Review

Digital Phenotyping in Health Research: Scoping Review of Methods, Gaps, and Opportunities

Nolwenn Badier[1], MCS; Alice Lafitte[1], MCS; Audrey Difernand[1], PhD; Gloria A Aguayo[2], PhD; Guy Fagherazzi[2], PhD; Jukka-Pekka Onnela[3], PhD; Benjamin Vittrant[4], PhD; Bastien Lechat[5], PhD; Quentin De Larochelambert[1], PhD; Jean-François Toussaint[1,6], MD, PhD; Lidia Delrieu[1], PhD

[1]Institut de Recherche bio-Médicale et d'Epidémiologie du Sport, Paris, Île-de-France, France

[2]Deep Digital Phenotyping Research Unit, Department of Precision Health, Luxembourg Institute of Health, Strassen, Luxembourg, Luxembourg

[3]Department of Biostatistics, Harvard University, Boston, MA, United States

[4]Withings (France), Issy-les-Moulineaux, Île-de-France, France

[5]Sleep Health, Flinders Health and Medical Research Institute, Flinders University, Adelaide, South Australia, Australia

[6]Center for Investigations in Medicine and Sports, Hôtel-Dieu de Paris, Paris, Île-de-France, France

Corresponding Author:
Nolwenn Badier, MCS
Institut de Recherche bio-Médicale et d'Epidémiologie du Sport
11 avenue du Tremblay
Paris, Île-de-France, 75012
France
Phone: 33 0141744376
Email: nolwenn.badier@insep.fr

Abstract

Background: Wearable and connected digital devices continuously generate large volumes of real-world behavioral, physiological, and environmental data, offering new opportunities for health monitoring and personalized care. Digital phenotyping has emerged as a promising paradigm; yet, there is limited consensus regarding how such data should be analyzed. Inconsistent analytical practices and insufficient methodological reporting may compromise reproducibility, comparability, and the validity of findings.

Objective: This scoping review examined current analytical practices in digital phenotyping in health research, identified methodological gaps, and highlighted the need for transparency and standardized approaches.

Methods: Literature searches were conducted in PubMed, Scopus, Embase, and Web of Science, with Google Scholar used as a complementary source. Eligible studies were published up to December 31, 2024, involved human populations, and used wearable digital devices in longitudinal health research. Studies were screened independently by multiple reviewers using predefined eligibility criteria. Data extraction focused on 6 methodological domains: sample size determination, variable selection and definition, data cleaning and preprocessing, digital phenotyping methods, predictive modeling, and statistical significance handling in large-scale data contexts.

Results: A total of 162 studies were included. Most studies were published from 2018 onward (n=144, 89%) and were conducted primarily in North America (n=92, 57%). Activity trackers (n=81, 50%), smartphones (n=35, 22%), accelerometers (n=26, 16%), and smartwatches (n=24, 15%) were the most frequently used devices. The most common outcomes were activity level (n=67, 41%), sleep (n=65, 40%), step counts (n=63, 39%), and heart rate (n=46, 28%). Heterogeneity and limited reporting were observed across all methodological domains. Preprocessing was the most frequently reported component (n=88, 54%), although specific aspects such as missing-data handling remained inconsistently described (n=35, 22%). Sample size determination methods were reported in only 30% (n=48) of studies, and variable selection methods were described in 27% (n=43). Digital phenotyping approaches were identified in 30% (n=48) of studies and predominantly used regression-based models. Predictive modeling approaches were reported in 17% (n=28) of studies, with substantial diversity in algorithms and limited reporting of validation procedures. Only 4.3% (n=7) of studies explicitly discussed statistical challenges related to large or high-dimensional datasets. Practices varied across health fields, with no domain consistently demonstrating comprehensive reporting across all methodological components.

Conclusions: Digital phenotyping research is expanding rapidly, but methodological practices remain heterogeneous and insufficiently standardized. By providing a cross-domain overview of how wearable-derived longitudinal data are currently processed and analyzed in health research, this review highlights recurring gaps in the reporting and justification of analytical choices, particularly regarding sample size determination, preprocessing, variable definition, predictive modeling, and statistical inference. These findings emphasize the need for clearer analytical frameworks and more consistent reporting practices to improve transparency, reproducibility, and methodological rigor in wearable-based digital health research.

Trial Registration: PROSPERO CRD420251234000; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251234000

(J Med Internet Res 2026;28:e91747) doi: 10.2196/91747

KEYWORDS

data analysis; digital health; phenotype; review; wearable devices

Introduction

Rationale

Digital devices, such as smartphones, smartwatches, fitness trackers, and other wearable devices, have become deeply integrated into everyday life. Beyond convenience, they continuously generate vast streams of social, behavioral, and physiological data in real-world environments. This data provides unique opportunities to gain deeper insights into human behavior and health, with the potential to transform diagnosis, monitoring, and treatment strategies [1].

Digital phenotyping was initially defined as the "moment-by-moment quantification of individual-level human behavior and physiology using data collected from personal digital devices" [2,3]. Since then, the concept has expanded beyond various active and passive data streams to also include digital assessments, such as digital cognitive tests [1,4]. Initially applied in psychiatry, digital phenotyping is now being explored across chronic disease management, general health monitoring, and public health research [5,6]. Unlike traditional approaches that rely on population averages, digital phenotyping focuses on individual changes over time, offering a more personalized perspective. Because most data are collected passively in real-world contexts, participant burden is minimized while ecological validity is preserved. These technologies thus enable large-scale, long-term monitoring, making them valuable tools for both personalized medicine and population health research [1,2].

Wearable devices have significantly advanced health research by enabling continuous data collection outside clinical settings [1]. Their affordability and ease of use make them particularly attractive for large-scale studies and research in resource-limited environments [7,8]. By providing longitudinal, real-life valid datasets, they advance both precision medicine and public health by fostering the identification of new phenotypes and the development of predictive algorithms for prevention and intervention [2,7,9,10].

Despite their potential, these technologies present significant challenges. Large, complex, and high-frequency datasets require innovative statistical methods to ensure valid interpretation. Missing data, for instance, can bias results if handled improperly, and simplistic methods such as linear interpolation may produce misleading conclusions [11]. Decisions about variable selection, preprocessing pipelines, and feature engineering are critical steps that can determine the quality and reproducibility of findings [1]. Recent studies have further emphasized the growing complexity of analytical pipelines and the lack of standardization in handling these data [12].

Conventional statistical approaches are also strained in this context. Issues such as determining appropriate sample sizes, addressing variability between individuals, and rethinking the role of statistical significance in big-data settings highlight the need for methodological innovation. Ethical considerations, including privacy and informed consent, must further guide the collection and use of personal digital data. As digital phenotyping continues to expand rapidly across disciplines, the need for structured methodological guidance has become increasingly evident [13]. Ultimately, the true promise of digital devices lies not in the scale of data collected, but in the ability to analyze and interpret the data rigorously to advance research and to deliver meaningful and clinically relevant health solutions.

Previous reviews have explored digital phenotyping across a range of specific contexts, particularly in focused domains such as mental health [14-16] or chronic diseases [17], or by examining specific types of devices or sensing modalities [12], or specific analytical approaches and processing pipelines [18]. These studies provide valuable insights into the use of digital phenotyping within well-defined applications. More recent work has also highlighted the broader clinical potential of digital phenotyping [13]. However, most reviews remain centered on a single clinical domain or methodological perspective. To our knowledge, few studies have taken a broader approach to systematically map the diversity of methodological practices used in digital phenotyping across health domains, identify cross-cutting methodological challenges, and highlight gaps in current analytical methods.

Despite rapid growth, digital phenotyping research remains relatively nascent, with limited consensus on best practices for data processing, modeling, and interpretation. Standardized guidelines are lacking, and robust frameworks for analysis are still under development. This fragmentation is further reinforced by the diversity of study designs, data sources, and analytical strategies used across the field [18]. Consolidating knowledge across methodological domains has therefore become increasingly important in order to guide future research and to ensure that digital phenotyping contributes meaningfully to precision medicine [10].

Objectives

This scoping review aims to fill these gaps by exploring how digital phenotyping is currently applied in health research, describing the analytical methods used, and highlighting their limitations. This work aims to contribute to the larger effort of harnessing the full potential of digital health data while laying the way for future research to address the challenges that remain.

Methods

Overview

Given the novelty, rapid evolution, and methodological diversity of this field, a scoping review was the most appropriate design to capture its current state and future directions [19]. The review followed the methodological framework of Arksey and O'Malley [20] and the Joanna Briggs Institute guidelines [21], and is reported in line with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist [22] (Multimedia Appendix 1). The search strategy was reported in accordance with the PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension) [23] (Multimedia Appendix 2).

Protocol and Registration

The methods were prespecified in the study protocol registered in the PROSPERO (International Prospective Register of Systematic Reviews) database on January 7, 2026 (CRD420251234000).

Eligibility Criteria

The eligibility criteria were defined in accordance with the Population-Concept-Context framework recommended by the Joanna Briggs Institute. Studies conducted in human populations were eligible, without restriction on age, sex, or health status. The review focused on the use of wearable digital devices in longitudinal health research, including studies collecting repeated or continuous data over time to investigate health-related outcomes, regardless of the methods used to analyze them. This approach allowed us to examine how such data are used in practice across studies. Studies limited to device or algorithm validation, feasibility, adherence, or protocol descriptions were excluded, as the aim was to capture applied uses of wearable data rather than methodological or technical developments. No restrictions were applied regarding the setting or geographic location, and studies conducted in both clinical and community settings were included. Both observational (eg, cohort) and interventional (eg, randomized controlled trial) study designs were considered. Studies reporting primary or secondary analyses were eligible, provided they met the inclusion criteria. Only peer-reviewed full-text articles published in English were included. Gray literature (eg, reports, theses, conference proceedings, and other non–peer-reviewed material) or unpublished studies were excluded to ensure methodological rigor. To capture the evolution of wearable devices used in health research, no start date restriction was applied, and all records published up to December 31, 2024, were considered.

Information Sources and Literature Search

The literature search was initially conducted between September 1, 2024, and January 31, 2025, using PubMed (National Library of Medicine). Google Scholar was also used as a complementary source to identify additional potentially relevant studies. Due to the large number of results and limited export functionality of Google Scholar, only the first 200 results sorted by relevance were screened. During the revision process, the search strategy was clarified, harmonized, and extended to additional databases to ensure comprehensive and consistent coverage of the literature. An updated search was conducted between March 24 and May 4, 2026, in PubMed (National Library of Medicine), Scopus and Embase (Elsevier), and Web of Science (Clarivate), using a consistent search framework across all databases.

The search strategy combined controlled vocabulary and free-text keywords related to wearable digital devices (eg, "Wearable Electronic Devices," "Wearable Devices," "Mobile Applications," and "Fitness Trackers"), health (eg, "Health," "Public Health," and "Chronic Disease"), and data analysis and methodological approaches (eg, "Data Analysis," "Statistical Power," "Processing," "Imputation," and "Predictive Models").

The PubMed search strategy was developed using both MeSH and free-text terms, structured across multiple components to comprehensively capture the different methodological dimensions of interest. To account for the methodological diversity of the field, multiple complementary search components were developed, each targeting specific methodological aspects. This conceptual framework was then consistently adapted for the other databases by adjusting syntax and indexing terms while preserving the same combination of concepts and level of detail. In Scopus, Embase, and Web of Science, comprehensive search strings were constructed using free-text terms applied to title and abstract fields, allowing all components of the search strategy to be integrated into a single combined query per database. This approach ensured that the scope and level of detail of the search were consistent across databases. The search strings used are presented in Multimedia Appendix 3.

Date limits were applied at the search stage to include only studies published up to December 31, 2024. Other eligibility criteria (eg, language and publication type) were applied during the screening process. The search strategy was developed by the authors, was not based on a previously published strategy, and was not peer-reviewed. No study registries were searched. No additional studies were identified through expert consultation. Relevant reviews identified through the search were screened to detect additional eligible studies, and both backward (reference list) and forward (citation tracking) snowballing was applied to ensure comprehensive coverage.

Selection of Sources of Evidence

All identified articles were imported into the reference management system Zotero [24] and transferred to CADIMA [25], a platform designed to support systematic and scoping reviews [26]. After duplicate removal, a two-step screening process was performed in accordance with PRISMA-ScR guidelines. In the first step, titles and abstracts were screened against predefined eligibility criteria, which had been refined through an initial pilot test. In the second step, full texts of potentially eligible studies were reviewed, with reasons for exclusion recorded.

Three reviewers independently assessed each record (NB, AL, and AD), and disagreements were resolved through discussion or, when necessary, adjudication by a fourth reviewer (LD).

Data Charting and Data Items

A data-charting form was jointly developed by two reviewers (LD and NB) to determine which variables to extract. Based on the needs identified during their respective research and the lack of available methodological guidance, the form was designed to capture key information relevant to the use of wearable digital devices in health research and to map the methodological approaches used in this field.

A pilot test was conducted on a subset of studies to refine the extraction fields and procedures before applying them consistently across all included studies. The data-charting process was iterative, allowing adjustments to the extraction form as needed during the review. Data extraction was performed by 1 reviewer (NB) using the CADIMA platform, and the extracted data were subsequently verified and approved by a second reviewer (LD) to ensure accuracy and consistency. The resulting extraction table was exported to Microsoft Excel for subsequent analyses.

Extracted variables included (1) bibliographic information (title, authors, and year of publication), (2) study objective (the primary objective when multiple objectives were reported), (3) country where the study was conducted, (4) study population (eg, sample size and key characteristics such as sex- or age-specific populations, when applicable), (5) study duration, (6) type of wearable digital device used, (7) health research field, and (8) relevant methodological approaches. These included details related to six domains: (1) sample size determination (eg, reported parameters such as alpha level, statistical power, or assumptions used for calculation), (2) variable selection and definition (eg, type and number of variables or features considered for subsequent analyses), (3) data cleaning, preprocessing, and transformation (eg, handling of missing data and outliers, normalization, and feature engineering), (4) digital phenotyping methods (eg, handling of longitudinal data to construct behavioral or digital biomarkers or to characterize individual-level patterns over time), (5) predictive modeling (eg, type of models or algorithms used), and (6) significance testing in big-data context (eg, use of *P* values, correction methods, or thresholds for statistical significance). Each methodological domain was considered independently from the others. Their presence or absence was recorded for all included studies, allowing the identification of gaps in reporting and analytical practices.

Synthesis of Results

A descriptive analysis was conducted to summarize the characteristics of the included studies. Variables examined included publication year, country of origin (based on study population or, if not specified, corresponding author), study population, duration of follow-up, type of wearable device, outcomes assessed, disease focus, and reported methodological approaches.

