Jou. Spo. Sci. Med. Volume No. 2, 6-10, 2019
Publisher: Swedish Scientific Publications, Gothenburg, Sweden

ISSN (2003-1521)

PHYSICAL ACTIVITY IN THE ERA OF MHEALTH BIG DATA: CONSIDERATIONS ON ACCURACY AND BIAS

Dr. Manolis Adamakis

School of Education, University College Cork. Ireland

***Corresponding Author Email:** *manosadam@phed.uoa.gr*

Abstract: Recently smartphones have been considered a powerful tool with which to study large-scale population health on a global scale. Landmark Nature publications highlight the global uptake of mobile technology and the staggering potential for big data to promote large observational studies with big datasets from electronic health records. However, these data have to be generated through unbiased, accurate and validated measurement techniques. This paper discusses in detail some of the inconsistencies in smartphone health apps' data collection and the issues arising from the increasing availability of low-quality health data. These threats to valid inference from observational studies using big data remain a great challenge to overcome and future research should focus on the development and validation of more accurate algorithms for health-related smartphone apps.

INTRODUCTION

Regular physical activity (PA), in addition to disease prevention benefits, provides a variety of extra benefits that help individuals sleep better, feel better, and perform daily tasks more easily (Physical Activity Guidelines Advisory Committee, 2018). These benefits can be achieved in a variety of ways, and walking interventions have shown clinically relevant improvements for many cardiovascular disease risk factors (Oja et al., 2018). In order to promote regular PA through walking, the 2018 Scientific Report stated that information and communication technologies, including wearable activity monitors and smartphone applications (apps), can be helpful (Physical Activity Guidelines Advisory Committee, 2018).

Recently, smartphones have been considered a powerful tool with which to study large-scale population health on a global scale, due to their vast adoption among adults in developed and developing countries (Anthes, 2016). Based on this approach, a new research published in Nature collected 68 million days of minute-by-minute step recordings from 717,527 individuals across 111 countries, in which step count was measured using the smartphone app Argus by Azumio (Althoff et al., 2017). And this is how the debate between big versus accurate data in PA begins!

(NON) VALIDATION OF THE APP USED

Initially, the researchers cited in their paper two studies to demonstrate that smartphones provide accurate step counts and reliable activity estimates in both laboratory and free-living settings (Case et al., 2015; Hekler et al., 2015). However, there are omissions in these citations.

The first cited paper concluded that Android smartphones can provide comparable PA estimates to an ActiGraph in both a laboratory-based and free-living context for estimating sedentary and moderate-to-vigorous-PA (Hekler et al., 2015). This paper did not validate any smartphone apps, only the raw values from three different Android smartphones [and not iOS ones that were used in Althoff et al.'s (2017) study], and further validated sedentary and moderate-to-vigorous-PA estimates, rather than step counts.

The second cited paper validated three iOS apps (Fitbit, Health Mate and Moves) and one Android app (Moves) in 14 healthy adults aged of 28.1±6.2 years, by using a protocol on a treadmill set at 3.0 mph for 500 and 1500 steps (Case et al., 2015). The results showed that, compared with direct observation, the relative difference in mean step count ranged from -6.7% to 6.2% for the smartphone apps, which cannot be considered an insignificant difference based on the study's limitations (young, healthy volunteers in a controlled setting of limited steps' number, with a convenience sample of a small number of apps, none of which was Argus Azumio). Also, in this study steps were measured only during treadmill walking in controlled laboratory settings, which may raise another bias towards apps' validity, because outside of the laboratory individuals walk with both slow and fast cadences, have more variable gait and complete far more short walks than long walks (Brodie et al., 2017; Brodie et al., 2016).

Based on the above considerations, the authors failed to take into account more researches on the validity of smartphone apps that provide evidence of low accuracy and unacceptable error percentage of similar apps in both laboratory and free-living settings (Boyce et al., 2012; Leong & Wong, 2017; Orr et al., 2015). On the other hand, a recent research found that the apps installed in an iPhone SE are accurate for step counting in different age groups and during various walking conditions (Höchsmann et al., 2018). However, all the above-mentioned studies did not include in the validation the Argus Azumio app, whose validity was examined by the latest study of Brodie et al. (2018). Their data revealed significant (15%-66%) undercounting by iPhones, as well as extraordinarily large (0% to 200% of steps taken) error ranges for both Android and iPhones running Argus Azumio app. These results are currently prohibitive for the use of pedometer apps, and more specifically Argus Azumio, in large-scale observational studies and step data collection due to validity discrepancies, and more research is definitely needed to further evaluate the accuracy of these apps.

