

A framework of digital biomarkers for neurodegenerative diseases

Favour Nerrise^{1,2}, Narayan Schütz^{1,3}, Qingyu Zhao⁴, Christine Gould¹, Arnold Milstein³, Kevin Schulman³, Victor W. Henderson^{1,5,6,7}, James Landay^{8,9}, Li Fei-Fei^{8,9}, Feng Vankee Lin¹ & Ehsan Adeli^{1,3,7,8,9,10}✉

Abstract

Digital biomarkers (DBMs) are a new class of health indicators derived from digital technologies – including smartphones, wearable devices and ambient sensors – that enable continuous, real-time monitoring of signals in everyday settings. By providing richer and more dynamic data than conventional, point-in-time measurements, DBMs offer fresh opportunities for remote patient assessment, personalized care and large-scale biomedical research. Importantly, DBMs function as powerful complementary tools to traditional biomarkers that can screen candidates for more invasive tests and provide contextual data between clinical visits. This Review provides a standardized classification of DBMs focused on neurodegenerative diseases, including Alzheimer disease, Parkinson disease, mild cognitive impairment, Huntington disease, multiple sclerosis, frontotemporal dementia, spinocerebellar ataxia and dementia with Lewy bodies, centred around three questions: what is being measured (the concept of interest), how it is measured (the sensing technologies) and why it is measured (the application areas). By examining these dimensions, we highlight the potential of DBMs to transform clinical monitoring, early detection and therapeutic interventions in these disorders.

Sections

Introduction

Rethinking biomarkers for digital health

Categorization by health aspects captured

Categorization by sensing technologies

Categorization by utility and application

Outlook

¹Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA. ²Department of Electrical Engineering, Stanford University, Stanford, CA, USA. ³Clinical Excellence Research Center (CERC), Stanford University, Stanford, CA, USA. ⁴Weill Cornell Medicine, Cornell University, New York, NY, USA. ⁵Department of Epidemiology and Population Health, Stanford University, Stanford, CA, USA. ⁶Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA, USA. ⁷Stanford Alzheimer's Disease Research Center (ADRC), Stanford University, Stanford, CA, USA. ⁸Department of Computer Science, Stanford University, Stanford, CA, USA. ⁹Stanford Institute for Human-Centered Artificial Intelligence (HAI), Stanford University, Stanford, CA, USA. ¹⁰Department of Biomedical Data Science, Stanford University, Stanford, CA, USA.

✉e-mail: eadeli@stanford.edu

Key points

- Digital biomarkers complement traditional biomarkers by providing continuous, real-world data that contextualizes sparse clinical measurements and screens candidates for more invasive testing.
- A three-dimensional framework classifies digital biomarkers by sensing technology (how), clinical application (why) and the health aspect captured (what), offering a structured foundation for the field.
- Parkinson disease has the most mature digital biomarker ecosystem, driven by quantifiable motor symptoms and clear utility for medication titration, unlike Alzheimer disease and other neurodegenerative diseases.
- Ambient and contactless sensors are emerging as critical tools for later-stage neurodegenerative diseases, where cognitive impairment limits adherence to wearable or active-task approaches.
- Regulatory frameworks, standardized validation protocols and transparent reporting are essential to translating digital biomarkers from research tools into clinically accepted end points.
- Multimodal sensor fusion and artificial intelligence-driven analytics can capture the multidomain nature of neurodegenerative diseases but require diverse training data, longitudinal validation and interpretable algorithms.

Introduction

Medicine has long relied on traditional biomarkers, measurable indicators of a biological state or condition, to diagnose and monitor diseases. Typically derived from blood, urine or tissue samples, these point-in-time measurements have been fundamental in guiding clinical decisions and advancing our understanding of numerous health conditions^{1,2}. However, they often require costly and invasive procedures that are not easily scalable, and their sensitivity or specificity can vary substantially depending on the timing of collection, patient heterogeneity and differences in clinical practice². Additionally, biomarkers are limited as surrogate end points due to their indirect measurement nature and the complexities involved in meeting the requirements for candidacy and validation. For example, there are currently no validated surrogate biomarkers for Alzheimer disease, and many putative surrogates result in conflicting clinical trial outcomes (NCT03367403, NCT02477800, NCT04437511)³. These challenges are even more pronounced in neurodegenerative diseases (NDDs), where a gradual decline in brain function requires ongoing assessment to capture subtle changes that might otherwise be missed⁴. NDDs affect millions of people worldwide⁵, and traditional approaches, ranging from patient diaries and pen-and-paper tests to sporadic biological or imaging markers, can be prohibitively subjective, intrusive, time-consuming or expensive. As a result, there is an urgent need for objective, low-cost and continuous measures that can be collected in real-world environments. Digital biomarkers (DBMs) have emerged as a promising solution, offering a way to capture ecologically valid data and reach communities that are often underserved or underrepresented in research⁶.

DBMs are digital measures quantified using digital health technologies (DHTs)⁷, particularly remote monitoring technologies, such

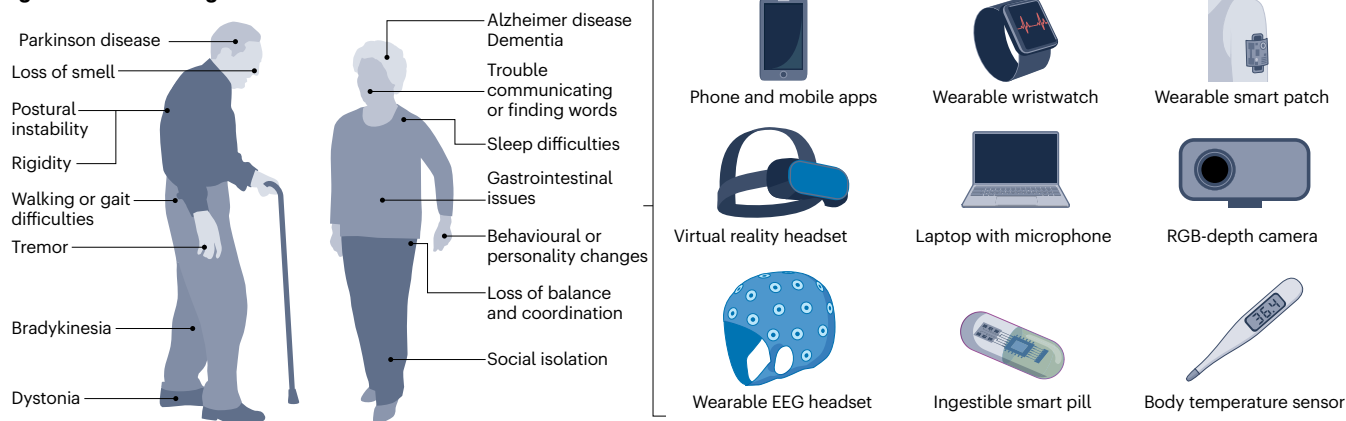
as wearables, portables and inner-body devices, as well as ambient and contactless sensing devices (Fig. 1). In this Review, we use the umbrella term DHTs to refer to technologies used outside the clinic. The term 'digital biomarkers' has evolved over the years owing to the broad popularization and adoption of DHTs, but the concept can be traced back to the advent of electronic medical devices and the digitization of medical records in the late twentieth century⁸. The field experienced great development in response to the miniaturization of digital technologies, advances in connectivity and display technologies, and robust app development ecosystems, which made remote monitoring technologies faster and cost-effective enough to be used by consumers and the broader research community⁷. Although the idea of using technologies to monitor health surrounding NDDs like Alzheimer disease had already been demonstrated in the early 2000s⁹, it was arguably the launch of the iPhone in 2007 and subsequent widespread smartphone use that spurred large-scale interest in DHTs, initiating a new digital health data collection era. The following explosion of mainstream fitness devices and wearables, such as Fitbit and Apple Watch, together with research-grade devices, like the *Axivity AX3* triaxial accelerometer and *Empatica E4* wireless wristband, further brought patient-generated health data measurement into the spotlight. Here, we first introduce a formal definition of DBMs and provide a structured overview of DBMs for NDDs. We structure this discussion around three key pillars (Fig. 2): (1) how the biological signal is measured, (2) why it is measured, and (3) what is being measured. By examining these dimensions, we offer a holistic understanding of how DBMs are personalized and used to monitor, diagnose and track the progression of NDDs. Using this framework, we aim to present a nuanced, multi-faceted view of DBMs, making the subject matter accessible and insightful to a general audience while highlighting the significance of these tools in advancing the understanding and treatment of NDDs.

Rethinking biomarkers for digital health

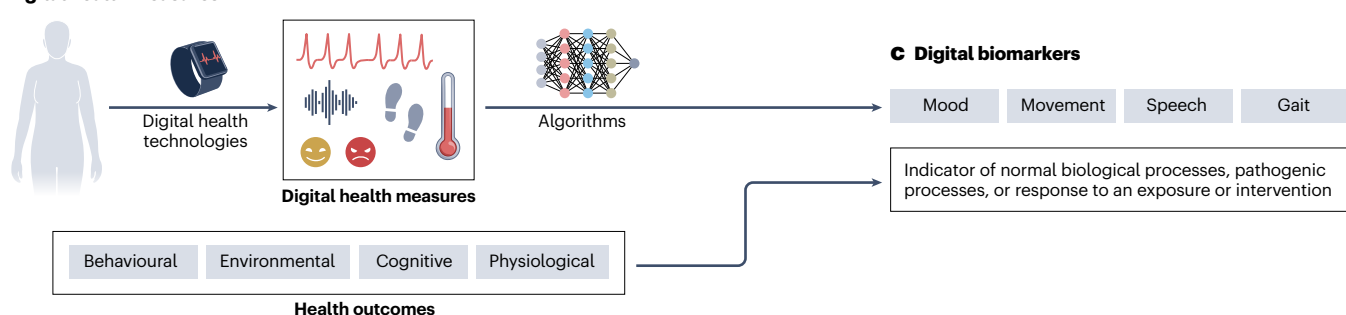
Despite growing use of the term 'digital biomarker', more precise terminology has evolved over time^{2,4,10,11}; the current consensus seems to refer to a specific DHT-derived quantity collected as a 'digital measure'^{12,13}. With regards to nomenclature, an exact definition of the term 'digital biomarker' that encompasses all use cases is difficult to find. For example, DBMs can be treated as equivalent to traditional biomarkers, only differing in the mode of data collection (via DHTs)¹². From this perspective, a DBM is measured as an indicator of biological processes, pathogenic processes or responses to interventions. However, this definition might be too narrow as it fails to encompass the full spectrum of DHT-captured data, particularly environmental or behavioural biomarkers that measure external factors or latent constructs rather than direct biological processes. A different and broader definition proposes DBMs to be: (1) direct measurements of biological processes (such as heart rate from photoplethysmography (PPG) sensors, glucose levels from continuous monitors or SpO₂ from smartwatch pulse oximeters, all of which are physiological signals with established clinical correlates) or (2) indirect measurements of latent variables such as typing speed as a proxy for fine motor control (potentially indicating Parkinson disease (PD) progression), GPS-derived social interaction patterns as depression indicators or voice acoustic features signalling cognitive decline¹⁴. Notably, measures like heart rate variability straddle both categories, representing both a direct physiological signal and a latent measure of autonomic balance. Similarly, actigraphy data from wearables directly measure movement but indirectly assess sleep quality, physical activity levels and circadian rhythmicity,

Review article

a Digital health technologies



b Digital health measures



c Digital biomarker applications

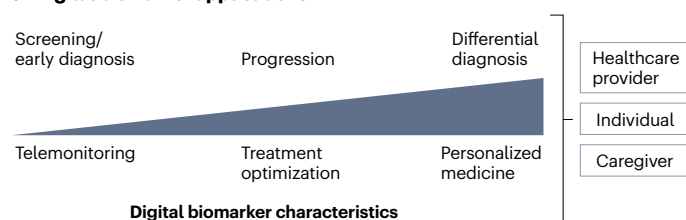


Fig. 1 | Digital health technologies, measures, biomarkers and clinical applications. **a**, Examples of symptoms and the sensing tools used to monitor them. Phone and mobile apps¹⁰², wearable wristwatch¹²⁷, electronic ‘smart’ patch⁹⁹, virtual reality headset¹⁵⁵, laptop with microphone³⁶, RGB-depth camera¹¹³, wearable electroencephalogram (EEG) headset⁹⁵, smart ingestible pill¹⁰⁷,

and body temperature sensor⁷⁶. **b**, Digital health measures collected from digital health technologies and analysed using algorithms. **c**, Examples of digital biomarkers and commonly linked health outcomes. **d**, Digital biomarker characteristics and applications.

all of which serve as proxies for various health conditions ranging from insomnia to bipolar disorder. Importantly, researchers and regulatory bodies widely agree that a DBM is not an assessment of how a patient feels, functions or survives (that is, not a digital outcome assessment), as addressed by the US Food and Drug Administration (FDA)⁷. However, even this distinction can be ambiguous for digital measures¹⁵ like step count during a timed walk test, which could be interpreted as either a biomarker of motor function or a clinical outcome assessment of functional capacity, and the FDA’s current guidance on DHTs does not yet provide definitive criteria for resolving such boundary cases. Moreover, continuous measurement enables the capture of

entirely new biomarkers (such as diurnal variability) to provide new insights into health and disease. DBMs also support a timelier, more objective, comprehensive and personalized analysis of an individual’s health and its changes or responses to intervention. Finally, some DBMs have no traditional counterparts such as those collected during active digital clinical assessments – for example, cadence measured during a walk test.

Traditional biomarker classification

Traditionally, biomarkers are divided into different categories based on their clinical applications, characterized by the FDA–National

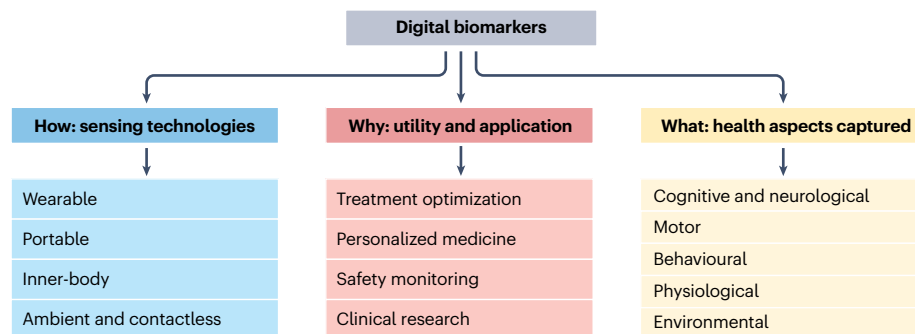


Fig. 2 | Categorization of digital biomarkers in neurodegenerative diseases.

The framework is organized across three fundamental dimensions: ‘How’ (sensing technologies) covers the technological modalities used for data acquisition, including wearable devices, portable sensors, inner-body sensors, and ambient/contactless systems; ‘why’ (utility and application) defines the

clinical and research purposes, spanning treatment optimization, personalized medicine approaches, safety monitoring, and clinical research applications; and ‘what’ (health aspects captured) outlines the measurable domains such as cognitive/neurological, motor, behavioural, physiological and environmental biomarkers.

Institutes of Health (NIH) Biomarkers, EndpointS and Other Tools (BEST) classification¹⁶:

1. Diagnostic biomarkers can detect whether a disease or other condition exists; for example, blood glucose levels are a diagnostic biomarker for type 2 diabetes mellitus¹⁷.
2. Predictive biomarkers help determine which individuals are more likely to experience a favourable or unfavourable outcome from a specific treatment. For example, the rs1080985G allele at *CYP2D6* is a predictive biomarker of poor response to cholinesterase inhibitors in donepezil, an Alzheimer disease medication¹⁸.
3. Prognostic biomarkers indicate how likely a clinical event, disease recurrence or progression will occur in patients with a particular disease. For example, elevated plasma pTau-181/217 levels in individuals with mild cognitive impairment (MCI) is a prognostic biomarker of progression to Alzheimer disease-related dementia and cognitive decline¹⁹.
4. Susceptibility or risk biomarkers indicate an individual’s potential to develop a disease or condition without a specific intervention. *APOE* (encoding apolipoprotein E) polymorphisms are susceptibility biomarkers for PD dementia²⁰.
5. Safety biomarkers indicate the likelihood, presence or extent of toxicity of a therapeutic intervention. Liver enzyme levels, such as alanine aminotransferase and aspartate aminotransferase, are safety biomarkers for hepatotoxicity²¹.
6. Pharmacodynamic biomarkers demonstrate the biological response to a therapeutic intervention without evaluation of efficacy. For example, circulating tumour DNA levels serve as a pharmacodynamic biomarker to observe the response to radiotherapy in patients with lung cancer²².
7. Monitoring biomarkers continuously track the status of a disease, condition or therapeutic response. For example, serum neurofilament light chain concentration levels are a monitoring biomarker for axonal damage and neurodegeneration in multiple sclerosis²³.

Following the general idea of treating DBMs like traditional biomarkers, the former can be grouped using established regulatory frameworks to facilitate clinical translation and regulatory

approval pathways (Table 1 and Box 1). Moreover, DBMs can be further grouped based on their data collection method, that is, sensing modalities (such as wearable devices, portable devices, invasive devices, and ambient and contactless devices) or the health aspect they capture (such as physiological, behavioural, motor or cognitive markers). A commonly used classification scheme differentiates between two types of sensing: passive (no conscious effort required from the user) and active (intentional user interaction required)²⁴. In the following sections, we extend these classifications to provide a more comprehensive overview of the various modalities captured by DBMs (Fig. 3).

Categorization by health aspects captured

DBMs can be grouped based on the health or physiological parameters they monitor such as behavioural, environmental, cognitive and physiological biomarkers (Fig. 1c). The health aspect targeted by the biomarker determines its relevance to certain medical conditions or wellness areas (Fig. 3). Based on clinician expert knowledge, we categorize each biomarker by cost, invasiveness and data collection frequency. Cost assessments are based on post-ownership measurement expenses, that is, focusing on the marginal cost per clinically meaningful measurement episode rather than initial device acquisition costs, with categories ranging from low (US \$0–50) to high cost (over US \$500) based on consumables, professional time, laboratory processing and use of infrastructure. Invasiveness was assessed on a four-tier progressive scale representing increasing breach of the body’s natural barriers, from non-invasive procedures that require no physical intervention to highly invasive approaches that require surgical procedures or substantial tissue disruption. Data frequency categories are distinguished between episodic measurements (infrequent, typically months to years apart), periodic measurements (regular intervals from daily to monthly) and continuous measurements (real-time or high-frequency monitoring capabilities), based on the technological and practical limitations of each biomarker approach (Table 1 and Supplementary Table S1).