Quantitative variables were summarized using descriptive statistics (minimum, maximum, mean, SD, median, and IQR), and qualitative variables were reported as counts and frequencies. A gap map was constructed to summarize the reporting of key methodological domains and their associated subdomains. Each item was categorized as "well reported," "partially reported," or "rarely reported" based on the proportion of studies reporting each item (>50%, 25%-50%, and <25%, respectively).

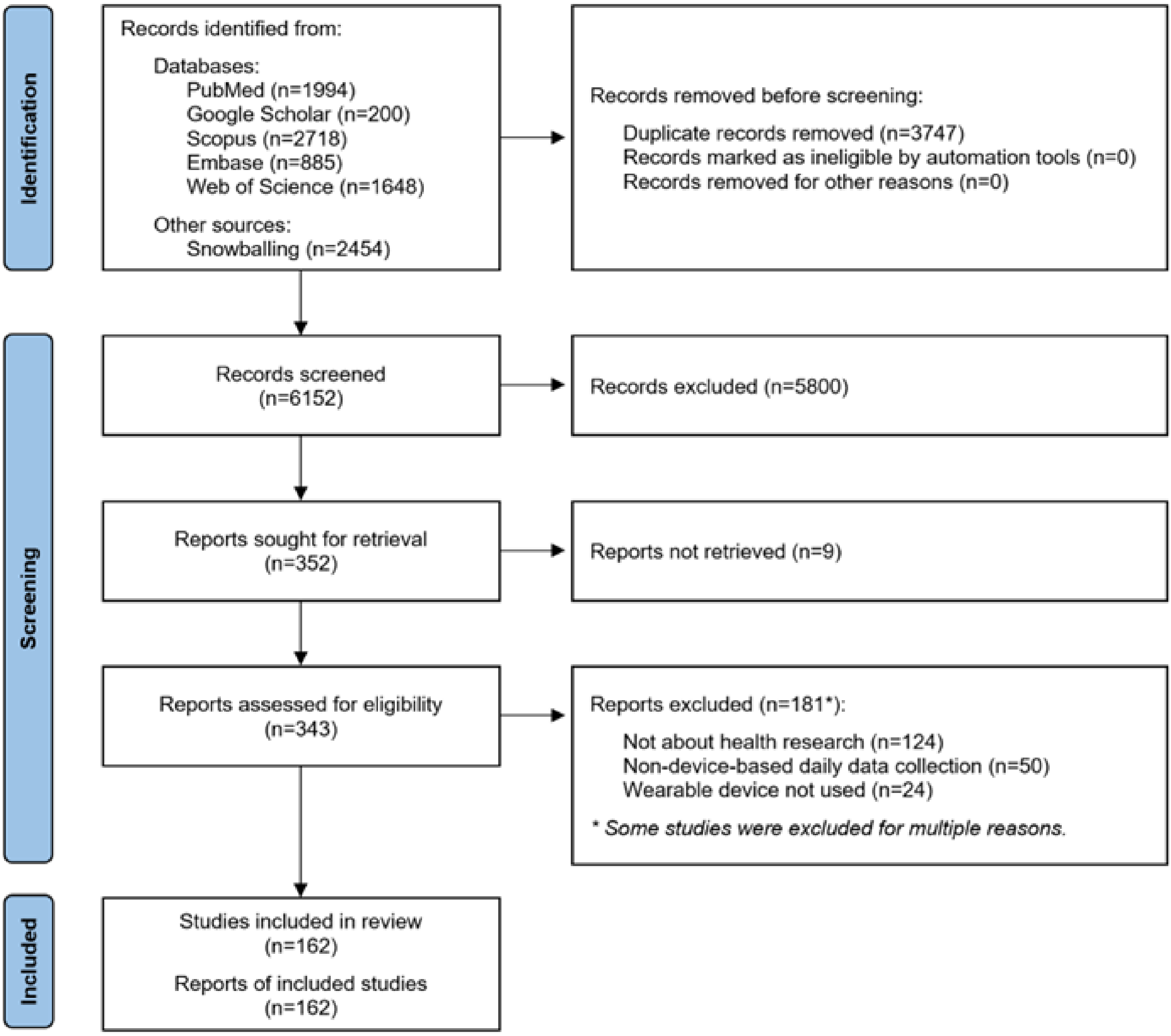
Analyses were performed in R software (version 4.5.0, R Foundation for Statistical Computing).

Results

Selection of Sources of Evidence

The search identified 1994 (20%) records from PubMed, 2718 (27%) from Scopus, 885 (8.9%) from Embase, and 1648 (17%) from Web of Science. In addition, the first 200 results from Google Scholar were screened as a complementary source, and 2454 records were identified through snowballing based on 52 relevant reviews, resulting in a total of 9899 records screened. After removing 3747 (38%) duplicates, 6152 abstracts were screened, leading to the exclusion of 5800 (94%) studies. The remaining 352 records were sought for retrieval, of which 9 were not retrieved. A total of 343 full-text articles were assessed for eligibility, and 181 (53%) were excluded. In total, 162 studies met the inclusion criteria and were included in the analysis. The study selection process is shown in Figure 1, and the main characteristics of the included studies are summarized in the Multimedia Appendix 4 [27-188].

Figure 1. PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flow diagram of the study selection process, including identification, screening, eligibility, and inclusion stages.



Characteristics of Sources of Evidence

Publication Year

The included studies were published between 2005 and 2024 (Table 1). Most studies (n=144, 89%) were published from 2018 onward. The average number of studies per year increased from 4 between 2014 and 2017, to 18 between 2018 and 2023. Publication output further increased in 2024, with 35 (22%) studies.

Table 1. Distribution of included studies (N=162) by year of publication (2005-2024).

Year	Value, n (%)
2005	1 (0.6)
2006	0 (0)
2007	0 (0)
2008	0 (0)
2009	0 (0)
2010	1 (0.6)
2011	0 (0)
2012	0 (0)
2013	0 (0)
2014	2 (1.2)
2015	3 (1.9)
2016	5 (3.1)
2017	6 (3.7)
2018	17 (10)
2019	19 (12)
2020	15 (9.3)
2021	18 (11)
2022	23 (14)
2023	17 (10)
2024	35 (22)

Countries

The studies were conducted across 23 countries (Table 2), most commonly in North America (n=92, 57%), followed by Europe (n=36, 22%), Asia (n=28, 17%), and Australia (n=6, 3.7%).

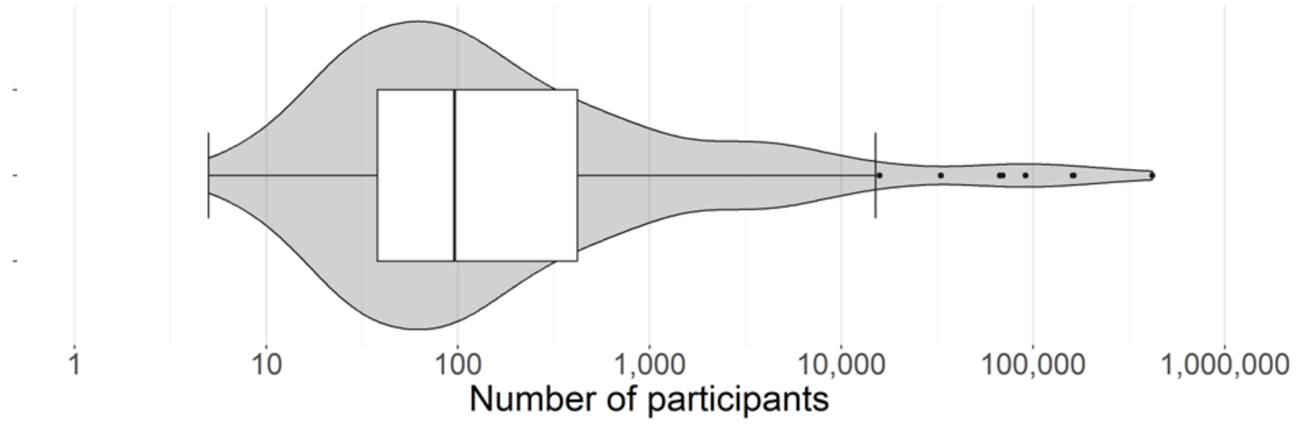
Table 2. Geographic distribution of included studies (N=162) by country, based on study population or corresponding author when not specified.

Countries	Value, n (%)
United States	85 (53)
Singapore	10 (6.2)
United Kingdom	10 (6.2)
Canada	7 (4.3)
Australia	6 (3.7)
China	5 (3.1)
Japan	5 (3.1)
Republic of Korea	5 (3.1)
Denmark	4 (2.5)
Finland	4 (2.5)
The Netherlands	3 (1.9)
Austria	2 (1.2)
France	2 (1.2)
Ireland	2 (1.2)
Norway	2 (1.2)
Sweden	2 (1.2)
Taiwan	2 (1.2)
Belgium	1 (0.6)
Czechia	1 (0.6)
Germany	1 (0.6)
India	1 (0.6)
Poland	1 (0.6)
Spain	1 (0.6)

Populations

Across the 162 studies, a total of 1,132,068 participants were included, with a median of 96 (IQR 38-420) participants per study (Figure 2). Sample sizes ranged from 5 to 419,093 participants. Most studies enrolled fewer than 100 participants (n=84, 52%), while 49 (30%) included 100-999 participants, 19 (12%) included 1000-9999, 7 (4.3%) included 10,000-99,999, and 3 (1.9%) included more than 100,000 participants.

Figure 2. Distribution of sample sizes across included studies (N=162). Sample sizes are displayed on a logarithmic scale to account for variability across studies.



Most studies (n=99, 61%) included adults aged 18 years and older without targeting a specific subgroup. Others focused on specific populations, including older adults aged 50 years and older (n=20, 12%), workers (n=17, 10%), women (n=10, 6.2%), students (n=10, 6.2%), and children (n=5, 3.1%). Only 1 (0.6%) study specifically focused on men.

Devices in Health Research

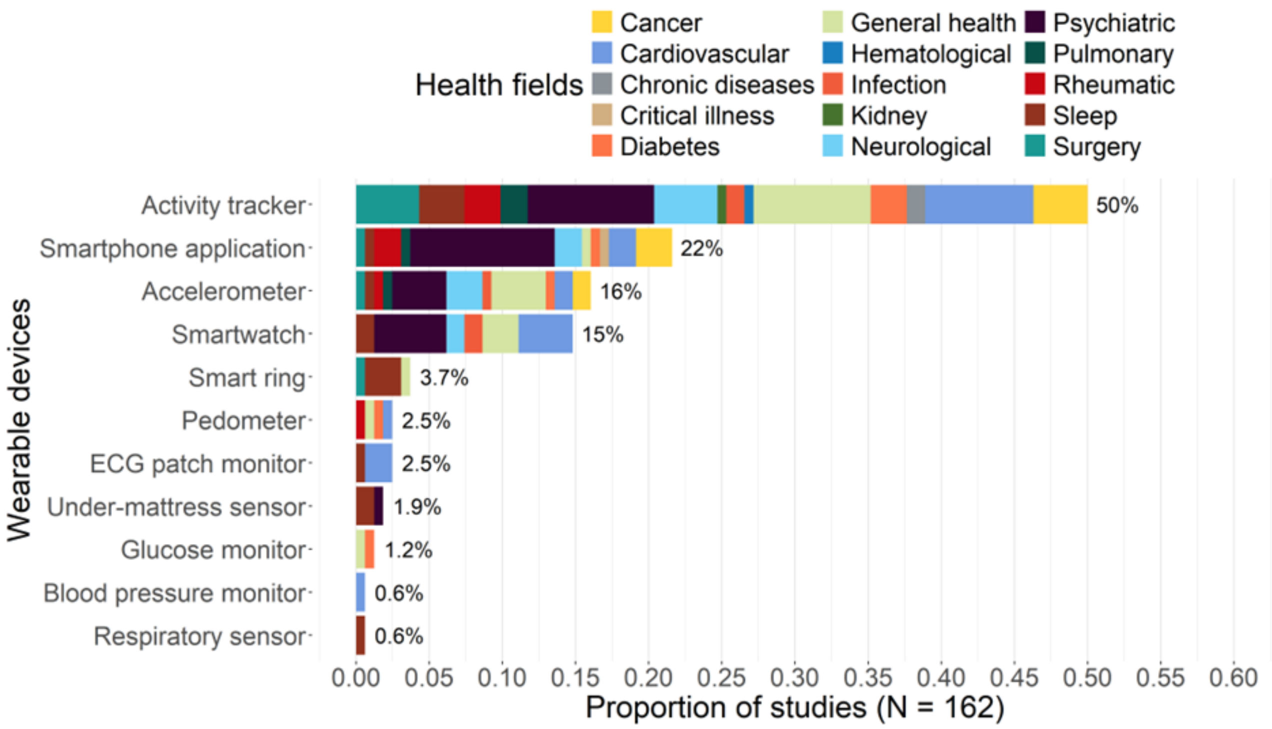
Most studies used a single wearable device (n=141, 87%). A total of 17 (10%) studies combined 2 devices, and 4 (2.5%) studies included 3 devices.

Across all studies, 11 types of wearable devices were identified. The most common were activity trackers (n=81, 50%), smartphones (n=35, 22%), accelerometers (n=26, 16%), and smartwatches (n=24, 15%). Device categories reflect the terminology used in the original studies. Because reporting standards vary widely, some labels correspond to hardware components (eg, accelerometers), whereas others describe the commercial form factor (eg, smartwatches and activity trackers). As a result, categories may partially overlap. Reported wearing time ranged from 1 day to 16.6 years, with a mean of 39.9 (SD 87.9) weeks.

A total of 10 categories of outcomes were reported. The most frequent were activity level (n=67, 41%), sleep (n=65, 40%), step counts (n=63, 39%), and heart rate (n=46, 28%). Other outcomes included movement patterns (n=14, 8.6%), time of device use (n=8, 4.9%), cardiorespiratory fitness and body composition (n=5, 3.1% each), light exposure (n=2, 1.2%), and blood pressure (n=1, 0.6%).

Health fields were grouped into 15 categories (Multimedia Appendix 5). The most represented categories were psychiatric disorders (n=41, 25%), general health (n=25, 15%), and cardiovascular diseases (n=23, 14%). Figure 3 shows the distribution of wearable devices across health research domains.

Figure 3. Distribution of wearable device types (n=11) across health research domains (n=15). Bars represent the proportion of studies (N=162) using each type of device, stratified by health field. Multiple device types could be reported within a single study.



Results of Individual Sources of Evidence

Sample Size Requirements

Of the 162 included studies, 48 (30%) reported a method for determining sample size. Among these, 19 (40%) described the parameters used in the calculation [30,51,55-57,93,105,108,109,123,149,153,164,166,177,183], while 3 provided justification based on references [62,118,157]. A total of 17 (35%) studies included as many participants as possible [28,40,45,59,63,65,71,76,84,128,132,136,168,170,184,185,188], and 12 (25%) relied on sample sizes reported in previous studies [34,43,44,104,110,111,116,131,140,142,176,180].

