cho¹ BA; Jaehan Jeremy Yi¹; Ethan Ho¹; Yen Hai Dinh¹ BSc; Mobashir Hasan Shandhi¹ BSc, MSc, PhD; Aneesha Patil¹; Leatrice Martin² BA, MPA, MBA; Geetika Singh¹ BEng, MSc; John Owens³; Brinnae Bent¹ BSc, MSc, PhD; Geoffrey Steven Ginsburg⁴ BSc, MD, PhD; Matthew Snuck⁵ BSc, MD; Veronica Palacios-Grandes³; Christopher Woods⁴,⁶,⁷ BA, MPA, MD; Ryan Shaw⁴,⁸ BSc, MSc, PhD; Jessilyn Pearl Dunn¹,⁹ BSc, MSc, PhD

¹Department of Biomedical Engineering Duke University Durham US
²Clinical and Translational Science Institute Duke University Durham US
³Duke Center for Applied Genomics & Precision Medicine Duke University Medical Center Durham US
⁴Department of Physical Medicine and Rehabilitation The Spine Program University of Michigan Ann Arbor US
⁵Division of Infectious Diseases Duke University Medical Center Durham US
⁶Duke VA Medical Center Durham US
⁷Duke University of Nursing Durham US
⁸Department of Biostatistics & Bioinformatics Duke University Durham US

Abstract

Digital health technologies such as smartphones and wearable devices promise to revolutionize disease prevention, detection, and treatment. Recently, there has been a surge of digital health studies where data is collected through a Bring-Your-Own-Device (BYOD) approach, in which participants who already own a specific technology may voluntarily sign up for the study and provide their digital health data. BYOD study design accelerates the collection of data on a larger number of participants than cohort design because researchers are not limited in the study population size based on the number of devices afforded by their budget. However, the BYOD study design may not support collecting data from a representative random sample of the target population where digital health technologies are intended to be deployed. This may result in biased study results and biased downstream technology development. In this viewpoint, we describe demographic imbalances discovered in existing BYOD studies, including our own, and we propose a Demographic Improvement Guideline to offset these imbalances.

DOI: https://doi.org/10.2196/preprints.29510
Demographic Imbalances Resulting from Bring-Your-Own-Device Study Design

Abstract

Digital health technologies such as smartphones and wearable devices promise to revolutionize disease prevention, detection, and treatment. Recently, there has been a surge of digital health studies where data is collected through a Bring-Your-Own-Device (BYOD) approach, in which participants who already own a specific technology may voluntarily sign up for the study and provide their digital health data. BYOD study design accelerates the collection of data on a larger number of participants than cohort design because researchers are not limited in the study population size based on the number of devices afforded by their budget or the number of people familiar with the technology. However, the BYOD study design may not support collecting data from a representative random sample of the target population where digital health technologies are intended to be deployed. This may result in biased study results and biased downstream technology development, as has occurred in other fields. In this viewpoint, we describe demographic imbalances discovered in existing BYOD studies, including our own, and we propose a Demographic Improvement Guideline to address these imbalances.

Keywords
bring-your-own-device; wearable device; mHealth

Introduction

Digital health tools, including mobile health (mHealth) and wearable devices, can provide researchers with high-frequency data that is more representative of a person’s health state during their daily life than what can be collected in a clinical setting [1], [2]. The enormous benefits of acquiring data outside of the clinic have led researchers to adopt new study designs to incorporate digital health data collection tools. Digital biomarkers are a type of biomarker that is developed from digitally collected data, such as wearable devices and smartphones, to evaluate functions and processes in the body and that can typically be recorded outside of a lab setting to provide continuous and more holistic information [3]. In particular, the “Bring-Your-Own Device” (BYOD) study design has become increasingly popular because it gives researchers the ability to collect large scale data at a low cost from participants who already own personal electronic devices such as smartphones and wearable devices. During the SARS-CoV-2 (COVID-19) pandemic, there has been a growing interest in using digital health data to track illness, either for COVID-19 detection or to support telemedicine [4]–[6]. To process all this data, artificial intelligence (AI) or machine learning (ML) algorithms are developed to detect health conditions by learning from previously collected data.

