RaY

Research at the University of York St John

For more information please contact RaY at ray@yorksj.ac.uk

1   **Should we use consumer-grade activity tracker data to understand population**

2   **physical activity patterns?**

3

4   Submitted 24 February 2021

5

6   Revised 28 June 2021

7

8   **ABSTRACT**

9

10   Researchers, practitioners, and public health organisations from around the world are

11   becoming increasingly interested in using data from wearable activity trackers, from

12   companies such as Fitbit Inc, Garmin Ltd, Xiaomi, and Apple Inc, to measure physical activity.

13   Indeed, large-scale, easily accessible, and autonomous data collection concerning physical

14   activity as well as other health behaviours is becoming ever more attractive. There are several

15   benefits of using wearable activity trackers to collect physical activity data, including the

16   ability to obtain big data, retrospectively as well as prospectively, to understand individual-

17   level physical activity patterns over time and in response to natural events. However, there

18   are challenges related to representativeness, data access, and proprietary algorithms that, at

19   present, limit the utility of this data in understanding population-level physical activity. In this

20   brief report we aim to highlight the benefits, as well as the limitations, of using existing data

21   from wearable activity trackers to understand large-scale physical activity patterns and

22   stimulate discussion amongst the scientific community on what the future holds with respect

23   to physical activity measurement and surveillance.

24

25   **KEY WORDS**

26   m-Health; quantified-self; big data; surveillance; wearables

INTRODUCTION

Physical activity (PA) and exercise have pronounced positive effects on physical, mental, and social health and wellbeing and, according to recent estimates, prevent 3.9 million premature deaths worldwide annually (Strain et al., 2020). Accordingly, global PA guidelines recommend all adults to undertake 150–300 min of moderate-intensity, or 75–150 min of vigorous-intensity PA, or some equivalent combination per week (Bull et al., 2020). Such guidelines rely on population level surveillance methods to regularly monitor PA indicators and inform public health policy, and the most common approach in this regard is to assess PA using self-report methods. Self-report remains an accepted method of large-scale data collection due to its cost effectiveness, unobtrusiveness and adaptability to different country contexts (Troiano et al., 2020). This is despite the accepted limitations of self-reporting with respect to accuracy, recall bias and social-desirability (Brenner & DeLamater, 2014; Prince et al., 2008). Advances in technology over the last two decades have, however, created new possibilities for PA measurement, not only for population-level surveillance but at an individual-level in terms of cohort studies, intervention research, and the evaluation of public health promotion programs. Research-grade devices such as wrist, hip, and thigh worn accelerometers have been used widely in such studies as they remove the biases associated with self-reporting and are able to provide a more granular quantification of PA. Nevertheless, research-grade accelerometers are costly to use at scale and cannot assess the domain or context in which PA takes place. Furthermore, accelerometers, as with self-report methods, only offer a 'snapshot in time' to infer usual PA behaviour (typically a 7-day period) meaning assessment of longer-term dynamic patterns of PA, particularly in response to natural events, is either not possible or not feasible. Over the last 15 years, the emergence of consumer-grade devices, such as smartphones and wearable activity trackers, has opened new doors in the

51 field of PA measurement. These devices gather rich activity data continuously in a free-living

52 setting thus providing large-scale and low-cost datasets that could advance our

53 understanding of PA patterns in a way that was never possible before. While this seems an

54 exciting prospect, as with any other PA measurement tool, the use of wearable activity tracker

55 data should be carefully considered before being used in PA research.

56

57 Comparing and contrasting all the available PA measurement methods available to

58 researchers, practitioners, and public health professionals is beyond the scope of this brief

59 report. Instead, we focus on the emerging opportunities offered by consumer grade devices,

60 including smartphones and wearable activity trackers, and how these may be utilised in the

61 fields of PA surveillance, cohort studies, intervention research, and evaluation of public health

62 promotion programs.

63

64 **Consumer-grade devices: too good to ignore?**

65 Technological advancements over the last decade, allied to the rapid proliferation of

66 smartphone use in both developed and developing regions globally (Deloitte, 2017), have

67 provided new possibilities in monitoring, understanding, and influencing human movement

68 at scale. Compared to traditional approaches, objective, real-world PA data sets with very

69 large sample sizes have become relatively low cost for researchers to collect or access.