Cognitive and neurological biomarkers

Cognitive and neurological DBMs can be used to monitor and quantify neurological functions and cognitive processes through continuous,

Review article

Table 1 | Examples of digital biomarkers

Biomarker category	Utility in NDDs	DBM examples		
		Study objective	Sensor	DBM measure
Diagnostic (presence of a disease)	Differentiate NDD A (such as AD) among a group of NDDs with similar symptoms (such as cognitive impairment often observed in AD, FTD and VD)	Differentiate MCI from healthy controls using recorded voice quality and speech fluency in a semantic memory test ³⁵	Portable recorder Speech data during a semantic memory test Researcher	Acoustic aspects of voice quality, namely cepstral peak prominence, centre of gravity and shimmer Speech fluency, namely articulation rate and average speaking time
Prognostic (likelihood of a clinical event, disease recurrence or progression)	Predict the future course of a NDD (for example, the onset of motor symptoms in PD)	Measure individual, longitudinal change in cognition in people with MCI or AD and healthy controls using 'digital neurosignature' from smartphone ³³	Altoida DNS (smartphone/tablet-based active digital assessment) 320+ cognitive and functional activity measures Participant	DNS metrics of intraindividual variability and dispersion (longitudinal digital cognitive change)
Monitoring (track disease or therapeutic response over time)	Track NDD progression (such as symptom changes in HD) or NDD treatment response (such as medication impact in PD)	Monitor limb movement to measure gait impairment in people with PD using a wearable accelerometer ⁴²	Accelerometer Limb movement while walking Researcher, clinician and participant	Triaxial gait variables (such as stride length, cadence, regularity index, symmetry index, and mechanical powers) Walking speed
Predictive (more likely to experience a favourable or unfavourable effect from treatment)	Predict response to specific NDD treatments (such as the efficacy of a drug in managing AD cognition or tremor responsiveness to motor neurorehabilitation therapy in PD)	Localize neural spectral power changes from high-density scalp EEG recording of people with PD, while modulating STN activity with DBS ¹⁰	Closed-loop implantable with sensing channels, neural stimulator, triaxial accelerometer and embedded algorithm Neural spectral power and synchrony Participant, clinician, researcher	Resting state corresponded to relatively constant beta band power AD state corresponded to increased beta band power Suppression state corresponded to decreased beta band power Movement patterns Muscle response Neural data Stimulus trains Classifier detections
Response/ pharmacodynamics (biological response to a therapeutic intervention)	Detect NDD response to an intervention (such as neuroprotective effects in HD treatment)	Detect changes in PD tremor and bradykinesia in response to levodopa medication using a wrist-worn accelerometer ⁸¹	Wrist-worn accelerometer Resting tremor and bradykinesia Researcher and participant	Amplitude of hand movements (amplitude and slowness) Smoothness of hand movements (hesitancy) Percentage of time spent with no hand movement (poverty or absence) Average length of bouts with no hand movement (poverty or absence)
Safety (likelihood or presence of intervention toxicity)	Assess potential adverse effects and risks associated with NDD treatments (such as risk/observation of muscle weakness in ALS treatments)	Identify bradykinesia fluctuations as a manifestation of a 'wearing-off' medication state in PD using a KinetiGraph ⁸²	PD KinetiGraph Bradykinesia and dyskinesia Participant, clinician, researcher	BKS Continuous BKS Median BKS Active BKS Median dyskinesia score Percentage time in tremor Percentage time in bradykinesia Percentage time in dyskinesia Dosage
Susceptibility/risk (potential for disease without intervention)	Identify specific hallmarks linked to a higher likelihood of developing NDD (such as repeated head injuries linked to CTE)	Detect postural and gait instability patterns associated with increased risk of PD and movement-related NDDs using wearable inertial sensors ⁴¹	Chest and leg-mounted IMUs Fall frequency, near-fall frequency, daily activities Participant	Frequency of lying down Number of ambulatory bouts Frequency of sitting Walking frequency Peak acceleration of the chest IMU Distribution of ambulatory bouts in time

Based on US Food and Drug Administration (FDA)–National Institutes of Health (NIH) Biomarkers, EndpointS and other Tools (BEST). It applies the FDA–NIH BEST classification system to digital biomarkers in neurodegenerative diseases (NDDs) and summarizes how these new measurements align with traditional biomarkers based on their intended clinical utility. Each row is a distinct biomarker category, and the columns provide further context with digital biomarker (DBM) examples specific to various NDDs, as well as technological implementation details such as the sensors used for data collection, the type of information collected and the users interacting with the sensors. AD, Alzheimer disease; ALS, amyotrophic lateral sclerosis; BKS, Bradykinesia Score; CTE, chronic traumatic encephalopathy; DBS, deep brain stimulation; DNS, digital neurosignature; EEG, electroencephalogram; FTD, frontotemporal dementia; HD, Huntington disease; IMU, inertial measurement unit; MCI, mild cognitive impairment; PD, Parkinson disease; STN, subthalamic nucleus; VD, vascular dementia.

Box 1 | Translational considerations

Several regulatory pathways now guide the development and approval of digital biomarkers. The US Food and Drug Administration (FDA) has established the [Digital Health Innovation Action Plan](#) and the Digital Health Software Precertification ([Pre-Cert](#)) programme, and supports the Clinical Trials Transformation Initiative for developing novel digital end points. The FDA–National Institutes of Health (NIH) Biomarkers, EndpointS and Other Tools (BEST) framework¹⁶ provides standardized terminology for classifying biomarker types across both traditional and digital domains. In Europe, the [European Medicines Agency Qualification of Novel Methodologies for Medicine Development](#) pathway offers a route for formal acceptance of digital measures in medicine development, whereas the [National Institute for Health and Care Excellence \(NICE\) Evidence Standards Framework for Digital Health Technologies](#) outlines tiered evidence requirements for digital health technologies in the UK. The [Digital Medicine Society's V3 framework](#) (Verification, Analytical Validation, Clinical Validation) provides a structured methodology for determining whether a digital measure is fit for its intended purpose.

Nonetheless, proving clinical validity remains extremely challenging. Some digital biomarkers have achieved analytical validation for monitoring or diagnostic purposes; for example, [Empatica's Health Monitoring Platform](#) holds FDA 510(k) clearance for six digital biomarkers, including electrodermal activity, SpO₂, pulse rate and respiratory rate collected via a wrist-worn wearable. However, regulatory agencies are cautious when evaluating digital measures as primary efficacy end points, mirroring their scrutiny of traditional biomarkers that could be used to approve drugs that ultimately fail to deliver meaningful clinical benefits, as in the case of amyloid PET in Alzheimer disease receiving only “reasonably

likely surrogate endpoint” status after two decades of research³. Thus, even if a new drug gains approval based on one or two pivotal trials, a biomarker must clear a substantially higher evidentiary bar, requiring years of independent replication across multiple centres, populations, settings, device types, software versions and real-world conditions.

A central tension in this process is the ecological validity paradox: clinical ‘gold-standard’ assessments are conducted under controlled conditions that differ fundamentally from the uncontrolled free-living environments in which digital biomarkers operate. One proposed solution is stepwise validation, progressing from clinical gold-standard comparisons to deployable ‘silver standard’ references such as mobile polysomnography for real-world sleep staging.

For example, Stride Velocity 95th Centile (SV95C) for Duchenne Muscular Dystrophy¹⁵⁴ (an ankle-worn accelerometer-derived measure of stride velocity) was accepted by the European Medicines Agency (EMA) as a secondary and later primary end point in pivotal trials (NCT03039686, NCT05096221, NCT05524883). It took years of comprehensive clinical studies, including patient and caregiver surveys and healthcare professional perspectives, to validate that this digital measure correlated with gold-standard clinical assessments like the 6-minute walk distance test. Notably, despite SV95C demonstrating earlier sensitivity to disease progression, regulators classified it as a performance outcome assessment rather than a biomarker, probably because the act of walking itself constitutes the clinical outcome, rather than serving as a surrogate for an underlying biological process. This distinction highlights the nuanced regulatory boundaries that digital measures must navigate even after achieving clinical acceptance.

high-dimensional data streams that capture both macro-scale brain activity patterns and micro-level cognitive performance metrics.

Brain activity patterns. There are different approaches to monitoring brain activity patterns; electroencephalogram (EEG) is amongst the most popular and is a portable, non-invasive and low-cost diagnostic tool for measuring neural oscillations and event-related potentials associated with cognitive tasks. For example, a study of older adults at risk of dementia ($n = 100$) used machine learning algorithms to analyse prefrontal EEG activity and observed that slow intrinsic EEG oscillation is associated with MCI caused by Alzheimer disease²⁵. A similar study of older adults at risk of MCI ($n = 15$) identified multiscale entropy EEG patterns associated with dementia prognosis²⁶. Portable functional near-infrared spectroscopy (fNIRS) systems provide complementary haemodynamic measurements for detecting altered brain connectivity and brain oxygenation that correlate with cognitive function in individuals with MCI²⁷. These biomarkers of brain activity require minimal patient effort compared to performance-based assessments. Wearable and portable magnetoencephalography systems are still emerging but could offer higher spatial resolution than EEG and other tools for detecting subtle changes in brain activity²⁸.

Cognitive response time. The time taken to respond to specific tasks, especially in gamified cognitive assessment apps, can indicate

cognitive decline or other neurodegenerative conditions²⁹. For example, the self-administered integrated cognitive assessment has been used to screen individuals at risk for Alzheimer disease ($n = 38$) by measuring cognitive impairment via a person's accuracy and response time in categorizing images (ISRCTN18112405)³⁰. The study found that integrated cognitive assessment performance correlated with elevated serum phosphorylated tau181 levels, which suggests that combining digital cognitive response time biomarkers with blood-based biomarkers, among other covariates, can modestly predict Alzheimer disease-related pathology. Additionally, an observational study of 15 individuals with PD-related freezing of gait revealed that impaired executive control (including lower attention) correlates with slower step execution in individuals with PD and freezing of gait³¹.

Eye tracking and movement. Devices that track eye movements can be used to diagnose and monitor conditions such as age-related macular degeneration, glaucoma and even certain neurological disorders, for example, in individuals with motor and verbal impairments³². Rapid eye movement (REM) patterns during daytime and sleep have also been associated with NDDs such as MCI, Alzheimer disease³³ and PD³⁴. In particular, in a prospective cohort with ~40 months of follow-up ($n = 213$ with MCI/prodromal Alzheimer disease; $n = 283$ cognitively normal/preclinical Alzheimer disease), a smartphone/tablet-based digital assessment (Altoida DNS) using eye tracking showed high diagnostic discrimination

(AUC 0.94 in MCI and 0.91 in cognitively normal at-risk participants)³³ (NCT02843529). An exploratory research study also tracked eye movement (initial response time) detected by a generic RGB camera to screen for dementia and other brain dysfunctions in a group of older adults ($n = 18$)³⁵. The study observed high measurement accuracy (root mean square error ≤ 0.028 s for older individuals), sufficient to detect the >0.1 s delay in initial response time characteristic of individuals with dementia.

Motor biomarkers
Motor biomarkers are measurable indicators of an individual’s movement and coordination. They can be continuously monitored to better evaluate disease progression and other health conditions.

Speech, linguistic and vocal changes. Speech is a complex motor function that involves the coordination of multiple muscles and brain regions.

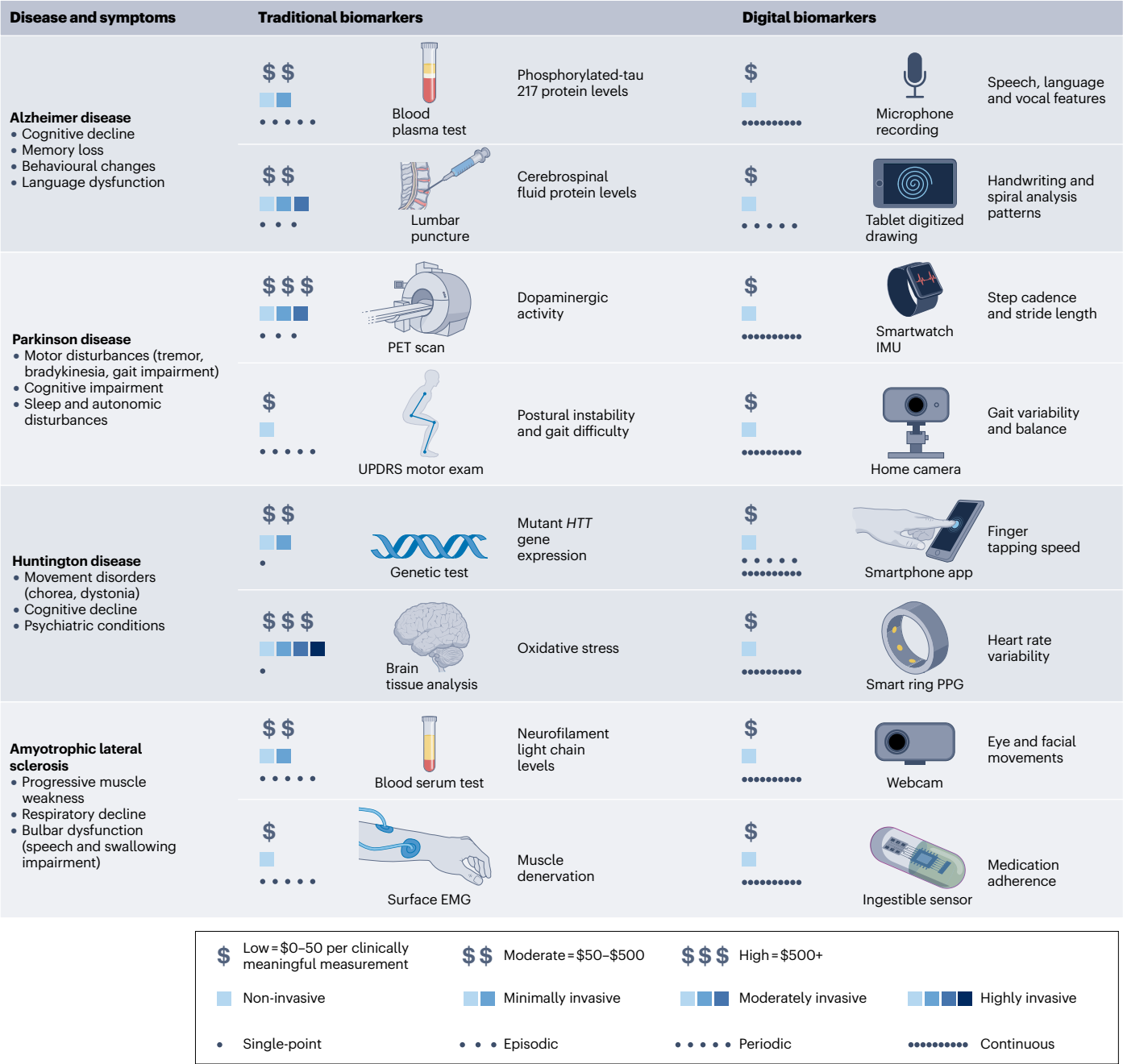


Fig. 3 | Comparison of traditional and digital biomarkers for neurodegenerative diseases. Traditional methods are characterized by invasive, costly and episodic measurements with variable outcomes, whereas digital approaches offer non-invasive, continuous, real-time data collection enhanced by artificial intelligence-driven analytics for better clinical use. EMG, electromyography; IMU, inertial measurement unit; PPG, photoplethysmography; UPDRS, Unified Parkinson’s Disease Rating Scale.

Impairment in speech, linguistic and vocal patterns, such as changes in pitch, volume, rate, syntax, semantics and articulation, can be early indicators of neurodegeneration (for example, articulation rate and averaged speaking time in people with amnesic MCI³⁶). In a multicentre study of individuals with PD ($n = 149$), microphone recordings were used to identify several speech biomarkers associated with PD, namely harsh voice, monotonic pitch and prolonged pauses³⁷. These speech biomarkers could differentiate both prodromal (idiopathic REM sleep behaviour disorder) and early-stage PD from healthy controls with up to 72% sensitivity. Similarly, machine learning algorithms have been used to derive speech and voice biomarkers (such as spontaneous speech and lexical-semantic patterns) as non-invasive methods for early Alzheimer disease detection³⁸, cognitive impairment progression³⁹ and multi-linguistic patterns in PD⁴⁰, with promising results.

Gait, balance and speed. Gait, the manner of walking, balance and speed are also important motor biomarkers that can provide valuable information about neurological and musculoskeletal health. Gait impairment features, such as slower walking speed, shorter stride length and increased variability in step timing, are common DBMs in NDDs and can be detected through wearable sensors^{41,42}, vision-based sensors^{43–45} and contactless sensors^{46,47}. For passive remote monitoring of Alzheimer disease, smart carpet sensors have been used to extract gait parameters and detect fall risk with high accuracy (95% sensitivity, 85% specificity) in a feasibility study of nine individuals with Alzheimer disease⁴⁸. Another study used a low-cost durable mat and smartphone camera to estimate gait impairment patterns such as freezing of gait, reduced stride length and lower step velocity in an investigative study of 44 people with PD⁴⁹. The stride length measurement captured by the study showed a mean absolute error of 0.62 cm, which is comparable with the ‘gold standard’ walking mat system [GAITrite](#).

Handwriting and finger tapping dynamics. Handwriting dynamics, including the speed, pressure and rhythm of writing, are informative motor biomarkers for diseases such as PD. Differences in geometrical writing patterns, velocity-based patterns and kinematic-based movement patterns from handwriting and drawing tasks collected from graphic tablets, smartphones and other devices can help diagnose PD^{50–52}. For example, a diagnostic pen with a soft magnetoelastic tip and ferrofluid ink was used to measure on-surface and in-air writing motions to diagnose PD in a pilot human study ($n = 16$)⁵⁰. Similarly, combining finger tapping, voice and gait parameters extracted from multimodal sensors in 6,418 individuals with and without PD revealed that models using the tapping coordinates (AUC 0.933) outperformed both gait/rest-based and voice-based models from the same cohort, and the assembled multimodal model (AUC 0.944) outperformed patient self-reported symptom scores⁵¹. Touchscreen dynamic typing can also be used to detect a decline in fine motor skills in people with early-stage PD⁵².

