<u>Viewpoint</u>

Charting a Course for Smartphones and Wearables to Transform Population Health Research

William G Dixon[1], PhD; Sabine N van der Veer[2], PhD; Syed Mustafa Ali[1], MPH; Lynn Laidlaw[1], BA; Richard J B Dobson[3], PhD; Cathie Sudlow[4], DPhil; Tim Chico[4], MD; Jacqueline A L MacArthur[4], PhD; Aiden Doherty[5], PhD

[1]Centre for Epidemiology Versus Arthritis, Manchester Academic Health Science Centre, University of Manchester, Manchester, United Kingdom

[2]Centre for Health Informatics, Manchester Academic Health Science Centre, Manchester, United Kingdom

[3]Department of Biostatistics and Health Informatics, King's College London, London, United Kingdom

[4]British Heart Foundation Data Science Centre, Health Data Research UK, London, United Kingdom

[5]Big Data Institute, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

Corresponding Author:
Aiden Doherty, PhD
Big Data Institute
Nuffield Department of Population Health
University of Oxford
Richard Doll Building, Old Road Campus
Oxford, OX3 7LF
United Kingdom
Phone: 44 01865 617794
Email: aiden.doherty@ndph.ox.ac.uk

Abstract

The use of data from smartphones and wearable devices has huge potential for population health research, given the high level of device ownership; the range of novel health-relevant data types available from consumer devices; and the frequency and duration with which data are, or could be, collected. Yet, the uptake and success of large-scale mobile health research in the last decade have not met this intensely promoted opportunity. We make the argument that digital person-generated health data are required and necessary to answer many top priority research questions, using illustrative examples taken from the James Lind Alliance Priority Setting Partnerships. We then summarize the findings from 2 UK initiatives that considered the challenges and possible solutions for what needs to be done and how such solutions can be implemented to realize the future opportunities of digital person-generated health data for clinically important population health research. Examples of important areas that must be addressed to advance the field include digital inequality and possible selection bias; easy access for researchers to the appropriate data collection tools, including how best to harmonize data items; analysis methodologies for time series data; patient and public involvement and engagement methods for optimizing recruitment, retention, and public trust; and methods for providing research participants with greater control over their data. There is also a major opportunity, provided through the linkage of digital person-generated health data to routinely collected data, to support novel population health research, bringing together clinician-reported and patient-reported measures. We recognize that well-conducted studies need a wide range of diverse challenges to be skillfully addressed in unison (eg, challenges regarding epidemiology, data science and biostatistics, psychometrics, behavioral and social science, software engineering, user interface design, information governance, data management, and patient and public involvement and engagement). Consequently, progress would be accelerated by the establishment of a new interdisciplinary community where all relevant and necessary skills are brought together to allow for excellence throughout the life cycle of a research study. This will require a partnership of diverse people, methods, and technologies. If done right, the synergy of such a partnership has the potential to transform many millions of people's lives for the better.

XSL•FO

RenderX

Introduction

Consumer digital devices provide a major opportunity to transform our understanding of the mechanisms, determinants, and consequences of diseases, including arthritis, dementia, and heart disease [1-3]. Most people in high- and low-income societies now own and regularly use consumer digital devices. Around 9 in 10 people own a smartphone in the United Kingdom [4], while one-fifth of US adults own wearable technologies, like smartwatches and fitness trackers [5]. This high level of device ownership means that many people could contribute to health research from the comfort of their home by offering small amounts of time to share data and help address clinical questions that matter to them.

Considering the wide range of the types of data available and the frequency and duration with which they are, or could be, collected, a significant step toward changes in how we conduct health research is within reach. Such data provide a much clearer picture of the daily rhythms of health, well-being, and disease, as well as the environment in which these occur. The touch screens, motion sensors, microphones, cameras, location sensors, and other technologies within these devices allow us to rethink how we measure things that are important and relevant to health research. Physical activity, for example, is an important risk factor for many diseases that is also negatively impacted when living with a condition, such as arthritis or stroke. Wrist-worn devices offer an opportunity to shift from the use of subjective questionnaires (eg, those asking "In a typical week, on how many days did you do 10 minutes or more of moderate physical activities like carrying light loads, cycling at normal pace?" [6]) to the continuous objective measurement of physical activity patterns [7]. One can easily see the differences in granularity, validity, reliability, and data collection burden between these two methods.