70 Consequently, we are now beginning to witness the emergence of big data on PA in the

71 literature. For example, Althoff and colleagues (2017) recently used minute-by-minute step

72 count data, collected from the smartphone's inbuilt inertial unit, from over 700,000

73 individuals across 111 countries, to identify variability in PA levels across the world. With this

74   data they revealed city walkability as a factor associated with PA levels as well as associations

75   between PA inequalities and obesity.

76

77   The questions that could be addressed, and the new insights afforded, by large-scale PA data

78   from smartphones is an exciting prospect. However, concerns remain over the quality of the

79   PA data that can be obtained from smartphones, including the validity and reliability of step

80   detection, the restriction to only ambulatory activity, and their reliance on the individual

81   carrying the smartphone (Brodie et al., 2018). Wearable activity trackers, although currently

82   less prevalent than smartphones, are growing in popularity (Deloitte, 2017; Thompson, 2019)

83   and can address some of these pitfalls. Activity trackers have progressed beyond simple

84   pedometers and can now provide data on pulse rate, distance covered, moderate-to-vigorous

85   PA minutes, stair flights climbed, energy expenditure, and sleep. Unlike research grade

86   devices such as accelerometers, they also blend attractive design with invisible and effort-

87   free data capture. This combination often results in high adherence (in terms of daily wear

88   time) for extended periods of time. Additionally, while synchronisation of the activity tracker

89   to a smartphone application displays summary information to users, these summary data are

90   calculated from extensive intra-day data gathered at high frequency (e.g., 1 Hz), which are

91   also stored. High adherence alongside high frequency data capture means individuals

92   accumulate an extensive data resource that could be utilised to answer important PA

93   questions.

94

95   Big data analyses from smartphones and wearable activity trackers have, thus far, been cross-

96   sectional which limits our understanding of PA to a single point in time. However, longitudinal

97   data from smartphones and wearable activity tracker could also be analysed prospectively or

98   retrospectively, given its perpetual collection and long-term storage. Of note is the unique

99   opportunity offered by smartphones and wearable activity trackers to analyse data

100   retrospectively and in response to natural events without the need for foresight. This has had

101   obvious implications during the coronavirus pandemic when objective accelerometery was

102   not possible, or feasible, due to the speed and variability with which restrictions were

103   imposed across the world. As a result, the need for remote and scalable means to both

104   measure and support PA has become more prominent since the coronavirus pandemic.

105   Although the use of wearable activity trackers is not without inherent limitations (discussed

106   further below), we feel they are unique in their ability to be utilised in retrospective cohort

107   study designs (i.e., when the start of the study is only known after the event).

108

109   Interestingly, to date, there has been almost no independent large-scale reporting of existing

110   data from wearable activity trackers. This might be due to the complexity of large-scale data

111   access and processing from commercial wearables. A cross-sectional study of pulse rate data

112   from over 8 million Fitbit users was recently published by the Fitbit Research team (Natarajan

113   et al., 2020), reporting a positive relationship between heart rate variability and step count.

114   Regardless of the study findings, it seems that collecting, processing, and interpreting this

115   volume of data is possible, but requires an interdisciplinary team including data scientists,

116   database analysts, and cardiovascular and behavioural scientists, which has so far been

117   limited to large proprietary companies such as Fitbit. Furthermore, accessing this volume of

118   data is, so far, only feasible for companies such as Fitbit because they require all users to give

119   them permission to use the data collected by the device. For independent researchers it is

120   possible to request access to the data from the user directly. But this would be on an

121 individual basis, therefore, to amass a dataset of 8 million users would require 8 million

122 individual access requests. The issue of data access is discussed on more detail below.

123

124 **Balancing feasibility against validity, reliability, and sensitivity**

125 When choosing a PA data collection tool or methodology, researchers must balance validity,

126 reliability, and sensitivity of the approach with the costs and feasibility of its deployment in

127 the target population. Despite their known limitations (Brenner & DeLamater, 2014; Prince

128 et al., 2008), self-report methods remain an accepted means by which to collect large-scale

129 and population-level PA data, particularly where cost and sample size make accelerometery

130 an unfeasible approach. However, the volume and detail of information that can be obtained

131 from self-report surveys can be limited, preventing more nuanced analysis of PA patterns.