MISSING DATA

Another consideration to take into account is the missing data of Althoff et al.'s (2017) study. Even though the authors stated that they performed complete-case analyses accompanied by sample correction, stratification, outlier, and balance testing to verify that their conclusions were robust to missing data, limitations arise by their definition of wear time of the activity-tracking smartphones, as well as the lack to capture activities when users do not carry the smartphone on their body. The authors mentioned as a research's limitation that their dataset may fail to capture time spent in activities where it is impractical to carry a phone (e.g., playing soccer or swimming) or steps are not a major component of the activity (e.g., bicycling), and, because users must carry their phone for steps to be recorded, there may exist systematic differences in wear time based on gender and age.

The above-mentioned limitations are extremely important for missing data. The definition used of wear time identifies the lack of the study to capture steps taken before the first recorded step (also mentioned as incorrect time zero) (Gill & Prasad, 2018) and after the last step every day, that is for a period of 10 hours per day. Also, the length of the PA that are impractical to carry a phone could fall between 40- to 90-minute timeframes. Using a non-wear definition that is too long increases the risk of overestimating sedentary time by misclassifying non-wear time (i.e., device removals) as sedentary (Clain et al., 2018). Furthermore, different non-wear definitions are needed for children, adolescents and adults (Clain et al., 2018; Evenson & Terry, 2009).

Another practical limitation of non-wear time is that smartphones, unlike wearable activity trackers worn on the wrist, are not carried all wearing time on users' bodies. This means that during the 14 hours of wear time definition, there might be timeslots that users did not have their smartphone attached to the body, resulting in further failure to capture steps and the increase of missing data. In fact, a recent study that examined women's smartphone location under several circumstances concluded that the most common location for the phone during the day, for all except active use, was in a bag or somewhere else off the body (Redmayne, 2017). In order to deal with the absence of clear documentation of device removals, researchers must make a reasonable estimation of duration and frequency of device removals to filter potential PA and sedentary misclassifications.

IPHONE BIASED SELECTION

The last major limitation of the Althoff et al.'s study (2017), which is somehow neglected, is that all retrospective data were collected only from Apple iPhone smartphone users between July 2013 and December 2014. Initially, it is possible that the choice of smartphones influenced the result as all apps access the phone's built-in accelerometer and its algorithms for step assessment and, at that time, most users owned the iPhone 4s and 5, both without a motion coprocessor (Höchsmann et al., 2018). Additionally, even though the authors mention that in high-income countries iPhones are used by a larger proportion of the population and do not belong to the wealthiest in the population, this is not the case. A recent report by the University of Chicago found that, across all years in their data, no individual brand was as informative about someone's education and as predictive of being high-income as owning an Apple iPhone in 2016 (Bertrand & Kamenica, 2018). Definitely iPhone is a luxury product that is usually priced higher than competing smartphones. While some low-end Android phones retail for as little as $100 or less, Apple recently raised the price of its lowest-end newer iPhone (XR) to $749 or more. Based on this information, the choice of including only data from iPhones, and not Android smartphones, can be considered biased towards the inclusion of data from higher-income individuals.

CONCLUSION

Clearly, large observational studies with big datasets from electronic health records have grown popular in recent years and are the next big thing in PA and exercise science. However, data have to be valid and reliable in order to extract meaningful and practically useful conclusions, without risking any type of bias. These threats to valid inference from observational studies using PA big data mentioned in this paper, coupled with more common threats (i.e. confounding, time-zero issues, multiplicity) remain a great challenge to overcome. Also, extracting research-grade data from electronic and mobile health records, which are built upon low quality and non-validated data derived from non-validated commercial smartphone apps, is a potential pitfall and has to be recognized as such.

Furthermore, non-wear definitions, can have a substantial effect on sedentary and PA estimates: this will result in sedentary behaviour being misclassified as non-wear time, hence underestimating or overestimating sedentary time in vulnerable population sub-groups across the world. Extreme caution should be taken when using low quality PA data to generate correlations with age, gender, body mass index groups and to predict obesity prevalence. Exercise and PA have many health benefits, however low-quality data risk reinforcing the often-inflated link between physical inactivity and obesity.

Future research should focus on the development, training, testing and validation of more accurate algorithms for large-scale research-based PA monitoring (smartphone apps included), before they are used in health research. Lastly, when using smartphone apps, more information should be included, such as type of smartphone and operating system, version of motion coprocessor, software app version, smartphone placement on the body and so on.