ML algorithms rely on data to train models and are susceptible to biases that result in poor predictions for segments of the population if the training data are not representative of the population where the model is intended to be used [7]. Therefore, one key aspect of ML is the data collection process, whereby researchers identify the target population and select a representative, random sample of the population from which to collect unbiased data [8]. BYOD studies are particularly susceptible to bias in the data collection process because the recruitment pool is limited to people who already own a device, and this population is generally not the only population where the tools are ultimately intended to be used in practice. In BYOD studies, a non-representative study population excludes key socioeconomic and physiologic circumstances that can co-vary with race.

and/or ethnicity. For example, disease prevalence and pathophysiology often vary by race and ethnicity (e.g., COVID-19 infection and mortality rates [9–11], manifestation of metabolic disease [12–14], cardiovascular disease [15], [16], and sleep irregularities [17], [18]) which can result in differences in how the newly developed technologies will perform. As a specific example, optical sensors to measure blood oxygen saturation may fail in people with more melanin and for people with sickle cell trait [19–21], both characteristics common in Black populations, for two completely separate reasons. To address such problems, representative digital health data are needed for building generalizable ML models that are as accurate under deployment as they are in the initial testing phase. Like other fields which have discovered that bias in data used to train models has led to biased models, we fear that digital health will face similar challenges if the bias inherent in BYOD studies is left unaddressed [22], [23]. In this viewpoint, we seek to raise awareness of demographic imbalances in several BYOD studies and propose a guideline based on our own BYOD case study to directly improve the demographic balance of BYOD digital health studies.

Demographic Imbalances in Existing BYOD Studies

BYOD is a term used to describe studies in which participants contribute data from self-owned personal electronic devices. We collected a sample of BYOD studies using multiple search terms on PubMed and Web of Science, including “bring your own device”, “(consumer) wearable device study”, “remote mobile study”, and “mHealth”, and performed manual review and filtering which generated 15 relevant studies. Although we note that this is not a systematic approach to compiling all published BYOD studies, and we acknowledge the potential to overlook studies that have successfully recruited a representative study population, here we demonstrate that many existing BYOD studies have gender and race imbalances when compared to the broader U.S. demographic. Of the 15 studies identified, four did not report any demographics on ethnicity or race, and none of the remaining 11 studies achieved participant demographic proportions representative of the overall US demographics (Supplementary Table 1).

One of the most preeminent BYOD studies is a sub-study of the US National Institute of Health’s All of Us research initiative [24]. In this sub-study, wearable device data was collected from participants owning Fitbit devices between 2008 and 2019 who consented to share their data. More than 80% of participants in the overarching All of Us study are from historically underrepresented groups in biomedical research. As a strong juxtaposition, 70% of the participants identify as white non-Hispanic in the All of Us BYOD sub-study, while only 4% and 3% identify as Black and Asian, respectively (Figure 1). The ethnicity distribution tells a similar story with over 90% identifying as not Hispanic or Latino and only 6% as Hispanic or Latino. Even in this diverse, large-scale study that had specifically targeted recruitment toward underrepresented groups, equitable demographic representation was limited by the BYOD study design.

Similar to other BYOD studies, we discovered demographic imbalances in our own CovIdentify study. The unique circumstances of the evolving COVID-19 pandemic led to a rapid launch of our study, where we aimed to develop an intelligent testing strategy using digital biomarkers extracted from personally owned commercial wearable devices under resource-constrained conditions (limited testing, rural areas, etc.). However, after our rapid study launch, we found that the communities most vulnerable to COVID-19 [9], [25] were not well-represented in our study.