Among the 19 studies that specified calculation parameters, expected outcomes were reported in 13 (68%). Among these, effect size was used in 5 (38%), difference rate in 5 (38%), Cohen *d* in 2, and the κ coefficient in 1 study. The type I error rate α was set at .05 in all 14 (74%) studies that reported it. Statistical power was reported in 15 (79%) studies, most often set at 80% (n=12, 80%) or 90% (n=2, 13%), with 1 study using 81%. Attrition was accounted for in 7 (37%) studies, with assumed rates of 10%, 20%, or 25% in 2 (29%) studies each. One study arbitrarily added 8 participants to its final sample size.

Overall, reporting of sample size justification was inconsistent, with variability in the parameters considered and limited reporting of underlying assumptions.

Selection and Definition of Variables

Of the 162 included studies, 43 (27%) reported methods for variable selection. Among these, 30 (70%) aggregated longitudinal data into summary measures (eg, means or categories) [41,42,44,45,50,63,75,76,78,87,100,104,109,111,117,121,137,142,149,154,156,158,159,167,168,171,178,180,182,187], 6 (14%) selected variables based on prior literature [59,81,98,127,170,183], 4 (9.3%) used factorial analysis

[92,110,134,164], and 2 (4.7%) applied Pearson correlations [61,138]. One study combined data aggregation with subsequent Pearson correlation analysis [29].

Overall, variable selection approaches were heterogeneous, with a predominant reliance on data aggregation methods, reducing the longitudinal dimension of the data.

Data Cleaning, Preprocessing, and Transformations

Of the 162 included studies, 6 (3.7%) cleaned their datasets by removing the first and last day of recording [49,96,109,152,160,165], and 2 others excluded the first 14 days (in 9-month [151] and 2-year studies [183]). Sleep studies applied context-specific rules: 2 studies retained the main sleep episode when multiple episodes were recorded in the same day [84,170], and another excluded specific calendar periods (eg, New Year, Thanksgiving, Christmas, and daylight-saving transitions) [135].

Wear-time validation was reported in 60 (37%) studies. Among these, 38 (63%) required a minimum number of hours per day [32,40,41,44,63,67,73,74,77,81,87,88,90,92,94,96,103,109,110,112,116,121,122,127,135,138-140,145,150,158,160,167,168,174,178,180,184], ranging from 1 to 22 hours (mean 9.5, SD 6.0). A total of 32 (53%) studies required a minimum number of valid days [32,43,44,49,54,56,59,62,67,70,77,78,82,87,90,92,100,110,128,133-135,139,140,142,145,148,165,167,174,183,185], defined based on a prior minimum number of hours, ranging from 1 to 83 days (mean 12.9, SD 19.2). In 3 studies, this requirement depended on study duration (eg, 15 days per month, 24 days during the first 60 days, or 7 days during a 28-day window). In addition, 13 (22%) studies used step counts to confirm wear [28,44,46,48,60,62,67,77,103,128,148,152,162], most often applying a minimum threshold of 100 daily steps (n=8, 62%). Among these, 2 studies required a maximum of 45,000 steps per day, and 2 others a maximum of 50,000 steps per day. Two studies required at least 1000 steps, with 1 also setting a maximum at 20,000 steps per day. Two other studies required at least 1 step per day, and 1 study required 500 steps per day. Two studies based their thresholds on prior literature [81,87], and 2 others used intraclass correlation coefficients ≥0.90 [49,82].

Missing data were addressed in 35 (22%) studies. Among these, 16 (46%) excluded missing data from the analysis [37,41,52,59,78,98,101-104,109,121,125,134,170,181], and 2 (5.7%) applied interpolation methods [143,152], including a 2-value surrounding method. Two other studies replaced missing values with the mean [29,71], and 1 replaced them with the median [132]. In addition, 14 (40%) studies applied imputation methods [31,47,57,61,76,81,86,87,105,119,120,131,155,179], such as multiple imputation (n=4, 12%) or sparse online Gaussian process (n=2, 6%). Five (15%) other studies used regression models, maximum likelihood estimation with the expectation-maximization algorithm, percentile-based methods (97th percentile), k-nearest neighbors, or similar time-of-day segments (1 study each).

A total of 10 (6.2%) studies reported explicit outlier management methods [49,71,96,102,121,132,139,165,180,188]. These included trimming values outside mean ±1 SD (n=3, 30%), z-scores >3 (n=3, 30%), mean ±2 SD (n=1, 10%), and Tukey's interquartile rule (n=1, 10%). Two studies explicitly retained all data points.

Overall, preprocessing practices were highly heterogeneous, with a wide range of approaches and limited consistency in reporting.

Digital Phenotyping Analysis Methods

Of the 162 included studies, 48 (30%) applied digital phenotyping methods. Regression models were the most commonly used methods, reported in 39 (81%) studies [27,36,37,41,46,50,53,61,72,74-77,79,80,84,94,100,102,104,107,111,117,119,127,137,139,149,156,157,164,165,167,172,176,179,180,182,185]. Mixed-effects models were the most frequent (n=25, 52%), followed by generalized linear models (n=11, 23%). Other approaches included polynomial models (n=2, 4.2%), as well as generalized estimating equations, survival analysis, cosinor model, and spline-based models (1 study each). Clustering methods were used in 7 (15%) studies [34,65,91,96,134,169,170], including k-means algorithms (n=3, 6.3%), latent class approaches (n=2, 4.2%), and density-based spatial clustering of applications with noise (n=1, 2.1%). One study did not report the clustering algorithm used. Finally, 2 (4.2%) studies applied mixed-effects models to the output of k-means clustering [140,141].

Overall, digital phenotyping analyses relied predominantly on regression-based approaches, with limited use of unsupervised methods.

Predictive Models

Of the 162 included studies, 28 (17%) reported the use of predictive modeling approaches, applying a total of 60 models. Most studies (n=16, 57%) relied on a single method [29,40,47,58,69,116,131,139,145,151,154,169,179,181,186,187], while others combined 2 (n=2, 7.1%) [143,170] or 3 (n=6, 21%) methods [43,74,77,78,138,163]. Four additional studies applied 4 [144], 5 [132], 6 [61], and 7 methods [79], respectively.

Among the 28 studies, decision tree-based approaches were the most frequently used, reported in 16 (57%) studies. Random forest was used in 12 (43%) studies, extreme gradient boosting in 11 (39%), gradient boosted trees in 2 (7.1%), and adaptive boosting (n=1, 3.6%). Distance-based methods were reported in 10 studies (36%), including support vector machines (n=8, 29%), k-nearest neighbors (n=6, 21%), and Gaussian processes (n=2, 7.1%). Predictive regression models were used in 4 (14%) studies, including ridge regression, least absolute shrinkage and selection operator, elastic net, and regularized regression (1 study each). Neural network architectures were also reported in 4 (14%) studies and included long-short-term memory and recurrent neural networks (n=2, 7.1% each), as well as gated recurrent units and multilayer perceptrons (1 study each). Sequential models and time-series approaches were each reported in 3 (11%) studies each. All sequential models included Markov chain transitions, while time-series approaches included autoregressive integrated moving average, seasonal autoregressive integrated moving average, and Bayesian structural time series models (1 study each). One study reported the use of an unspecified internal algorithm.

Validation procedures were detailed in 17 (61%) of the 28 studies. A total of 7 (25%) studies used leave-one-subject-out cross-validation [61,80,127,138,143,145,179], and 10 (36%) applied repeated k-fold cross-validation [58,74,77-79,116, 131,139,156,187], with k sometimes specified as 3 (n=1), 4 (n=1), 5 (n=2), or 10 (n=2).

Performance metrics were reported in 10 (36%) studies [58,61,74,77-79,116,138,141,179]. The area under the receiver operating characteristic curve was the most common metric (n=8, 29%). Other metrics included bootstrapped feature stability and F_1 score (n=3, 11% each), confusion matrices, negative predictive value, and κ statistic (n=2, 7.1% each), as well as Shapley values, root-mean-square deviation, Pearson and Spearman correlations, Brier metrics, and mean absolute error (1 study each).

Finally, only 2 (7.1%) studies explicitly discussed model limitations [139,143], particularly the distinction between prediction and causality.

Overall, predictive modeling approaches were diverse but inconsistently reported, with limited transparency regarding validation, performance evaluation, and model limitations.

Statistical Significance in Big-Data Contexts

Only 7 (4.3%) studies addressed the limitations of statistical significance testing in a big-data context [50,71,78,108,121,184,185]. Four studies emphasized the need for appropriate adjustment of *P* values, and 1 highlighted that a large number of analyses may increase the risk of spurious findings. Two studies discussed the limitations of their conclusions and stressed the importance of assessing clinical impact beyond statistical significance. They also acknowledged the difficulty of distinguishing relevant from irrelevant correction methods.

Overall, the implications of large-scale data on statistical significance were rarely addressed.

Methodological Practices Across Health Domains

Methodological practices were further examined across health domains to explore potential differences in reporting patterns (Multimedia Appendix 6). To account for differences in the number of studies across domains, results are presented as proportions within each health domain.

Preprocessing was the most frequently reported methodological component across domains, ranging from 29% (2/7) in surgery-related studies to 64% (9/14) in neurological studies, reaching 60% (3/5) in diabetes and pulmonary-related studies, and 59% (24/41) in psychiatric studies. Sample size planning was inconsistently reported, with proportions ranging from 21% (3/14) in neurological studies to 43% (3/7) in surgical studies, and 40% (2/5) in both diabetes and infection-related studies. It was also reported in 39% (9/23) of cardiovascular studies. Similarly, variable definition was reported in 36% (9/25) of general health and neurological studies (5/14), while lower proportions were observed in cancer (1/10, 10%), surgery (1/7, 14%), and rheumatism (1/6, 17%). Digital phenotyping methods were reported in 44% (11/25) of general health studies, 40% (2/5) in infection-related studies, and 37% (15/41) in psychiatric studies, while remaining below 20% (1/5) in several other domains, including neurological (2/14, 14%) and cancer (1/10, 10%). Predictive modeling was reported in 43% (3/7) of surgical studies and 34% (14/41) of psychiatric studies, but was rare in several domains, including general health (2/25, 8%) and cardiovascular (1/23, 4.3%), and absent in cancer, diabetes, and pulmonary studies. Statistical considerations related to large or high-dimensional data were rarely reported across all domains.

Overall, these results highlight substantial variability in methodological reporting across health domains, with no field consistently demonstrating high levels of reporting across all methodological components.

Synthesis of Results

The synthesis of methodological practices across included studies highlights several recurring patterns and gaps in the application of digital phenotyping in health research.

Overall, reporting of key methodological components was inconsistent across domains. Sample size justification was rarely described, and when reported, relied on heterogeneous parameters with limited transparency. Similarly, variable selection approaches were highly diverse, with a predominant reliance on data aggregation methods, often reducing the longitudinal dimension of the data. Preprocessing strategies showed substantial variability, with a wide range of approaches for wear-time validation, missing data handling, and outlier management, and no clear consensus on best practices. These inconsistencies were further reflected in analytical choices. While digital phenotyping analyses were predominantly based on regression models, the use of more advanced or unsupervised approaches remained limited. In studies applying predictive modeling, a wide diversity of algorithms was observed, but reporting of validation procedures, performance metrics, and model limitations remained incomplete. In particular, only a small proportion of studies explicitly addressed issues related to overfitting, model generalizability, or the distinction between prediction and causal inference. Finally, the implications of large-scale and high-dimensional data on statistical inference were rarely discussed. Only a few studies addressed challenges related to multiple testing or the interpretation of statistical significance, despite their importance in big-data contexts.

To provide a structured overview of these methodological gaps, Figure 4 presents a gap map summarizing the extent and consistency of reporting across key methodological domains.

Figure 4. Gap map of methodological reporting across key domains in studies using wearable digital devices in health research (N=162). Reporting levels were categorized as "well reported" (>50% of studies), "partially reported" (25%-50%), and "rarely reported" (<25%). For each domain, the "Total" category indicates whether at least 1 methodological component within the domain was reported in the study.

Taken together, these findings highlight a lack of standardization and transparency in methodological practices, as well as important gaps in the reporting and justification of analytical choices. Despite the richness and high frequency of longitudinal data collected through wearable devices, many studies relied on aggregated summary measures, thereby limiting the ability to fully capture temporal dynamics and individual variability. These patterns suggest the need for clearer methodological guidance to support the robust and reproducible use of digital phenotyping in health research.

Discussion

Summary of Evidence

This scoping review aimed to examine how digital phenotyping is currently applied in health research, with a focus on analytical methods and their reporting. Overall, the findings reveal substantial heterogeneity and limited transparency across key methodological domains, including sample size determination, variable selection, data preprocessing, and analytical approaches. Despite the increasing availability of rich longitudinal data, many studies relied on simplified or insufficiently described methods, and explicit digital phenotyping approaches were only applied in a minority of studies. These findings highlight important gaps in current methodological practices and emphasize the need for clearer guidance to support robust and reproducible use of wearable-derived data in health research.

Sample size determination was often insufficiently reported and lacked transparency across studies. When described, the underlying assumptions and parameters varied substantially, reflecting the absence of methodological standardization in the field. This heterogeneity may reflect the exploratory nature of many studies, but it also raises important concerns regarding the robustness, reproducibility, and comparability of findings in digital health research [189]. Improving the reporting of sample size justification is therefore essential. Clear documentation of effect sizes, statistical power, significance thresholds, and underlying assumptions is necessary to ensure appropriate study design and interpretation of results [190]. The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines provide a useful framework to improve transparency in observational studies and should be more consistently applied in this field [191]. Wider adoption of such frameworks would enhance consistency and support the generation of more robust and generalizable evidence in digital phenotyping research. However, in the context of digital phenotyping, these considerations are further complicated by the longitudinal and high-frequency nature of the data, where sample size requirements also depend on follow-up duration and within-subject variability [192].

Variable selection and definition approaches were highly heterogeneous and often insufficiently described across studies. In many cases, the process used to select or construct variables was not clearly reported, limiting the transparency and interpretability of the analyses [193]. This lack of detail makes it difficult to assess methodological rigor or to compare findings across studies. A notable pattern was the frequent use of data aggregation strategies, where longitudinal data were reduced to summary measures such as means or categories. While these approaches simplify analysis, they may also limit the ability to capture temporal dynamics and within-individual variability, which are central to digital phenotyping, and may lead to misinterpretation of effects [194]. More generally, variable selection and definition should be guided by clearly defined research questions rather than by the constraints of device-derived metrics alone. Improved reporting and more thoughtful use of longitudinal data are therefore essential to fully leverage the richness of wearable-derived datasets.