Facial expression and oculomotor movement. Facial expression analysis using computer vision techniques can identify reduced facial expressiveness, a symptom associated with several NDD conditions, often referred to as facial masking (hypomimia) in PD⁵³. For example, a contactless video-based system and computer vision algorithms enabled automatic analysis of facial expressions and revealed how individuals with PD ($n = 17$) had difficulty acting (that is, making up) and imitating facial gestures upon request⁵⁴. Smartphone cameras can also detect hypomimia as a self-screening tool for PD⁵⁵ as abnormalities in

eye movement (such as saccadic dysmetria, smooth pursuit deficits or nystagmus) have been closely linked with Alzheimer disease³² and PD³³. Similarly, video-based eye-tracking DHTs have been used for earlier detection of MCI⁵⁶, to distinguish between people with PD and healthy individuals⁵⁷, and to better predict PD motor and cognitive scores⁵⁸. The ViewMind DBM was developed to detect MCI progression to Alzheimer disease ($n = 65$ individuals with MCI) by drawing cognitive-related patterns of eye movement while participants perform a visual short-term memory binding task³².

Behavioural biomarkers

Beyond physiological measures, behaviours, daily routines, and mental and emotional states can also be hallmarks of NDD conditions.

Sleep activity and quality. Sleep has a role in cognitive function, mood regulation and overall health. Wearable inertial sensors⁵⁸, bed sensors^{46,59} and environmental sensors^{60,61} are being tested for less invasive monitoring of sleep disorders and abnormal movements in individuals with NDD. These technologies aim to assess sleep quality, detect disturbances and analyse abnormal motions during sleep. A pilot study of seven people with PD used the [Emerald device](#) (a radio-wave home sensor to passively track time in bed, among other measures) and observed high day-to-day variability in comparison to healthy controls⁴⁶.

Physical activity. Sedentary behaviour is a risk factor for numerous conditions, from depression to dementia. A cross-sectional study of individuals with PD ($n = 96$) used an actigraph to conclude that those who met World Health Organization recommendations for 150 weekly minutes of moderate-intensity physical activity showed better cognition and higher functional brain connectivity compared to less active people⁶². Similarly, clinical trials like EngagePD and Preactive-PD have leveraged actigraphs to monitor changes in motor and cognitive function in individuals with PD in response to physical activity interventions (NCT03696589, NCT04922190). These results suggest that physical activity can be a candidate biomarker for cognitive and functional decline associated with PD.

Mood and emotion tracking. Devices equipped with automated facial recognition⁶³, voice analysis algorithms³⁶ or human–computer interactions² can detect micro-expressions, tonal changes or behaviour profiles linked to specific moods or emotions⁶⁴ in Alzheimer disease, MCI and dementia. PD Manager is a mobile health platform that monitors mood fluctuations (among other DBMs) associated with PD to assist patients in actively managing their disease⁶⁵.

Stress levels. Psychological stress is a well-established risk factor for various mental and physical health problems, including NDDs⁶⁶. Various markers, such as cortisol levels⁶⁶, galvanic skin response (electrodermal activity or skin conductance)⁶⁷, heart rate variability⁶⁸ and voice markers³⁶, can indicate stress levels. Wearable devices can continuously monitor these indicators; for example, a wrist-worn device consisting of a galvanic skin response sensor revealed a correlation between sensor-derived stress measurements and clinical staff observations in six nursing home residents with dementia, with accuracy up to 75.9% using personalized stress level thresholds from the device⁶⁷.

Physiological biomarkers

Physiological biomarkers directly measure body functions and reflect the operational status of an individual’s physiological systems.

Heart rate and blood pressure. Heart rate and blood pressure serve as key indicators of stress, cardiovascular health and overall cardiac function. Wearable devices and smart patches have facilitated continuous monitoring of these parameters for conditions like MCI⁶⁹ and PD⁷⁰. For example, a wrist-worn device combined with machine learning detected circadian disruptions in sleep and motor functions in a cross-sectional study of 12 individuals with PD and age-matched controls. External sensor devices within an in-home Internet of Things system could also be used to detect blood pressure variations during daily activities⁷¹.

Respiration rate. Changes in respiration or breathing rate can signal respiratory distress, anxiety disorders and metabolic dysfunction. Radio-wave and other passive sensors have been used to monitor respiratory patterns associated with PD and dementia such as irregular breathing rhythms and sleep-disordered breathing^{60,72}. For instance, ultra-wideband impulse radars can unobtrusively monitor and reconstruct nocturnal respiration rate patterns in 47 individuals with Alzheimer disease⁷². Contact-based and contactless sensors, such as smart garments and chest bands, can also be used to capture respiration features⁷³. A study of 95 community-dwelling older adults ($n = 33$ with amnesic MCI) identified a novel sleep biometric, the time latency associated with the tight gap between sleep movements and respiratory coupling, to distinguish diagnosed MCI versus normal cognition with 87% sensitivity⁷⁴.

Environmental biomarkers

Environmental signals could also provide clinically relevant information in the context of NDDs.

Ambient light. The quality and quantity of light exposure have important roles in mood disorders and sleep quality; for example, nocturnal blue light exposure increases the risk of insomnia⁷⁵. A pilot study validated an ambulatory circadian monitoring device that integrates light exposure sensors, wrist temperature, actimetry, and body position to detect sleep–wake states in individuals with PD ($n = 15$) and found substantially reduced sleep efficiency, lower total sleep time and worse circadian functioning compared to age-matched controls ($n = 15$)⁷⁶. Additionally, a meta-analysis of older adults with dementia found that light therapy improves sleep quality and reduces nocturnal behavioural restlessness⁷⁷.

Location tracking. Location patterns offer many insights into a person's behaviour, such as the complexity of typical daily routines and their level of independence. For example, a GPS-based lower-limb ambulatory sensor to detect voice, activity and location monitoring enabled profiling of modifiable lifestyle risk factors related to Alzheimer disease and related dementia such as social and geographic activity⁷⁸. Location eigenbehaviour (that is, the structure in indoor location information routines obtained from contactless ambient sensors) and GPS driving tracking have also been proposed as DBMs for detecting cognitive levels and preclinical Alzheimer disease in older adults^{79,80}.

Categorization by sensing technologies

A sensor is a device, module, machine or subsystem that detects physical phenomena or environmental changes and produces an output signal to communicate this information. We propose sensing as an umbrella term referring to the modality in which DHTs acquire (or sense)

data (Fig. 1a). This section provides an overview of various sensing modalities, including wearable, portable, inner-body, and ambient and contactless sensors (Fig. 4 and Table 2).

Wearable sensors

Body-worn inertial and optical devices. Wearable sensors positioned on the chest, upper limb (such as forearm, wrist or finger), trunk (waist) and lower limb (such as thigh, calf or ankle) are the most prevalent sensing devices for obtaining DBMs owing to their low cost, convenience and versatility. These devices can capture motor function, sleep patterns and cardiovascular function^{25,26,69}.

Inertial measurement units (IMUs), consisting of accelerometers and gyroscopes, are instrumental for measuring the motor impairment features (such as tremors, gait disturbances and bradykinesia) that are hallmarks of various NDDs that emerge at varying stages of disease progression^{41,42,81–83}, for example, at early stages in PD and multiple sclerosis, and often later in Alzheimer disease⁸⁴. IMUs are commonly placed on the wrist, ankle, trunk or lower back, depending on what is being measured and the signal-to-noise ratio, with wrist placement being the most common owing to user comfort. Feet and waist-mounted IMUs can measure gait variability and balance ability in individuals with Alzheimer disease, revealing increased stance and swing time, and larger average sway speed compared to healthy individuals⁸⁴. In an age-matched, medium-sized study of individuals with multiple sclerosis ($n = 44$) and healthy controls, a small, body-fixed inertial sensor (Axivity AX3) was placed on the lower back of participants to measure physical activity and real-world, community ambulation for a week⁸³ (NCT02427997). Their findings revealed reduced gait speed, cadence and complexity (sample entropy) compared to controls, as well as lower physical activity levels, particularly in the free-living community ambulation settings.

Actigraphy is used for evaluating sleep changes, such as insomnia and excessive daytime sleepiness, which are known markers that precede motor and cognitive NDD symptoms^{62,85}. Actigraphs are devices integrated with accelerometers and other ambient sensors that monitor changes in physical motion, for example, to quantify restless sleep. Although the gold standard for sleep monitoring is polysomnography measured via video or EEG, actigraphs provide similar measurements, can be easily placed on various body positions, and are convenient for routine and at-home use⁸⁶. In a randomized controlled trial of the LOCK sleep intervention (NCT04533815), wrist-worn actigraphy measured total sleep time changes in nursing home residents with Alzheimer disease and related dementias ($n = 23$) as a primary end point⁸⁵.

PPG and optical sensing devices are other prevalent wearables primarily used to measure cardiovascular and autonomic dysfunction in NDDs^{86,87}. PPG sensors harness light-emitting diodes and photodetectors to capture blood volume changes in microvascular tissue and measure heart rate, heart rate variability, blood oxygen saturation and pulse wave velocity⁸⁸. These sensors are typically located in wrist-worn devices like smartwatches^{89,90} or finger-mounted configurations, such as smart rings⁸⁷, owing to better vascular access and reduced motion artefact. Similar to sleep disturbances, autonomic dysfunction manifests early in disease progression (such as in PD), often preceding preclinical symptoms by years. For example, an at-home, wrist-worn smartwatch with PPG sensors was used to associate resting heart rate during day and night with the Unified Parkinson's Disease Rating Scale (UPDRS) Part 1b scores (assessing non-motor experiences of daily living such as sleep and fatigue) in a large longitudinal study in individuals with PD ($n = 484$)⁹¹. The study revealed increasing resting heart rate

Review article

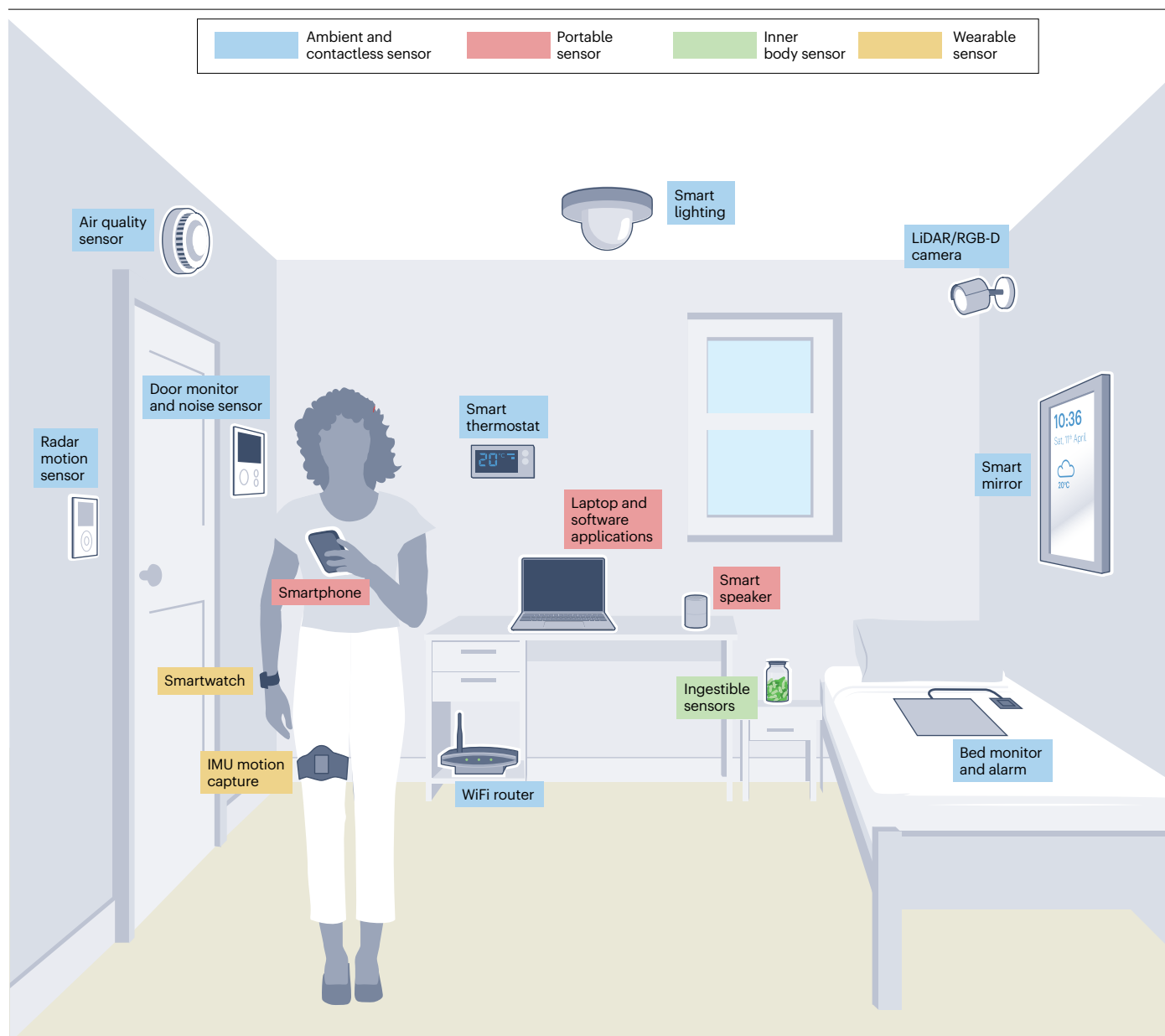


Fig. 4 | Different types of sensing modalities. Wearables, portables, inner-body, ambient and contactless. IMU, inertial measurement unit; LiDAR, light detection and ranging.

with the scores during day and night, using discriminative machine learning classifiers with over 93% accuracy. Similarly, finger-mounted PPG sensors⁸⁷, such as the Oura ring (Oura Health Ltd, Oulu, Finland) and wireless pulse oximeters, can reliably estimate nocturnal heart rate and heart rate variability and distinguish between individuals with MCI and healthy controls ($n = 31$, classification accuracy 90%) in alignment with the gold-standard electrocardiography⁹².

Electronic textiles. Electronic textiles (e-textiles) or smart clothing can integrate electronic components into the fabric to track symptoms of NDD conditions such as chronic pain in Alzheimer disease and sensory

deprivation in dementia^{93,94}. For example, sensorized shirts have been used to detect heart rate and respiratory rate in PD patients for wellness monitoring⁷⁰, and smart headbands such as MUSE can extract EEG signals that are coupled with fine motor control analysis to diagnose MCI⁹⁵ and for early-onset prognosis of possible dementia⁹⁶. In addition, IMUs have been mounted on belts and bands worn on limbs to predict symptoms such as bradykinesia or reduced physical activity in individuals with NDDs⁴². Similarly, bracelets embedded with inertial sensors have been used to track physical movements to detect hand tremors and bradykinesia in individuals with PD⁹⁷. In a more complex example, a smart sleep mask with integrated physiological sensors,

Table 2 | Digital health sensing technologies for neurodegenerative disease monitoring

Sensing technology	Advantages	Disadvantages	Maturity	User friction
(W) Body-worn inertial measurement units ^{41,42,81–84}	Continuous motion tracking Small and lightweight Proven validity for varied biomechanics	Sensor drift Require alignment and calibration Placement influences accuracy Susceptible to noise Limited context	TRL 8–9: Mature	Daily charging Consistent placement needed Compliance Independent donning/doffing Aesthetics
(W) Wrist-based actigraphy ^{62,85}	Prolonged sleep monitoring Strong correlation with sleep measures Longer battery life	Less precise for sleep staging Environmental confounders Limited physiological data	TRL 8–9+: Mature and clinically active	Wrist discomfort Skin irritation Software maintenance Aesthetics Device durability Minimal user feedback
(W) Photoplethysmography rings/watches ^{86–92}	Continuous heart rate/heart rate variability Small form factor	Skin tone bias Motion artefacts Temperature sensitivity	TRL 7–8: Mature	Finger swelling Limited battery life Comfort and fit Notification fatigue
(W) Smartwatches ^{86,89,90}	Multimodal sensing High acceptance Existing ecosystem Emergency assistance Direct user feedback	Proprietary algorithms Inter-device variability Power consumption Dependent on connectivity	TRL 7–8: Ready	Daily charging Skin irritation Environmental conditions Information overload Lifestyle compatibility
(W) Electronic textiles (e-textiles/smart clothing) ^{42,70,93–98}	Unobtrusive integration Larger, multimodal sensing area Higher comfort and wearability Customization	Washing durability Connectivity dependence Power consumption Flexibility Hardware–textile integration	TRL 5–6: Emerging	Limited clothing/designs Washability Physical wear Stretchability and bending
(W) Electronic skin patches ^{99–101}	Localized, targeted measurements Flexible Single-use Discreet form factor Multimodal sensing	Adhesive degradation Sweat interference Disposable components Biocompatibility Stretchability Thickness	TRL 5–7: Trials	Skin irritation/allergies Limited reusability Hygiene and maintenance Water resistance Placement
(W) Ear and eye devices (wearables/smart glasses) ^{102–104}	Socially acceptable Multimodal sensing Stability Vestibular and balance monitoring Discreet form factor	Few channels Ear variance Cerumen Power management Motion and acoustic interference Sensor crosstalk	TRL 4–5: Research	Ear fatigue Hygiene Audio interference Occlusion effect Weight and pressure points Heat and irritation
(P) Smartphones ^{102,104,105} (P) Tablets ^{36,103}	Ubiquitous Multi-sensor Local processing User engagement and feedback Large display Longer battery life Accessibility Multi-user	Battery drain Operating system limits Bring-your-own-device variance Environmental interference Less portable Limited sensor suite Hardware variability	TRL 7–8: Mature	Battery anxiety Privacy Carrying and placement Compliance Accessibility Size/weight Shared device Inconsistent availability Ergonomics
(I) Ingestibles (smart pills) ^{106–108}	Direct gastrointestinal access Medication adherence Novel gut–brain axis markers	Single-use Limited transmission Variable transit time Biocompatibility	TRL 4–5: Early	Swallowing Psychological discomfort Bowel preparation Limited reuse
(I) Implantables ^{109–112}	Continuous, internal signals High signal fidelity Therapeutic capability	Minimally invasive Biocompatibility Stability Device miniaturization	TRL 3–6: Research	Surgery Infection risk Limited flexibility Scarring

Table 2 (continued) | Digital health sensing technologies for neurodegenerative disease monitoring

Sensing technology	Advantages	Disadvantages	Maturity	User friction
(A/C) Vision-based (cameras) ^{43–45,113}	Contactless Multiple people	Lighting Occlusion Computation	TRL 6–7: Clinical	Privacy Fixed install Blind spots
(A/C) Infrared light/radiation (fNIRS/PIR) ^{10,46,58,114,115}	Passive Years-long battery	Low resolution Environmental interference	TRL 8–9: Ready	Installation Room-specific
(A/C) LiDAR ¹¹⁶	Precise 3D maps Privacy-preserving	Limited field of view Computation	TRL 5–6: Emerging	Installation Coverage planning
(A/C) Environmental (air quality, noise, light and temperature) ^{76,117–120}	Contextual information Shared environment monitoring Internet of Things integration	Indirect measurement Multi-occupancy Calibration needs	TRL 7–8: Mature	Multiple devices Network setup Maintenance
(A/C) Radio-wave (ultra-wideband/mmWave) radar ^{59–61,72,121}	Privacy-preserving Through-wall	Signal processing Multi-occupancy	TRL 5–6: Emerging	Installation WiFi interference
(A/C) Bed/sleep sensors ^{46,59}	Unobtrusive Partner data	User distinction Motion transfer	TRL 6–7: Commercial	Bed-specific Portability

Sensing categories: (W) wearable, (P) portable, (I) inner body, (A/C) ambient/contactless. Technology Readiness Level (TRL; 1–2: basic research, 2–4: preclinical research, 6: phase I trials; 7: phase II trials, 8: phase III trials, 9: phase IV trials and deployment). fNIRS, functional near-infrared spectroscopy; LiDAR, light detection and ranging; PIR, passive infrared motion.

such as triaxial accelerometers, a respiratory acoustic sensor and an eye movement sensor, was developed to monitor eye movement frequency, direction and amplitude during sleep as candidate DBMs for NDDs⁹⁸. Nonetheless, e-textiles are still not common because integrating reliable electronics into flexible, washable fabrics presents considerable manufacturing challenges. Sensor performance can degrade with repeated use and fabric wear and tear, embedded components complicate laundering and maintenance, and production costs remain higher than for conventional wearables.