132 Device-based methods, such as accelerometery, provide a more valid and reliable estimate

133 of PA than do self-report measures (Dowd et al., 2018), but also have several limitations. Not

134 only are accelerometers costly, but their data also must be extracted from each device

135 individually, making them unfeasible for large-scale use. Data from wearable activity trackers

136 on the other hand should be considered feasible for large-scale use. Suitable activity trackers

137 are generally cheaper than accelerometers and their attractive design should translate into

138 greater wear time. Like accelerometers, they provide continuous data capture, but with the

139 additional advantage of these data being stored on a central server meaning data retrieval

140 and analysis can occur remotely and at scale. Thus, in principle, it is possible to analyse the

141 PA data of thousands of participants worldwide in a manner that is simply not possible with

142 current research grade accelerometers. Research has shown wearable activity trackers to

143 have high interdevice reliability for measuring steps, energy expenditure, and sleep (Evenson

144 et al., 2015), and despite ongoing concerns, the accuracy of wearable activity trackers also

145   continues to improve. In a recent systematic review of 67 studies, Fitbit devices were found

146   to provide a relatively accurate measure of free-living steps (within ± 10%, 50% of the time)

147   when compared to research-grade accelerometers (Feehan et al., 2018). Garmin activity

148   trackers are also reported to have good-to-excellent correlation coefficients and acceptable

149   (<10%) mean absolute percentage errors with respect to step count (Evenson & Spade, 2020).

150   While the accuracy of wearable activity trackers in measuring step count in free-living settings

151   is considered to be acceptable for normal walking pace (Evenson & Spade, 2020; Feehan et

152   al., 2018; Fokkema et al., 2017), they do not yet provide a valid measure of moderate-to-

153   vigorous PA (Redenius et al., 2019) or walking at very slow or very fast speeds (Fokkema et

154   al., 2017). However, considering this evidence is based on devices manufactured up to 2015,

155   refined algorithms over the past 5 years have likely further improved accuracy.

156

157   For intervention research the responsiveness, or sensitivity, to change in PA over time may

158   be a more important consideration than the validity of the tool. When examining the

159   effectiveness of an intervention in changing PA it is paramount that the measurement tool

160   employed is capable of detecting change. Research has shown the responsiveness indices for

161   self-report and device-based methods to vary not just by tool, or device, but by PA variable

162   measured. Reeves et al. (2010) compared the responsiveness of the Community Health

163   Activities Model Program for Seniors (CHAMPS) questionnaire, the Active Australia

164   Questionnaire (AAQ), and two items on exercise from the US National Health Interview Survey

165   (USNHIS), and reported responsiveness indices ranging from 0.15 (AAQ) to 0.27 (USNHIS) for

166   walking duration and 0.25 (AAQ) to 0.32 (CHAMPS) for moderate to vigorous intensity PA

167   duration per week. Swartz et al. (2014) compared two research-grade accelerometers, the

168   Actigraph GTX3 (ActiGraph LLC, Pensacola, Florida, USA) and the activPAL (PAL Technologies

169   Ltd, Glasgow, Scotland, UK) and found both to have comparable responsiveness to change

170   across a range of free living physical activity and sedentary behaviour variables (standardised

171   response mean values between 0.159 – 0.436). Donnachie and colleagues (2020) compared

172   a self-report PA measure (the International Physical Activity Questionnaire; IPAQ) and an

173   accelerometer (activPAL), and found both to have comparable and moderate standardised

174   response mean values of 0.54 (activPAL) and 0.59 (IPAQ) for total PA duration per day. There

175   appears to be no evidence on the responsiveness to change of wearable activity trackers. This

176   surprisingly under-researched topic warrants further attention by the PA research

177   community.

178

179   We have an array of options to measure elements of PA (such as duration, intensity, type,

180   domain, context, and quality), but no single tool can fully capture the complexity of PA

181   behaviour. Consumer-grade devices offer new opportunities for combining PA data collection

182   methods. For example, passive sensing of movement using a smartphone or wearable activity

183   tracker, combined with synchronised 'smart' self-report techniques, such as ecological

184   momentary assessment, could address many of the issues outlined previously. With further

185   evidence to support the validity, reliability and sensitivity of such methods, this approach

186   could provide powerful insights into PA patterns and help us better understand PA behaviour.

187

188   **The issue of data harmonisation**

189   Another issue researchers must consider when evaluating device-based PA measurement

190   tools is the harmonisation or comparability between devices from different manufacturers.