Data cleaning, preprocessing, and transformation practices were highly heterogeneous across studies, with no clear consensus on how wearable-derived data should be processed. This variability was reflected in key steps such as wear-time validation, inclusion criteria, and analytical procedures, limiting comparability across studies and complicating evidence synthesis [8,195]. A major source of heterogeneity relates to the use of proprietary vs raw data. Many studies relied on device-derived metrics, which are often based on proprietary algorithms that may vary across manufacturers and over time, thereby limiting transparency and reproducibility [195,196]. In contrast, access to raw sensor data enables more transparent preprocessing, facilitates harmonization across devices, and supports data pooling, thereby improving reproducibility [195,197]. Explicit reporting of whether analyses are based on raw signals or device-derived summaries should therefore be encouraged in future digital phenotyping research. Handling of missing data was also inconsistent. While some studies applied advanced imputation methods, many relied on simpler approaches such as deletion, averaging, or interpolation, which may introduce bias or lead to information loss [198,199]. In addition, key aspects of missing data handling were often poorly reported, including the level at which imputation was performed and the type of data concerned. This lack of detail limits the interpretability and reproducibility of analytical strategies. More broadly, few studies discussed the underlying mechanisms of missingness or considered alternative approaches such as sensitivity analyses, despite their importance in longitudinal data analysis [198,199]. Outlier handling was rarely described and lacked standardization when reported. Although classical approaches such as threshold-based methods or interquartile rules were occasionally used, their application remained inconsistent [200,201]. Clear and consistent reporting of preprocessing steps, including missing data and outlier management, is therefore essential to improve the reliability and reproducibility of digital health research.

Despite the growing availability of complex digital health data, digital phenotyping remains inconsistently applied in health research. In this review, most studies relied on regression-based models, which are primarily designed to assess associations rather than to capture dynamic behavioral patterns over time [1]. As a result, these approaches may not fully exploit the richness and temporal complexity of high-frequency data generated by wearable devices [1,202]. There is a gap between the definition of digital phenotyping and its implementation in practice. Digital phenotyping is intended to capture individual-level patterns over time, yet in many studies, it is effectively reduced to the use of standard statistical models applied to repeated measurements. Overall, these findings suggest that the field remains exploratory and heterogeneous. Analytical choices often appear to be driven by feasibility and familiarity rather than by clearly defined objectives. This highlights the need for clearer methodological frameworks and an opportunity to develop approaches better suited to the dynamic nature of digital phenotyping data [1].

Predictive modeling approaches were used across studies, with a wide range of methods including tree-based models, distance-based algorithms, regression-based approaches, and neural networks. This diversity reflects the growing interest in machine learning techniques for analyzing wearable-derived data [203]. However, model development and reporting were often incomplete. Key aspects such as validation procedures and performance metrics were not consistently described, limiting the interpretability and reproducibility of the results [204,205]. In addition, the distinction between prediction and other analytical objectives, such as association or causal inference, was rarely discussed. Only a small number of studies explicitly addressed model limitations, including the interpretation of predictive outputs. This lack of clarification raises concerns about how these models are understood and used in practice [206]. Overall, these observations suggest that the main challenge is not the lack of available methods, but how they are applied and reported. Recent guidelines provide a useful framework for the development and validation of prediction models [207], but clearer reporting standards and more consistent application are essential to ensure that predictive models are appropriately developed, validated, and interpreted in digital health research.

Only a small number of studies addressed the limitations of statistical significance testing in a big-data context. When discussed, concerns mainly related to the increased risk of false-positive findings due to multiple testing and the difficulty of applying appropriate correction methods when the number of comparisons is large or not clearly defined [208]. In this context, classical approaches such as Bonferroni correction may be overly conservative, potentially reducing statistical power and obscuring meaningful effects. Some studies also emphasized that statistical significance does not necessarily translate into clinical relevance, highlighting the importance of considering the practical impact of findings alongside *P* values [209]. Overall, these observations suggest that the implications of large-scale data on statistical inference remain insufficiently explored.

Taken together, these findings highlight substantial variability in methodological practices across health domains, with no domain consistently demonstrating high levels of reporting across all components. This suggests that the observed limitations are not confined to particular clinical fields, but rather reflect broader, systemic challenges in the application of digital phenotyping in health research.

Limitations

This scoping review has several limitations. Despite efforts to maximize coverage, some relevant studies may have been missed, particularly those published in languages other than English. Data extraction was performed by a single reviewer, which may have introduced bias, although verification steps were implemented to limit this risk. No formal quality appraisal or risk-of-bias assessment was conducted, in line with the objectives of scoping reviews, which aim to map the available literature rather than to evaluate study rigor. In addition, the unit of analysis was the individual study rather than the underlying dataset. As a result, multiple publications based on the same data source may have been included. However, this approach is consistent with the scoping nature of the review, which aimed to capture the diversity of methodological practices across studies. Furthermore, this review focused on methodological practices related to data processing and analysis. Other important aspects, such as code and data availability, reproducible pipelines, and benchmark datasets, were not assessed, although they represent key dimensions of standardization and reproducibility in digital phenotyping research. Finally, the substantial heterogeneity in study designs, technologies, and analytical approaches limited the ability to synthesize findings quantitatively or to identify definitive best practices. These limitations should be considered when interpreting the results.

Conclusions

By exploring methodological practices across health fields, this scoping review provides a broad cross-domain overview of how wearable-derived longitudinal data are currently processed and analyzed in digital phenotyping research. By structuring these approaches into key domains and examining their reporting across studies, this review highlights recurring gaps and substantial heterogeneity in current analytical practices. Despite the increasing availability of high-frequency real-world data, many studies relied on simplified or insufficiently reported methods, particularly regarding sample size determination, preprocessing, variable definition, predictive modeling, and statistical inference. Overall, the rapid expansion of digital phenotyping research appears to outpace the methodological standards needed to support robust and reproducible analyses. These findings emphasize the need for clearer analytical frameworks and more consistent reporting practices to improve transparency, reproducibility, and interpretability in digital phenotyping research. Beyond identifying current limitations, this review provides a foundation for future methodological guidance and may help support the development of more robust and standardized approaches for the analysis of wearable-derived data in health research.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Funding

This study was funded by the Institut National du Cancer (France) through a PhD fellowship granted to Nolwenn Badier (INCa_19571). The funder had no role in the study.

Authors' Contributions

NB, AL, and AD independently assessed each record, under the supervision of LD. NB extracted, cleaned, and analyzed the data, with methodological supervision from GAA. NB drafted the manuscript. GF, JPO, BV, JFT, and LD contributed to the interpretation of results. All authors critically reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 50 KB-Multimedia Appendix 1]

Multimedia Appendix 2

PRISMA-S checklist.
[PDF File (Adobe PDF File), 35 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Search strings used for the records identification.
[PDF File (Adobe PDF File), 35 KB-Multimedia Appendix 3]

Multimedia Appendix 4

Main characteristics of the 162 included studies.
[XLSX File (Microsoft Excel File), 61 KB-Multimedia Appendix 4]

Multimedia Appendix 5

Distribution of health research topics in studies.
[PDF File (Adobe PDF File), 59 KB-Multimedia Appendix 5]

Multimedia Appendix 6

Distribution of methodological practices across health domains.
[PDF File (Adobe PDF File), 120 KB-Multimedia Appendix 6]

References

1. Onnela J. Opportunities and challenges in the collection and analysis of digital phenotyping data. Neuropsychopharmacology. 2021;46(1):45-54. [FREE Full text] [doi: 10.1038/s41386-020-0771-3] [Medline: 32679583]
2. Torous J, Kiang MV, Lorme J, Onnela J. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Ment Health. 2016;3(2):e16. [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]
3. Onnela J, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. Neuropsychopharmacology. 2016;41(7):1691-1696. [FREE Full text] [doi: 10.1038/npp.2016.7] [Medline: 26818126]
4. Insel TR. Digital phenotyping: technology for a new science of behavior. JAMA. 2017;318(13):1215-1216. [doi: 10.1001/jama.2017.11295] [Medline: 28973224]
5. Cornet VP, Holden RJ. Systematic review of smartphone-based passive sensing for health and wellbeing. J Biomed Inform. 2018;77:120-132. [FREE Full text] [doi: 10.1016/j.jbi.2017.12.008] [Medline: 29248628]
6. Mohr DC, Shilton K, Hotopf M. Digital phenotyping, behavioral sensing, or personal sensing: names and transparency in the digital age. NPJ Digit Med. 2020;3:45. [FREE Full text] [doi: 10.1038/s41746-020-0251-5] [Medline: 32219186]
7. Majumder S, Mondal T, Deen MJ. Wearable sensors for remote health monitoring. Sensors (Basel). 2017;17(1):130. [FREE Full text] [doi: 10.3390/s17010130] [Medline: 28085085]
8. Huhn S, Axt M, Gunga H, Maggioni MA, Munga S, Obor D, et al. The impact of wearable technologies in health research: scoping review. JMIR Mhealth Uhealth. 2022;10(1):e34384. [FREE Full text] [doi: 10.2196/34384] [Medline: 35076409]
9. Steinhubl SR, Mehta RR, Ebner GS, Ballesteros MM, Waalen J, Steinberg G, et al. Rationale and design of a home-based trial using wearable sensors to detect asymptomatic atrial fibrillation in a targeted population: the mHealth screening to prevent strokes (mSToPS) trial. Am Heart J. 2016;175:77-85. [FREE Full text] [doi: 10.1016/j.ahj.2016.02.011] [Medline: 27179726]

10. Dunn J, Runge R, Snyder M. Wearables and the medical revolution. Per Med. 2018;15(5):429-448. [doi: 10.2217/pme-2018-0044] [Medline: 30259801]
11. Barnett I, Onnela JP. Inferring mobility measures from GPS traces with missing data. Biostatistics. 2020;21(2):e98-e112. [FREE Full text] [doi: 10.1093/biostatistics/kxy059] [Medline: 30371736]
12. Linardon J, Chen K, Gajjar S, Eadara A, Wang S, Flathers M, et al. Smartphone digital phenotyping in mental health disorders: a review of raw sensors utilized, machine learning processing pipelines, and derived behavioral features. Psychiatry Res. 2025;348:116483. [doi: 10.1016/j.psychres.2025.116483] [Medline: 40187059]
13. Zhang Y, Wang J, Zong H, Singla RK, Ullah A, Liu X, et al. The comprehensive clinical benefits of digital phenotyping: from broad adoption to full impact. NPJ Digit Med. 2025;8(1):196. [FREE Full text] [doi: 10.1038/s41746-025-01602-5] [Medline: 40195396]
14. Mendes JPM, Moura IR, Van de Ven P, Viana D, Silva FJS, Coutinho LR, et al. Sensing apps and public data sets for digital phenotyping of mental health: systematic review. J Med Internet Res. 2022;24(2):e28735. [FREE Full text] [doi: 10.2196/28735] [Medline: 35175202]
15. Bufano P, Laurino M, Said S, Tognetti A, Menicucci D. Digital phenotyping for monitoring mental disorders: systematic review. J Med Internet Res. 2023;25:e46778. [FREE Full text] [doi: 10.2196/46778] [Medline: 38090800]
16. Moura I, Teles A, Viana D, Marques J, Coutinho L, Silva F. Digital phenotyping of mental health using multimodal sensing of multiple situations of interest: a systematic literature review. J Biomed Inform. 2023;138:104278. [FREE Full text] [doi: 10.1016/j.jbi.2022.104278] [Medline: 36586498]
17. Tan CYH, Koh JYJ, Ang WW, Tan XM, Koh SWC, Lin W, et al. State-of-the-art digital phenotyping methods for cardiometabolic risk prevention and management: a scoping review. Int J Med Inform. 2026;206:106133. [doi: 10.1016/j.ijmedinf.2025.106133] [Medline: 41086640]
18. Heckler WF, Feijó LP, de Carvalho JV, Barbosa JLV. Digital phenotyping for mental health based on data analytics: a systematic literature review. Artif Intell Med. 2025;163:103094. [doi: 10.1016/j.artmed.2025.103094] [Medline: 40058310]
19. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol. 2018;18(1):143. [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
20. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8(1):19-32. [doi: 10.1080/1364557032000119616]
21. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth. 2020;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
22. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med. 2018;169(7):467-473. [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
23. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S Group. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev. 2021;10(1):39. [FREE Full text] [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
24. Your personal research assistant. Zotero. URL: https://www.zotero.org/ [accessed 2024-12-10]
25. CADIMA. URL: https://www.cadima.info/ [accessed 2024-12-10]
26. Kohl C, McIntosh EJ, Unger S, Haddaway NR, Kecke S, Schiemann J, et al. Correction to: online tools supporting the conduct and reporting of systematic reviews and systematic maps: a case study on CADIMA and review of existing tools. Environ Evid. 2018;7(1):8. [doi: 10.1186/s13750-018-0124-4]
27. Alzueta E, de Zambotti M, Javitz H, Dulai T, Albinni B, Simon KC, et al. Tracking sleep, temperature, heart rate, and daily symptoms across the menstrual cycle with the Oura ring in healthy women. Int J Womens Health. 2022;14:491-503. [FREE Full text] [doi: 10.2147/IJWH.S341917] [Medline: 35422659]
28. Bassett DR, Wyatt HR, Thompson H, Peters JC, Hill JO. Pedometer-measured physical activity and health behaviors in U.S. adults. Med Sci Sports Exerc. 2010;42(10):1819-1825. [FREE Full text] [doi: 10.1249/MSS.0b013e3181dc2e54] [Medline: 20305579]
29. Baumgartner M, Grössl M, Haumer R, Poimer K, Prantl F, Weick K, et al. Neural network-based prediction of perceived sleep quality through wearable device data. Stud Health Technol Inform. 2024;313:221-227. [doi: 10.3233/SHTI240041] [Medline: 38682534]
30. Berrouiguet S, Barrigón ML, Castroman JL, Courtet P, Artés-Rodríguez A, Baca-García E. Combining mobile-health (mHealth) and artificial intelligence (AI) methods to avoid suicide attempts: the Smartcrises study protocol. BMC Psychiatry. 2019;19(1):277. [FREE Full text] [doi: 10.1186/s12888-019-2260-y] [Medline: 31493783]
31. Beukenhorst AL, Collins E, Burke KM, Rahman SM, Clapp M, Konanki SC, et al. Smartphone data during the COVID-19 pandemic can quantify behavioral changes in people with ALS. Muscle Nerve. 2021;63(2):258-262. [FREE Full text] [doi: 10.1002/mus.27110] [Medline: 33118628]
32. Bevier W, Glantz N, Hoppe C, Morales Glass J, Larez A, Chen K, et al. Self-reported and objectively measured physical activity levels among Hispanic/Latino adults with type 2 diabetes. BMJ Open Diabetes Res Care. 2020;8(1):e000893. [FREE Full text] [doi: 10.1136/bmjdrc-2019-000893] [Medline: 32169933]