Electronic skin patches and body network sensors. Electronic skin patches are wearable-like devices that adhere directly to the skin. Although they could be equipped with conventional IMUs, they are particularly suited to measuring electrophysiological signals, including EEG, electrooculography, electromyography, electrocardiography and bodily fluids such as sweat. In an exploratory, pilot study ($n = 25$), the NIMBLE biosensor skin patch, which contains an accelerometer and EMG sensors, was placed on various locations on the body to record movements at home and in the clinic in patients with PD⁹⁹. Results revealed a correlation between clinically observed and sensor algorithm-predicted UPDRS motor symptom assessment scores with predicted scores within a ± 1 range 91% of the time. Similarly, a wearable, microfluidic patch-work was developed for near real-time continuous measurement of at-rest thermoregulatory sweat composition and rate on various body sites¹⁰⁰. Tested in a feasibility study with four participants, the device tracked sweat secretion and sweat levodopa, the primary treatment for PD. Temporary-tattoo EEGs can also be used to identify sleep behaviour disorders in individuals with NDDs¹⁰¹.

Portable sensors

Smartphones and tablets. Smartphones and tablets have built-in sensors that offer a platform for collecting a variety of digital measures, including accelerometer, GPS, microphone and other sensor-derived metrics, to monitor different health-related behaviours and cognitive functions². In this Review, we do not explicitly focus on ‘active’ tasks, which make up most smartphone-based use cases and often enable instrumented assessments. Instead, we emphasize what has been

measurable in ‘passive’ everyday settings, as these measurements capture more natural behaviours with no additional user burden, allow for continuous data collection without requiring sustained compliance from patients who could experience cognitive or motor difficulties, and offer the most clinically viable approach for early detection and longitudinal monitoring as diseases progress. For example, smartphone typing and tablet handwriting or drawing dynamics can be measured to monitor fine motor decline in individuals with PD and MCI^{102,103}. Similarly, speech samples collected from smartphones and tablets can help distinguish healthy individuals from those with MCI³⁶ and detect subtypes of primary progressive aphasia and frontotemporal dementia¹⁰⁴. Additionally, DBMs of cognitive performance in personalized games have been repeatedly used to detect PD¹⁰⁵, which could be considered passive tasks owing to their entertainment component. Overall, smartphones and tablets are powerful for collecting longitudinal, high-frequency data that reflect fluctuating symptoms in real-world settings. This data will be especially useful as complementary information to traditional clinical assessments and more detailed patient monitoring between clinical visits.

Inner-body sensors

Ingestible sensors. Ingestible sensors are non-invasive electronic devices, often designed as pills or capsules (digital medicine systems) that can be swallowed and activated to release a solution or mechanism within the gastrointestinal tract^{106–108}. This approach comes with direct access to measure otherwise hard-to-obtain quantities throughout the gastrointestinal tract, such as pH, temperature and pressure, while taking images of the passage. For example, an ingestible pill was used to monitor and transmit core body temperature dysregulation in individuals with PD ($n = 12$) experiencing REM sleep behaviour disorder symptoms¹⁰⁶. The study revealed that people with PD showed reduced core body temperature patterns that correlate with REM sleep behaviour disorder severity.

Ingestibles are commonly used to administer medication and monitor medication adherence in patients, whereby the sensor is co-encapsulated with the medication and transmits a signal upon dissolution in the stomach. The FDA recently approved the first ingestible

sensor (MIND1) with aripiprazole, [Abilify MyCite](#), for individuals with dementia-related psychosis (NCT02219009). Its performance was compared alongside a wearable sensor and mobile application for medication management in a medium-sized study ($n = 67$) of individuals with schizophrenia¹⁰⁷. Over half of participants could independently pair and apply the patch by the end of the study, which suggests improvements in usability. Ingestible sensors are also being used in NDDs to monitor the gut–brain axis, a new area of research that links gut dynamics (such as motility and dysbiosis) to conditions such as PD¹⁰⁸.

Implantable sensors. Implantable sensors are surgically placed inside the body to continuously monitor specific health parameters. We specifically refer to implantable sensors that can be used in at-home settings for tracking everyday activities such as implantable brain–computer (or brain–machine) interfaces, neuroprostheses and cardioverter-defibrillators. Although brain–computer interfaces have been used to administer clinical interventions for long-term cognitive decline^{109,110} or to analyse and interpret brain signals into useful commands that can replace lost function¹¹¹, they are not actively being used as DBMs for disease monitoring. As an example, a minimally invasive, degradable, implantable ultrasound metagel sensor was used for monitoring intracranial biomarkers in vivo, such as pressure, temperature, pH and blood flow velocity, to identify the onset and progression of NDDs associated with injuries like chronic traumatic encephalopathy¹¹².

Ambient and contactless sensors

Ambient sensors often operate in the background and are integrated into an individual's surrounding environment to collect and infer passive data. Once installed, ambient sensors require very little maintenance and are therefore less affected by limited patient adherence.

Moreover, they are applicable in cases where a person might not be able to reliably maintain most other types of devices such as during later stages of various NDDs.

Vision-based sensors. Although not commonly used at home owing to privacy concerns (Box 2) and limitations with environment illumination, vision-based sensors, such as RGB-depth cameras, infrared, thermal, and light detection and ranging (LiDAR) sensors provide a non-intrusive, low-cost modality of ambient sensing that captures high-resolution, visual data from the environment, including gait and posture^{43–45,113}. Portable fNIRS systems have been widely used to assess brain function in individuals with MCI^{114,115}. For example, fNIRS and gait measurements revealed that an MCI cohort ($n = 57$) exhibited lower activation in the primary motor cortex, secondary motor cortex and parietal brain regions compared to the healthy group during the motor execution stage of a prepared walking task¹¹⁴. Similarly, assessment of gait impairment via an fNIRS device in a dual-task walking experiment revealed substantial between-group variability ($n = 16$ cognitively impaired; $n = 38$ cognitively healthy) on functional connectivity strength change values and region-of-interest connection strength change values¹¹⁵. Additionally, LiDAR technology, which uses light in the form of a pulsed laser to measure distances, has been used to create precise 2D and 4D maps of the environment to detect gait parameters from usual walking patterns to predict the risk of falls and neurocognitive performance in older adults¹¹⁶. This accuracy enables a detailed assessment of physical activity and coordination, which can be valuable for detecting early signs of neurodegenerative disorders.

Other ambient and smart environment sensors. Since the mid-2010s, simple passive infrared motion sensors to detect activity in equipped

Box 2 | Data considerations

Digital biomarker development generates sensitive health data that falls under regulatory frameworks such as the Health Insurance Portability and Accountability Act (HIPAA) in the USA and the General Data Protection Regulation (GDPR) in Europe. However, a critical gap persists where most early-stage digital biomarker research occurs in academic or pilot settings without rigorous privacy oversight, and protections are often retrofitted rather than built in from the start. Addressing this limitation requires privacy-by-design principles, an approach in which data protection is embedded into the architecture of a system from its inception, rather than added as an afterthought. In practice, this means adopting data minimization (collecting only what is strictly necessary), on-device processing where feasible and dynamic consent management systems that allow participants to adjust their data-sharing preferences as studies evolve.

Algorithmic fairness presents equally pressing concerns. Machine learning models trained on data that underrepresent marginalized communities risk encoding and amplifying existing health disparities. Populations without reliable digital access might be excluded entirely, and algorithms that fail to account for social determinants of health can produce systematically biased outputs. These risks are not hypothetical; for example, in Parkinson disease, women more commonly present with dyskinesias and depression, while men more often exhibit rigidity¹⁵⁶. This illustrates how demographic

differences can confound models that treat patient populations as homogeneous. Mitigating these biases requires algorithmic impact assessments across demographic subgroups, community advisory boards with diverse stakeholder representation, and disaggregated performance reporting by race, ethnicity, gender, socioeconomic status, and digital literacy.

Transparency in reporting is essential for building trust and enabling independent scrutiny. Studies should document detailed participant demographics, including language proficiency and digital literacy, stratified model performance metrics, and inclusion/exclusion criteria alongside their equity implications. Where possible, researchers should release open-source code and synthetic datasets to facilitate independent validation and provide plain-language summaries accessible to patient communities. Several frameworks and checklists support these goals, including the [National Institute for Health Care and Excellence \(NICE\) Evidence Standards Framework for Digital Health Technologies](#), the Consolidated Standards of Reporting Trials–Artificial Intelligence (CONSORT-AI)¹⁵⁷, the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT-AI) reporting guidelines¹⁵⁸, the Minimum Information about Clinical Artificial Intelligence Modelling (MI-CLAIM) checklist¹⁵⁹, and the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD+AI) checklist^{160,161}.

rooms or door contact sensors to detect opening and closing patterns in individuals with MCI and Alzheimer disease have been used^{10,46,58}. These sensors can extract a wide array of potential DBMs, including gait speed and physical activity measured during clinical visits or at home. Sleep or presence sensors placed near or under the bed accurately and unobtrusively quantify sleep-related and movement markers such as sleep duration or heart rate patterns during sleep that can be associated with MCI and Alzheimer disease⁵⁸. Noise, air quality, light and ambient room temperature sensors also hold clinical relevance for NDD populations. Long-term community noise exposure has been associated with a higher risk of both MCI and Alzheimer disease in older adults¹¹⁷ whereas inadequate ambient light disrupts circadian rhythms and exacerbates behavioural symptoms in individuals with dementia^{76,118}. Indoor environmental quality sensors monitoring temperature, humidity and air quality can be integrated into Internet of Things systems to unobtrusively track daily living and improve care for individuals with MCI¹¹⁹, and similar applications in dementia care facilities have shown utility in detecting conditions that may worsen symptoms or compromise safety¹²⁰. Additionally, radio-wave-based sensors like ultra-wideband and millimetre-wave (mmWave) radar (low-cost privacy-preserving sensors that emit radio waves to detect motion and body position without cameras) have been deployed in at-home settings to recognize nocturnal breathing patterns and daily human activities such as walking^{72,121}. In a year-long observational study of 50 participants, passively monitored gait speed from radio waves correlated strongly with MDS-UPDRS Part III scores, declined nearly twice as fast in individuals with PD versus controls, and captured intraday medication response fluctuations⁶¹. A large, cross-sectional study including individuals with PD ($n = 25$) used a contactless, low-power, wireless radio device to collect radio signals from the environment and extract nocturnal breathing signals. Using one night of nocturnal breathing signals analysed with an artificial intelligence model could predict PD with an AUC of 0.906 and a sensitivity of 86.23%, as well as a strong correlation between the model's severity prediction and the MDS-UPDRS Part III ($r = 0.93$)⁶⁰. Similar radio-signal-based sensors, such as the Emerald Device, are also being explored to monitor gait, home activity and time in bed in the context of PD⁵⁹.

Categorization by utility and application

DBMs can redefine health care by providing precise, personalized and timely data. Their applications span various domains, such as improving treatment optimization, enabling personalized medicine, and ensuring continuous at-home, remote and clinical monitoring (Fig. 1).

Treatment optimization

DBMs can directly optimize treatment by transforming episodic symptom reports into time-resolved measures that map onto actionable decisions such as medication timing, dose adjustment, therapy selection or trigger-based interventions. For instance, in PD, wrist-worn devices such as the Parkinson KinetiGraph or Personal KinetiGraph (PKG) quantify longitudinal bradykinesia, dyskinesia, tremor, fluctuations and sleep-related features outside the clinic, which allows clinicians to distinguish 'wearing-off' patterns from baseline impairment and to adjust levodopa schedules accordingly^{82,122}. In an observational study of people with PD ($n = 70$), PKG review changed the initial treatment plan in 31.8% of participants, and clinician assessment frequently differed from PKG-derived profiles for bradykinesia/wearing-off, dyskinesia and sleep¹²². In a separate routine-care report, PKG data contributed to treatment changes in a large fraction of visits/patients (reported

as 84% of visits and 79% of patients in that clinic cohort), highlighting how DBMs can shift management beyond what is inferred from a short exam and self-report alone¹²³.

DBM-informed optimization is also well suited for semi-automated care pathways. For example, continuous monitoring can support standardized escalation rules (for example, sustained increases in bradykinesia burden or dyskinesia time) that prompt clinician review and shorten delays between symptom worsening and treatment adjustment. In cognitive decline, performance-based DBMs derived from computerized cognitive tasks (such as reaction time distributions, error profiles, learning curves and fatigue slopes)¹²⁴ can similarly be used to adapt training intensity and content, enabling protocols to be titrated to the individual's response rather than delivered as a fixed 'dose' (this is consistent with the broader concept of DBMs enabling tailored, longitudinal monitoring and treatment).

Personalized medicine using artificial intelligence

Artificial intelligence-enabled DBMs can help support the transition to personalized medicine and to better understand and treat complex neurological conditions. At its core, this integration enables what has been termed 'deep phenotyping' – the extraction of fine-grained, longitudinal phenotypic components that describe how disease manifests and evolves within a specific individual rather than an 'average' patient^{125,126}. In NDDs, this is particularly valuable because clinical syndromes are heterogeneous across motor and non-motor domains, and progression trajectories vary substantially. A practical instantiation of 'deep phenotyping' is now feasible with commodity devices like the WATCH-PD study, a multicentre observational effort showing that a commercially available smartwatch paired with a smartphone research application can quantify key motor and non-motor features in early, untreated PD¹²⁷. The study captured longitudinal motor (such as tremor, gait and arm movement) and non-motor features (such as speech and cognitive tasks) in people with early, untreated PD ($n = 82$) and monitored patient-specific trajectories over 12 months.

Personalized medicine applications can also integrate DBMs with high-dimensional biological and health-system data to refine risk estimation and therapeutic matching¹²⁵. At the systems level, artificial intelligence-enabled precision medicine platforms (for example, [Helix's](#) clinical interpretation stack; Syapse, now within [N-Power Medicine](#)) show how genomic or real-world evidence infrastructures can be paired with digital measures to stratify risk and guide individualized care or research pathways¹²⁸.

Adverse events and safety monitoring

DBMs support safety monitoring by enabling continuous, remote detection of clinically meaningful risk states and by shortening the time between risk escalation and intervention¹⁰. In older adults and people living with NDDs, unobtrusive wearables and ambient sensors can track safety-relevant functions, such as gait stability, turning smoothness and activity fragmentation, without requiring frequent clinic visits³⁰. Safety monitoring becomes materially more important when DBMs are tied to predictive signals and actionable care pathways. In the ISAAC and ORCATECH living laboratory studies, unobtrusive in-home assessment of walking speed was evaluated for identifying future falls ($n = 125$)¹²⁹. Results demonstrated how continuous gait-speed trajectories can function as an early warning signal that supports proactive interventions (for example, home safety modifications, physical therapy referral, and medication review) before an adverse event occurs. In NDDs, analogous approaches can be extended to cognitive and functional safety

markers (routine stability, nighttime activity, medication adherence proxies)¹⁰ that allow for care plan adjustments and resource allocation responsive to real-world decline patterns (consistent with the broader role of DBMs in continuous monitoring).

Clinical research

In clinical research, DBMs are increasingly operated as digital end points that complement or, in some settings, potentially substitute clinical scales by providing higher-frequency, ecologically valid measures of function. A concrete example is the use of smartphone-based active tests, such as finger tapping speed and rhythm, and passive mobility measures (such as step-to-step variability) collected repeatedly during longitudinal studies to quantify motor fluctuations and track symptom trajectories between visits¹³⁰. In a phase III PD clinical trial sub-study (NCT02642393), these mobility measures correlated with MDS-UPDRS scores and reduced participant burden outcomes relative to solely in-clinic assessments, thereby enabling more efficient trial designs and remote follow-up¹³⁰. In parallel, vision-based biomarkers^{43–45} extracted from standardized videos (such as recorded MDS-UPDRS assessments) can provide objective kinematic features that reduce rater variability and scale clinician time, while preserving clinically interpretable links to established motor outcomes. Together, these DBM-enabled end points support decentralized data collection, richer characterization of treatment response and, when validated, can reduce required sample sizes for prevention and progression trials by increasing measurement precision^{131,132}, as shown in conditions like dementia (NCT05385913).

Outlook

DBMs are transforming how diseases are monitored, diagnosed and treated, with a strong focus on detecting the earliest pathological changes in pre-symptomatic and prodromal disease stages. Wearable sensors, smartphones and ambient monitoring systems have enabled continuous assessment of motor, cognitive and behavioural symptoms outside clinical settings to identify subtle, functional health alterations that might precede clinical manifestations. This fundamental change provides a deeper understanding of disease trajectories and treatment responses, which are particularly relevant for neurodegenerative disorders characterized by heterogeneous symptom presentations and fluctuations that traditional, episodic assessments often fail to capture.

Current evidence reveals clear patterns in DBM effectiveness across different health aspects and NDDs. Motor-related symptom quantification consistently demonstrates the highest success rates across multiple NDDs. Within motor symptoms, gait analysis has emerged as the most universally effective technology, showing robust performance across multiple NDDs^{133–135}. Disease-specific motor-related DBMs are also promising; motor subtyping⁴⁵, hand tremor and bradykinesia quantification⁹⁷, and keystroke dynamics⁵¹ can be used to identify individuals with PD from healthy controls, whereas speech analysis can classify both individuals with Alzheimer disease and PD (although speech arguably reflects more than purely motor dysfunction)³⁸.