To mitigate this demographic imbalance and to ensure the inclusion of participants from underserved communities, we developed a Demographic Improvement Guideline and correspondingly altered our recruitment process. Although we developed this guideline retrospectively after the discovery of demographic imbalances in our study, we are calling for future research to take proactive measures.
during the BYOD study design as well as responsive measures during participant recruitment and retention efforts. Many BYOD studies acknowledge demographic imbalance as a limitation, and we believe a concerted effort is needed to enact change to reduce bias in digital health data used in research.

**Demographic Improvement Guideline**

The goal of BYOD studies is to develop new device-based technologies and interventions to improve the health or well-being of populations. In order for these interventions to be fit-for-purpose, the research and development should include the populations where the technologies will ultimately be used or their exclusion should be well-justified and not introduce bias or harm. There may be cases where demographic imbalances (i.e., sampling bias) would not be problematic in a BYOD study, for example, if both 1) the manifestation of the disease as measured by the wearable device does not differ across race, ethnicity, or age, and 2) the technology works the same for all people across the entire population. In such cases, the researcher would not need to focus efforts and resources on obtaining a representative demographic distribution in the study population.

However, it is known that 1) most diseases do not manifest in the same way across different demographics, and 2) smart devices do not work equally well or are not equally accessible across all demographics. As such, we cannot conclude that a biased sampling strategy is generally an appropriate choice. There will certainly be exceptions where the study population is appropriately focused on a certain group (e.g., only including females in pregnancy-related studies), which can be achieved by applying inclusion and exclusion criteria for study participation. In this case, before a biased sampling strategy is chosen, the researcher is obligated to prove the null for differences in disease manifestation and device function between the biased sample and the population where the technology is intended to be used. We purport that this step is often an even larger barrier than designing an equitably sampled study population. Next, we describe two concrete examples of times when equitable sampling was not prioritized and resulted in incorrect study conclusions.

The first condition for considering sampling bias is when domain knowledge from a field gives a priori indication that disease prevalence and pathophysiology may vary by race and/or ethnicity, e.g., COVID-19 infection and mortality rates, metabolic disease, cardiovascular disease, sleep disturbances, etc.[9]–[18]. Coronary artery disease (CAD) is a historical example of the unintended and harmful effects that biased study populations can have on study conclusions: most early CAD studies consisted of homogenous male populations, and as a result, differences in symptoms for men and women were not discovered until follow-up studies included female populations [26], [27]. Often there is insufficient a priori knowledge of differences in disease manifestation across age, gender, race, ethnicity, etc. As a result, the effect of biased sampling on study conclusions is frequently unknown. Furthermore, data quality and sampling often vary across demographics, particularly in mobile and wearable device studies, which further complicates determining the most appropriate sampling method. For example, adults over aged 50 have lower use of wearable devices, which range from commercial wearable devices to more sophisticated health monitoring devices, than adults aged 18 and 34. Young, healthy, and more educated individuals are more likely to own wearable devices [28].

The second condition for considering sampling bias is when the technology used for data collection has not been validated systematically, i.e., it has either not reported demographic distributions of the test population or has uncertainty in its measurements. For example, a study published last year in the New England Journal of Medicine compared values of blood oxygen saturation in occult hypoxia
measured via pulse oximetry versus arterial oxygen saturation in arterial blood gas, the gold standard, to determine the validity of the pulse oximetry measurements [19]. The study obtained nearly 50,000 pairs of measurements from more than 8,000 white patients and 1,000 Black patients. The frequency of occult hypoxia detection via pulse oximetry was three times lower for Black patients than white patients. Given the prevalent use of pulse oximetry in medical decision making, the implication of sensors reporting varying results based on an individual’s skin tone is concerning. The study points to the need for manufacturers of optical heart rate and blood oxygen saturation sensors to disclose the demographic distribution of the populations that were used to test the sensors. Since the primary sensor on most of the wrist-worn commercial wearable devices (e.g., Apple Watch, Fitbit, Garmin, etc.) relies on a similar optical sensor, ensuring that the measurements are accurate for anyone who wears the device is crucial. Because the performance of technology across demographic characteristics is not systematically evaluated and published for most commercial wearable devices, the technology and models derived from it may fail to generalize across demographic characteristics [29], [30].