33. Bian J, Guo Y, Xie M, Parish AE, Wardlaw I, Brown R, et al. Exploring the association between self-reported asthma impact and Fitbit-derived sleep quality and physical activity measures in adolescents. JMIR Mhealth Uhealth. 2017;5(7):e105. [FREE Full text] [doi: 10.2196/mhealth.7346] [Medline: 28743679]
34. Bini SA, Shah RF, Bendich I, Patterson JT, Hwang KM, Zaid MB. Machine learning algorithms can use wearable sensor data to accurately predict six-week patient-reported outcome scores following joint replacement in a prospective trial. J Arthroplasty. 2019;34(10):2242-2247. [doi: 10.1016/j.arth.2019.07.024] [Medline: 31439405]
35. Bjørgaas M, Vik JT, Saeterhaug A, Langlo L, Sakshaug T, Mohus RM, et al. Relationship between pedometer-registered activity, aerobic capacity and self-reported activity and fitness in patients with type 2 diabetes. Diabetes Obes Metab. 2005;7(6):737-744. [doi: 10.1111/j.1463-1326.2004.00464.x] [Medline: 16219018]
36. Brunet J, Tulloch HE, Wolfe Phillips E, Reid RD, Pipe AL, Reed JL. Motivation predicts change in nurses' physical activity levels during a web-based worksite intervention: results from a randomized trial. J Med Internet Res. 2020;22(9):e11543. [FREE Full text] [doi: 10.2196/11543] [Medline: 32915158]
37. Buckholz A, Clarke L, Paik P, Jesudian A, Schwartz R, Krieger A, et al. Evaluating sleep in covert encephalopathy with wearable technology: results from the WATCHES study. Hepatol Commun. 2023;7(2):e0002. [FREE Full text] [doi: 10.1097/HC9.0000000000000028] [Medline: 36724117]
38. Cai Y, Zheng YJ, Cheng CM, Strohl KP, Mason AE, Chang JL. Impact of hypoglossal nerve stimulation on consumer sleep technology metrics and patient symptoms. Laryngoscope. 2024;134(7):3406-3411. [doi: 10.1002/lary.31398] [Medline: 38516821]
39. Čermák J, Pietrucha S, Nawka A, Lipone P, Ruggieri A, Bonelli A, et al. An observational pilot study using a digital phenotyping approach in patients with major depressive disorder treated with trazodone. Front Psychiatry. 2023;14:1127511. [FREE Full text] [doi: 10.3389/fpsyt.2023.1127511] [Medline: 37032913]
40. Chapman BP, Gullapalli BT, Rahman T, Smelson D, Boyer EW, Carreiro S. Impact of individual and treatment characteristics on wearable sensor-based digital biomarkers of opioid use. NPJ Digit Med. 2022;5(1):123. [FREE Full text] [doi: 10.1038/s41746-022-00664-z] [Medline: 35995825]
41. Chen C, Lai F, Huang L, Guo YL. Short- and medium-term cumulative effects of traffic-related air pollution on resting heart rate in the elderly: a wearable device study. Ecotoxicol Environ Saf. 2024;285:117140. [FREE Full text] [doi: 10.1016/j.ecoenv.2024.117140] [Medline: 39368154]
42. Chen J, Wang J, Qi Z, Liu S, Zhao L, Zhang B, et al. Smartwatch measures of outdoor exposure and myopia in children. JAMA Netw Open. 2024;7(8):e2424595. [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.24595] [Medline: 39136948]
43. Choudhary S, Thomas N, Ellenberger J, Srinivasan G, Cohen R. A machine learning approach for detecting digital behavioral patterns of depression using nonintrusive smartphone data (Complementary Path to Patient Health Questionnaire-9 Assessment): prospective observational study. JMIR Form Res. 2022;6(5):e37736. [FREE Full text] [doi: 10.2196/37736] [Medline: 35420993]
44. Cochrane SK, Chen S, Fitzgerald JD, Dodson JA, Fielding RA, King AC, et al. LIFE Study Research Group. Association of accelerometry-measured physical activity and cardiovascular events in mobility-limited older adults: the LIFE (lifestyle interventions and independence for elders) study. J Am Heart Assoc. 2017;6(12):e007215. [FREE Full text] [doi: 10.1161/JAHA.117.007215] [Medline: 29197830]
45. Concheiro-Moscoso P, Groba B, Martínez-Martínez FJ, Miranda-Duro MDC, Nieto-Riveiro L, Pousada T, et al. Use of the Xiaomi Mi Band for sleep monitoring and its influence on the daily life of older people living in a nursing home. Digit Health. 2022;8:20552076221121162. [FREE Full text] [doi: 10.1177/20552076221121162] [Medline: 36060611]
46. Cormack F, McCue M, Skirrow C, Cashdollar N, Taptiklis N, van Schaik T, et al. Characterizing longitudinal patterns in cognition, mood, and activity in depression with 6-week high-frequency wearable assessment: observational study. JMIR Ment Health. 2024;11:e46895. [FREE Full text] [doi: 10.2196/46895] [Medline: 38819909]
47. Cote DJ, Barnett I, Onnela J, Smith TR. Digital phenotyping in patients with spine disease: a novel approach to quantifying mobility and quality of life. World Neurosurg. 2019;126:e241-e249. [FREE Full text] [doi: 10.1016/j.wneu.2019.01.297] [Medline: 30797933]
48. Dang K, Ritvo P, Katz J, Gratzer D, Knyahnytska Y, Ortiz A, et al. The role of daily steps in the treatment of major depressive disorder: secondary analysis of a randomized controlled trial of a 6-month internet-based, mindfulness-based cognitive behavioral therapy intervention for youth. Interact J Med Res. 2023;12:e46419. [FREE Full text] [doi: 10.2196/46419] [Medline: 38064262]
49. DasMahapatra P, Chiauzzi E, Bhalerao R, Rhodes J. Free-living physical activity monitoring in adult US patients with multiple sclerosis using a consumer wearable device. Digit Biomark. 2018;2(1):47-63. [FREE Full text] [doi: 10.1159/000488040] [Medline: 32095756]
50. Deng H, Abouzeid CA, Shepler LJ, Ni P, Slavin MD, Barron DS, et al. Moderation effects of daily behavior on associations between symptoms and social participation outcomes after burn injury: a 6-month digital phenotyping study. Arch Phys Med Rehabil. 2024;105(9):1700-1708. [doi: 10.1016/j.apmr.2024.05.011] [Medline: 38754720]
51. Dohrn I, Hagströmer M, Hellénius M-L, Ståhle A. Short- and long-term effects of balance training on physical activity in older adults with osteoporosis: a randomized controlled trial. J Geriatr Phys Ther. 2017;40(2):102-111. [FREE Full text] [doi: 10.1519/JPT.0000000000000077] [Medline: 26859463]

52. Doryab A, Villalba DK, Chikersal P, Dutcher JM, Tumminia M, Liu X, et al. Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: statistical analysis, data mining and machine learning of smartphone and Fitbit data. JMIR Mhealth Uhealth. 2019;7(7):e13209. [FREE Full text] [doi: 10.2196/13209] [Medline: 31342903]
53. Ensari I, Caceres BA, Jackman KB, Suero-Tejeda N, Shechter A, Odlum ML, et al. Digital phenotyping of sleep patterns among heterogenous samples of Latinx adults using unsupervised learning. Sleep Med. 2021;85:211-220. [FREE Full text] [doi: 10.1016/j.sleep.2021.07.023] [Medline: 34364092]
54. Fagherazzi G, El Fatouhi D, Bellicha A, El Gareh A, Affret A, Dow C, et al. An international study on the determinants of poor sleep amongst 15,000 users of connected devices. J Med Internet Res. 2017;19(10):e363. [FREE Full text] [doi: 10.2196/jmir.7930] [Medline: 29061551]
55. Faurholt-Jepsen M, Ritz C, Frost M, Mikkelsen RL, Margrethe Christensen E, Bardram J, et al. Mood instability in bipolar disorder type I versus type II-continuous daily electronic self-monitoring of illness activity using smartphones. J Affect Disord. 2015;186:342-349. [doi: 10.1016/j.jad.2015.06.026] [Medline: 26277270]
56. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as an electronic biomarker of illness activity in bipolar disorder. Bipolar Disord. 2015;17(7):715-728. [doi: 10.1111/bdi.12332] [Medline: 26395972]
57. Finkelstein EA, Haaland BA, Bilger M, Sahasranaman A, Sloan RA, Nang EEK, et al. Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial. Lancet Diabetes Endocrinol. 2016;4(12):983-995. [doi: 10.1016/S2213-8587(16)30284-4] [Medline: 27717766]
58. Ganglberger W, Bucklin AA, Tesh RA, Da Silva Cardoso M, Sun H, Leone MJ, et al. Sleep apnea and respiratory anomaly detection from a wearable band and oxygen saturation. Sleep Breath. 2022;26(3):1033-1044. [FREE Full text] [doi: 10.1007/s11325-021-02465-2] [Medline: 34409545]
59. Golbus JR, Gosch K, Birmingham MC, Butler J, Lingvay I, Lanfear DE, et al. Association between wearable device measured activity and patient-reported outcomes for heart failure. JACC Heart Fail. 2023;11(11):1521-1530. [FREE Full text] [doi: 10.1016/j.jchf.2023.05.033] [Medline: 37498273]
60. Golbus JR, Pescatore NA, Nallamothu BK, Shah N, Kheterpal S. Lancet Digit Health. 2021;3(11):e707-e715. [FREE Full text] [doi: 10.1016/S2589-7500(21)00138-2] [Medline: 34711377]
61. Greenberg JK, Frumkin M, Xu Z, Zhang J, Javeed S, Zhang JK, et al. Preoperative mobile health data improve predictions of recovery from lumbar spine surgery. Neurosurgery. 2024;95(3):617-626. [doi: 10.1227/neu.0000000000002911] [Medline: 38551340]
62. Gresham G, Hendifar AE, Spiegel B, Neeman E, Tuli R, Rimel BJ, et al. Wearable activity monitors to assess performance status and predict clinical outcomes in advanced cancer patients. NPJ Digit Med. 2018;1:27. [FREE Full text] [doi: 10.1038/s41746-018-0032-6] [Medline: 31304309]
63. Griffiths EA, Min JS, Lee W, Yu JC, Patel Y, Myren K, et al. Patient-reported outcomes and daily activity assessed with a digital wearable device in patients with paroxysmal nocturnal hemoglobinuria treated with ravulizumab: REVEAL, a prospective, observational study. Health Qual Life Outcomes. 2024;22(1):62. [FREE Full text] [doi: 10.1186/s12955-024-02279-2] [Medline: 39123253]
64. Groat D, Kwon HJ, Grando MA, Cook CB, Thompson B. Comparing real-time self-tracking and device-recorded exercise data in subjects with type 1 diabetes. Appl Clin Inform. 2018;9(4):919-926. [FREE Full text] [doi: 10.1055/s-0038-1676458] [Medline: 30586673]
65. Gupta V, Kariotis S, Rajab MD, Errington N, Alhathli E, Jammeh E, et al. Unsupervised machine learning to investigate trajectory patterns of COVID-19 symptoms and physical activity measured via the MyHeart Counts app and smart devices. NPJ Digit Med. 2023;6(1):239. [FREE Full text] [doi: 10.1038/s41746-023-00974-w] [Medline: 38135699]
66. Han M, Williams S, Mendoza M, Ye X, Zhang H, Calice-Silva V, et al. Quantifying physical activity levels and sleep in hemodialysis patients using a commercially available activity tracker. Blood Purif. 2016;41(1-3):194-204. [doi: 10.1159/000441314] [Medline: 26765515]
67. Hardcastle SJ, Jiménez-Castuera R, Maxwell-Smith C, Bulsara MK, Hince D. Fitbit wear-time and patterns of activity in cancer survivors throughout a physical activity intervention and follow-up: exploratory analysis from a randomised controlled trial. PLoS One. 2020;15(10):e0240967. [FREE Full text] [doi: 10.1371/journal.pone.0240967] [Medline: 33075100]
68. Hartman SJ, Chen R, Tam RM, Narayan HK, Natarajan L, Liu L. Fitbit use and activity levels from intervention to 2 years after: secondary analysis of a randomized controlled trial. JMIR Mhealth Uhealth. 2022;10(6):e37086. [FREE Full text] [doi: 10.2196/37086] [Medline: 35771607]
69. Hedrick TL, Hassinger TE, Myers E, Krebs ED, Chu D, Charles AN, et al. Wearable technology in the perioperative period: predicting risk of postoperative complications in patients undergoing elective colorectal surgery. Dis Colon Rectum. 2020;63(4):538-544. [doi: 10.1097/DCR.0000000000001580] [Medline: 32015289]
70. Henson P, Rodriguez-Villa E, Torous J. Investigating associations between screen time and symptomatology in individuals with serious mental illness: longitudinal observational study. J Med Internet Res. 2021;23(3):e23144. [FREE Full text] [doi: 10.2196/23144] [Medline: 33688835]