The dominance of motor-related DBMs probably stems from multiple factors; these symptoms are easier to measure using widely available IMUs and often correlate directly with disease processes and progression. By contrast, more abstract concepts, such as cognition, are more challenging to measure, particularly when assessed passively. Among NDDs, PD demonstrates the most mature DBM ecosystem owing to its prominent motor manifestations. However, conditions like MCI or Alzheimer disease present different challenges as treatment

options remain limited even after diagnosis, restricting the real-world utility of monitoring outside clinical trial contexts.

This Review highlights several key advances: taxonomic organization of DBMs according to sensing modalities (wearable, portable, inner body and ambient) clarifies the technical landscape; categorization by health aspect (cognitive, motor, behavioural, physiological, environmental and molecular) helps match technologies to specific symptom domains; classification by utility (diagnosis, monitoring and treatment optimization) establishes a clear intended use for DBMs. Despite these advances, knowledge gaps and challenges remain.

Existing sensing technologies come with limitations as to what can be measured; even in the best-case scenario, they are limited by available sensing technologies (Box 3). For example, new molecular wearable sensing, which leverages sweat, interstitial fluid and breath analysis, could enable the monitoring of various 'traditional' biomarkers in digital settings, including inflammation, glucose levels or stress hormones¹³⁶. In the future, it might be possible to directly monitor blood-based biomarkers like plasma amyloid- $\beta_{42:40}$ ratio and phosphorylated tau181 and phosphorylated tau217 in the context of Alzheimer disease¹⁹. Innovations such as the 'electronic nose', which mimics olfactory sensing to detect volatile organic compounds in breath or sebum, could be used for early disease detection in neurological disorders such as Alzheimer disease and PD¹³⁷. Additionally, miniaturized implantable sensors¹³⁸, optical biosensing with flexible technologies¹³⁹ and bioimpedance monitoring¹⁴⁰ support the tracking of cardiovascular, neurological and metabolic health in real-time. Integrating artificial intelligence-driven analysis with these sensing modalities can further empower new sensing capabilities.

Future research should develop robust frameworks for integrating data across multiple sensing modalities. Currently, most studies leverage single sensors focusing on isolated symptoms but NDDs manifest across multiple domains (for example, cognitive decline, motor dysfunction, sleep disturbances and emotional dysregulation). Integrating independent sensors comes with various challenges (Table 2) such as sensors running out of sync, variability in sensor quality across devices, battery life limitations and computing power restrictions⁴ (Supplementary Table S1). Although these issues can be corrected mathematically to some degree¹⁴¹, manufacturers are often not incentivized to provide synchronization capabilities for interoperability because proprietary signal-processing algorithms are usually protected as core intellectual property and resources are typically directed towards consumer-facing features like user experience and battery life. Computational methods that combine data from wearables, smartphones and ambient sensors could provide a more holistic disease profile; for example, composite DBMs that integrate motor assessments (such as gait parameters from wearable accelerometers) with cognitive measures (such as attention metrics from smartphone interactions) and physiological parameters (such as heart rate variability from smartwatches and sleep patterns from bed sensors)^{142,143}.

Despite artificial intelligence algorithms having shown promise in analysing DBMs, most current approaches apply uniform models across patients. The heterogeneity of NDDs would benefit from personalized models that account for individual baseline characteristics and disease subtypes while remaining interpretable. Interpretable algorithms that adapt to personal baselines and detect subtle deviations could help indicate each individual's disease progression or treatment response.

Moreover, **foundation models** (models pre-trained on billions or trillions of samples on vast multimodal datasets including omics data,

Box 3 | Technology transfer considerations

Every digital biomarker study confronts a fundamental trade-off between measurement quality and practical feasibility. Research-validated devices, such as the ActiGraph, offer superior reliability but come at a higher cost and limited scalability, whereas bring-your-own-device approaches using participants' smartphones and smartwatches reduce cost and support larger recruitment but introduce complex validation requirements arising from hardware and software heterogeneity. A parallel tension exists between obtrusiveness and adherence: high-precision, high-burden methods, such as 12-lead electrocardiogram or polysomnography, require clinic or laboratory visits, moderate-burden options like single-lead electrocardiogram patches and wearable sleep trackers balance precision with home use, and low-burden solutions, such as wrist-based photoplethysmography or contactless smart mattress sensors, maximize long-term adherence at the expense of some measurement fidelity. Matching technology to study design is therefore critical. Long-term studies spanning months to years should prioritize ambient sensing and passive solutions to minimize participant fatigue, whereas short pivotal trials can tolerate more obtrusive approaches. Population characteristics also matter. Cognitive impairment might preclude technologies requiring active engagement, making passive and ambient systems essential for later-stage neurodegenerative disease research. Emerging ambient sensors, such as the Emerald device, combined with artificial intelligence and edge computing, are narrowing the performance gap with wearables while maintaining minimal user burden.

Care-assisting technologies for patients with neurodegenerative diseases are expanding rapidly, fuelled by an ageing population on the demand side and the convergence of artificial intelligence with ubiquitous smartphones and wearables on the supply side. One model to support innovation involves strategic partnerships between technology firms, pharmaceutical companies and research institutions, exemplified by Biogen and Apple's 'Intuition' brain health study using the Apple Watch for cognitive monitoring (NCT05058950)¹⁴⁹. However, reliance on large technology platforms introduces 'platform risk', whereby platform holders maintain control over core algorithms, sensors, and data infrastructure and their business priorities might migrate away from regulated healthcare applications. For example, when Apple modified its HealthKit data access policies in 2019, several digital health companies experienced

operational disruptions, a cautionary example for neurodegenerative disease applications where long-term longitudinal data collection is essential.

Innovative startups such as [Rune Labs](#), [Altoida](#) and [Indivi](#) are attempting to mitigate platform risk by pursuing business-to-business models in which they provide pharmaceutical companies with validated digital end points and remote monitoring services for clinical trials. This strategy generates near-term revenue while building clinical evidence for future regulatory approvals. However, it creates its own constraints: pharmaceutical companies typically seek digital biomarkers that can serve as surrogate end points to accelerate drug development timelines or improve trial efficiency, which might not align with the broader utility these technologies could provide in real-world clinical practice.

The direct-to-consumer market offers an increasingly viable alternative pathway. Companies like [MindX Sciences](#) engage consumers around holistic lifestyle improvement, whereas disease-specific platforms such as [Rune Labs' StrivePD](#) and the Michael J. Fox Foundation's [Fox Insight](#) demonstrate how patient-driven data collection creates value both for individuals managing their conditions and for researchers requiring real-world evidence. This expansion from preventive into secondary care signals maturation in healthcare technology adoption that could accelerate digital biomarker uptake through consumer demand rather than insurance-based payment models.

A broader challenge is that regulatory and reimbursement systems were not designed for technology-based service delivery. Moreover, technology development is inherently iterative: tools available today will improve with wider adoption and continued development, yet assessing performance too strictly at an early stage risks foreclosing the opportunity for a better tool tomorrow. Supporting this iterative trajectory would represent a paradigm shift in how the commercialization of new technologies is approached. Until reimbursement frameworks evolve to accommodate these realities, digital biomarker systems must simultaneously pursue multiple commercialization channels (business-to-business partnerships, direct-to-consumer offerings and clinical integration pilots) to sustain development and build the evidence base necessary for long-term adoption.

medical imaging, continuous wearable sensor streams and electronic health records) could enable large-scale modelling of complex biological and physiological data and help uncover hidden patterns and new biomarkers. Although interpretability remains an important issue and subject of active research, the sheer performance of foundation models could still bring benefits for lower-stakes use cases like population-level screening and enriching clinical trial cohorts. Foundation models support self-supervised learning (learning without supervision/ground truth), which could be further explored in the context of biomedical time series data and DBMs, where large quantities of unlabelled data are the norm. Furthermore, with advancements in cross-modal and transfer learning, foundation models could support adaptation to diverse populations and conditions, making DBM discovery more robust, scalable and clinically meaningful. Nonetheless, foundation

models such as large language models suffer from hallucination, which requires caution and extensive domain expertise to be exercised in the context of biomedical research.

Another critical limitation is the predominance of cross-sectional studies with small sample sizes. Large-scale, longitudinal studies are essential to establishing the prognostic value of DBMs. The field would benefit from multicentre initiatives that track participants from pre-symptomatic stages through disease progression over 5–10 years, like the Alzheimer's Disease Neuroimaging Initiative¹⁴⁴, the Parkinson's Progression Markers Initiative¹⁴⁵, and the African American Dementia and Aging Project¹⁴⁶, but focus on digital assessments. These studies should include diverse populations across age, ethnicity, socioeconomic status and technology literacy to ensure generalizability. Apart from major initiatives, such as [Mobilise-D](#) and [RADAR-CNS](#), typical

Alzheimer disease studies include fewer than 100 participants and lack external validation^{147,148}. A focus on ubiquitous and easily scalable mobile and wearable consumer devices could help facilitate this transition, as demonstrated by a collaboration between Biogen and Apple¹⁴⁹. Open-source digital health frameworks for consumer operating systems (like iOS and Android) could further foster the implementation and standardization of larger longitudinal studies^{150,151}.

The cost-effectiveness of implementing DBMs remains largely unexamined, which hinders DBM adoption. For example, high family-related out-of-pocket costs are associated with dementia care and probably increase with dementia severity, therefore supporting the need for home-based or telehealth interventions¹⁵². Moreover, initial evidence indicates that DBMs could drive down sample sizes for clinical trials in NDDs, which would translate to substantial cost savings in clinical trials^{132,153,154}. Health economics research should quantify the potential value of these technologies more broadly in terms of earlier diagnosis, optimized treatment, reduced hospitalization and improved quality of life. These analyses should consider various stakeholder perspectives, including healthcare systems, insurers and patients, to better inform reimbursement policies and implementation strategies.

Published online: 23 April 2026

References

- Babrack, L. M. et al. Traditional and digital biomarkers: two worlds apart? *Digit. Biomark.* **3**, 92–102 (2019).
- Dagum, P. Digital biomarkers of cognitive function. *npj Digit. Med.* **1**, 10 (2018).
- Tharick, A. P. et al. Insights into the use of biomarkers in clinical trials in Alzheimer's disease. *EBioMedicine* **108**, 105322 (2024).
- Dorsey, E. R. et al. The first frontier: digital biomarkers for neurodegenerative disorders. *Digit. Biomark.* **1**, 6–13 (2017).
- Feigin, V. L. et al. Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* **18**, 459–480 (2019).
- Weiner, M. W. et al. Increasing participant diversity in AD research: plans for digital screening, blood testing, and a community-engaged approach in the Alzheimer's Disease Neuroimaging Initiative 4. *Alzheimers Dement.* **19**, 307–317 (2023).
- Vasudevan, S. et al. Digital biomarkers: convergence of digital health technologies and biomarkers. *npj Digit. Med.* **5**, 36 (2022).
- Abernethy, A. et al. The promise of digital health: then, now, and the future. *NAM Perspectives* <https://doi.org/10.31478/202206e> (2022).
- Juan, M. C. et al. Intelligent environment for monitoring Alzheimer patients, agent technology for health care. *Decis. Support Syst.* **44**, 382–396 (2008).
- Piau, A. et al. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild Alzheimer disease and implications for clinical care: systematic review. *J. Med. Internet Res.* **21**, e12785 (2019).
- Au, R., Kolachalama, V. B. & Paschalidis, I. C. Redefining and validating digital biomarkers as fluid, dynamic multi-dimensional digital signal patterns. *Front. Digit. Health* **3**, 751629 (2022).
- Leyens, L. et al. Why language matters in digital endpoint development: harmonized terminology as a key prerequisite for evidence generation. *Digit. Biomark.* **8**, 1–12 (2024).
- Coravos, A., Khozin, S. & Mandl, K. D. Developing and adopting safe and effective digital biomarkers to improve patient outcomes. *npj Digit. Med.* **2**, 14 (2019).
- Montag, C., Elhai, J. D. & Dagum, P. On blurry boundaries when defining digital biomarkers: how much biology needs to be in a digital biomarker? *Front. Psychiatry* **12**, 740292 (2021).
- Izmailova, E. S., Demanuele, C. & McCarthy, M. Digital health technology derived measures: biomarkers or clinical outcome assessments? *Clin. Transl. Sci.* **16**, 1113–1120 (2023).
- FDA-NIH Biomarker Working Group. BEST (Biomarkers, Endpoints, and other Tools). Glossary <https://www.ncbi.nlm.nih.gov/books/NBK338448/> (FDA, 2016).
- Ortiz-Martinez, M. et al. Recent developments in biomarkers for diagnosis and screening of type 2 diabetes mellitus. *Curr. Diab. Rep.* **22**, 95–115 (2022).
- Albani, D. et al. Replication study to confirm the role of CYP2D6 polymorphism rs1080985 on donepezil efficacy in Alzheimer's disease patients. *J. Alzheimers Dis.* **30**, 745–749 (2012).
- Mielke, M. M. & Fowler, N. R. Alzheimer disease blood biomarkers: considerations for population-level use. *Nat. Rev. Neurol.* **20**, 495–504 (2024).
- Sun, R. et al. Apolipoprotein E polymorphisms and Parkinson disease with or without dementia: a meta-analysis including 6453 participants. *J. Geriatr. Psych. Neurol.* **32**, 3–15 (2019).
- Navarro, V. J. et al. Liver injury from herbals and dietary supplements in the US Drug-Induced Liver Injury Network. *Hepatology* **60**, 1399–1408 (2014).
- de Kock, R. et al. Circulating biomarkers for monitoring therapy response and detection of disease progression in lung cancer patients. *Cancer Treat. Res. Commun.* **28**, 100410 (2021).
- Kapoor, R. et al. Serum neurofilament light as a biomarker in progressive multiple sclerosis. *Neurology* **95**, 436–444 (2020).
- Mohd, J. et al. Sensors for daily life: a review. *Sensors Int.* **2**, 100121 (2021).
- Choi, J. et al. Prefrontal EEG slowing, synchronization, and ERP peak latency in association with prodromal stages of Alzheimer's disease. *Front. Aging Neurosci.* **15**, 1131857 (2023).
- Rutkowski, T. M., Abe, M. S., Komendzinski, T. & Otake-Matsuura, M. Older adult mild cognitive impairment prediction from multiscale entropy EEG patterns in reminiscent interior image working memory paradigm. In *43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* 6345–6348 (IEEE, 2021).
- Talamonti, D. et al. Exploring cognitive and brain oxygenation changes over a 1-year period in physically active individuals with mild cognitive impairment: a longitudinal fNIRS pilot study. *BMC Geriatr.* **22**, 648 (2022).
- Boto, E. et al. Wearable neuroimaging: combining and contrasting magnetoencephalography and electroencephalography. *NeuroImage* **201**, 116099 (2019).
- Kourtis, L. C. et al. Digital biomarkers for Alzheimer's disease: the mobile/wearable devices opportunity. *npj Digit. Med.* **2**, 9 (2019).
- Aghaei, M. et al. Association between a computerized, self-administered cognitive assessment and phosphorylated Tau in Alzheimer's disease. Preprint at medRxiv <https://doi.org/10.1101/2024.11.23.24317846> (2024).
- Bayot, M. et al. The contribution of executive control dysfunction to freezing of gait in Parkinson's disease. *Clin. Neurophysiol.* **152**, 75–89 (2023).
- Parra, M. A., Granada, J. & Fernández, G. Memory-driven eye movements prospectively predict dementia in people at risk of Alzheimer's disease. *Alzheimers Dement.* **14**, e12386 (2022).
- Meier, I. B. et al. Using a digital neuro signature to measure longitudinal individual-level change in Alzheimer's disease: the Altsoida large cohort study. *npj Digit. Med.* **4**, 101 (2021).
- Antoniades, C. A. & Spering, M. Eye movements in Parkinson's disease: from neurophysiological mechanisms to diagnostic tools. *Trends Neurosci.* **47**, 71–83 (2023).
- Endo, T. et al. Initial response time measurement in eye movement for dementia screening test. In *2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)* 262–265 (IEEE, 2017).
- Themistocleous, C., Eckerström, M. & Kokkinakis, D. Voice quality and speech fluency distinguish individuals with mild cognitive impairment from healthy controls. *PLoS ONE* **15**, e0236009 (2020).
- Rusz, J. et al. Speech biomarkers in rapid eye movement sleep behavior disorder and Parkinson disease. *Ann. Neurol.* **90**, 62–75 (2021).
- López-de-Ipiña, K. et al. On automatic diagnosis of Alzheimer's disease based on spontaneous speech analysis and emotional temperature. *Cognitive Comput.* **7**, 44–55 (2015).
- Hajjar, I. et al. Development of digital voice biomarkers and associations with cognition, cerebrospinal biomarkers, and neural representation in early Alzheimer's disease. *Alzheimers Dement.* **15**, e12393 (2023).
- Laganas, C. et al. Parkinson's disease detection based on running speech data from phone calls. *IEEE Trans. Biomed. Eng.* **69**, 1573–1584 (2021).
- Nouriani, A. et al. Real world validation of activity recognition algorithm and development of novel behavioral biomarkers of falls in aged control and movement disorder patients. *Front. Aging Neurosci.* **15**, 1117802 (2023).
- Demonceau, M. et al. Contribution of a trunk accelerometer system to the characterization of gait in patients with mild-to-moderate Parkinson's disease. *IEEE J. Biomed. Health Inform.* **19**, 1803–1808 (2015).
- Lu, M. et al. Quantifying Parkinson's disease motor severity under uncertainty using MDS-UPDRS videos. *Med. Image Anal.* **73**, 102179 (2021).
- Endo, M. et al. GaitForeMer: self-supervised pre-training of transformers via human motion forecasting for few-shot gait impairment severity estimation. *Med. Image Comput. Comput. Assist. Interv.* **13438**, 130–139 (2022).
- Endo, M. et al. Data-driven discovery of movement-linked heterogeneity in neurodegenerative diseases. *Nat. Mach. Intell.* **6**, 1034–1045 (2024).
- Schütz, N. et al. A systems approach towards remote health-monitoring in older adults: introducing a zero-interaction digital exhaust. *npj Digit. Med.* **5**, 116 (2022).
- Muheidat, F. & Tawalbeh, L. A. In-home floor based sensor system-smart carpet-to facilitate healthy aging in place (AIP). *IEEE Access* **8**, 178627–178638 (2020).
- Zhu, W., Anderson, B., Zhu, S. & Wang, Y. A computer vision-based system for stride length estimation using a mobile phone camera. In *Proc. 18th International ACM SIGACCESS Conference on Computers and Accessibility* 121–130 (ACM, 2016).
- Guorui, C. et al. Neural network-assisted personalized handwriting analysis for Parkinson's disease diagnostics. *Nat. Chem. Eng.* **2**, 358–368 (2025).
- Deng, K. et al. Heterogeneous digital biomarker integration out-performs patient self-reports in predicting Parkinson's disease. *Commun. Biol.* **5**, 58 (2022).
- Dimitrios, I. et al. Motor impairment estimates via touchscreen typing dynamics toward Parkinson's disease detection from data harvested in-the-wild. *Front. ICT* **5**, 28 (2018).
- Skibińska, J. & Hosek, J. Computerized analysis of hypomimia and hypokinetic dysarthria for improved diagnosis of Parkinson's disease. *Heliyon* **9**, 11 (2023).