This guideline is relevant to studies in which sampling bias resulting in demographic imbalance could challenge the validity and generalizability of a BYOD study’s conclusions. The method can be implemented iteratively in the study design and execution process, and includes the following steps:

1. Identify the population(s) at risk of being omitted from the study for whom the technology may ultimately be used and determine if BYOD study design is appropriate for the research question
2. If the BYOD study design is insufficient for addressing issues associated with demographic imbalance, modify the study design using internal and external resources to improve dissemination of information and improve engagement with the target population(s)
3. Launch study and monitor study demographics in real time to adjust downstream efforts accordingly

**Identify Populations of Interest**

To identify the populations that are at risk of being omitted, a literature review can reveal a baseline expectation of demographic distributions from prior studies using similar devices and advertisement strategies. To support this, it is necessary for publication venues and funding agencies to require detailed demographic reporting of BYOD studies. In addition to the proactive measures mentioned above, researchers should conduct early and ongoing systematic analyses of their study demographics and iteratively adapt their strategies accordingly.

**Modify Study Design**

Capabilities to disseminate information, provide physical resources, and improve engagement with the target population can be assessed and developed internally and externally. Internal resources may involve organizations within the research institution that have experience recruiting underrepresented populations or have ties with the underrepresented groups. Another internal approach may be choosing the devices to be included in the study, considering whether these devices have widespread use in the underrepresented groups, and augmenting/utilizing capabilities of available devices (e.g., utilizing sensors of more widely available smartphones to capture physiology and activity characteristics instead of consumer wearables) [31]. External resources may include community groups that are representative of the target population, government and non-governmental organizations that work with the underrepresented communities, or donors who can donate devices that can be distributed to the target population. Other external resources may include clinician
referral, which can help establish trust with the target population and improve retention of study participants [32].

For national or international studies, researchers can partner with nation-wide organizations and/or launch social media advertising campaigns [33]–[35]. For regional studies, researchers can connect with institutional and local resources to learn about and connect with community groups. Researchers may recruit a liaison to aid in establishing partnerships with external organizations and establish a community advisory board to interact directly with community groups to contribute to the study design and advocate for participation in the study.

One existing challenge is a dearth of funding to support equitable digital health study design, including the purchase of personal electronic devices such as wearables and smartphones. Funders should be aware of this challenge and develop funding opportunities to support equitable digital health research. Given that acquiring funding is a well-known challenge, there are methods by which researchers can alter study design to reduce costs. This can include altering the study to follow a subset of the study population based on specific cases and controls. One method is the implementation of a randomized withdrawal design where participants who are inactive at the beginning of the study are excluded so that participants who are active are presented in the actual randomized study [36]. This could limit the number of participants researchers need to follow up on and provide a more robust dataset but may also skew demographics. Another challenge is to ensure that the distributed devices are used as intended, given that they require a certain level of digital literacy and need to be consistently worn and charged over time. As such, participant compensation amount and timing, and wear tracking should be included as considerations in the study design and should be supported by funding agencies [37].

**Launch and Reassess Study Population**

After study modifications, the study is ready for deployment. Depending on the funding received and study design changes, the resulting study may be a hybrid study where devices are provided for the target population. During this launch phase, researchers should regularly assess their demographic distributions to determine whether their strategy is effective. If the strategy is not effective, researchers can re-strategize with their internal and external community partners and identify aspects of the study that may be deterring participant engagement or recruitment. Finally, study teams should continue to engage with the study population after the end of the study to convey findings and future opportunities for participation in research.