71. Hinterberger A, Eigl E, Schwemlein RN, Topalidis P, Schabus M. Investigating the subjective and objective efficacy of a cognitive behavioural therapy for insomnia (CBT-I)-based smartphone app on sleep: a randomised controlled trial. J Sleep Res. 2024;33(4):e14136. [doi: 10.1111/jsr.14136] [Medline: 38156655]
72. Hirten RP, Danieletto M, Tomalin L, Choi KH, Zweig M, Golden E, et al. Factors associated with longitudinal psychological and physiological stress in health care workers during the COVID-19 pandemic: observational study using Apple Watch data. J Med Internet Res. 2021;23(9):e31295. [FREE Full text] [doi: 10.2196/31295] [Medline: 34379602]
73. Hirten RP, Danieletto M, Tomalin L, Choi KH, Zweig M, Golden E, et al. Use of physiological data from a wearable device to identify SARS-CoV-2 infection and symptoms and predict COVID-19 diagnosis: observational study. J Med Internet Res. 2021;23(2):e26107. [FREE Full text] [doi: 10.2196/26107] [Medline: 33529156]
74. Hirten RP, Suprun M, Danieletto M, Zweig M, Golden E, Pyzik R, et al. A machine learning approach to determine resilience utilizing wearable device data: analysis of an observational cohort. JAMIA Open. 2023;6(2):ooad029. [FREE Full text] [doi: 10.1093/jamiaopen/ooad029] [Medline: 37143859]
75. Horwitz A, Czyz E, Al-Dajani N, Dempsey W, Zhao Z, Nahum-Shani I, et al. Utilizing daily mood diaries and wearable sensor data to predict depression and suicidal ideation among medical interns. J Affect Disord. 2022;313:1-7. [FREE Full text] [doi: 10.1016/j.jad.2022.06.064] [Medline: 35764227]
76. Huang Q, Lin H, Xiao H, Zhang L, Chen D, Dai X. Sleeping more than 8 h: a silent factor contributing to decreased muscle mass in Chinese community-dwelling older adults. BMC Public Health. 2024;24(1):1246. [FREE Full text] [doi: 10.1186/s12889-024-18520-y] [Medline: 38711104]
77. Hurwitz E, Butzin-Dozier Z, Master H, O'Neil ST, Walden A, Holko M, et al. Harnessing consumer wearable digital biomarkers for individualized recognition of postpartum depression using the All of Us Research Program data set: cross-sectional study. JMIR Mhealth Uhealth. 2024;12:e54622. [FREE Full text] [doi: 10.2196/54622] [Medline: 38696234]
78. Ikäheimonen A, Luong N, Baryshnikov I, Darst R, Heikkilä R, Holmen J, et al. Predicting and monitoring symptoms in patients diagnosed with depression using smartphone data: observational study. J Med Internet Res. 2024;26:e56874. [FREE Full text] [doi: 10.2196/56874] [Medline: 39626241]
79. Jacobson NC, Lekkas D, Huang R, Thomas N. Deep learning paired with wearable passive sensing data predicts deterioration in anxiety disorder symptoms across 17-18 years. J Affect Disord. 2021;282:104-111. [FREE Full text] [doi: 10.1016/j.jad.2020.12.086] [Medline: 33401123]
80. Jacobson NC, Weingarden H, Wilhelm S. Using digital phenotyping to accurately detect depression severity. J Nerv Ment Dis. 2019;207(10):893-896. [doi: 10.1097/NMD.0000000000001042] [Medline: 31596769]
81. Johnson SA, Karas M, Burke KM, Straczkiewicz M, Scheier ZA, Clark AP, et al. Wearable device and smartphone data quantify ALS progression and may provide novel outcome measures. NPJ Digit Med. 2023;6(1):34. [FREE Full text] [doi: 10.1038/s41746-023-00778-y] [Medline: 36879025]
82. Jonasdottir S, Minor K, Lehmann S. Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study. Sleep. 2021;44(2):zsaa169. [doi: 10.1093/sleep/zsaa169] [Medline: 32886772]
83. Jonathan GK, Abitante G, McBride A, Bernstein-Sandler M, Babington P, Dopke CA, et al. LiveWell, a smartphone-based self-management intervention for bipolar disorder: intervention participation and usability analysis. J Affect Disord. 2024;350:926-936. [doi: 10.1016/j.jad.2024.01.099] [Medline: 38246280]
84. Kalmbach DA, Fang Y, Arnedt JT, Cochran AL, Deldin PJ, Kaplin AI, et al. Effects of sleep, physical activity, and shift work on daily mood: a prospective mobile monitoring study of medical interns. J Gen Intern Med. 2018;33(6):914-920. [FREE Full text] [doi: 10.1007/s11606-018-4373-2] [Medline: 29542006]
85. Kańtoch E, Kańtoch A. Cardiovascular and pre-frailty risk assessment during shelter-in-place measures based on multimodal biomarkers collected from smart telemedical wearables. J Clin Med. 2021;10(9):1997. [FREE Full text] [doi: 10.3390/jcm10091997] [Medline: 34066571]
86. Karas M, Huang D, Clement Z, Millner AJ, Kleiman EM, Bentley KH, et al. Smartphone screen time characteristics in people with suicidal thoughts: retrospective observational data analysis study. JMIR Mhealth Uhealth. 2024;12:e57439. [FREE Full text] [doi: 10.2196/57439] [Medline: 39392706]
87. Karas M, Olsen J, Straczkiewicz M, Johnson SA, Burke KM, Iwasaki S, et al. Tracking amyotrophic lateral sclerosis disease progression using passively collected smartphone sensor data. Ann Clin Transl Neurol. 2024;11(6):1380-1392. [FREE Full text] [doi: 10.1002/acn3.52050] [Medline: 38816946]
88. Kaura A, Sztriha L, Chan FK, Aeron-Thomas J, Gall N, Piechowski-Jozwiak B, et al. Early prolonged ambulatory cardiac monitoring in stroke (EPACS): an open-label randomised controlled trial. Eur J Med Res. 2019;24(1):25. [FREE Full text] [doi: 10.1186/s40001-019-0383-8] [Medline: 31349792]
89. Kim DH, Nam KH, Choi BK, Han IH, Jeon TJ, Park SY. The usefulness of a wearable device in daily physical activity monitoring for the hospitalized patients undergoing lumbar surgery. J Korean Neurosurg Soc. 2019;62(5):561-566. [FREE Full text] [doi: 10.3340/jkns.2018.0131] [Medline: 31337197]
90. Kim Y, Wijndaele K, Sharp SJ, Strain T, Pearce M, White T, et al. Specific physical activities, sedentary behaviours and sleep as long-term predictors of accelerometer-measured physical activity in 91,648 adults: a prospective cohort study. Int J Behav Nutr Phys Act. 2019;16(1):41. [FREE Full text] [doi: 10.1186/s12966-019-0802-9] [Medline: 31064403]

91. Konda S, Ogasawara I, Fujita K, Aoyama C, Yokoyama T, Magome T, et al. Variability in physical inactivity responses of university students during COVID-19 pandemic: a monitoring of daily step counts using a smartphone application. Int J Environ Res Public Health. 2022;19(4):1958. [FREE Full text] [doi: 10.3390/ijerph19041958] [Medline: 35206149]
92. Kringle EA, Tucker D, Wu Y, Lv N, Kannampallil T, Barve A, et al. Associations between daily step count trajectories and clinical outcomes among adults with comorbid obesity and depression. Ment Health Phys Act. 2023;24:100512. [FREE Full text] [doi: 10.1016/j.mhpa.2023.100512] [Medline: 37206660]
93. Lee J, Kong S, Shin S, Lee G, Kim HK, Shim YM, et al. Wearable device-based intervention for promoting patient physical activity after lung cancer surgery: a nonrandomized clinical trial. JAMA Netw Open. 2024;7(9):e2434180. [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.34180] [Medline: 39302678]
94. Liao Y, Robertson MC, Winne A, Wu IHC, Le TA, Balachandran DD, et al. Investigating the within-person relationships between activity levels and sleep duration using Fitbit data. Transl Behav Med. 2021;11(2):619-624. [FREE Full text] [doi: 10.1093/tbm/ibaa071] [Medline: 32667039]
95. Lim SER, Dodds R, Bacon D, Sayer AA, Roberts HC. Physical activity among hospitalised older people: insights from upper and lower limb accelerometry. Aging Clin Exp Res. 2018;30(11):1363-1369. [FREE Full text] [doi: 10.1007/s40520-018-0930-0] [Medline: 29542070]
96. Lim WK, Davila S, Teo JX, Yang C, Pua CJ, Blöcker C, et al. Beyond fitness tracking: the use of consumer-grade wearable data from normal volunteers in cardiovascular and lipidomics research. PLoS Biol. 2018;16(2):e2004285. [FREE Full text] [doi: 10.1371/journal.pbio.2004285] [Medline: 29485983]
97. Liu G, Henson P, Keshavan M, Pekka-Onnela J, Torous J. Assessing the potential of longitudinal smartphone based cognitive assessment in schizophrenia: a naturalistic pilot study. Schizophr Res Cogn. 2019;17:100144. [FREE Full text] [doi: 10.1016/j.scog.2019.100144] [Medline: 31024801]
98. Low CA, Bovbjerg DH, Ahrendt S, Choudry MH, Holtzman M, Jones HL, et al. Fitbit step counts during inpatient recovery from cancer surgery as a predictor of readmission. Ann Behav Med. 2018;52(1):88-92. [FREE Full text] [doi: 10.1093/abm/kax022] [Medline: 29538623]
99. Lowe SS, Danielson B, Beaumont C, Watanabe SM, Baracos VE, Courneya KS. Associations between objectively measured physical activity and quality of life in cancer patients with brain metastases. J Pain Symptom Manage. 2014;48(3):322-332. [FREE Full text] [doi: 10.1016/j.jpainsymman.2013.10.012] [Medline: 24630754]
100. Luong N, Mark G, Kulshrestha J, Aledavood T. Sleep during the COVID-19 pandemic: longitudinal observational study combining multisensor data with questionnaires. JMIR Mhealth Uhealth. 2024;12:e53389. [FREE Full text] [doi: 10.2196/53389] [Medline: 39226100]
101. Mack DL, DaSilva AW, Rogers C, Hedlund E, Murphy EI, Vojdanovski V, et al. Mental health and behavior of college students during the COVID-19 pandemic: longitudinal mobile smartphone and ecological momentary assessment study, part II. J Med Internet Res. 2021;23(6):e28892. [FREE Full text] [doi: 10.2196/28892] [Medline: 33900935]
102. Martinez GJ, Grover T, Mattingly SM, Mark G, D'Mello S, Aledavood T, et al. Alignment between heart rate variability from fitness trackers and perceived stress: perspectives from a large-scale in situ longitudinal study of information workers. JMIR Hum Factors. 2022;9(3):e33754. [FREE Full text] [doi: 10.2196/33754] [Medline: 35925662]
103. Massar S, Ong J, Lau T, Ng B, Chan L, Koek D, et al. Working-from-home persistently influences sleep and physical activity 2 years after the COVID-19 pandemic onset: a longitudinal sleep tracker and electronic diary-based study. Front Psychol. 2023;14:1145893. [FREE Full text] [doi: 10.3389/fpsyg.2023.1145893] [Medline: 37213365]
104. Matcham F, Carr E, Meyer N, White K, Oetzmann C, Leightley D, et al. RADAR-CNS consortium. The relationship between wearable-derived sleep features and relapse in major depressive disorder. J Affect Disord. 2024;363:90-98. [FREE Full text] [doi: 10.1016/j.jad.2024.07.136] [Medline: 39038618]
105. McDermott MM, Spring B, Berger JS, Treat-Jacobson D, Conte MS, Creager MA, et al. Effect of a home-based exercise intervention of wearable technology and telephone coaching on walking performance in peripheral artery disease: the HONOR randomized clinical trial. JAMA. 2018;319(16):1665-1676. [FREE Full text] [doi: 10.1001/jama.2018.3275] [Medline: 29710165]
106. McDonald SM, Yeo S, Liu J, Wilcox S, Sui X, Pate RR. Association between change in maternal physical activity during pregnancy and infant size, in a sample overweight or obese women. Women Health. 2020;60(8):929-938. [FREE Full text] [doi: 10.1080/03630242.2020.1779904] [Medline: 32588785]
107. McGinnis EW, Lunna S, Berman I, Loftness BC, Bagdon S, Danforth CM, et al. Discovering digital biomarkers of panic attack risk in consumer wearables data. Annu Int Conf IEEE Eng Med Biol Soc. 2023;2023:1-4. [doi: 10.1109/EMBC40787.2023.10339982] [Medline: 38083448]
108. McNeil J, Brenner DR, Stone CR, O'Reilly R, Ruan Y, Vallance JK, et al. Activity tracker to prescribe various exercise intensities in breast cancer survivors. Med Sci Sports Exerc. 2019;51(5):930-940. [doi: 10.1249/MSS.0000000000001890] [Medline: 30694978]
109. Mehl CV, Benum SD, Aakvik KAD, Kongsvold A, Mork PJ, Kajantie E, et al. Physical activity and associations with health-related quality of life in adults born small for gestational age at term: a prospective cohort study. BMC Pediatr. 2023;23(1):430. [FREE Full text] [doi: 10.1186/s12887-023-04256-y] [Medline: 37641030]