53. Bandini, A. et al. Analysis of facial expressions in parkinson's disease through video-based automatic methods. *J. Neurosci. Methods* **281**, 7–20 (2017).
54. Yury, S. et al. Hypomimia detection with a smartphone camera as a possible self-screening tool for Parkinson disease (P3.047). *Neurology* **90**, P3–047 (2018).
55. Greco, D., Masulli, F., Rovetta, S., Cabri, A. & Daffonchio, D. A cost-effective eye-tracker for early detection of mild cognitive impairment. In *2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON)* 1141–1146 (IEEE, 2022).
56. Agop, A., Dimashki, N., Hage-Diab, A., Abdul-Nabi, S. & Sharif-Askari, E. Feasibility of developing an in-home Parkinson's disease early diagnosing system. In *2021 Sixth International Conference on Advances in Biomedical Engineering (ICABME)* 103–107 (IEEE, 2021).
57. Brien, D. C. et al. Classification and staging of Parkinson's disease using video-based eye tracking. *Parkinsonism Relat. Disord.* **110**, 105316 (2023).
58. Schutz, N. et al. A sensor-driven visit detection system in older adults' homes: towards digital late-life depression marker extraction. *IEEE J. Biomed. Health Inform.* **26**, 1560–1569 (2021).
59. Kabelac, Z. et al. Passive monitoring at home: a pilot study in Parkinson disease. *Digital Biomarkers* **3**, 22–30 (2019).
60. Yang, Y. et al. Artificial intelligence-enabled detection and assessment of Parkinson's disease using nocturnal breathing signals. *Nat. Med.* **28**, 2207–2215 (2022).
61. Liu, Y. et al. Monitoring gait at home with radio waves in Parkinson's disease: a marker of severity, progression, and medication response. *Sci. Transl. Med.* **14**, eadc9669 (2022).
62. Donahue, E. K. et al. Physical activity intensity is associated with cognition and functional connectivity in Parkinson's disease. *Parkinsonism Relat. Disord.* **104**, 7–14 (2022).
63. Jiang, Z. et al. Automated analysis of facial emotions in subjects with cognitive impairment. *PLoS ONE* **17**, e0262527 (2022).
64. Reichmann, H. et al. The use of wearables for the diagnosis and treatment of Parkinson's disease. *J. Neural Transm.* **130**, 783–791 (2023).
65. Tsiouris, K. M. et al. PD Manager: an mHealth platform for Parkinson's disease patient management. *Healthcare Technol. Lett.* **4**, 102–108 (2017).
66. Knezevic, E. et al. The role of cortisol in chronic stress, neurodegenerative diseases, and psychological disorders. *Cells* **12**, 2726 (2023).
67. Kikhia, B. et al. Utilizing a wristband sensor to measure the stress level for people with dementia. *Sensors* **16**, 1989 (2016).
68. Veldkamp, K. et al. Real-life heart rate assessment: a digital biomarker for autonomic dysfunction in early Parkinson's disease? *Mov. Disord.* **38**, S796 (2023).
69. Alharbi, E. A., Jones, J. M. & Alomainy, A. Non-invasive solutions to identify distinctions between healthy and mild cognitive impairments participants. *IEEE J. Transl. Eng. Health Med.* **10**, 2700206 (2022).
70. Lieber, B. et al. Motion sensors to assess and monitor medical and surgical management of Parkinson disease. *World Neurosurg.* **84**, 561–566 (2015).
71. Park, Y. J. et al. Self-organizing IoT device-based smart diagnosing assistance system for activities of daily living. *Sensors* **21**, 785 (2021).
72. Yin, M. et al. Unobtrusive sleep health assessment using impulse radar: a pilot study in older people. *IEEE Trans. Biomed. Eng.* **72**, 2663–2674 (2025).
73. Massaroni, C. et al. Contact-based methods for measuring respiratory rate. *Sensors* **19**, 908 (2019).
74. Khosroozad, S. et al. Sleep movements and respiratory coupling as a biobehavioral metric for early Alzheimer's disease in independently dwelling adults. *BMC Geriatr.* **23**, 252 (2023).
75. Zimmerman, M. E. et al. Neuropsychological function response to nocturnal blue light blockage in individuals with symptoms of insomnia: a pilot randomized controlled study. *J. Int. Neuropsychol. Soc.* **25**, 668–677 (2019).
76. Madrid-Navarro, C. J. et al. Validation of a device for the ambulatory monitoring of sleep patterns: a pilot study on Parkinson's disease. *Front. Neurol.* **10**, 356 (2019).
77. Prins, A. J. et al. Sensory stimulation for nursing-home residents: systematic review and meta-analysis of its effects on sleep quality and rest-activity rhythm in dementia. *Dement. Geriatr. Cogn. Disord.* **49**, 219–234 (2020).
78. Tung, J. Y. et al. Ambulatory assessment of lifestyle factors for Alzheimer's disease and related dementias. In *2011 AAAI Spring Symposium Series. No. 4: Computational Physiology* 50–54 (AAAI, 2011).
79. Botros, A. A. et al. Eigenbehaviour as an indicator of cognitive abilities. *Sensors* **22**, 2769 (2022).
80. Bayat, S. et al. GPS driving: a digital biomarker for preclinical Alzheimer disease. *Alzheimers Res. Ther.* **13**, 115 (2021).
81. Mahadevan, N. et al. Development of digital biomarkers for resting tremor and bradykinesia using a wrist-worn wearable device. *npj Digital Med.* **3**, 5 (2020).
82. Farzanehfard, P., Woodrow, H. & Horne, M. Assessment of wearing off in Parkinson's disease using objective measurement. *J. Neurol.* **268**, 914–922 (2021).
83. Shema-Shiratzky, S. et al. A wearable sensor identifies alterations in community ambulation in multiple sclerosis: contributors to real-world gait quality and physical activity. *J. Neurol.* **267**, 1912–1921 (2020).
84. Celik, Y. et al. Gait analysis in neurological populations: progression in the use of wearables. *Med. Eng. Phys.* **87**, 9–29 (2021).
85. Snow, A. L. et al. Enhancing sleep quality for nursing home residents with dementia: a pragmatic randomized controlled trial of an evidence-based frontline huddling program. *BMC Geriatr.* **21**, 281 (2021).
86. Vitazkova, D. et al. Transforming sleep monitoring: review of wearable and remote devices advancing home polysomnography and their role in predicting neurological disorders. *Biosensors* **15**, 117 (2025).
87. Cao, R. et al. Accuracy assessment of oura ring nocturnal heart rate and heart rate variability in comparison with electrocardiography in time and frequency domains: comprehensive analysis. *J. Med. Internet Res.* **24**, e27487 (2022).
88. Almarshad, M. A. et al. Diagnostic features and potential applications of PPG signal in healthcare: a systematic review. *Healthcare* **10**, 547 (2022).
89. Lydia, D. B. et al. Selecting a smartwatch for trials involving older adults with neurodegenerative diseases: a researcher's framework to avoid hidden pitfalls. *J. Biomed. Inform.* **162**, 104781 (2025).
90. Schalkamp, A. K. et al. Digital outcome measures from smartwatch data relate to non-motor features of Parkinson's disease. *npj Parkinsons Dis.* **10**, 110 (2024).
91. Veldkamp, K. I. Towards Digital Biomarkers to Assess Autonomic Dysfunction in Parkinson's Disease: Feasibility of Home-based Photoplethysmographic Signals for Heart Rate Analysis. MA thesis, Univ. Twente (2023).
92. Gwak, M., Woo, E. & Sarrafzadeh, M. The role of PPG in identification of mild cognitive impairment. In *Proc. 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments* 32–35 (ACM, 2019).
93. Li, Y. et al. Rethinking pain communication of patients with Alzheimer's disease through E-textile interaction design. *Front. Physiol.* **14**, 1248893 (2023).
94. Anke, J. & Lesley, C. Sensory design for dementia care — the benefits of textiles. *J. Textile Design Res. Pract.* **5**, 232–250 (2017).
95. Chai, J. et al. Classification of mild cognitive impairment based on handwriting dynamics and qEEG. *Comput. Biol. Med.* **152**, 106418 (2023).
96. Rutkowski, T. M., Abe, M. S., Tokunaga, S., Komendziński, T. & Otake-Matsuura, M. Dementia digital neuro-biomarker study from theta-band eeg fluctuation analysis in facial and emotional identification short-term memory oddball paradigm. In *44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* 4056–4059 (IEEE, 2022).
97. Channa, A. et al. A-WEAR bracelet for detection of hand tremor and bradykinesia in Parkinson's patients. *Sensors* **21**, 981 (2021).
98. Dang, B. et al. Development of a smart sleep mask with multiple sensors. In *43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* 7058–7062 (IEEE, 2021).
99. Boroojerdi, B. et al. Clinical feasibility of a wearable, conformable sensor patch to monitor motor symptoms in Parkinson's disease. *Parkinsonism Relat. Disord.* **61**, 70–76 (2019).
100. Nyein, H. Y. Y. et al. A wearable patch for continuous analysis of thermoregulatory sweat at rest. *Nat. Commun.* **12**, 1823 (2021).
101. Shustak, S. et al. Home monitoring of sleep with a temporary-tattoo EEG, EOG and EMG electrode array: a feasibility study. *J. Neural Eng.* **16**, 026024 (2019).
102. Alfallahi, H. et al. Diagnostic accuracy of keystroke dynamics as digital biomarkers for fine motor decline in neuropsychiatric disorders: a systematic review and meta-analysis. *Sci. Rep.* **12**, 7690 (2022).
103. Dentamaro, V., Impedovo, D. & Pirlo, G. An analysis of tasks and features for neuro-degenerative disease assessment by handwriting. In *International Conference on Pattern Recognition* 536–545 (Springer, 2021).
104. Jefferson, S. A. et al. Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. *Pattern Recognit. Lett.* **125**, 55–62 (2019).
105. Nef, T. et al. Development and evaluation of maze-like puzzle games to assess cognitive and motor function in aging and neurodegenerative diseases. *Front. Aging Neurosci.* **12**, 87 (2020).
106. Zhong, G. et al. The relationship between thermoregulation and REM sleep behaviour disorder in Parkinson's disease. *PLoS ONE* **8**, e72661 (2013).
107. Peters-Strickland, T. et al. Usability of a novel digital medicine system in adults with schizophrenia treated with sensor-embedded tablets of aripiprazole. *Neuropsychiatr. Dis. Treat.* **12**, 2587–2594 (2016).
108. Sharma, S. et al. Location-aware ingestible microdevices for wireless monitoring of gastrointestinal dynamics. *Nat. Electron.* **6**, 242–256 (2023).
109. Yeo, B. S. Y. et al. Association of hearing aids and cochlear implants with cognitive decline and dementia: a systematic review and meta-analysis. *JAMA Neurol.* **80**, 134–141 (2023).
110. Afshar, P. et al. A translational platform for prototyping closed-loop neuromodulation systems. *Front. Neural Circuits* **6**, 117 (2013).
111. Willett, F. R. et al. A high-performance speech neuroprosthesis. *Nature* **620**, 1031–1036 (2023).
112. Li, J. et al. Toward long-term and high-precision multimodal intracranial biomarker monitoring. *Med-X* **2**, 13 (2024).
113. Galna, B. et al. Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. *Gait Posture* **39**, 1062–1068 (2014).
114. Li, H. et al. Brain activation during motor preparation and execution in patients with mild cognitive impairment: an fNIRS study. *Brain Sci.* **15**, 333 (2025).
115. Wang, Z. et al. Assessment of brain function in patients with cognitive impairment based on fNIRS and gait analysis. *Front. Aging Neurosci.* **14**, 799732 (2022).
116. Yoon, S. et al. Development and validation of 2D-LiDAR-based gait analysis instrument and algorithm. *Sensors* **21**, 414 (2021).
117. Weuve, J. et al. Long-term community noise exposure in relation to dementia, cognition, and cognitive decline in older adults. *Alzheimer Dement.* **17**, 525–533 (2021).
118. Jao, Y. L. et al. Effect of ambient bright light on behavioral and psychological symptoms in people with dementia: a systematic review. *Innov. Aging* **6**, igac018 (2022).

119. Himanshu, T., Rajdeep, K. N. & Saraju, P. M. Smart home environment for mild cognitive impairment population: solutions to improve care and quality of life. *IEEE Consum. Electron. Mag.* **7**, 68–76 (2017).
120. Chuan, M. et al. Supporting dementia care by monitoring indoor environmental quality in a nursing home. *Indoor Built Environ.* **32**, 1843–1861 (2023).
121. Yuheng, W. et al. m-Activity: accurate and real-time human activity recognition via millimeter wave radar. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 8298–8302 (IEEE, 2021).
122. Sundgren, M. et al. Does information from the Parkinson KinetiGraph™ (PKG) influence the neurologist's treatment decisions? — An observational study in routine clinical care of people with Parkinson's disease. *J. Personal. Med.* **11**, 519 (2021).
123. Joshi, R. et al. PKG movement recording system use shows promise in routine clinical care of patients with Parkinson's disease. *Front. Neurol.* **10**, 1027 (2019).
124. Petrella, J. R. et al. Impact of computerized cognitive training on default mode network connectivity in subjects at risk for Alzheimer's disease: a 78-week randomized controlled trial. *J. Alzheimers Dis.* **91**, 483–494 (2023).
125. Reicher, L. et al. Deep phenotyping of health–disease continuum in the Human Phenotype Project. *Nat. Med.* **31**, 3191–3203 (2025).
126. Schalkamp, A. K. et al. Deep phenotyping for precision medicine in Parkinson's disease. *Dis. Models Mech.* **15**, dmm049376 (2022).
127. Adams, J. L. et al. Using a smartwatch and smartphone to assess early Parkinson's disease in the WATCH-PD study over 12 months. *npj Parkinsons Dis.* **10**, 112 (2024).
128. Phillips, K. A. et al. "What goes around comes around": lessons learned from economic evaluations of personalized medicine applied to digital medicine. *Value Health* **20**, 47–53 (2017).
129. Piau, A. et al. When will my patient fall? Sensor-based in-home walking speed identifies future falls in older adults. *J. Gerontol. Ser. A* **75**, 968–973 (2020).
130. Page, A. et al. A smartphone application as an exploratory endpoint in a phase 3 Parkinson's disease clinical trial: a pilot study. *Digit. Biomark.* **6**, 1–8 (2022).
131. Gold, M. et al. Digital technologies as biomarkers, clinical outcomes assessment, and recruitment tools in Alzheimer's disease clinical trials. *Alzheimers Dement.* **4**, 234–242 (2018).
132. Dodge, H. H. et al. Use of high-frequency in-home monitoring data may reduce sample sizes needed in clinical trials. *PLoS ONE* **10**, e0138095 (2015).
133. Nerrise, F. et al. An explainable geometric-weighted graph attention network for identifying functional networks associated with gait impairment. *Med. Image Comput. Comput. Assist. Interv.* **14221**, 723–733 (2023).
134. Mc Ardle, R. et al. Differentiating dementia disease subtypes with gait analysis: feasibility of wearable sensors? *Gait Posture* **76**, 372–376 (2020).
135. Ellis, R. J. et al. A validated smartphone-based assessment of gait and gait variability in Parkinson's disease. *PLoS ONE* **10**, e0141694 (2015).
136. Brasier, N. et al. Applied body-fluid analysis by wearable devices. *Nature* **636**, 57–68 (2024).
137. Li, Y. et al. Research progress of electronic nose technology in exhaled breath disease analysis. *Microsyst. Nanoeng.* **9**, 129 (2023).
138. Neely, R. M. et al. Recent advances in neural dust: towards a neural interface platform. *Curr. Opin. Neurobiol.* **50**, 64–71 (2018).
139. Chen, Y. et al. How is flexible electronics advancing neuroscience research? *Biomaterials* **268**, 120559 (2021).
140. Everitt, A. et al. A bioimpedance-based monitor for real-time detection and identification of secondary brain injury. *Sci Rep.* **11**, 15454 (2021).
141. Preview : Journal Type Schütz, N. et al. Deep canonical correlation alignment for sensor signals. Preprint at <https://doi.org/10.48550/arXiv.2106.03637> (2021).
142. Vairavan, S. et al. A multimodal digital biomarker of functional deficits in early-stage Alzheimer's disease: results of the RADAR-AD study. *Alzheimers Dement.* **19**, e071136 (2023).
143. Qi, W. et al. Alzheimer's disease digital biomarkers multidimensional landscape and AI model scoping review. *npj Digit. Med.* **8**, 366 (2025).
144. Mueller, S. G. et al. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement.* **1**, 55–66 (2005).
145. Marek, K. et al. Parkinson Progression Marker Initiative. The Parkinson Progression Marker Initiative (PPMI). *Prog. Neurobiol.* **95**, 629–635 (2011).
146. Lindauer, A. et al. The African American Dementia and Aging Project: an Oregon-based longitudinal study. *Front. Dement.* **3**, 1498835 (2024).
147. Kirk, C. et al. Mobilise-D insights to estimate real-world walking speed in multiple conditions with a wearable device. *Sci Rep.* **14**, 1754 (2024).
148. Ranjan, Y. et al. RADAR-base: open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. *JMIR Mhealth Uhealth* **7**, e11734 (2019).
149. Butler, P. M. et al. Smartwatch- and smartphone-based remote assessment of brain health and detection of mild cognitive impairment. *Nat. Med.* **31**, 829–839 (2025).
150. Langer, P. et al. CLAIID: Closing the Loop on AI & Data Collection — a cross-platform transparent computing middleware framework for smart edge-cloud and digital biomarker applications. *Future Gener. Comput. Syst.* **159**, 505–521 (2024).
151. Aalami, O. et al. CardinalKit: open-source standards-based, interoperable mobile development platform to help translate the promise of digital health. *JAMIA Open* **6**, ooad044 (2023).
152. Dawson, W. D. et al. Ascertaining out-of-pocket costs of dementia care: feasibility study of a web-based weekly survey. *JMIR Form. Res.* **8**, e56878 (2024).
153. Lavine, J. S. et al. Impacts on study design when implementing digital measures in Parkinson's disease-modifying therapy trials. *Front. Digit. Health* **6**, 1430994 (2024).
154. Joerg, G., Noé, B. & Lindsay, K. Digitalizing health trials by the Clinical Trials Transformation Initiative. *Nat. Rev. Bioeng.* **2**, 717–718 (2024).
155. de Oliveira, A., Khamis, M. & Esteves, A. GaitWear: a smartwatch application for in-the-wild gait normalisation based on a virtual field study assessing the effects of visual and haptic cueing. *Behav. Inf. Technol.* **40**, 1292–1309 (2021).
156. Servais, L. et al. Evidentiary basis of the first regulatory qualification of a digital primary efficacy endpoint. *Sci Rep.* **14**, 29681 (2024).
157. Miller, I. N. & Cronin-Golomb, A. Gender differences in Parkinson's disease: clinical characteristics and cognition. *Move. Disord.* **25**, 2695–2703 (2010).
158. Liu, X. et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Lancet Digit. Health* **2**, e537–e548 (2020).
159. Rivera, S. C. et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Lancet Digit. Health* **2**, e549–e560 (2020).
160. Norgeot, B. et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. *Nat. Med.* **26**, 1320–1324 (2020).
161. Collins, G. S. et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* **385**, e078378 (2024).