191   Data harmonisation is an essential step if researchers wish to conduct analyses on data

192   derived from different sources (Pearce et al., 2020). While all activity tracking devices gather

193   raw uni- or tri-axial accelerations, each manufacturer applies different algorithms to process

194   the data into its summary form thereby influencing the comparability of the data gathered.

195   Therefore, researchers who wish to use data from multiple devices/manufacturers to

196   increase sample representativeness and reach will need to consider data harmonisation using

197   statistical models derived from validation studies (Pearce et al., 2020). This could be

198   problematic when algorithms change, and validation data are no longer available.

199   Manufacturers of research-grade devices publish open source algorithms allowing

200   researchers to evaluate the impact of changes on measurement properties (Evenson et al.,

201   2015), however consumer-grade device manufacturers keep this information proprietary.

202   The use of different proprietary algorithms by each consumer-grade device manufacturer is

203   undoubtedly an issue for harmonisation too. In the longer term, this would be solved by

204   manufacturers making raw data counts available or at least allowing researchers to apply to

205   access this information. However, due to the proprietary nature of data processing, it is

206   unclear if raw data or only processed data are available. In the short-term however,

207   comparative validation between devices should enable statistical techniques that allow for

208   between device data pooling without compromising data quality. Finally, it is also worth

209   noting that there is a small but growing sector of 'hackable' wearables. These devices are

210   usually based on small form factor processing boards (e.g., small Raspberry Pi or Arduino

211   boards) which include tri-axial accelerometers, heart rate measurement, WIFI and Bluetooth.

212   These devices also support the remote storage of raw data signals, which would overcome

213   the limitations of unknown and proprietary algorithms. Although useful for research studies,

214   it seems unlikely that such devices will achieve the market penetration of larger

215   manufacturers.

216

The issue of representativeness

Given the widespread use of smartphones and the growing use of activity trackers, we should not ignore the possibility that in the near future wearable activity tracker data could also be used as a population PA surveillance tool. However, at present the primary challenge relating to such data is that it likely over-represents individuals who are more physically active and more proactive in setting and meeting activity goals relative to the general population who may not be tracking their activity level (Omura et al., 2017; Strain et al., 2019). Therefore, any cohort or surveillance research exclusively involving participants who own, and wear, activity trackers will introduce selection bias. The issue of representativeness is, however, not necessarily limited to wearable activity trackers. Selection bias might also occur in data derived from public calls to self-report PA or participate in cohort studies involving self-report or device-based measures of PA. Indeed, it has previously been suggested that selection bias is a significant issue in many cohort studies including those with objective assessments (Barreto et al., 2013; Folley et al., 2018; Stamatakis et al., 2021). Nevertheless, in such cohort and surveillance studies it is possible to use weighting to adjust for non-responders. This is not currently possible for data from wearable activity tracker and future research should focus on statistical approaches to estimate the population effect, and the effect in those with trackers, to help overcome this limitation.

While activity tracker sales and usage are increasing, the demographic reach appears, so far, to be constrained to young adults from more affluent backgrounds (Omura et al., 2017; Strain et al., 2019). Nevertheless, the cost of activity trackers has decreased significantly in recent years making them more affordable and accessible. This, combined with the increasing interest in activity trackers as behaviour change tools, may reduce this constrained

241   demographic reach over time. For example, recent initiatives to provide activity trackers as

242   part of health care (NHS England, 2019), health improvement (Yao et al., 2020) or health

243   insurance (Buckle et al., 2020) may serve to increase the breadth of the population using the

244   devices. The more initiatives and interventions utilising activity trackers, the more they could

245   be adopted by individuals from underserved populations, such as older adults and those with

246   lower incomes.

247

248   **The issue of data access**

249   Finally, it is worth noting some of the challenges inherent in accessing data from consumer-

250   grade activity trackers. To access data, researchers can establish an industry agreement with

251   a relevant company (e.g. Fitbit or Garmin) whose terms of service for collecting research data

252   are different from those governing commercial access (Hicks et al., 2019). While the specific

253   manufacturers control access to the data repositories, the data remains the property of the

254   individual user, therefore to access any data collected by the device, each individual user must

255   consent and agree to share the data. Managing thousands, and possibly tens of thousands, of

256   data sharing requests to individual users, and subsequently also having to manage their

257   authorisation and access details, brings its own logistic challenges. The most effective

258   approach is for participants to be directed to a project website which manages participant

259   information, consent, and authorisation requests via the specific manufacturers API.