110. Mesquita R, Spina G, Pitta F, Donaire-Gonzalez D, Deering BM, Patel MS, et al. Physical activity patterns and clusters in 1001 patients with COPD. Chron Respir Dis. 2017;14(3):256-269. [FREE Full text] [doi: 10.1177/1479972316687207] [Medline: 28774199]
111. Mezlini A, Shapiro A, Daza EJ, Caddigan E, Ramirez E, Althoff T, et al. Estimating the burden of influenza-like illness on daily activity at the population scale using commercial wearable sensors. JAMA Netw Open. 2022;5(5):e2211958. [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.11958] [Medline: 35552722]
112. Minhas J, Shou H, Hershman S, Zamanian R, Ventetuolo CE, Bull TM, et al. Physical activity and its association with traditional outcome measures in pulmonary arterial hypertension. Ann Am Thorac Soc. 2022;19(4):572-582. [FREE Full text] [doi: 10.1513/AnnalsATS.202105-560OC] [Medline: 34473938]
113. Mishra R, Park C, York MK, Kunik ME, Wung S, Naik AD, et al. Decrease in mobility during the COVID-19 pandemic and its association with increase in depression among older adults: a longitudinal remote mobility monitoring using a wearable sensor. Sensors (Basel). 2021;21(9):3090. [FREE Full text] [doi: 10.3390/s21093090] [Medline: 33946664]
114. Mobbs RJ, Phan K, Maharaj M, Rao PJ. Physical activity measured with accelerometer and self-rated disability in lumbar spine surgery: a prospective study. Global Spine J. 2016;6(5):459-464. [FREE Full text] [doi: 10.1055/s-0035-1565259] [Medline: 27433430]
115. Morcos MW, Teeter MG, Somerville LE, Lanting B. Correlation between hip osteoarthritis and the level of physical activity as measured by wearable technology and patient-reported questionnaires. J Orthop. 2020;20:236-239. [FREE Full text] [doi: 10.1016/j.jor.2019.11.049] [Medline: 32071522]
116. Morimoto M, Nawari A, Savic R, Marmor M. Exploring the potential of a Smart Ring to predict postoperative pain outcomes in orthopedic surgery patients. Sensors (Basel). 2024;24(15):5024. [FREE Full text] [doi: 10.3390/s24155024] [Medline: 39124071]
117. Mousavi ZA, Lai J, Simon K, Rivera AP, Yunusova A, Hu S, et al. Sleep patterns and affect dynamics among college students during the COVID-19 pandemic: intensive longitudinal study. JMIR Form Res. 2022;6(8):e33964. [FREE Full text] [doi: 10.2196/33964] [Medline: 35816447]
118. Moyle W, Jones C, Murfield J, Thalib L, Beattie E, Shum D, et al. Effect of a robotic seal on the motor activity and sleep patterns of older people with dementia, as measured by wearable technology: a cluster-randomised controlled trial. Maturitas. 2018;110:10-17. [FREE Full text] [doi: 10.1016/j.maturitas.2018.01.007] [Medline: 29563027]
119. Nawabi NLA, Emedom-Nnamdi P, Kilgallon JL, Gerstl JVE, Cote DJ, Jha R, et al. Assessing mobility in patients with glioblastoma using digital phenotyping-piloting the digital assessment in neuro-oncology. Neurosurgery. 2025;96(1):183-192. [doi: 10.1227/neu.0000000000003051] [Medline: 38912791]
120. Nelson BW, Flannery JE, Flournoy J, Duell N, Prinstein MJ, Telzer E. Concurrent and prospective associations between Fitbit wearable-derived RDoC arousal and regulatory constructs and adolescent internalizing symptoms. J Child Psychol Psychiatry. 2022;63(3):282-295. [doi: 10.1111/jcpp.13471] [Medline: 34184767]
121. Niela-Vilén H, Auxier J, Ekholm E, Sarhaddi F, Asgari Mehrabadi M, Mahmoudzadeh A, et al. Pregnant women's daily patterns of well-being before and during the COVID-19 pandemic in Finland: longitudinal monitoring through smartwatch technology. PLoS One. 2021;16(2):e0246494. [FREE Full text] [doi: 10.1371/journal.pone.0246494] [Medline: 33534854]
122. Orme MW, Harvey-Dunstan TC, Boral I, Chaplin EJL, Hussain SF, Morgan MDL, et al. Changes in physical activity during hospital admission for chronic respiratory disease. Respirology. 2019;24(7):652-657. [FREE Full text] [doi: 10.1111/resp.13513] [Medline: 30845363]
123. Östlind E, Eek F, Stigmar K, Sant'Anna A, Hansson EE. Promoting work ability with a wearable activity tracker in working age individuals with hip and/or knee osteoarthritis: a randomized controlled trial. BMC Musculoskelet Disord. 2022;23(1):112. [FREE Full text] [doi: 10.1186/s12891-022-05041-1] [Medline: 35114983]
124. Pakhomov SVS, Thuras PD, Finzel R, Eppel J, Kotlyar M. Using consumer-wearable technology for remote assessment of physiological response to stress in the naturalistic environment. PLoS One. 2020;15(3):e0229942. [FREE Full text] [doi: 10.1371/journal.pone.0229942] [Medline: 32210441]
125. Panda N, Solsky I, Cauley CE, Lipsitz S, Desai EV, Huang EJ, et al. Smartphone-based assessment of preoperative decision conflict and postoperative physical activity among patients undergoing cancer surgery: a prospective cohort study. Ann Surg. 2022;276(1):193-199. [doi: 10.1097/SLA.0000000000004487] [Medline: 32941270]
126. Panda N, Solsky I, Huang EJ, Lipsitz S, Pradarelli JC, Delisle M, et al. Using smartphones to capture novel recovery metrics after cancer surgery. JAMA Surg. 2020;155(2):123-129. [FREE Full text] [doi: 10.1001/jamasurg.2019.4702] [Medline: 31657854]
127. Pellegrini AM, Huang EJ, Staples PC, Hart KL, Lorme JM, Brown HE, et al. Estimating longitudinal depressive symptoms from smartphone data in a transdiagnostic cohort. Brain Behav. 2022;12(2):e02077. [FREE Full text] [doi: 10.1002/brb3.2077] [Medline: 35076166]
128. Perry AS, Annis JS, Master H, Nayor M, Hughes A, Kouame A, et al. Association of longitudinal activity measures and diabetes risk: an analysis from the National Institutes of Health All of Us Research Program. J Clin Endocrinol Metab. 2023;108(5):1101-1109. [FREE Full text] [doi: 10.1210/clinem/dgac695] [Medline: 36458881]

129. Pozehl BJ, Mcguire R, Duncan K, Hertzog M, Deka P, Norman J, et al. Accelerometer-measured daily activity levels and related factors in patients with heart failure. J Cardiovasc Nurs. 2018;33(4):329-335. [FREE Full text] [doi: 10.1097/JCN.0000000000000464] [Medline: 29538050]
130. Pradhan S, Kelly VE. Quantifying physical activity in early Parkinson disease using a commercial activity monitor. Parkinsonism Relat Disord. 2019;66:171-175. [FREE Full text] [doi: 10.1016/j.parkreldis.2019.08.001] [Medline: 31420310]
131. Price GD, Heinz MV, Song SH, Nemesure MD, Jacobson NC. Using digital phenotyping to capture depression symptom variability: detecting naturalistic variability in depression symptoms across one year using passively collected wearable movement and sleep data. Transl Psychiatry. 2023;13(1):381. [FREE Full text] [doi: 10.1038/s41398-023-02669-y] [Medline: 38071317]
132. Qirtas MM, Zafeiridi E, White EB, Pesch D. The relationship between loneliness and depression among college students: mining data derived from passive sensing. Digit Health. 2023;9:20552076231211104. [FREE Full text] [doi: 10.1177/20552076231211104] [Medline: 38025106]
133. Ranjan T, Melcher J, Keshavan M, Smith M, Torous J. Longitudinal symptom changes and association with home time in people with schizophrenia: an observational digital phenotyping study. Schizophr Res. 2022;243:64-69. [doi: 10.1016/j.schres.2022.02.031] [Medline: 35245703]
134. Rao C, Di Lascio E, Demanse D, Marshall N, Sopala M, De Luca V. Association of digital measures and self-reported fatigue: a remote observational study in healthy participants and participants with chronic inflammatory rheumatic disease. Front Digit Health. 2023;5:1099456. [FREE Full text] [doi: 10.3389/fdgth.2023.1099456] [Medline: 37426890]
135. Rezaei N, Grandner MA. Changes in sleep duration, timing, and variability during the COVID-19 pandemic: large-scale Fitbit data from 6 major US cities. Sleep Health. 2021;7(3):303-313. [FREE Full text] [doi: 10.1016/j.sleh.2021.02.008] [Medline: 33771534]
136. Robbins R, Affouf M, Seixas A, Beaugris L, Avirappattu G, Jean-Louis G. Four-year trends in sleep duration and quality: a longitudinal study using data from a commercially available sleep tracker. J Med Internet Res. 2020;22(2):e14735. [FREE Full text] [doi: 10.2196/14735] [Medline: 32078573]
137. Rowan SP, Lilly CL, Claydon EA, Wallace J, Merryman K. Monitoring one heart to help two: heart rate variability and resting heart rate using wearable technology in active women across the perinatal period. BMC Pregnancy Childbirth. 2022;22(1):887. [FREE Full text] [doi: 10.1186/s12884-022-05183-z] [Medline: 36451120]
138. Rykov YG, Patterson MD, Gangwar BA, Jabar SB, Leonardo J, Ng KP, et al. Predicting cognitive scores from wearable-based digital physiological features using machine learning: data from a clinical trial in mild cognitive impairment. BMC Med. 2024;22(1):36. [FREE Full text] [doi: 10.1186/s12916-024-03252-y] [Medline: 38273340]
139. Rykov Y, Thach T, Bojic I, Christopoulos G, Car J. Digital biomarkers for depression screening with wearable devices: cross-sectional study with machine learning modeling. JMIR Mhealth Uhealth. 2021;9(10):e24872. [FREE Full text] [doi: 10.2196/24872] [Medline: 34694233]
140. Rykov Y, Thach T, Dunleavy G, Roberts AC, Christopoulos G, Soh CK, et al. Activity tracker-based metrics as digital markers of cardiometabolic health in working adults: cross-sectional study. JMIR Mhealth Uhealth. 2020;8(1):e16409. [FREE Full text] [doi: 10.2196/16409] [Medline: 32012098]
141. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res. 2015;17(7):e175. [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]
142. Sakal C, Li T, Li J, Yang C, Li X. Association between sleep efficiency variability and cognition among older adults: cross-sectional accelerometer study. JMIR Aging. 2024;7:e54353. [FREE Full text] [doi: 10.2196/54353] [Medline: 38596863]
143. Sano A, Taylor S, McHill AW, Phillips AJ, Barger LK, Klerman E, et al. Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: observational study. J Med Internet Res. 2018;20(6):e210. [FREE Full text] [doi: 10.2196/jmir.9410] [Medline: 29884610]
144. Sarda A, Munuswamy S, Sarda S, Subramanian V. Using passive smartphone sensing for improved risk stratification of patients with depression and diabetes: cross-sectional observational study. JMIR Mhealth Uhealth. 2019;7(1):e11041. [FREE Full text] [doi: 10.2196/11041] [Medline: 30694197]
145. Sarhaddi F, Azimi I, Niela-Vilen H, Axelin A, Liljeberg P, Rahmani AM. Maternal social loneliness detection using passive sensing through continuous monitoring in everyday settings: longitudinal study. JMIR Form Res. 2023;7:e47950. [FREE Full text] [doi: 10.2196/47950] [Medline: 37556183]
146. Scott H, Lechat B, Guyett A, Reynolds AC, Lovato N, Naik G, et al. Sleep irregularity is associated with hypertension: findings from over 2 million nights with a large global population sample. Hypertension. 2023;80(5):1117-1126. [doi: 10.1161/HYPERTENSIONAHA.122.20513] [Medline: 36974682]
147. Scott H, Naik G, Lechat B, Manners J, Fitton J, Nguyen DP, et al. Are we getting enough sleep? Frequent irregular sleep found in an analysis of over 11 million nights of objective in-home sleep data. Sleep Health. 2024;10(1):91-97. [FREE Full text] [doi: 10.1016/j.sleh.2023.10.016] [Medline: 38071172]

148. Semaan S, Dewland TA, Tison GH, Nah G, Vittinghoff E, Pletcher MJ, et al. Physical activity and atrial fibrillation: data from wearable fitness trackers. Heart Rhythm. 2020;17(5 Pt B):842-846. [FREE Full text] [doi: 10.1016/j.hrthm.2020.02.013] [Medline: 32354448]
149. Shin S, Kim S. Rotating between day and night shifts: factors influencing sleep patterns of hospital nurses. J Clin Nurs. 2021;30(21-22):3182-3193. [doi: 10.1111/jocn.15819] [Medline: 34046951]
150. Silverman-Lloyd LG, Kianoush S, Blaha MJ, Sabina AB, Graham GN, Martin SS. mActive-Smoke: a prospective observational study using mobile health tools to assess the association of physical activity with smoking urges. JMIR Mhealth Uhealth. 2018;6(5):e121. [FREE Full text] [doi: 10.2196/mhealth.9292] [Medline: 29752250]
151. Song YM, Jeong J, de Los Reyes AA, Lim D, Cho C, Yeom JW, et al. Causal dynamics of sleep, circadian rhythm, and mood symptoms in patients with major depression and bipolar disorder: insights from longitudinal wearable device data. EBioMedicine. 2024;103:105094. [FREE Full text] [doi: 10.1016/j.ebiom.2024.105094] [Medline: 38579366]
152. Sprint G, Cook D, Weeks D, Dahmen J, La Fleur A. Analyzing sensor-based time series data to track changes in physical activity during inpatient rehabilitation. Sensors (Basel). 2017;17(10):2219. [FREE Full text] [doi: 10.3390/s17102219] [Medline: 28953257]
153. Steinhubl SR, Waalen J, Edwards AM, Ariniello LM, Mehta RR, Ebner GS, et al. Effect of a home-based wearable continuous ECG monitoring patch on detection of undiagnosed atrial fibrillation: the mSToPS randomized clinical trial. JAMA. 2018;320(2):146-155. [FREE Full text] [doi: 10.1001/jama.2018.8102] [Medline: 29998336]
154. Stewart C, Ranjan Y, Conde P, Sun S, Zhang Y, Rashid Z, et al. Physiological presentation and risk factors of long COVID in the UK using smartphones and wearable devices: a longitudinal, citizen science, case-control study. Lancet Digit Health. 2024;6(9):e640-e650. [FREE Full text] [doi: 10.1016/S2589-7500(24)00140-7] [Medline: 39138096]
155. Strain T, Wijndaele K, Dempsey PC, Sharp SJ, Pearce M, Jeon J, et al. Wearable-device-measured physical activity and future health risk. Nat Med. 2020;26(9):1385-1391. [FREE Full text] [doi: 10.1038/s41591-020-1012-3] [Medline: 32807930]
156. Su JG, Vuong V, Shahriary E, Aslebagh S, Yakutis E, Sage E, et al. Health effects of air pollution on respiratory symptoms: a longitudinal study using digital health sensors. Environ Int. 2024;189:108810. [FREE Full text] [doi: 10.1016/j.envint.2024.108810] [Medline: 38875815]
157. Takano K, Oba T, Katahira K, Kimura K. Deconstructing Fitbit to specify the effective features in promoting physical activity among inactive adults: pilot randomized controlled trial. JMIR Mhealth Uhealth. 2024;12:e51216. [FREE Full text] [doi: 10.2196/51216] [Medline: 38996332]
158. Teo JX, Davila S, Yang C, Hii AA, Pua CJ, Yap J, et al. Digital phenotyping by consumer wearables identifies sleep-associated markers of cardiovascular disease risk and biological aging. Commun Biol. 2019;2:361. [FREE Full text] [doi: 10.1038/s42003-019-0605-1] [Medline: 31602410]
159. Téoule J, Woll C, Ray J, Sütterlin M, Filsinger B. The effectiveness of integrated online health-coaching on physical activity and excessive gestational weight gain: a prospective randomized-controlled trial. Arch Gynecol Obstet. 2024;310(1):307-314. [FREE Full text] [doi: 10.1007/s00404-023-07296-y] [Medline: 38217763]
160. Thijs I, Fresiello L, Oosterlinck W, Sinnaeve P, Rega F. Assessment of physical activity by wearable technology during rehabilitation after cardiac surgery: explorative prospective monocentric observational cohort study. JMIR Mhealth Uhealth. 2019;7(1):e9865. [FREE Full text] [doi: 10.2196/mhealth.9865] [Medline: 30702433]
161. Thomas S, Yingling L, Adu-Brimpong J, Mitchell V, Ayers CR, Wallen GR, et al. Mobile health technology can objectively capture physical activity (PA) targets among African-American women within resource-limited communities-the Washington, D.C. cardiovascular health and needs assessment. J Racial Ethn Health Disparities. 2016. [FREE Full text] [doi: 10.1007/s40615-016-0290-4] [Medline: 27913983]
162. Thorup C, Hansen J, Grønkjær M, Andreasen JJ, Nielsen G, Sørensen EE, et al. Cardiac patients' walking activity determined by a step counter in cardiac telerehabilitation: data from the intervention arm of a randomized controlled trial. J Med Internet Res. 2016;18(4):e69. [FREE Full text] [doi: 10.2196/jmir.5191] [Medline: 27044310]
163. Tsai CH, Christian M, Kuo YY, Lu CC, Lai F, Huang WL. Sleep, physical activity and panic attacks: a two-year prospective cohort study using smartwatches, deep learning and an explainable artificial intelligence model. Sleep Med. 2024;114:55-63. [doi: 10.1016/j.sleep.2023.12.013] [Medline: 38154150]
164. Turakhia MP, Desai M, Hedlin H, Rajmane A, Talati N, Ferris T, et al. Rationale and design of a large-scale, app-based study to identify cardiac arrhythmias using a smartwatch: the apple heart study. Am Heart J. 2019;207:66-75. [FREE Full text] [doi: 10.1016/j.ahj.2018.09.002] [Medline: 30392584]
165. Twiggs J, Salmon L, Kolos E, Bogue E, Miles B, Roe J. Measurement of physical activity in the pre- and early post-operative period after total knee arthroplasty for osteoarthritis using a Fitbit flex device. Med Eng Phys. 2018;51:31-40. [doi: 10.1016/j.medengphy.2017.10.007] [Medline: 29117912]
166. Van der Walt N, Salmon LJ, Gooden B, Lyons MC, O'Sullivan M, Martina K, et al. Feedback from activity trackers improves daily step count after knee and hip arthroplasty: a randomized controlled trial. J Arthroplasty. 2018;33(11):3422-3428. [doi: 10.1016/j.arth.2018.06.024] [Medline: 30017217]
167. van Ekris E, Wijndaele K, Altenburg TM, Atkin AJ, Twisk J, Andersen LB, et al. International Children's Accelerometry Database (ICAD) Collaborators. Tracking of total sedentary time and sedentary patterns in youth: a pooled analysis using