Smartphones and wearables have, however, not been used for research at scale beyond a handful of high-profile studies. Among the better examples of large-scale studies is the COVID Zoe study, which demonstrated that the mass collection of digital person-generated health data is both feasible and valuable, providing important early evidence for public health that anosmia is a key symptom of COVID-19 [8]. Further, a study on the Apple Watch (Apple Inc) proved that smartwatches can detect clinically meaningful heart rhythm patterns, like atrial fibrillation [9]. However, despite these studies illustrating digital devices' major potential for answering important research questions at speed and scale, this opportunity has yet to be fully exploited. Furthermore, no large-scale study has yet established the linkage of longitudinal wearable data to major clinical outcomes. Such linkage is important, as it brings together key ingredients for important population health research questions; for example, it would allow us to understand whether digital interventions for improving physical activity result in improvements in hard clinical outcomes, like a reduction in myocardial infarctions or a reduction in the number of people who develop diabetes.

In this viewpoint paper, we make the case that there remains a critical need to collect and link digital person-generated health data at scale by illustrating that such data are *required and necessary* to answer many vital research questions that matter to patients, clinicians, and policy makers, and we describe the requirements for collecting and linking such data. We then summarize what is needed to advance progress in this important and emerging field.

Opportunities

To illustrate the importance of and need for digital person-generated health data, we reviewed priority research questions for a number of common conditions. The James Lind Alliance is a UK initiative that brings together patients, carers, clinicians, and researchers in priority setting partnerships to identify and prioritize the top 10 most important unanswered questions or uncertainties for a given disease area [10]. Although there are other means for identifying research priorities, the James Lind Alliance follows a standardized process that is common across diseases, plus it brings together the views of different stakeholders. We reviewed the lists of the top 10 questions for the following six common disorder areas: arthritis, diabetes, chronic obstructive pulmonary disease, inflammatory bowel disease, stroke, and mental health. Each disorder area contained at least one question (often several questions) that would be optimally addressed with digital person-generated health data, with or without additional linked clinical data. Textbox 1 contains some of these questions, showing the need to collect data on physical and mental health symptoms and environmental factors, such as diet and exercise.

Textbox 1. Examples of as yet unanswered questions that digital person-generated health data would optimally address as part of the solution. These questions come from James Lind Alliance Priority Setting Partnerships exercises for 6 common disorder areas.

Example questions

1. "Is regular exercise and physical activity effective at reducing disease progression [in hip and knee osteoarthritis]?" [11]

2. "How do stress and anxiety influence the management of type 2 diabetes and does a positive mental wellbeing have an effect?" [12]

3. "What is the best way to tell the start of an exacerbation [of chronic obstructive pulmonary disease] from day-to day variation in symptoms?" [13]

4. "What role does diet have in the management of mildly active or inactive ulcerative colitis or Crohn's Disease to achieve normal daily activities and symptom control?" [14]

5. "How common are psychological problems and what impact do they have on the lives of people affected by stroke?" [15]

6. "How do certain mental health conditions (e.g. depression) affect how people engage with technology?" [16]

A recent review of what happens after a priority setting exercise [17] noted that addressing a priority topic requires researchers to design a dedicated study. The opportunity to collect data directly from patients at scale via digital devices could now help researchers and the public to address many top priority questions more easily and robustly. However, before we can harness this potential, we need to chart a course to overcome the barriers to conducting such large-scale population health research well.

We ran 2 parallel and complementary initiatives in 2021 to investigate possible solutions for successfully using smartphones and wearable data in population health research. The first was a British Heart Foundation Data Science Centre workshop, which focused on wearables for cardiovascular research [18]. The second was a roundtable event that considered the future of digital person-generated health data for UK health research and was hosted by the Centre for Epidemiology Versus Arthritis [19]. Both initiatives brought together multiple stakeholders, including patients, health care professionals, researchers, funders, policy makers, governance experts, and industry representatives, reflecting the importance of widespread consultation. The reports on these two initiatives [18,19] underline the major opportunities for population health research using digital person-generated health data. They both also recognize that countries such as the United Kingdom are in a particularly strong position, given the possibility of linking person-generated health data with routinely collected health data, such as those from the National Health Service, which has universal access to health care and cradle-to-grave health records. There is a pressing need for national-scale studies in which large numbers of smartphone and wearable users are invited to consent to the sharing of their device data to allow these data to be linked to their routinely collected health care information for research. These mobile data could enhance population health research if they could be integrated into emergent digital infrastructure to support health data research using routinely collected electronic health record data [20] and into large population cohort studies with genetic and deep phenotypic information, such as information from UK Biobank [21] and Our Future Health [22].

Requirements

Well-conducted population health research must consider potential challenges during study design and how to navigate them—a key area of discussion in both aforementioned initiatives. The recruitment of study participants based on device ownership would be skewed, as not everyone owns a device, introducing possible selection bias; for example, people who use wearable activity trackers are more active, younger, and more affluent than those who do not [23]. Study results must be useful and ideally generalizable to a wider population. It is vital that research does not worsen already existing health, social, and racial inequalities [24]. Researchers need to be able to set up studies easily and efficiently, use high-quality study designs, and have access to the right data collection tools that are both stable and flexible [25]. Data harmonization and interoperability are important challenges; the proliferation of devices with different proprietary software algorithms for determining measures like step count has resulted in researchers being unable to trust the outputs of consumer devices. Different devices provide different step counts for the same activity and vary greatly in accuracy [26]. There is a need to generate reproducible digital phenotypes from raw sensor data and low-level features (eg, measures of mobility or sleep), as well as the need to understand the environment and context in which data are generated, which may need more qualitative approaches. There is also a need for the harmonization of self-reported information, such as symptoms within and across diseases, especially as the number of people with multiple long-term conditions increases [27]. Public trust, engagement, and involvement are essential from the earliest point. These involve defining and prioritizing the most important, relevant, and feasible questions to address; designing the most appropriate studies; co-designing user-friendly devices and apps [28]; inviting people to join a study through the remote consent process [29]; and keeping them motivated to optimize ongoing engagement [30]. It is also important to enable participants to maintain and feel in control of where and how their data are used and to share the benefits and results of their contributions [31].

Proposed Solutions

Realizing the potential of patient-generated data in health care research requires a new interdisciplinary community to be established. Academics from diverse areas, such as epidemiology, software development, data science and biostatistics, psychometrics, and behavioral and social science, need to work with patients and health care professionals, alongside colleagues from industry who could contribute skills such as hardware and software engineering, user interface design, cybersecurity, and data management. Only by operating across disciplinary boundaries can we develop the foundations for future high-quality research and in turn support a wider group of interested, but so far relatively inexperienced, researchers. This can be done by defining and supporting best practices and providing access to the tools and methods needed to address the highest priority questions.

In countries such as the United Kingdom, a crucial requirement is understanding how we can best link digital person-generated health data with national health care data sets for research in a way that is understandable, feasible, and acceptable to participants and provides them with the option of retaining control over how and by whom their data are used. This linkage should use existing national infrastructure, such as trustworthy research environments [32]. In addition to the technical infrastructure, such linkage also requires the development and evaluation of a range of approaches and methods, such as determining how best to recruit and remotely consent participants, securely storing and linking the different data types across different geographical areas, ensuring the validity and harmonization of data across devices, engaging participants through feedback, and providing them with control to ensure that we maintain trust. In this context, the prominent involvement of patients and the public is the most vital factor as we proceed; we can only undertake large-scale population health research if people are willing to participate in and consent

to the collection and sharing of their data repeatedly over time. Before asking this of patients and the public, we must ensure that research is done in a way that is acceptable, is valuable, and has meaning and relevance to them [33,34].

We believe that the time is right to create the partnerships, platforms, tools, and methods that will allow us to collect data directly from patients via digital devices; securely link these data to their routinely collected health care data in a trustworthy way; and answer many more questions that matter to patients, health care professionals, policy makers, and the wider public.

Acknowledgments

Authors' Contributions

WGD prepared the first draft and shared it with coauthors for feedback. WGD, SNvdV, and SMA reviewed the James Lind Alliance lists of top 10 questions. All coauthors provided feedback, and WGD prepared the final draft version. All coauthors reviewed and approved the final version before submission.

Conflicts of Interest

The authors declare no competing nonfinancial interests but declare the following competing financial interests: (1) WGD has received consultancy fees from Google, unrelated to this work, and (2) AD is the recipient of research funding from Novo Nordisk and Swiss Re.

References

1. Deering MJ. Issue brief: Patient-generated health data and health IT. The Office of the National Coordinator for Health Information Technology. 2013 Dec 20. URL: https://www.healthit.gov/sites/default/files/pghd_brief_final122013.pdf [accessed 2023-01-03]

2. The Academy of Medical Sciences. Our data-driven future in healthcare: People and partnerships at the heart of health related technologies. The Academy of Medical Sciences. 2018. URL: https://acmedsci.ac.uk/file-download/74634438 [accessed 2023-01-03]

3. The Economist. Wearable technology promises to revolutionise health care. The Economist. 2022 May 05. URL: https://www.economist.com/leaders/2022/05/05/wearable-technology-promises-to-revolutionise-health-care [accessed 2023-01-03]

4. Taylor P. Share of adults who own a smartphone in the United Kingdom (UK) in 2008 and 2019 to 2021, by demographics. Statista. 2022 May 23. URL: https://www.statista.com/statistics/956297/ownership-of-smartphones-uk/ [accessed 2022-07-27]

5. Vogels EA. About one-in-five Americans use a smart watch or fitness tracker. Pew Research Centre. 2020 Jan 09. URL: https://www.pewresearch.org/fact-tank/2020/01/09/about-one-in-five-americans-use-a-smart-watch-or-fitness-tracker/ [accessed 2022-11-23]

6. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]

7. Doherty A, Jackson D, Hammerla N, Plötz T, Olivier P, Granat MH, et al. Large scale population assessment of physical activity using wrist worn accelerometers: The UK Biobank study. PLoS One 2017 Feb 01;12(2):e0169649 [FREE Full text] [doi: 10.1371/journal.pone.0169649] [Medline: 28146576]

8. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med 2020 Jul;26(7):1037-1040 [FREE Full text] [doi: 10.1038/s41591-020-0916-2] [Medline: 32393804]

9. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, Apple Heart Study Investigators. Large-scale assessment of a smartwatch to identify atrial fibrillation. N Engl J Med 2019 Nov 14;381(20):1909-1917 [FREE Full text] [doi: 10.1056/NEJMoa1901183] [Medline: 31722151]

10. The James Lind Alliance. James Lind Alliance Priority Setting Partnerships. URL: https://www.jla.nihr.ac.uk/ [accessed 2022-11-23]

11. James Lind Alliance Priority Setting Partnerships. Hip and knee replacement for osteoarthritis top 10. James Lind Alliance Priority Setting Partnerships. URL: https://www.jla.nihr.ac.uk/priority-setting-partnerships/hip-and-knee-replacement-for-osteoarthritis/top-10-priorities/ [accessed 2023-01-03]

12. Finer S, Robb P, Cowan K, Daly A, Robertson E, Farmer A. Top ten research priorities for type 2 diabetes: results from the Diabetes UK-James Lind Alliance Priority Setting Partnership. Lancet Diabetes Endocrinol 2017 Dec;5(12):935-936. [doi: 10.1016/S2213-8587(17)30324-8] [Medline: 29092776]

13. Alqahtani JS, Aquilina J, Bafadhel M, Bolton CE, Burgoyne T, Holmes S, et al. Research priorities for exacerbations of COPD. Lancet Respir Med 2021 Aug;9(8):824-826. [doi: 10.1016/S2213-2600(21)00227-7] [Medline: 34000234]

14. Hart AL, Lomer M, Verjee A, Kemp K, Faiz O, Daly A, et al. What are the top 10 research questions in the treatment of inflammatory bowel disease? A priority setting partnership with the James Lind Alliance. J Crohns Colitis 2017 Feb;11(2):204-211 [FREE Full text] [doi: 10.1093/ecco-jcc/jjw144] [Medline: 27506537]

15. James Lind Alliance Priority Setting Partnerships, Stroke Association. Priorities in stroke rehabilitation and long-term care. James Lind Alliance Priority Setting Partnerships. URL: https://www.jla.nihr.ac.uk/priority-setting-partnerships/Stroke/downloads/Stroke-PSP-Top-Ten-list-of-results-rehab.pdf [accessed 2023-01-03]

16. James Lind Alliance Priority Setting Partnerships. Digital technology for mental health top 10. James Lind Alliance Priority Setting Partnerships. URL: https://www.jla.nihr.ac.uk/priority-setting-partnerships/digital-technology-for-mental-health/top-10-priorities.htm [accessed 2023-01-03]

17. Staley K, Crowe S, Crocker JC, Madden M, Greenhalgh T. What happens after James Lind Alliance Priority Setting Partnerships? A qualitative study of contexts, processes and impacts. Res Involv Engagem 2020 Jul 11;6:41 [FREE Full text] [doi: 10.1186/s40900-020-00210-9] [Medline: 32670611]

18. British Heart Foundation Data Science Centre. Workshop report – How can consumer wearables transform our understanding of cardiovascular disease? Zenodo. 2022 Jan. URL: https://zenodo.org/record/5827260 [accessed 2023-01-03]

19. Dixon WG, McBeth J, Mackey E, Peek N, van der Veer S, Ainsworth J. The future of patient-generated data for UK health research. The University of Manchester. 2022. URL: https://figshare.manchester.ac.uk/articles/report/The_Future_of_Patient-Generated_Data_for_UK_Health_Research/20066132/2 [accessed 2022-11-23]

20. Department of Health and Social Care, NHS England. A plan for digital health and social care. GOV.UK. 2022. URL: https://www.gov.uk/government/publications/a-plan-for-digital-health-and-social-care/a-plan-for-digital-health-and-social-care [accessed 2022-07-27]

21. UK Biobank. UK Biobank. URL: https://www.ukbiobank.ac.uk/ [accessed 2022-11-23]

22. Our Future Health. Our Future Health. URL: https://ourfuturehealth.org.uk/ [accessed 2022-11-23]

23. Strain T, Wijndaele K, Brage S. Physical activity surveillance through smartphone apps and wearable trackers: Examining the UK potential for nationally representative sampling. JMIR Mhealth Uhealth 2019 Jan 29;7(1):e11898 [FREE Full text] [doi: 10.2196/11898] [Medline: 30694198]

24. Vicary P, Mathie E. Access to research findings: Digital exclusion and health inequalities. NIHR Applied Research Collaboration East of England. 2022 Jun 15. URL: https://arc-eoe.nihr.ac.uk/news-insights/news-latest/access-research-findings-digital-exclusion-and-health-inequalities [accessed 2023-01-03]

25. Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeeck D, Hyve, RADAR-CNS Consortium. RADAR-Base: Open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. JMIR Mhealth Uhealth 2019 Aug 01;7(8):e11734 [FREE Full text] [doi: 10.2196/11734] [Medline: 31373275]

26. Germini F, Noronha N, Debono VB, Philip BA, Pete D, Navarro T, et al. Accuracy and acceptability of wrist-wearable activity-tracking devices: Systematic review of the literature. J Med Internet Res 2022 Jan 21;24(1):e30791 [FREE Full text] [doi: 10.2196/30791] [Medline: 35060915]

27. The Academy of Medical Sciences. Cross-funder multimorbidity research framework. The Academy of Medical Sciences. 2019 Nov. URL: https://acmedsci.ac.uk/file-download/49628715 [accessed 2023-01-03]

28. Polhemus AM, Novák J, Ferrao J, Simblett S, Radaelli M, Locatelli P, et al. Human-centered design strategies for device selection in mHealth programs: Development of a novel framework and case study. JMIR Mhealth Uhealth 2020 May 07;8(5):e16043 [FREE Full text] [doi: 10.2196/16043] [Medline: 32379055]

29. Sage Bionetworks. The elements of informed consent: A toolkit. Sage Bionetworks. 2020 Jan. URL: https://sagebionetworks.org/wp-content/uploads/2020/01/SageBio_EIC-Toolkit_V3_21Jan20_final.pdf [accessed 2023-01-03]

30. White KM, Williamson C, Bergou N, Oetzmann C, de Angel V, Matcham F, et al. A systematic review of engagement reporting in remote measurement studies for health symptom tracking. NPJ Digit Med 2022 Jun 29;5(1):82 [FREE Full text] [doi: 10.1038/s41746-022-00624-7] [Medline: 35768544]

31. NHS Health Research Authority. Make it public: transparency and openness in health and social care research. NHS Health Research Authority. URL: https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/research-transparency/make-it-public-transparency-and-openness-health-and-social-care-research/ [accessed 2023-01-03]

32. Sudlow C. The what and why of trusted research environments. Understanding Patient Data. 2021 Sep 08. URL: https://understandingpatientdata.org.uk/news/what-why-trusted-research-environments [accessed 2023-01-03]

33. Laidlaw L. Lynn Laidlaw: How do we bring data and its outputs to life? BMJ Opinion. 2021 Apr 08. URL: https://blogs.bmj.com/bmj/2021/04/08/lynn-laidlaw-how-do-we-bring-data-and-its-outputs-to-life/ [accessed 2022-07-27]

34. Tully MP, Bozentko K, Clement S, Hunn A, Hassan L, Norris R, et al. Investigating the extent to which patients should control access to patient records for research: A deliberative process using citizens' juries. J Med Internet Res 2018 Mar 28;20(3):e112 [FREE Full text] [doi: 10.2196/jmir.7763] [Medline: 29592847]

XSL•FO
RenderX