260   Following successful authorisation, access codes for each user can be securely sent to the

261   research team for subsequent processing. It is worth noting that, even with successful

262   authorisation, there remain additional challenges. Authorisation is usually limited to a

263   maximum of 12 months before the user must re-approve access, which may limit follow-up

264   assessments in very large cohorts where direct contact with participants is limited. In

265   addition, it is users, not researchers, who define the scope of the data that can be accessed;

266   therefore, users may allow access to all or only some of their data (e.g. only pulse rate, or

267   step count, or some combination thereof), resulting in incomplete data sets. Additionally,

268   most devices allow users to manually add activity to account for any activity not passively

269   detected by the device (e.g. swimming or cycling). At present, it is unclear if such self-reported

270   estimates affect validity. Most databases separate device collected (passive) data from user

271   added (self-reported) data, meaning the research team have to make a decision regarding

272   which should be regarded as the 'canonical' source of users' PA.

273

274   Clearly, these challenges are not trivial, and future research teams will require multi-

275   disciplinary skills, including specialists in behavioural science, PA, data science, and software

276   and web development to successfully manage such projects. Nevertheless, if accessed and

277   interpreted appropriately, these data may allow understanding of PA behaviour at a scale

278   previously unimaginable. We are in the process of using this method at a national level to

279   understand the impact of coronavirus, but future research using this technique could examine

280   worldwide PA patterns, both prospectively and retrospectively, using multi-site and multi-

281   lingual research teams.

282

283   **CONCLUSIONS**

284   As with other device-based and self-report methods, we propose that consumer-grade

285   activity tracker data be considered with their limitations in mind rather than dismissed as a

286   flawed approach, particularly when the feasibility of large-scale accelerometery is prohibitive.

287   Given the rising popularity of wearable activity trackers, the volume of data collected, and

288   the possibilities in analysing data retrospectively, we believe data from wearable activity

289   trackers should be considered a viable PA measurement tool. To be clear, we are not

290   advocating that other tools, particularly self-report methods, should be consigned to history

291   or replaced by wearable activity tracker 'big data'. Quite the contrary, despite their limitations

292   self-report methods have provided critical insights into PA behaviour and are likely to remain

293   important in the future. Rather, our view is that if physical activity researchers, practitioners,

294   and public health professionals can use and interpret self-report data in light of their

295   limitations, the same should be possible for activity tracker data.

296

297

298   **LIST OF ABBREVIATIONS**

299   AAQ   Active Australia Questionnaire

300   API   Application Programming Interface

301   CHAMPS   Community Health Activities Model Program for Seniors

302   IPAQ   International Physical Activity Questionnaire

303   PA   Physical Activity

304   USNHIS   US National Health Interview Survey

305

306   **DECLARATIONS**

307   **Ethical Approval**

308   Not Applicable

309

310   **Consent for Publication**

311   Not Applicable

312

313   **Competing Interests**

314   None to declare.

315

319

320   **Authors' Contributions**

321   All authors contributed equally to the writing of this manuscript.

322

323   **Acknowledgements**

324   Not Applicable

325

326   **REFERENCES**

327   Althoff, T., Sosič, R., Hicks, J. L., King, A. C., Delp, S. L., & Leskovec, J. (2017). Large-scale

328   physical activity data reveal worldwide activity inequality. *Nature*, *547*(7663), 336–

329   339. https://doi.org/10.1038/nature23018

330   Barreto, P. de S., Ferrandez, A.-M., & Saliba-Serre, B. (2013). Are Older Adults Who

331   Volunteer to Participate in an Exercise Study Fitter and Healthier Than

332   Nonvolunteers? The Participation Bias of the Study Population. *Journal of Physical*

333   *Activity and Health*, *10*(3), 359–367. https://doi.org/10.1123/jpah.10.3.359

334   Brenner, P. S., & DeLamater, J. D. (2014). Social Desirability Bias in Self-reports of Physical

335   Activity: Is an Exercise Identity the Culprit? *Social Indicators Research*, *117*(2), 489–

336   504. https://doi.org/10.1007/s11205-013-0359-y

337 Brodie, M. A., Pliner, E. M., Ho, A., Li, K., Chen, Z., Gandevia, S. C., & Lord, S. R. (2018). Big

338   data vs accurate data in health research: Large-scale physical activity monitoring,

339   smartphones, wearable devices and risk of unconscious bias. *Medical Hypotheses*,

340   *119*, 32–36. https://doi.org/10.1016/j.mehy.2018.07.015

341 Buckle, J., Hayward, T., Singhal, N., & Desai, K. (2020). *The role of wearables in private*

342   *medical insurance* (p. 24). Milliman.