the International Children's Accelerometry Database (ICAD). Int J Behav Nutr Phys Act. 2020;17(1):65. [FREE Full text] [doi: 10.1186/s12966-020-00960-5] [Medline: 32423404]
168. van Woudenberg TJ, Bevelander KE, Burk WJ, Buijzen M. The reciprocal effects of physical activity and happiness in adolescents. Int J Behav Nutr Phys Act. 2020;17(1):147. [FREE Full text] [doi: 10.1186/s12966-020-01058-8] [Medline: 33213465]
169. Vidal Bustamante CM, Coombs G, Rahimi-Eichi H, Mair P, Onnela J, Baker JT, et al. Fluctuations in behavior and affect in college students measured using deep phenotyping. Sci Rep. 2022;12(1):1932. [FREE Full text] [doi: 10.1038/s41598-022-05331-7] [Medline: 35121741]
170. Viswanath VK, Hartogenesis W, Dilchert S, Pandya L, Hecht FM, Mason AE, et al. Five million nights: temporal dynamics in human sleep phenotypes. NPJ Digit Med. 2024;7(1):150. [FREE Full text] [doi: 10.1038/s41746-024-01125-5] [Medline: 38902390]
171. von Coelln R, Dawe RJ, Leurgans SE, Curran TA, Truty T, Yu L, et al. Quantitative mobility metrics from a wearable sensor predict incident parkinsonism in older adults. Parkinsonism Relat Disord. 2019;65:190-196. [FREE Full text] [doi: 10.1016/j.parkreldis.2019.06.012] [Medline: 31272924]
172. Wang C, Lizardo O, Hachen DS. Using Fitbit data to monitor the heart rate evolution patterns of college students. J Am Coll Health. 2022;70(3):875-882. [FREE Full text] [doi: 10.1080/07448481.2020.1775610] [Medline: 32569509]
173. Watanabe K, Sugimura N, Shishido I, Konya I, Yamaguchi S, Yano R. Effects of 90 min napping on fatigue and associated environmental factors among nurses working long night shifts: a longitudinal observational study. Int J Environ Res Public Health. 2022;19(15):9429. [FREE Full text] [doi: 10.3390/ijerph19159429] [Medline: 35954787]
174. Webel AR, Perazzo J, Longenecker CT, Jenkins T, Sattar A, Rodriguez M, et al. The influence of exercise on cardiovascular health in sedentary adults with human immunodeficiency virus. J Cardiovasc Nurs. 2018;33(3):239-247. [FREE Full text] [doi: 10.1097/JCN.0000000000000450] [Medline: 29189426]
175. Weeks DL, Sprint GL, Stilwill V, Meisen-Vehrs AL, Cook DJ. Implementing wearable sensors for continuous assessment of daytime heart rate response in inpatient rehabilitation. Telemed J E Health. 2018;24(12):1014-1020. [FREE Full text] [doi: 10.1089/tmj.2017.0306] [Medline: 29608421]
176. Weng D, Ding J, Sharma A, Yanek L, Xun H, Spaulding EM, et al. Heart rate trajectories in patients recovering from acute myocardial infarction: a longitudinal analysis of Apple Watch heart rate recordings. Cardiovasc Digit Health J. 2021;2(5):270-281. [FREE Full text] [doi: 10.1016/j.cvdhj.2021.05.003] [Medline: 35265918]
177. Wile DJ, Ranawaya R, Kiss ZHT. Smart watch accelerometry for analysis and diagnosis of tremor. J Neurosci Methods. 2014;230:1-4. [doi: 10.1016/j.jneumeth.2014.04.021] [Medline: 24769376]
178. Xiao H, Zhou Z, Ma Y, Li X, Ding K, Dai X, et al. Association of wearable device-measured step volume and variability with blood pressure in older Chinese adults: mobile-based longitudinal observational study. J Med Internet Res. 2024;26:e50075. [FREE Full text] [doi: 10.2196/50075] [Medline: 39141900]
179. Yang Y, Hirdes JP, Dubin JA, Lee J. Fall risk classification in community-dwelling older adults using a smart wrist-worn device and the resident assessment instrument-home care: prospective observational study. JMIR Aging. 2019;2(1):e12153. [FREE Full text] [doi: 10.2196/12153] [Medline: 31518278]
180. Yao J, Brugger VK, Edney SM, Tai E, Sim X, Müller-Riemenschneider F, et al. Diet, physical activity, and sleep in relation to postprandial glucose responses under free-living conditions: an intensive longitudinal observational study. Int J Behav Nutr Phys Act. 2024;21(1):142. [FREE Full text] [doi: 10.1186/s12966-024-01693-5] [Medline: 39696319]
181. Yokoyama S, Kagawa F, Takamura M, Takagaki K, Kambara K, Mitsuyama Y, et al. Day-to-day regularity and diurnal switching of physical activity reduce depression-related behaviors: a time-series analysis of wearable device data. BMC Public Health. 2023;23(1):34. [FREE Full text] [doi: 10.1186/s12889-023-14984-6] [Medline: 36604656]
182. Zhang Y, Folarin AA, Sun S, Cummins N, Bendayan R, Ranjan Y, et al. RADAR-CNS Consortium. Relationship between major depression symptom severity and sleep collected using a wristband wearable device: multicenter longitudinal observational study. JMIR Mhealth Uhealth. 2021;9(4):e24604. [FREE Full text] [doi: 10.2196/24604] [Medline: 33843591]
183. Zhang Y, Folarin AA, Sun S, Cummins N, Ranjan Y, Rashid Z, et al. RADAR-CNS consortium. Longitudinal assessment of seasonal impacts and depression associations on circadian rhythm using multimodal wearable sensing: retrospective analysis. J Med Internet Res. 2024;26:e55302. [FREE Full text] [doi: 10.2196/55302] [Medline: 38941600]
184. Zhang Y, Wang X, Pathiravasan CH, Spartano NL, Lin H, Borrelli B, et al. Association of smartwatch-based heart rate and physical activity with cardiorespiratory fitness measures in the community: cohort study. J Med Internet Res. 2024;26:e56676. [FREE Full text] [doi: 10.2196/56676] [Medline: 38870519]
185. Zheng NS, Annis J, Master H, Han L, Gleichauf K, Ching JH, et al. Sleep patterns and risk of chronic disease as measured by long-term monitoring with commercial wearable devices in the All of Us Research Program. Nat Med. 2024;30(9):2648-2656. [doi: 10.1038/s41591-024-03155-8] [Medline: 39030265]
186. Zhou S, Ogihara A, Nishimura S, Jin Q. Analysis of health changes and the association of health indicators in the elderly using TCM pulse diagnosis assisted with ICT devices: a time series study. Eur J Integr Med. 2019;27:105-113. [doi: 10.1016/j.eujim.2019.02.010]

187. Zhou W, Chan YE, Foo CS, Zhang J, Teo JX, Davila S, et al. High-resolution digital phenotypes from consumer wearables and their applications in machine learning of cardiometabolic risk markers: cohort study. J Med Internet Res. 2022;24(7):e34669. [FREE Full text] [doi: 10.2196/34669] [Medline: 35904853]
188. Zhuo K, Gao C, Wang X, Zhang C, Wang Z. Stress and sleep: a survey based on wearable sleep trackers among medical and nursing staff in Wuhan during the COVID-19 pandemic. Gen Psychiatr. 2020;33(3):e100260. [FREE Full text] [doi: 10.1136/gpsych-2020-100260] [Medline: 32596641]
189. Charles P, Giraudeau B, Dechartres A, Baron G, Ravaud P. Reporting of sample size calculation in randomised controlled trials: review. BMJ. 2009;338:b1732. [FREE Full text] [doi: 10.1136/bmj.b1732] [Medline: 19435763]
190. Althubaiti A. Sample size determination: a practical guide for health researchers. J Gen Fam Med. 2023;24(2):72-78. [FREE Full text] [doi: 10.1002/jgf2.600] [Medline: 36909790]
191. von E. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. 2007. URL: https://core.ac.uk/reader/33050540?utm_source=linkout [accessed 2025-09-03]
192. Barnett I, Torous J, Reeder H, Baker J, Onnela J-P. Determining sample size and length of follow-up for smartphone-based digital phenotyping studies. J Am Med Inform Assoc. 2020;27(12):1844-1849. [FREE Full text] [doi: 10.1093/jamia/ocaa201] [Medline: 33043370]
193. Pressat-Laffouilhère T, Jouffroy R, Leguillou A, Kerdelhue G, Benichou J, Gillibert A. Variable selection methods were poorly reported but rarely misused in major medical journals: literature review. J Clin Epidemiol. 2021;139:12-19. [doi: 10.1016/j.jclinepi.2021.07.006] [Medline: 34280475]
194. Curran PJ, Bauer DJ. The disaggregation of within-person and between-person effects in longitudinal models of change. Annu Rev Psychol. 2011;62:583-619. [FREE Full text] [doi: 10.1146/annurev.psych.093008.100356] [Medline: 19575624]
195. Cho S, Ensari I, Weng C, Kahn MG, Natarajan K. Factors affecting the quality of person-generated wearable device data and associated challenges: rapid systematic review. JMIR Mhealth Uhealth. 2021;9(3):e20738. [FREE Full text] [doi: 10.2196/20738] [Medline: 33739294]
196. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth. 2020;8(9):e18694. [FREE Full text] [doi: 10.2196/18694] [Medline: 32897239]
197. Ortiz BL, Gupta V, Kumar R, Jalin A, Cao X, Ziegenbein C, et al. Data preprocessing techniques for AI and machine learning readiness: scoping review of wearable sensor data in cancer care. JMIR Mhealth Uhealth. 2024;12:e59587. [FREE Full text] [doi: 10.2196/59587] [Medline: 38626290]
198. Ren W, Liu Z, Wu Y, Zhang Z, Hong S, Liu H, et al. Missing Data in Electronic health Records (MINDER) Group. Moving beyond medical statistics: a systematic review on missing data handling in electronic health records. Health Data Sci. 2024;4:0176. [FREE Full text] [doi: 10.34133/hds.0176] [Medline: 39635227]
199. Di J, Demanuele C, Kettermann A, Karahanoglu FI, Cappelleri JC, Potter A, et al. Considerations to address missing data when deriving clinical trial endpoints from digital health technologies. Contemp Clin Trials. 2022;113:106661. [FREE Full text] [doi: 10.1016/j.cct.2021.106661] [Medline: 34954098]
200. Smiti A. A critical overview of outlier detection methods. Comput Sci Rev. 2020;38:100306. [doi: 10.1016/j.cosrev.2020.100306]
201. Olteanu M, Rossi F, Yger F. Meta-survey on outlier and anomaly detection. Neurocomputing. 2023;555:126634. [doi: 10.1016/j.neucom.2023.126634]
202. Xiao X, Yin J, Xu J, Tat T, Chen J. Advances in machine learning for wearable sensors. ACS Nano. 2024;18(34):22734-22751. [doi: 10.1021/acsnano.4c05851] [Medline: 39145724]
203. Shamout F, Zhu T, Clifton DA. Machine learning for clinical outcome prediction. IEEE Rev Biomed Eng. 2021;14:116-126. [doi: 10.1109/RBME.2020.3007816] [Medline: 32746368]
204. Collins GS, Dhiman P, Ma J, Schlussel MM, Archer L, Van Calster B, et al. Evaluation of clinical prediction models (part 1): from development to external validation. BMJ. 2024;384:e074819. [FREE Full text] [doi: 10.1136/bmj-2023-074819] [Medline: 38191193]
205. Riley RD, Archer L, Snell KIE, Ensor J, Dhiman P, Martin GP, et al. Evaluation of clinical prediction models (part 2): how to undertake an external validation study. BMJ. 2024;384:e074820. [FREE Full text] [doi: 10.1136/bmj-2023-074820] [Medline: 38224968]
206. Ramspek CL, Steyerberg EW, Riley RD, Rosendaal FR, Dekkers OM, Dekker FW, et al. Prediction or causality? A scoping review of their conflation within current observational research. Eur J Epidemiol. 2021;36(9):889-898. [FREE Full text] [doi: 10.1007/s10654-021-00794-w] [Medline: 34392488]
207. Moons KGM, Damen JAA, Kaul T, Hooft L, Andaur Navarro C, Dhiman P, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ. 2025;388:e082505. [doi: 10.1136/bmj-2024-082505] [Medline: 40127903]
208. Althouse AD. Adjust for multiple comparisons? It's not that simple. Ann Thorac Surg. 2016;101(5):1644-1645. [doi: 10.1016/j.athoracsur.2015.11.024] [Medline: 27106412]

209. Harris JD, Brand JC, Cote MP, Faucett SC, Dhawan A. Research pearls: the significance of statistics and perils of pooling. Part 1: clinical versus statistical significance. Arthroscopy. 2017;33(6):1102-1112. [doi: 10.1016/j.arthro.2017.01.053] [Medline: 28454999]

Abbreviations

PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
PROSPERO: International Prospective Register of Systematic Reviews
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by S Brini; submitted 20.Jan.2026; peer-reviewed by T Edison, A Teles; comments to author 24.Mar.2026; revised version received 18.May.2026; accepted 03.Jun.2026; published 27.Aug.2026

Please cite as:
Badier N, Lafitte A, Difernand A, Aguayo GA, Fagherazzi G, Onnela J-P, Vittrant B, Lechat B, De Larochelambert Q, Toussaint J-F, Delrieu L
Digital Phenotyping in Health Research: Scoping Review of Methods, Gaps, and Opportunities
J Med Internet Res 2026;28:e91747
URL: https://www.jmir.org/2026/1/e91747
doi: 10.2196/91747
PMID: