BJPsych | The British Journal of Psychiatry (2026)
1–3. doi: 10.1192/bjp.2026.10878

Guest Editorial

Data donation as an engine for passive sensing research in psychiatry

Philipp Herzog, Sylvie Blairy and Philip S. Santangelo

Summary

Passive sensing research in psychiatry often yields only small effects due to reliability and heterogeneity issues caused by underpowered, unrepresentative samples. We argue that patient-driven data donation, leveraging widespread health self-tracking, offers an innovative complementary method for collecting larger, more diverse data-sets and improving the validity of inferences from passive measures.

Keywords

Big data; ethics; medical technology; precision medicine; sampling methods.

Copyright and usage

© The Author(s), 2026. Published by Cambridge University Press on behalf of Royal College of Psychiatrists.

Technological advancements such as wearable devices (e.g. smartphones, smartwatches) now enable the collection of large volumes of personal physiological and behavioural data (e.g. geolocation from the global positioning system (GPS), heart rate variability using photo-plethysmography), providing a valuable resource for clinical science. These technologies collect continuously and unobtrusively objective data streams in a person's naturalistic environment to reflect activities such as movement, sleep patterns and social interactions. Clinical researchers have tried to harness this potential systematically over the last decade using passive smartphone and wearable sensing approaches, aiming to identify contextually relevant markers associated with the development and persistence of psychopathology.[1] Integrating passive sensing data promises to advance 'digital phenotyping' in mental health research, paving the way towards precision mental healthcare.

Passive sensing is an ambulatory assessment method that leverages sensors in digital devices to collect intensive longitudinal data about human behaviour in situ without requiring active user input. Traditional methodological approaches often rely on time-limited loaner devices to answer research questions in prospective observational studies. However, research using passive sensing to detect and predict clinical targets (e.g. mental disorders such as post-traumatic stress disorder (PTSD)) has shown only small effects.[2] Two important challenges may explain this finding: reliability and heterogeneity. Although an advantage of passive sensing is to collect vast amounts of within-individual data, most studies have included only a small number of patients. Due to limited sample sizes, these studies underrepresent marginalised groups, which can lead to biased models and low generalisability across subgroups. Consequently, the research field has called to gather data from larger, more diverse samples to increase the validity of inferences based on passive measures. That said, the rapid uptake of wearable technologies for health self-tracking in the general public generates large amounts of digital data that have not yet been leveraged.

Using personal devices, a potential complement to traditional passive sensing research can be *data donations* – a new consent-based data collection method grounded in citizen science, in which participants (or patients) are asked to download personal digital data packages gathered through their own monitoring (e.g. wearables) and other tracking devices (e.g. smartphones) over months and years, and voluntarily donate them to academic research. While maintaining trust, transparency and privacy, this approach allows the efficient secondary use of real-world, hitherto unobtainable granular data at scale in an ethically justifiable way (i.e. without direct data access), potentially advancing passive sensing research in psychiatry.

Since this approach relies on informed consent and active participant cooperation, the *acceptability* and *feasibility* of data donations need to be determined before implementation. In the general UK population, the overall willingness to share commercial data for health research with the government and academic research institutions was estimated at 60%, depending on data type (32% for smartphone data, 31.8% wearable device data and 30.4% social media data), indicating *acceptability*.[3] Among German health self-trackers, 41% would donate their self-collected personal health data to research.[4] Self-reported reasons for unwillingness related primarily to expected technical issues, privacy concerns and perceived usefulness.[5] Because of a potential discrepancy between willingness to share and actual donation behaviour, *feasibility* is a second consideration. In a youth study combining data donation with ecological momentary assessment (EMA), 78.9% of UK-based participants donated Instagram data and 65.8% donated TikTok data.[6] Among older Dutch adults, 29% reported willingness to donate physical activity data, 17% started data donation and 12% successfully provided data,[5] suggesting an generation-related age effect.

To date, only Fitzgerald et al have applied data donation in mental health research, requesting screenshots from adolescents' social media accounts.[7] However, this study did not use passive sensing data. In fact, donated personal physiological and behavioural data can provide additional insights into a wide range of promising research questions. For example, an interesting question is which passively collected markers are most robustly associated with psychopathology in prospectively donated data, and whether donated retrospective longitudinal data can identify changes preceding symptom worsening, relapse or hospitalisation. Another methodological question is to compare directly sensor data gathered by smartphones and wearables, e.g. by investigating the relative benefit of each source for the same variables. Methodologically, setting up a data donation study for passive sensing research involves several challenges. These include selecting optimal wearables for a given data donation study and using a privacy-preserving donation platform.[8] Importantly, such studies should always be combined with clinical data (e.g. from electronic health records) and/or psychological outcome measures (e.g. symptom reports via EMA). Recommendations[6] and best practices[9] for digital data donation studies have recently been developed.

Despite enthusiasm for data donation, some inherent limitations in measurement quality should be considered when designing studies, including validity and reliability issues. For example, data can

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 21:51:14, subject to the Cambridge Core terms of use.

Table 1 Essential considerations for designing passive sensing data donation studies in psychiatry

Stage	Essential question	Recommended minimum requirements
1. Define the purpose and clinical target	What construct, dimension or event should the donated data explain, monitor or predict, for whom and over what period?	Pre-specify the clinical target using a transdiagnostic, dimensional framework (e.g. HiTOP) where appropriate; define intended use, population, time scale, inference and within-person versus between-person estimand.
2. Choose the design and data streams	Which design and data streams can address the clinical target with acceptable feasibility, burden and risk?	Justify retrospective, prospective or repeated donation and observational, experimental or quasi-experimental design; compare candidate data streams by construct relevance, exportability, resolution, burden and re-identification risk; select donation-suitable data streams.
3. Establish provenance and technical validity	How were the variables generated, transformed and exported, and are the relevant records stable and interpretable?	Record device and sensor specifications, operating system, software/firmware, algorithm and export versions, where available; distinguish raw, derived and proprietary variables; document configuration settings, preprocessing steps and schema changes.
4. Design the data sampling and donation procedure	What data and time period will participants be asked to donate? Who can and will donate, and can intended participants complete the procedure reliably?	Pre-define eligibility, device or account criteria and recruitment; develop/adapt/select a suitable data donation platform; standardise the device-specific donation instructions; provide accessible participant briefing materials (e.g. video) that account for varying levels of digital access and literacy, and map the participant journey from invitation to final data donation; specify how donation attempts, failures, donated files and covered time periods will be recorded; plan the assessment of participation bias; and pilot the full workflow.
5. Ensure participant protection and data governance	How will donors understand and control what data they share, and how will privacy and clinical risks be managed?	Determine requirements for ethics and data-protection review; where feasible, enable data preview before granular consent; apply data minimisation principles and implement secure data transfer and storage; address third-party and re-identification risks, data access and retention, participant control over withdrawal or deletion, monitoring, incidental findings and crisis response.
6. Assess coverage, linkage and data quality	Do donated data cover the target time window, align with relevant clinical records and permit reliable analysis?	Pre-define coverage, wear time and non-use criteria; distinguish absent behaviour from technical loss and selective donation; align timestamps across donated data and combine with clinical data (e.g. EHR), symptom reports (e.g. EMA) and other validated measures, treatment or life events; version-control data extraction and linkage procedures; conduct sensitivity analyses.
7. Validate inference and clinical utility	Do the derived features validly index the clinical target and improve clinically relevant inference across people, contexts and time?	Use multi-method construct validation and appropriate within-person or between-person models; separate general from target-specific effects; pre-specify power and analyses; test calibration, incremental validity, subgroup performance and external replication.

HiTOP, hierarchical taxonomy of psychopathology; EHR, electronic health record; EMA, ecological momentary assessment.
The stages are iterative rather than strictly linear.

be noisy and incomplete (particularly from older devices) and can present many challenges for analysis and reproducibility (e.g. obtaining data summaries based on hidden proprietary algorithms). Some data types (e.g. GPS-based location data) may be too sensitive and re-identifiable, even if 'anonymised', and therefore not suitable for this approach. In addition to gender and age differences, another important aspect is selection bias: data donors reported better health, fewer health-related limitations, fewer difficulties performing tasks and more physical activity,[5] highlighting the importance of a solid recruitment strategy to increase representativeness. Relatedly, wearables may be linked to specific subgroups (e.g. running devices versus smartwatches versus activity trackers).

Importantly, threats to measurement reliability arise on both sides: in passive sensing data but also in clinical assessment. A further methodological requirement for meaningful big-data passive sensing research in psychiatry therefore concerns the reliability and validity of the clinical constructs to which passive sensing features are linked. Currently, passive sensing data are usually associated with a clinical target commonly used in psychiatry research, such as a diagnosis (e.g. schizophrenia, bipolar disorder, major depression and PTSD) or a particular clinical state (e.g. anhedonia, stress and fatigue). However, the measurement of these constructs can introduce substantial problems with reliability and validity. Psychiatric diagnoses such as depression and PTSD can include considerable heterogeneity, as a diagnosis can be determined by a variety of symptom constellations, whereas clinical self-reported states such as stress or fatigue can introduce problems with construct validity and response consistency, making them imprecise and/or inadequate targets. While the detection of heterogeneous psychiatric disorders and higher-order complex states remains difficult, passive sensing data yield potential when linked to clearly defined, lower-level, reliably measurable clinical constructs – an approach that can map onto dimensional components of the Hierarchical Taxonomy of Psychopathology (HiTOP) or the behavioural and functional domains proposed within the Research Domain Criteria (RDoC) framework.

A HiTOP-aligned strategy would be to link passive sensing features to dimensional psychopathology spectra and symptom components rather than to heterogeneous diagnostic categories.[1] For the internalising spectrum, for example, sensor-derived indicators of sleep timing and fragmentation, circadian irregularity, reduced physical activity (e.g. step count), decreased mobility (e.g. GPS), autonomic arousal (e.g. heart rate variability) and avoidance-related location patterns may be especially relevant to distress (e.g. insomnia, dysphoria and trauma-related symptoms), fear and traits such as anxiousness. Similarly, these indicators may provide scalable physiological and behavioural markers of constructs within the arousal and regulatory system (i.e. sleep-wakefulness, circadian rhythms and arousal) and the negative valence system (e.g. acute threat/fear) in RDoC. Such targets can still be cross-walked to traditional diagnostic categories for clinical interpretation – for instance, depression, anxiety disorders, PTSD, psychosis-spectrum conditions and bipolar disorder – but the primary scientific target would be the more homogeneous dimensional clinical construct rather than the broad diagnosis itself.

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 21:51:14, subject to the Cambridge Core terms of use.

Despite its limitations, data donation is an innovative complementary research approach for using passive sensing data from personal devices with minimal legal and ethical risk, offering exciting new opportunities for psychiatry research. If implemented responsibly, it may help to better capture heterogeneity in psychopathology without placing significant additional burden on patients or requiring extensive additional resources from researchers. As a next step, the research field should critically discuss the suitability of different data types for data donation studies in mental health research, such as physical behaviour data, autonomic nervous system data, phone logs/screen usage data and mobile application usage data. Researchers should prioritise high-feasibility, low-risk data types to establish proof-of-concept associations with dimensional psychopathology constructs before extending to higher-risk data types under stronger governance safeguards. In this regard, ethical considerations should go beyond initial consent procedures and engage with questions of data ownership, participant autonomy and accountable data governance, including transparent consent management and auditable access control. Blockchain-based approaches may represent one possible infrastructure to support such patient-centred governance of donated data. In addition, future research should address methodological and practical considerations, including the choice of a donation platform, wearable selection, study design choices (retrospective versus prospective donation; observational versus quasi-experimental design) and data quality control. Given the many design considerations involved in data donation studies, Table 1 outlines a stage-based planning sequence for their design and implementation in mental health research, intended as a starting point for more detailed methodological guidance as the field advances. In particular, a step-by-step tutorial on how to design and implement data donation in mental health research would be a valuable next step.

Philipp Herzog, PhD, Department of Psychology, RPTU University Kaiserslautern–Landau, Germany; and Department of Psychology, University of Liège, Belgium; **Sylvie Blairy**, PhD, Department of Psychology, University of Liège, Belgium; **Philip S. Santangelo**, PhD, Department of Behavioural and Cognitive Sciences, University of Luxembourg, Luxembourg

Correspondence: Philipp Herzog. Email: philipp.herzog@rptu.de

First received 24 Apr 2026 UTC, final revision 8 Sep 2026 UTC, accepted 11 Oct 2026 UTC

Acknowledgement

The authors made use of ChatGPT (model GPT-5.6; OpenAI, San Francisco, California, USA; https://chatgpt.com), accessed through its web interface on macOS in September 2026, to assist with the development, structuring and refinement of Table 1. All AI-assisted content was subsequently critically reviewed, revised and verified by the authors, who take full responsibility for the final content.

Author contributions

All authors contributed to the conceptualisation of the editorial and the development of its central arguments. P.H. drafted the initial manuscript. S.B. and P.S.S. critically reviewed the manuscript for important intellectual content. All authors approved the final version and agree to be accountable for all aspects of the work.

Funding

Financial support for P.H. was provided by the University of the Greater Region (UniGR) through its UniGR-Guest Professorship programme.

Declaration of interest

None.

References

1 Ringwald WR, King G, Vize CE, Wright AGC. Passive smartphone sensors for detecting psychopathology. *JAMA Netw Open* 2025; **8**: e2519047.

2 Filip TF, Akhras S, Hellemann GS, Hsu T, McCleery A. Geolocation-derived mobility indices and the association with clinical symptoms and functioning in severe mental illness: a multivariate meta-analysis. *J Psychopathol Clin Sci* 2025; **134**: 935–49.

3 Hirst Y, Stoffel ST, Brewer HR, Timotijevic L, Raats MM, Flanagan JM. Understanding public attitudes and willingness to share commercial data for health research: survey study in the United Kingdom. *JMIR Publ Health Surveill* 2023; **9**: e40814.

4 Pilgrim K, Bohnet-Joschko S. Effectiveness of digital forced-choice nudges for voluntary data donation by health self-trackers in Germany: web-based experiment. *J Med Internet Res* 2022; **24**: e31363.

5 Keusch F, Struminskaya B, Mulder J, Jongerius S. Data donation as a method to measure physical activity in older adults: cross-sectional web survey assessing consent rates, donation success, and bias. *J Med Internet Res* 2025; **27**: e69799.

6 Yap VZY, Skeggs A, Ferguson AM, Leyland-Craggs A, Boeschoten L, Welbers K, et al. Digital data donation with adolescents. *Ann N Y Acad Sci* 2025; **1554**: 251.

7 Fitzgerald K, Vandenbosch L, Tabruyn T. Adolescent emotional expression on social media: a data donation study across three European countries. *Affect Sci* 2024; **5**: 436–48.

8 Hakobyan O, Hillmann P-J, Martin F, Böttinger E, Drimalla H. Development and evaluation of Dona, a privacy-preserving donation platform for messaging data from WhatsApp, Facebook, and Instagram. *Behav Res* 2025; **57**: 94.

9 Carrière TC, Boeschoten L, Struminskaya B, Janssen HL, de Schipper NC, Araujo T. Best practices for studies using digital data donation. *Qual Quant* 2025; **59**: 389–412.

Downloaded from https://www.cambridge.org/core. 15 Nov 2026 at 21:51:14, subject to the Cambridge Core terms of use.