343 Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C.,

344   Chaput, J.-P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J.,

345   Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., … Willumsen, J.

346   F. (2020). World Health Organization 2020 guidelines on physical activity and

347   sedentary behaviour. *British Journal of Sports Medicine*, *54*(24), 1451–1462.

348   https://doi.org/10.1136/bjsports-2020-102955

349 Deloitte. (2017). *Global mobile consumer trends, 2nd edition*.

350   https://www2.deloitte.com/content/dam/Deloitte/us/Documents/technology-

351   media-telecommunications/us-global-mobile-consumer-survey-second-edition.pdf

352 Donnachie, C., Hunt, K., Mutrie, N., Gill, J. M. R., & Kelly, P. (2020). Responsiveness of

353   Device-Based and Self-Report Measures of Physical Activity to Detect Behavior

354   Change in Men Taking Part in the Football Fans in Training (FFIT) Program. *Journal*

355   *for the Measurement of Physical Behaviour*, *3*(1), 67–77.

356   https://doi.org/10.1123/jmpb.2019-0018

357 Dowd, K. P., Szeklicki, R., Minetto, M. A., Murphy, M. H., Polito, A., Ghigo, E., van der Ploeg,

358   H., Ekelund, U., Maciaszek, J., Stemplewski, R., Tomczak, M., & Donnelly, A. E. (2018).

359   A systematic literature review of reviews on techniques for physical activity

360   measurement in adults: A DEDIPAC study. *International Journal of Behavioral*

361   *Nutrition and Physical Activity*, *15*(1), 15. https://doi.org/10.1186/s12966-017-0636-

362   2

363 Evenson, K. R., Goto, M. M., & Furberg, R. D. (2015). Systematic review of the validity and

364   reliability of consumer-wearable activity trackers. *International Journal of Behavioral*

365   *Nutrition and Physical Activity*, *12*(1), 159. https://doi.org/10.1186/s12966-015-

366   0314-1

367 Evenson, K. R., & Spade, C. L. (2020). Review of Validity and Reliability of Garmin Activity

368   Trackers. *Journal for the Measurement of Physical Behaviour*, *3*(2), 170–185.

369   https://doi.org/10.1123/jmpb.2019-0035

370 Feehan, L., Geldman, J., Sayre, E., Park, C., Ezzat, A., Yoo, J., Hamilton, C., & Li, L. (2018).

371   Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of

372   Quantitative Data. *JMIR Mhealth Uhealth*, *6*(8), e10527.

373   https://doi.org/10.2196/10527

374 Fokkema, T., Kooiman, T. J. M., Krijnen, W. P., Van Der Schans, C. P., & De Groot, M. (2017).

375   Reliability and Validity of Ten Consumer Activity Trackers Depend on Walking Speed.

376   *Medicine & Science in Sports & Exercise*, *49*(4), 793–800.

377   https://doi.org/10.1249/MSS.0000000000001146

378 Folley, S., Zhou, A., & Hyppönen, E. (2018). Information bias in measures of self-reported

379   physical activity. *International Journal of Obesity*, *42*(12), 2062–2063.

380   https://doi.org/10.1038/s41366-018-0223-x

381 Hicks, J. L., Althoff, T., Sosič, R., Kuhar, P., Bostjancic, B., King, A. C., Leskovec, J., & Delp, S. L.

382   (2019). Best practices for analyzing large-scale health data from wearables and

383   smartphone apps. *Npj Digital Medicine*, *2*. https://doi.org/10.1038/s41746-019-

384   0121-1

385 Natarajan, A., Pantelopoulos, A., Emir-Farinas, H., & Natarajan, P. (2020). Heart rate

386   variability with photoplethysmography in 8 million individuals: A cross-sectional

387   study. *The Lancet Digital Health*, *2*(12), e650–e657. https://doi.org/10.1016/S2589-

388   7500(20)30246-6

389 NHS England. (2019). *Digital diabetes prevention rolled out as part of NHS Long Term Plan*.

390   https://www.england.nhs.uk/2019/08/digital-diabetes-prevention-rolled-out-as-

391   part-of-nhs-long-term-plan/

392 Omura, J. D., Carlson, S. A., Paul, P., Watson, K. B., & Fulton, J. E. (2017). National physical

393   activity surveillance: Users of wearable activity monitors as a potential data source.

394   *Preventive Medicine Reports*, *5*, 124–126.

395   https://doi.org/10.1016/j.pmedr.2016.10.014

396 Pearce, M., Bishop, T. R. P., Sharp, S., Westgate, K., Venables, M., Wareham, N. J., & Brage,

397   S. (2020). Network Harmonization of Physical Activity Variables Through Indirect

398   Validation. *Journal for the Measurement of Physical Behaviour*, *3*(1), 8–18.

399   https://doi.org/10.1123/jmpb.2019-0001

400 Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Gorber, S. C., & Tremblay, M. (2008). A

401   comparison of direct versus self-report measures for assessing physical activity in

402   adults: A systematic review. *International Journal of Behavioral Nutrition and*

403   *Physical Activity*, *5*(1), 56. https://doi.org/10.1186/1479-5868-5-56

404 Redenius, N., Youngwon, K., & Wonwoo, B. (2019). Concurrent validity of the Fitbit for

405   assessing sedentary behavior and moderate-to-vigorous physical activity. *BMC*

406   *Medical Research Methodology*, *19*, 29. https://doi.org/10.1186/s12874-019-0668-1

407 Reeves, M. M., Marshall, A. L., Owen, N., Winkler, E. A. H., & Eakin, E. G. (2010). Measuring

408   Physical Activity Change in Broad-Reach Intervention Trials. *Journal of Physical*

409   *Activity and Health*, *7*(2), 194–202. https://doi.org/10.1123/jpah.7.2.194

410 Stamatakis, E., Owen, K. B., Shepherd, L., Drayton, B., Hamer, M., & Bauman, A. E. (2021). Is

411   Cohort Representativeness Passé? Poststratified Associations of Lifestyle Risk Factors

412   with Mortality in the UK Biobank. *Epidemiology (Cambridge, Mass.)*, *32*(2), 179–188.

413   https://doi.org/10.1097/EDE.0000000000001316

414 Strain, T., Brage, S., Sharp, S. J., Richards, J., Tainio, M., Ding, D., Benichou, J., & Kelly, P.

415   (2020). Use of the prevented fraction for the population to determine deaths

416   averted by existing prevalence of physical activity: A descriptive study. *The Lancet*

417   *Global Health*, *8*(7), e920–e930. https://doi.org/10.1016/S2214-109X(20)30211-4

418 Strain, T., Wijndaele, K., & Brage, S. (2019). Physical Activity Surveillance Through

419   Smartphone Apps and Wearable Trackers: Examining the UK Potential for Nationally

420   Representative Sampling. *JMIR MHealth and UHealth*, *7*(1), e11898.

421   https://doi.org/10.2196/11898

422 Swartz, A. M., Rote, A. E., Cho, Y. I., Welch, W. A., & Strath, S. J. (2014). Responsiveness of

423   motion sensors to detect change in sedentary and physical activity behaviour. *British*

424   *Journal of Sports Medicine*, *48*(13), 1043–1047. https://doi.org/10.1136/bjsports-

425   2014-093520

426 Thompson, W. R. (2019). WORLDWIDE SURVEY OF FITNESS TRENDS FOR 2020. *ACSM's*

427   *Health & Fitness Journal*, *23*(6), 10–18.

428   https://doi.org/10.1249/FIT.0000000000000526

429   Troiano, R. P., Stamatakis, E., & Bull, F. C. (2020). How can global physical activity

430   surveillance adapt to evolving physical activity guidelines? Needs, challenges and

431   future directions. *British Journal of Sports Medicine*, *54*(24), 1468–1473.

432   Yao, J., Tan, C. S., Chen, C., Tan, J., Lim, N., & Müller-Riemenschneider, F. (2020). Bright

433   spots, physical activity investments that work: National Steps Challenge, Singapore: a

434   nationwide mHealth physical activity programme. *British Journal of Sports Medicine*,

435   *54*(17), 1047–1048. https://doi.org/10.1136/bjsports-2019-101662

436