**A Case Study on Practically Implementing the Demographic Improvement Guideline**

The following subsections provide details about the responsive implementation of the Demographic Improvement Guideline after the launch of the CovIdentify study (Institutional Review Board# 2020-0412). The study aimed to develop machine learning algorithms to detect COVID-19 and influenza from wearable device data, with a long-term vision of developing an intelligent diagnostic testing strategy using digital biomarkers extracted from personally owned commercial wearable devices under resource-constrained conditions (limited testing, rural areas, etc.). In April 2020 CovIdentify began enrolling participants. Following informed consent, participants contributed their wearable device data (e.g., Fitbit, Garmin, and Apple Watch) and reported daily symptoms for 12 months via a downloadable app, e-mail, or text message.
Identify
Following the rapid launch, exploratory data analysis revealed major differences between CovIdentify demographics and the demographics of COVID-19 positive cases and deaths in the U.S. [31], as well as overall U.S. demographics based on the 2020 U.S. Census. The communities hardest hit by the COVID-19 pandemic, including Black/African American and Hispanic/Latinx communities, also had the lowest representation (Figure 2) [9], [38], [39].

Modify
To address this imbalance, we designed the demographic improvement guideline mitigation strategy in partnership with Duke’s Clinical and Translational Science Institute (CTSI), an NIH-funded center that connects researchers with local community partners and improves the reach and efficacy of research [40]. We piloted this method in the Durham, NC community from June through October of 2020. The CTSI facilitated connections between our research team and local community and faith leaders, including groups to address COVID-19-related health disparities affecting the Latinx and African American communities. This engagement enabled our team to learn directly from community members, researchers, and healthcare professionals to improve our advertisement and recruitment strategies to 1) increase awareness about the study, and 2) distribute wearables devices donated to our study and purchased through research funding. To expand awareness of the study, our team presented to community groups on the advantages of continuous health monitoring, the uses of participant information in this study, and imbalances in our current study population that would make it difficult for our team to develop generalizable study findings. We also recruited a liaison to the Latinx community, translated our website to four additional languages, ran multilingual social media advertisements featuring diverse images and videos, and posted messages about the study on various social media platforms.

Launch and Reassess Study Population
To support the purchase of devices and social media advertising campaigns, we applied to nearly 30 funding opportunities from government, non-profit, and industry sources. We were awarded a Duke Bass Connections Fellowship, a North Carolina Biotechnology Institute grant, and a Duke MEDx/CTSI award that enabled us to purchase 65 wearable devices for distribution. We also received a donation of 300 additional devices from a previously completed study. We attended 12 community events, including food and medication distribution events for low-income members of the Durham community, and distributed 250 free wearable devices in a socially distant manner to ensure safety during the COVID-19 pandemic. It should be noted that this was a hybrid BYOD design because participants were still required to own a smart phone to connect their wearable to the study. We also worked with wearables companies to set up reduced device pricing with a direct link from our study’s main webpage to improve accessibility. Together, these efforts resulted in a 250% increase in the representation of Black/African American participants and a 49% increase in the Latinx/Hispanic population within 4 months of the implementation of the guideline (Figure 2).

Discussion
In this viewpoint, we discuss the need for digital health studies to sample from populations representative of the target population to ensure equitable performance of predictive or machine-learning models. We explore demographic imbalances in BYOD studies and propose a Demographic Improvement Guideline to address these imbalances with an implementation example from the CovIdentify study. We believe that future efforts and funding in this space can further improve equitable digital health study design and data collection. Further, we recommend that researchers...
carefully consider the financial incentives and personal motivations provided to the participant by the study to identify driving factors for participation and engagement.

By developing the Demographic Improvement Guideline, we aim to facilitate improvement of future BYOD study design through fostering relationships and trust with local community groups. These methods are not limited to BYOD studies only. We can translate these methods to non-BYOD studies as well. In addition, we believe that implementing existing community-based engagement methods, such as training recruiters and providing face-to-face screening, can improve both recruitment methods and adherence to studies [41]–[44]. We also emphasize the need for increased funding opportunities in this area to enable the development of equitable algorithms and models that are representative of all individuals.