REFERENCES

Anthes E, 2016. Mental health: There's an app for that. Nature 532:20-23. DOI:10.1038/532020a.

Althoff T, Sosič R, Hicks JL, King AC, Delp SL, Leskovec J, 2018. Large-scale physical activity data reveal worldwide activity inequality. Nature 547(7663):336-339. DOI: 10.1038/nature23018.

Brodie MA, Coppens MJ, Ejupi A, Gschwind YJ, Annegarn J, Schoene D, et al., 2017. Comparison between clinical gait and daily-life gait assessments of fall risk in older people. Geriatr Gerontol Int. 17(11):2274-2282. DOI: 10.1111/ggi.12979.

Brodie MA, Coppens MJ, Lord SR, Lovell NH, Gschwind YJ, Redmond SJ, et al., 2016. Wearable pendant device monitoring using new wavelet-based methods shows daily life and laboratory gaits are different. Med Biol Eng Comput. 54(4):663-674. DOI: 10.1007/s11517-015-1357-9.

Boyce G, Padmasekara G, Blum M, 2012. Accuracy of mobile phone pedometer technology. J Mob Technol Med. 1(2):23-27. DOI: 10.7309/jmtm.13

Brodie MA, Pliner EM, Ho A, Li K, Chen Z, Gandevia SC, et al., 2018. Big data vs accurate data in health research: Large-scale physical activity monitoring, smartphones, wearable devices and risk of unconscious bias. Med Hypotheses 119:32-36. DOI: 10.1016/j.mehy.2018.07.015.

Bertrand M, Kamenica E, 2018. Coming apart? Cultural distances in the United States over time Available from https://faculty.chicagobooth.edu/emir.kamenica/documents/comingApartOnline.pdf [updated 2018 June 30; cited 2018 November 3]. DOI: 10.3386/w24771.

Cain KL, Bonilla E, Conway TL, Schipperijn J, 2018. Defining accelerometer nonwear time to maximize detection of sedentary time in youth. Pediatr Exerc Sci. 30(2):288-95. DOI: 10.1123/pes.2017-0132.

Case MA, Burwick HA, Volpp KG, Patel MS, 2015. Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA 313(6):625-626. DOI: 10.1001/jama.2014.17841.

Evenson K, Terry X, 2009. Assessment of differing definitions of accelerometer nonwear time. Res Q Exerc Sport 80(2):355-362. DOI: 10.1080/02701367.2009.10599570

Gill J, Prasad V, 2018. Improving observational studies in the era of big data. Lancet 392(10149):716-717. DOI: 10.1016/S0140-6736(18)31619-2

Hekler EB, Buman MP, Grieco L, Rosenberger M, Winter SJ, Haskell W, King AC, 2015. Validation of physical activity tracking via android smartphones compared to ActiGraph accelerometer: laboratory-based and free-living validation studies. JMIR MHealth UHealth 3(2): e36. DOI: 10.2196/mhealth.3505.

Höchsmann C, Knaier R, Eymann J, Hintermann J, Infanger D, Schmidt-Trucksäss A, 2018. Validity of activity trackers, smartphones, and phone applications to measure steps in various walking conditions. Scand J Med Sci Sports 28(7):1818-1827. DOI: 10.1111/sms.13074

Leong JY, Wong JE, 2017. Accuracy of three Android-based pedometer applications in laboratory and free-living settings. J Sports Sci. 35(1):14-21. DOI: 10.1080/02640414.2016.1154592.

Oja P, Kelly P, Murtagh EM, Murphy MH, Foster C, Titze S, 2018. Effects of frequency, intensity, duration and volume of walking interventions on CVD risk factors: a systematic review and meta-regression analysis of randomised controlled trials among inactive healthy adults. Br J Sports Med. 52:769-775. DOI: 10.1136/bjsports-2017-098558.

Orr K, Howe HS, Omran J, Smith KA, Palmateer TM, Ma AE, et al., 2015. Validity of smartphone pedometer applications. BMC Res Notes 8:733. DOI: 10.1186/s13104-015-1705.

Physical Activity Guidelines Advisory Committee, 2018. Physical Activity Guidelines Advisory Committee Scientific Report. Washington, DC: U.S. Department of Health and Human Services. Available from https://health.gov/paguidelines/second-edition/report/pdf/PAG_Advisory_Committee_Report.pdf [updated 2018 October 31; cited 2018 November 3].

Redmayne M, 2017. Where's your phone? A survey of where women aged 15-40 carry their smartphone and related risk perception: A survey and pilot study. PLoS ONE 12(1): e0167996. DOI: 10.1371/journal.pone.0167996.