One limitation of the Demographic Improvement Guideline was the condition under which it was developed. While the guideline is a proactive recommendation, the implementation within our case study was retrospective to our findings. Therefore, shifts in our study design, such as donating commercial wearable devices to underrepresented groups have resulted in a non-BYOD (hybrid) study. Furthermore, due to the evolving nature of the pandemic and the resulting rapid launch of the CovIdentify study, factors mentioned in the guideline may not be applicable to all study designs. Here, we intend to present the guideline as one unique potential solution of the many possible solutions to address demographic imbalances in BYOD studies. This experience underscores the importance of addressing potential demographic inequities prior to BYOD study launch.

Acknowledgements

This research was supported by the Bass Connections Fellowship, the Duke-Robert J. Margolis, MD, Center for Health Policy at Duke University, MEDx, NC Biotech awards, and Duke Clinical Translational Science Institute Community Engaged Research Initiative. We would like to acknowledge Rosa Gonzalez-Guarda, Leonor Corsino, Sabrena Mervin-Blake, Nicholas Eberlein, Jamie Roberts, and Kelly Keefe. Support was also provided, in part, by the Duke Clinical and Translational Science Institute by a US National Institutes of Health Clinical and Translational Science Award (grant number UL1TR002553).

Abbreviations

AI: artificial intelligence
BYOD: Bring-Your-Own-Device
mHealth: mobile health
ML: machine learning

Multimedia Appendix 1

Supplementary Table 1. Race and Ethnicity Breakdown from BYOD studies

Multimedia Appendix 2

Supplementary Table 2. Sex and Race/Ethnicity Reported by Various BYOD Studies, Census, and National Vital Statistics System

Multimedia Appendix 3

Supplementary Table 3. Wearable Device Ownership by Race or Ethnicity for CovIdentify Case Study
Supplementary Table 4. Wearable Device Ownership of iPhone (iOS) Users by Race or Ethnicity for CovIdentify Case Study

References:


Supplementary Files
Figures
Comparing demographic distributions from various BYOD studies, the national census, and the National Vital Statistics System. Studies with an asterisk did not separate ethnicity and race and therefore have percentages that sum up to be greater than 100. Other studies did not report a breakdown for all the ethnicity and race groups and therefore resulted in an aggregated percentage less than 100. (Further details are provided in Supplementary Table 2).
Visualization of Demographic Improvement Guideline: (1) Identify the population(s) at risk of being omitted from the study and for whom the technology may ultimately be used. (2) Modify study design based on internal and external resources to disseminate information and improve engagement with the target population(s). (3) Launch study and monitor study demographics in real-time and adjust downstream efforts accordingly. Researchers should reassess their study population demographics to ensure that target distributions are achieved and re-strategize accordingly. (The red dashed line in the bar charts are visual representations of the acceptable threshold for each subgroup population).
Percent Change Increase in CovIdentify Study Population Compared to June 21, 2020.
Multimedia Appendixes
Supplementary Table 1. Race and Ethnicity Breakdown from BYOD studies.
URL: http://asset.jmir.pub/assets/112d42f9e247eeb82bb1d205f19f3d4.docx

Supplementary Table 2. Sex and Race/Ethnicity Reported by Various BYOD Studies, Census, and National Vital Statistics System.
URL: http://asset.jmir.pub/assets/8b2f47892e26ccc13d072863b1f7895a.docx

Supplementary Table 3. Wearable Device Ownership by Race or Ethnicity for CovIdentify Case Study.
URL: http://asset.jmir.pub/assets/281fb72eb9f3d21945ecd8ac28eba24.docx

Supplementary Table 4. Wearable Device Ownership of iPhone (iOS) Users by Race or Ethnicity for CovIdentify Case Study.
URL: http://asset.jmir.pub/assets/29997a8a5230940f8b4b75a747fe6f.docx