Acknowledgements

This research was supported in part by the National Institutes of Health grant numbers AG089169 (E.A.), AG084471 (E.A., F.V.L.), AG073107 (E.A.), AG072701 (E.A., F.V.L.), AG047366 (V.W.H., E.A.) and AG066515 (V.W.H.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This study was also supported by the Stanford Graduate and NeuroTech Fellowships (F.N.).

Author contributions

F.N., N.S. and E.A. contributed substantially to the discussion, writing, and editing of this manuscript and managed the submission process. F.N. developed figures and tables for the manuscript, revised by N.S., F.V.L. and E.A. Q.Z., C.G., A.M., K.S., F.V.L. and E.A. all contributed important ideas for the manuscript text, tables and figures. All authors reviewed the manuscript and provided critical feedback that formed the ideas for the proposed framework.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44222-026-00433-7>.

Peer review information *Nature Reviews Bioengineering* thanks Bakul Patel, Graham Jones, Eric Thelin, Paul Dagum and the other, anonymous, reviewer for their contribution to the peer review of this work.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Related links

Abilify MyCite: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/207202Orig1s000Approv.pdf
Altoida: <https://altoida.com/>
Apple modified its HealthKit: <https://developer.apple.com/videos/play/wwdc2019/217/>
Axivity AX3: <https://axivity.com/>
Digital Health Innovation Action Plan: <https://www.fda.gov/media/106331/download>
Digital Medicine Society's V3 framework: <https://datacc.dimesociety.org/resources/the-v3-framework/>
Emerald device: <https://emeraldinno.com/>
Empatica E4: <https://www.empatica.com/research/e4/>
Empatica's Health Monitoring Platform: <https://www.empatica.com/blog/empatica-receives-new-fda-clearance-for-cardiac-digital-biomarkers>
European Medicines Agency Qualification of Novel Methodologies for Medicine Development: <https://www.ema.europa.eu/en/human-regulatory-overview/research-development/scientific-advice-protocol-assistance/qualification-novel-methodologies-medicine-development>
Foundation models: <https://crfm.stanford.edu/assets/report.pdf>
Fox Insight: <https://foxinsight.michaeljfox.org/>
GAIRite: <https://www.gairite.com/>
GDPR: <https://gdpr-info.eu/>
Helix: <https://www.helix.com/>
HIPAA: <http://www.eolusinc.com/pdf/hipaa.pdf>

Review article

Indivi: <https://indivi.io/>

MindX Sciences: <https://mindxsciences.com/>

Mobilise-D: <https://mobilise-d.eu/>

National Institute for Health Care and Excellence (NICE) Evidence Standards

Framework for Digital Health Technologies: <https://www.nice.org.uk/corporate/ecd7/resources/evidence-standards-framework-for-digital-health-technologies-pdf-1124017457605>

N-Power Medicine: <https://www.npowermedicine.com/>

Pre-Cert: <https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-software-precertification-pre-cert-pilot-program>

RADAR-CNS: <https://www.ih.europa.eu/projects-results/project-factsheets/radar-cns>

Rune Labs: <https://runelabs.io/>

StrivePD: <https://www.strive.group/>

© Springer Nature Limited 2026

Supplementary information

A framework of digital biomarkers for neurodegenerative diseases

**In the format provided by
the authors and unedited**

Supplementary Information for

A framework of digital biomarkers for neurodegenerative diseases

Favour Nerrise^{1,2}, Narayan Schütz^{1,3}, Qingyu Zhao⁴, Christine Gould¹, Arnold Milstein³, Kevin Schulman³, Victor W. Henderson^{5,6,10}, James Landay^{7,9}, Li Fei-Fei^{7,9}, Feng Vankee Lin¹, and Ehsan Adeli^{1,3,7,8,9,10*}

¹Department of Psychiatry and Behavioral Sciences, Stanford University

²Department of Electrical Engineering, Stanford University

³Clinical Excellence Research Center (CERC), Stanford University

⁴Weill Cornell Medicine, Cornell University

⁵Department of Epidemiology and Population Health, Stanford University

⁶Department of Neurology and Neurological Sciences, Stanford University

⁷Department of Computer Science, Stanford University

⁸Department of Biomedical Data Science, Stanford University

⁹Stanford Institute for Human-Centered Artificial Intelligence (HAI)

¹⁰Stanford Alzheimer's Disease Research Center (ADRC)

* email: eadeli@stanford.edu

SI 1. Systematic literature review and analysis

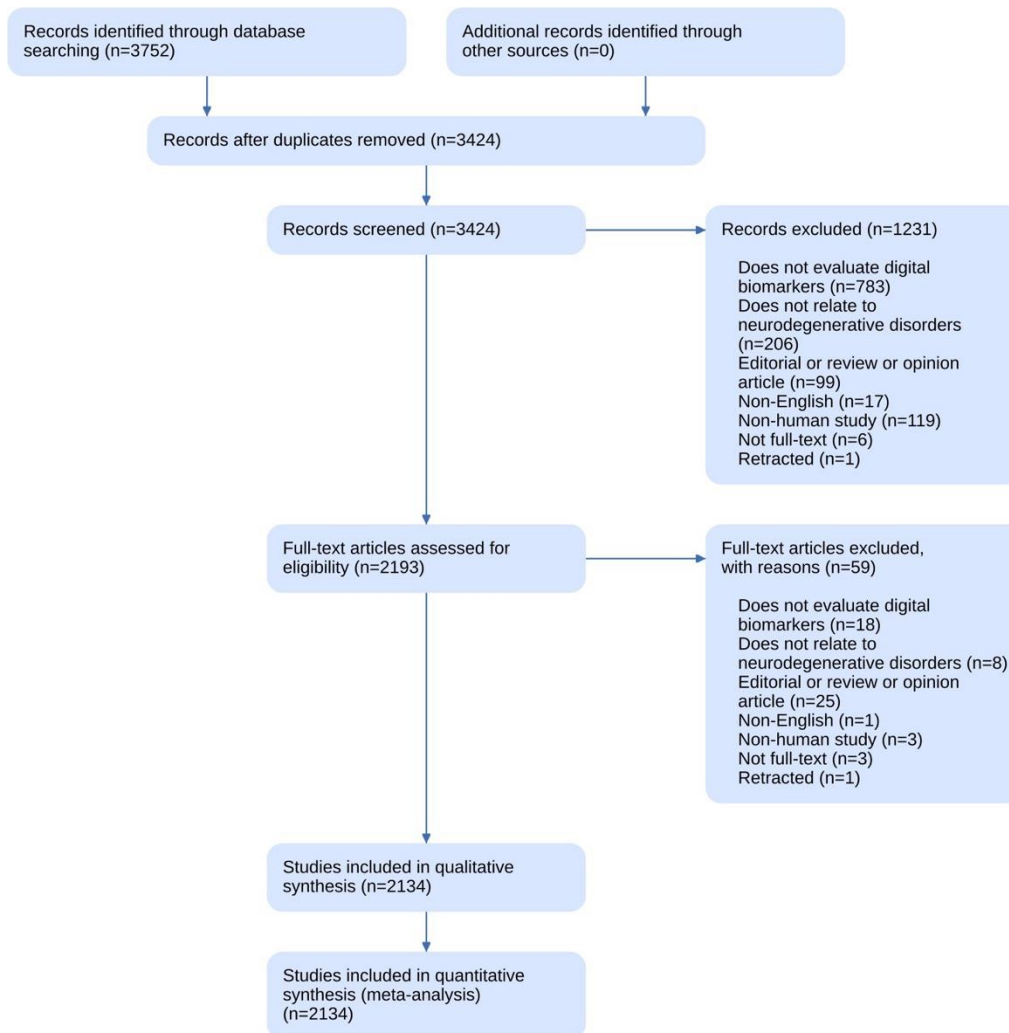


Figure S1. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow chart.

We conducted a comprehensive literature review in two steps using the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews)¹ guidelines to inform our narrative review framework (Figure S1). We retrieved 3,752 articles from searching multiple databases, removed 328 duplicates, and, after screening titles and abstracts, identified 2134 articles that informed our narrative synthesis and framework development. Our application of PRISMA-ScR guidelines served to ensure comprehensive coverage of the literature rather than to conduct systematic evidence synthesis or meta-analysis. The structured search process provided the

empirical foundation for identifying key themes, technological approaches, and clinical applications that we discuss throughout our narrative framework. We conducted a manual search of PubMed, Web of Science, Scopus, and Embase with the following keywords and MeSH terms:

```
("digital biomarker" OR "digital biomarkers") OR  
("implantable" OR "implantables" OR "wearable" OR "wearables" OR  
"portable"  
OR "portables" OR "inner-body" OR "ingestible" OR "ingestibles" OR  
"ambient"  
OR "contactless" OR "remote" AND "sensor") AND  
("neurodegenerative disease" OR "neurodegenerative diseases" OR  
"neurodegenerative disorder" OR "neurodegenerative disorders") AND  
("artificial intelligence" OR "AI" OR "machine learning" OR "ML" OR  
"deep learning" OR "statistical")
```

We performed a second search in the electronic databases with the search and MeSH terms:

```
("digital biomarker" OR "digital biomarkers") AND  
"Parkinson's disease" OR "Alzheimer's disease" OR "Mild Cognitive  
Impairment" OR "Dementia" OR "Tremor" OR "Huntington's disease" OR  
"Ataxia" OR "Amyotrophic Lateral Sclerosis" OR "Multiple Sclerosis"  
OR "Chronic Traumatic Encephalopathy" OR "neurodegenerative disease"  
OR "neurodegeneration" OR "PD" OR "AD" OR "MCI" OR "DLB" OR "LBD"  
OR "HD" OR "ALS" OR "MS" OR "CTE")
```

We screened the abstracts of all articles using the following inclusion criteria: evaluation of digital biomarker (DBM) study, human participants, written in English, and published between 2010 to 2025. The search excluded studies focused on traditional biomarkers (e.g., blood or cerebrospinal fluid markers) without integrating digital biomarkers (DBMs) or digital health technologies (DHTs), studies that evaluated digital biomarkers in non-neurodegenerative diseases or conditions, and studies that were review articles, case reports, conference abstracts, editorials, or commentaries. The extracted information included study design, study population, sample size, type of neurodegenerative disease (NDD), DBM technology employed, outcomes assessed, and key findings.

Table S1: Comparative Analysis of Digital vs. Traditional Biomarkers for Neurodegenerative Disease Assessment

Biomarker Domain	Traditional Gold Standard	Digital “Silver Standard”	Key Performance Metrics	Sample Population	Primary Limitations
Brain Activity	Medical-grade EEG (19-21 electrodes)	Sleep Profiler™ wireless EEG (8 channels) ²	Test-retest: >75% agreement for AD, LBD, and HCs, v. clinical diagnosis for NDD profiling	n=446 (HC=232, AD=46, LBD=28, iRBD=33, PD=29, MCI=78)	Limited spatial resolution; Lower signal quality; Subject to motion artifacts; Source variation due to different positioning, heads, and hairstyles; User discomfort with prolonged use
		LEGO-like, dry EEG headset (8 channels) ³	Sens: 61.1%, Spec: 68.6.8%, Acc: 66.3%	n=59 (HC=29, PD typical=23, PD-ICD=10)	
		Cumulus Neuroscience portable, dry EEG headset (16 channels) ⁴	Standardized measurement errors (SMEs): 2-4 sessions yield same SMEs as in-lab, wet-EEG for all tasks	n=30 (50 older adults, 30 younger male adults)	
Sleep Assessment	Medical-grade Polysomnography (PPG)	NIGHT-Recorder accelerometer ⁵	Pearson’s correlation: rolling over duration & UPDRS axial score (r=0.619, p=0.005), nocturia & UPDRS-III (r=0.579, p=0.009)	n=38 (19 PD, 19 spouses)	Accuracy affected by temperature (e.g., altered thermoregulation); Subject to motion artifacts; Environmental lighting effects; Battery-life constraints; Device positioning affects reliability; Limited clinical validation
		Oura Ring Generation 3 ⁶	Sleep/wake: Sens=94.5%, Acc=91.8%, bias-adjusted κ =0.84, PVS=96%	n=96 HC, 421k epochs	
		Wrist-worn actigraphs ⁷⁻⁹ Consumer smartwatches ¹⁰⁻¹²	Sleep Efficiency (SE): 72.1% w.r.t. low morning mobility Sleep stage: Fitbit Sense 2 Sens=61.7–78%; Apple watch Sens=50.5–86.1%; WHOOP 3.0 agreement=60% (κ = 0.44)	n=157 PD n=35/n=53 healthy adults	
Gait Analysis	3D Optical Motion Capture System (Vicon)	Microsoft Kinect v2 ¹³	Cadence: r=0.81, Speed: r=0.80, Stride Length: r=0.81, p<0.0001. vs. Vicon, Gait Score Classification: Acc=90.4%, Sens=86-97%+, Spec=80-94%	n=51 children (HC=20, Progressive Ataxia=15, Chronic Ataxia=16)	Reduced accuracy during turns; Device placement reduces reliability; Limited validation in severe impairment; Poor joint kinematics Clothing; Lighting; Occlusion
	IMU Motion Capture System (GAITRite walkway)	Axivity AX3 ¹⁴ , Mobilise-D ¹⁵	Step Time SD: (F/ χ^2 =11.0, p=0.004) between AD, LDB, and PDD vs. (F/ χ^2 =6.2, p=0.046 GAITRite)	n=74 (AD=32, DLB=28, PD=14)	
		Smartphone accelerometry (SmartMOVE ¹⁶)	Effect-size: Small (η^2 < .02) vs. GAITRite on Step Time/Length	n=24 (HC=12, PD=12)	

	Video-based analysis (MDS-UPDRS Clinician Rating)	Video-based analysis (VIBE + POSE + AI algorithm) ¹⁷⁻¹⁹	Inter-rater agreement: $\kappa=0.49$ vs. 0.35 clinician rating, Gait Score classification: AUC=0.83, F1=0.66, Pre.=0.63, BAcc=72%	n=35 PD	
Speech & Cognition	Auditory-perceptual assessment (Clinician Rating)	Automatic Speech Recognition (ASR) System (Speech and Pause Analysis software) ²⁰	SCC: Total Duration $p=0.93-0.98$, Speech Duration $p=0.70-0.96$, Pause Duration $p=0.62-0.82$, $p<0.001$ vs. ASR	n=72 (HCs=219, ALS=243)	Recording artifacts; Small sample sizes; Inconsistent distances between the mouth and recorder; Varied measurement metrics across datasets; Limited linguistic diversity; Susceptibility to biases
	Neuropsychological Battery	Digital Cognitive Screening (DCR) ²¹	Sens=90.0%, Spec=92.3% for Dementia/MCI, $p>0.05$ vs. MoCA	n=328 (NCI=209, MCI=99; Dementia=20)	
		Remote Smartphone Cognitive Test ²²	ICC: 0.77-0.95; Early symptoms AUC=0.82 vs. MoCA AUC=0.68	n=306 with FLTD	
Circadian Rhythm	1-12-lead ECG/EKG (Clinician Rating)	Kronowise Ambulatory Circadian Monitoring (ACM) device ²³	TST: $r=0.76$, bias=-1.05%, SE: $r=0.7$, bias=-0.48% vs. PPG	n=70 PD	Skin reliability; Low specificity

Abbreviations: Sens=Sensitivity; Spec=Specificity; Acc=Accuracy; ICC=Intraclass Correlation Coefficient; PPV=Positive Predictive Value; AFib=Atrial Fibrillation; TRL=Technology Readiness Level; ICD: Impulse Control Disorder; FLTD: Frontotemporal Lobar Degeneration.

Performance Metrics: Values represent validation studies under controlled conditions and at-home. Real-world performance often varies significantly based on environmental factors and user compliance.

All digital methods are non-invasive; traditional assessments range from non-invasive (clinical scales) to minimally invasive.

SI 2. Supplementary discussion

SI 2.1 Additional considerations on the regulatory validation of digital biomarkers

Translating digital biomarkers from research to clinical practice requires navigating evolving regulatory frameworks, which is very challenging. The U.S. Food & Drug Administration (FDA) and European Medicines Agency (EMA) are developing specific pathways for digital health technologies (DHTs), with considerable progress including the FDA's Digital Health Innovation Action Plan and Pre-Cert program for "Software as a Medical Device." These initiatives aim to balance innovation with patient safety while acknowledging the rapid development cycle of digital technologies compared to traditional medical devices²⁴. Regulatory considerations include not only approval processes but also data privacy, security protocols, and reference standards for validating digital biomarkers. The Clinical Trials Transformation Initiative has proposed recommendations for developing novel digital endpoints that emphasize transparency in algorithm development and validation methods²⁵.

The Digital Medicine Society has established multiple playbooks and frameworks, such as the V3 framework (Verification, Analytical Validation, and Clinical Validation), to help address standardization needs and guide development around DBM systems²⁶. Additionally, many DBM systems face practical implementation barriers such as user adherence, particularly for long-term monitoring applications in populations with cognitive impairments (Table S1). Validating digital biomarkers against established clinical measures presents unique challenges, particularly when capturing novel dimensions of a disease. To address this, researchers are adapting concepts from psychometrics, originally developed for psychological measurement. Approaches such as construct validity, convergent and discriminant validity, and reliability measures are being applied to candidate DBMs (e.g., EEG-based motor network biomarkers^{27,28} and eye-tracking biomarkers²⁹⁻³¹) to ensure they measure intended aspects and provide consistent results. However, challenges persist due to the lack of gold standards, the multi-dimensional nature of DBMs, and variability in real-world data. Innovative methods that combine psychometric principles with modern techniques are emerging, including ecological momentary assessment, multi-trait multi-method matrices, ML-driven analyses, and the incorporation of patient-reported outcomes. These approaches aim to develop validation methods for DBMs that strengthen their credibility and clinical use.

When considering technical validation (verification + analytical), it is important to be aware that measurable symptoms can differ significantly between healthy people and individuals with NDDs; thus, larger technical validation studies that go beyond small pilots are necessary, not just in healthy but also in diseased patient populations. A first step in this direction for mobility outcomes has been taken by the Mobilise-D project¹⁵. Here, it is also crucial to realize the ecological validity paradox that comes with technical validation. There is a significant domain gap between clinical/laboratory settings (where the gold-standard is usually established) and free-living conditions (with uncontrolled variables that may affect measurement precision), so validating in

clinical conditions may not always carry over to the real-world. While there is no perfect solution, the most reasonable approach seems to be to do a stepwise validation, such as validating against a clinical gold standard first, then using (or creating) a “silver standard” that can be deployed in the real world, such as a mobile polysomnography setup for sleep staging³².

Establishing both analytical and clinical validity tends to require years of independent replication across multiple academic centers, diverse patient populations, and varied clinical settings. Digital biomarkers do not make this easier; they even face additional complexity because they likely have to demonstrate this validity across different devices, software versions, and real-world conditions, introducing more variability than traditional laboratory-based biomarkers, while operating in a regulatory environment that lacks an established wealth of precedents for continuously evolving algorithms and sensor technologies. This also raises the question of whether it is feasible to achieve regulatory approval in consumer-facing bring-your-own-device (BYOD) settings (see Box 3 in the main text).

SI 2.2 Practical challenges in real-world deployment

Deploying DHTs to extract DBMs in NDD populations presents numerous practical challenges that require careful consideration. The initial step involves identifying which characteristics or outcomes to assess (see Section 3.1 in the main text). However, selecting appropriate sensing technology proves more complex, as it requires balancing multiple trade-offs. Cost considerations often drive initial decisions. Research-validated devices such as ActiGraph accelerometers, while more expensive than commercial fitness trackers, may offer superior reliability. Following the FDA’s fit-for-purpose approach, the optimal choice depends on the intended use of the DBM and its required validation level.

Equally critical is the complexity of the system architecture; the number of components (devices, sensor firmware, etc.) directly impacts validation difficulty. While BYOD approaches using smartphones or smartwatches present validation challenges due to their heterogeneity, they offer cost-efficiency and facilitate larger-scale recruitment. Generally, controlled DHTs suit smaller studies, particularly in rare diseases, where device provision costs are negligible relative to study size, similar to pivotal clinical trials. Technology obtrusiveness represents another crucial consideration in NDDs. A clear inverse relationship exists between obtrusiveness and participant adherence, despite more obtrusive technologies typically delivering superior performance. Clinical gold standards exemplify this trade-off: a 12-lead ECG provides optimal heart rate measurement but requires clinic visits, whereas single-lead ECG patches offer reasonable accuracy with reduced burden, and wrist-based PPG sensors maximize convenience at some performance cost. Sleep monitoring follows a similar pattern. Polysomnography remains the gold standard but necessitates laboratory stays, while wearable sleep trackers adequately capture overall patterns despite limitations in detecting complex sleep architecture. Contactless technologies, such as smart

mattress sensors, minimize friction and maximize long-term adherence but typically sacrifice measurement precision.

The selection of an appropriate DHT depends fundamentally on study requirements (Table S1). Long-term studies spanning months to years may benefit from ambient sensing solutions, while shorter pivotal trials might justify more obtrusive approaches if adherence remains acceptable. Importantly, friction manifests differently across technologies: wearables face challenges related to usability, physical comfort, and aesthetics that directly impact adherence, while ambient sensing devices encounter friction primarily through privacy concerns, affecting recruitment rather than adherence among enrolled participants. This technology-specific friction can introduce selection bias toward particular subpopulations. However, the traditional friction-performance trade-off is evolving. Emerging low-friction ambient sensing technologies, exemplified by the [Emerald device](#), or privacy-preserving camera solutions in combination with advanced AI and edge computing, are narrowing the performance gap with wearables. For instance, the Emerald device demonstrates accurate detection of sleep stages, gait characteristics, and respiration patterns. Concurrently, advances in monocular computer vision enable precise pose estimation, facilitating extraction and analysis of sophisticated kinematic features from single-camera setups.

SI 2.3 Data Privacy, algorithmic fairness, and transparency

Implementing DBMs in healthcare brings a complex web of challenges around data privacy, algorithmic fairness, and transparency that demand immediate attention. Addressing these concerns is essential for building trust and ensuring responsible inclusion of DBMs in clinical practice.

Digital biomarkers inherently capture highly sensitive health information that necessitate robust privacy protections such as regulatory compliance with the [United States Health Insurance Portability and Accountability Act \(HIPAA\)](#) and the [European Union General Data Protection Regulation \(GDPR\)](#). However, these regulations are not strictly enforced for independent developers until systems undergo formal regulatory review, thereby creating a critical gap where early-stage DBM development and user testing occur without rigorous privacy oversight; as such, participants are likely to be exposed to substantial security and privacy risks during the most experimental phases of technology development without their knowledge. While technical solutions like differential privacy, blockchain-based decentralized storage, and federated learning systems show promise, they must be coupled with meaningful informed consent processes that acknowledge communication gaps and trust barriers between participants and DBM developers³³⁻³⁵. DBM developers should adopt privacy-by-design principles³³, such as implementing data minimization processes and conducting regular privacy impact assessments. Additionally, deploying homomorphic encryption for cloud-based analytics and secure multi-party computation protocols could also allow for collaborative research while preserving participant privacy. Most critically, consent mechanisms must evolve beyond static forms to dynamic, granular consent management

systems that allow participants to understand and control how their continuous health information is used, shared, and retained over time³⁵.

The potential for DBMs to perpetuate and amplify healthcare inequities is perhaps the most pressing ethical challenge^{36,37}. Machine learning models trained on data that underrepresent marginalized communities³⁸, exclude those without digital access³⁹, or fail to account for social determinants of health^{36,40} risk creating systems that work well for privileged populations while failing others. For instance, gender-specific symptom presentations in Parkinson's disease—with women more commonly experiencing dyskinesias and depression while men manifest more rigidity⁴¹—show how biased training data can lead to delayed diagnosis and treatment for underrepresented groups. To address these challenges, DBM developers should implement concrete measures: conduct mandatory algorithmic impact assessments using fairness metrics across demographic subgroups⁴²; establish community advisory boards with diverse stakeholder representation; and invest in targeted recruitment strategies to ensure representative datasets. Furthermore, funding agencies and journals should incentivize reporting of disaggregated performance reporting across race, ethnicity, gender, socioeconomic status, and digital literacy levels in considerations for grant awards and publications^{43,44}.

Transparency in DBM development and deployment is essential for clinical adoption and regulatory approval, yet current practices often obscure critical information about participant demographics, model limitations, and performance variations across populations. Many studies fail to report basic demographic data or analyze differential performance across groups, which makes it difficult to assess generalizability⁴⁵. To address this, DBM developers conducting validation studies should adhere to established reporting guidelines such as the [National Institute for Clinical Excellence \(NICE\) Evidence Standards Framework \(ESF\)](#) for digital health technologies, and for AI-based systems, transparent reporting frameworks like the CONSORT-AI framework⁴⁶, SPIRIT-AI framework⁴⁷, MI-CLAIM checklist for clinical AI studies⁴⁸, and the TRIPOD+ AI statement for prediction model development⁴⁹. DBM validation studies should include standardized reporting of: (1) detailed participant demographics including often-overlooked factors like language proficiency and digital literacy; (2) model performance metrics stratified by all relevant demographic variables; (3) explicit documentation of inclusion/exclusion criteria and their potential impact on health equity; (4) open-source code and synthetic datasets where possible to enable independent validation; and (5) plain-language summaries accessible to patient communities. Regulatory bodies should mandate these transparency requirements as part of the approval process, while establishing post-market surveillance systems to monitor real-world performance across diverse populations. Only through such comprehensive transparency can we ensure that DBMs reduce rather than exacerbate existing health disparities⁴³.

SI 2.4 Challenges with the reimbursement economic model

The widespread clinical adoption of DBMs also faces significant challenges from outdated reimbursement structures, especially in the United States. Healthcare reimbursement relies on standardized billing codes that clinicians submit to insurance companies when claiming payment for medical services. These codes, structured around the time required to perform each service, specify the procedures performed, diagnoses made, and supplies used. While recent Current Procedural Terminology (CPT) codes for medical billing have begun incentivizing [remote patient monitoring](#) (RPM), the billing frameworks still require physician or technician supervision for diagnostic tools, which is fundamentally incompatible with the scalable, automated analyses that make DBMs economically viable⁵⁰. Additionally, the CPT coding system is largely based on the concept of episodic service delivery, with a few recent attempts at providing a reimbursement model for continuous services such as RPM that can be applied to patients with NDDs⁵¹. For NDDs where early detection of disease manifestations can prevent costly complications like falls or medication-related hospitalizations, DBMs align well with a disease management or population health emphasis on prevention. For instance, in Parkinson's disease, [Rune Labs](#)' Apple Watch app called [StrivePD](#) was shown to reduce emergency room admissions by 42% and reduce specialists' time by 18% in a large-scale research partnership with Kaiser Permanente⁵².

SI 2.5 Future considerations

The lack of standardization across sensing hardware, protocols, and reference datasets severely impedes meaningful comparisons between DBM studies. For instance, proprietary algorithms calculating seemingly basic metrics like step count can vary between devices, making cross-study comparisons challenging. This challenge is compounded when manufacturers release software and firmware updates that modify these algorithms without clear documentation, potentially invalidating previously established baselines and thresholds (a “black box” problem as found in the Apple Watch⁵³). While complex, these challenges are addressable in nature; as such, interdisciplinary research consortia and developers of digital health technologies should strive to develop universal digital biomarker protocols and establish open-access reference datasets with comprehensive documentation of calibration, preprocessing pipelines, sensor specifications, and software/firmware versions. These standards should encompass all aspects of the measurement chain: data collection parameters (sampling frequency, sensor placement, calibration procedures), pre-processing methods, feature extraction techniques, and validation procedures against clinical gold standards. Initiatives like the Parkinson's Disease Digital Biomarker DREAM Challenge demonstrate the value of community-driven standardization efforts⁵⁴. Additionally, novel incentive structures may be needed to encourage hardware manufacturers to provide researchers access to raw sensor signals, despite potential trade-offs with battery life or competitive advantages. Recent developments, such as Apple's SensorKit framework for research purposes, indicate a promising industry willingness to support scientific advancement in these areas. However, these initiatives must be expanded alongside requirements for transparent documentation of algorithmic changes

in software updates. Without such transparency, DBMs built on processed signals rather than raw data will remain vulnerable to undetected biases and shifts in performance, particularly for rare disease populations where validation efforts are already challenging and resource-intensive. While this is a potential market for new companies, it also limits potential scalability, particularly in more resource-constrained settings.

Current digital biomarkers often lack contextual information necessary for accurate interpretation. For example, changes in gait parameters might indicate disease progression or simply reflect different walking environments. Future research should incorporate contextual sensing that captures environmental factors (e.g., terrain, weather, social context, higher-level activities) alongside symptom measurements. Discerning pathological changes from environmental variations will be crucial to improving ecological validity and reducing noise. While current technology may struggle to achieve this, up-and-coming devices like smart glasses combined with recent developments in egocentric AI, like [Project Aria](#), could enable a new wealth of contextual information to truly understand in what context a specific marker has been measured.

For digital biomarkers to enter clinical practice, we must address barriers that fall into two distinct categories: (i) addressable technical gaps and (ii) fundamental adoption blockers. The first category includes challenges like algorithmic interpretability within predictive models, where frameworks like machine learning models often function as “black boxes” by making predictions without revealing their reasoning process. This opacity creates a justifiable hesitation among clinicians who need to understand why specific recommendations are made. However, emerging explainable AI techniques offer a constructive solution by creating systems that provide interpretable reasoning, quantify uncertainty in predictions, and communicate how predictions relate to clinical outcomes and disease mechanisms⁵⁵. Similarly, algorithmic bias and data privacy concerns are serious but technically solvable problems through methods such as differential privacy, federated learning, and bias detection algorithms that can be implemented with sufficient resources and expertise. Data standardization across devices, while challenging due to proprietary algorithms that create inconsistencies in basic metrics like step counting, can be addressed through industry consortia and regulatory frameworks that establish universal protocols for sensor calibration and feature extraction. On the other hand, there exist different fundamental adoption blockers that require systemic healthcare transformation beyond technical innovation alone. One of those blockers involves demonstrating clinical utility, i.e., proving that digital biomarkers actually improve patient outcomes, reduce costs, or improve diagnostic accuracy compared to existing standards through large-scale, long-term trials that face inherent limitations in study duration and patient recruitment that cannot be solved through solely technical advances. Another fundamental blocker centers on healthcare economics and reimbursement structures that remain oriented toward discrete medical events rather than continuous monitoring paradigms, thus requiring policy and economic change that surpasses the digital health sector. We suggest that while researchers should continue addressing technical gaps, the field’s ultimate success depends on demonstrating clear clinical value propositions and advocating for healthcare policy reforms that support preventive,

continuous care models. Implementation strategies should prioritize transparency tools that visualize confidence levels, underlying data quality, and specific factors influencing predictions. Accordingly, healthcare professionals require training programs to interpret digital biomarker data confidently and patients need accessible explanations of how these technologies benefit their care. Research should prioritize human-centered design approaches with multidisciplinary teams, including clinicians, patients, caregivers, and developers, to create intuitive systems that enhance rather than complicate care delivery.

References:

- 1 Tricco, A. C. *et al.* PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of internal medicine* **169**, 467–473 (2018).
- 2 Borhani, S. *et al.* Gauging Working Memory Capacity From Differential Resting Brain Oscillations in Older Individuals With A Wearable Device. *Front Aging Neurosci* **13**, doi:ARTN 625006 10.3389/fnagi.2021.625006 (2021).
- 3 Lin, W. C. *et al.* Electroencephalogram-Driven Machine-Learning Scenario for Assessing Impulse Control Disorder Comorbidity in Parkinson's Disease Using a Low-Cost, Custom LEGO-Like Headset. *IEEE T Neur Sys Reh* **31**, 4106–4114, doi:10.1109/Tnsre.2023.3323902 (2023).
- 4 Barbey, F. M. *et al.* Neuroscience from the comfort of your home: Repeated, self-administered wireless dry EEG measures brain function with high fidelity. *Frontiers in Digital Health* **4**, doi:ARTN 944753 10.3389/fdgth.2022.944753 (2022).
- 5 Sringean, J., Taechalerpaisarn, P., Thanawattano, C. & Bhidayasiri, R. How well do Parkinson's disease patients turn in bed? Quantitative analysis of nocturnal hypokinesia using multisite wearable inertial sensors. *Parkinsonism Relat D* **23**, 10–16, doi:10.1016/j.parkreldis.2015.11.003 (2016).
- 6 Svensson, T., Madhawa, K., Nt, H., Chung, U. I. & Kishi, A. Validity and reliability of the Oura Ring Generation 3 (Gen3) with Oura sleep staging algorithm 2.0 (OSSA 2.0) when compared to multi-night ambulatory polysomnography: A validation study of 96 participants and 421,045 epochs. *Sleep Med* **115**, 251–263, doi:10.1016/j.sleep.2024.01.020 (2024).
- 7 Kataoka, H., Saeki, K., Yamagami, Y., Sugie, K. & Obayashi, K. Quantitative associations between objective sleep measures and early-morning mobility in Parkinson's disease: cross-sectional analysis of the PHASE study. *Sleep* **43**, doi:ARTN zsz203 10.1093/sleep/zsz203 (2020).
- 8 Snow, A. L. *et al.* Enhancing sleep quality for nursing home residents with dementia: a pragmatic randomized controlled trial of an evidence-based frontline huddling program. *Bmc Geriatr* **21**, doi:ARTN 281 10.1186/s12877-021-02189-8 (2021).
- 9 Donahue, E. K. *et al.* Physical activity intensity is associated with cognition and functional connectivity in Parkinson's disease. *Parkinsonism Relat D* **104**, 7–14, doi:10.1016/j.parkreldis.2022.09.005 (2022).
- 10 Vitazkova, D. *et al.* Transforming Sleep Monitoring: Review of Wearable and Remote Devices Advancing Home Polysomnography and Their Role in Predicting Neurological Disorders (vol 15, 117, 2025). *Biosensors-Basel* **15**, doi:ARTN 508 10.3390/bios15080508 (2025).
- 11 Miller, D. J., Sargent, C. & Roach, G. D. A Validation of Six Wearable Devices for Estimating Sleep, Heart Rate and Heart Rate Variability in Healthy Adults. *Sensors-Basel* **22**, doi:ARTN 6317 10.3390/s22166317 (2022).
- 12 Robbins, R. *et al.* Accuracy of Three Commercial Wearable Devices for Sleep Tracking in Healthy Adults. *Sensors-Basel* **24**, doi:ARTN 6532 10.3390/s24206532 (2024).

- 13 Summa, S. *et al.* Validation of low-cost system for gait assessment in children with ataxia. *Comput Meth Prog Bio* **196**, doi:ARTN 105705 10.1016/j.cmpb.2020.105705 (2020).
- 14 Mc Ardle, R., Del Din, S., Galna, B., Thomas, A. & Rochester, L. Differentiating dementia disease subtypes with gait analysis: feasibility of wearable sensors? *Gait Posture* **76**, 372–376, doi:10.1016/j.gaitpost.2019.12.028 (2020).
- 15 Kirk, C. *et al.* Mobilise-D insights to estimate real-world walking speed in multiple conditions with a wearable device. *Scientific reports* **14**, 1754 (2024).
- 16 Ellis, R. J. *et al.* A Validated Smartphone-Based Assessment of Gait and Gait Variability in Parkinson's Disease. *Plos One* **10**, doi:ARTN e0141694 10.1371/journal.pone.0141694 (2015).
- 17 Lu, M. *et al.* Quantifying Parkinson's disease motor severity under uncertainty using MDS-UPDRS videos. *Med Image Anal* **73**, doi:ARTN 102179 10.1016/j.media.2021.102179 (2021).
- 18 Endo, M. *et al.* GaitForeMer: Self-supervised Pre-training of Transformers via Human Motion Forecasting for Few-Shot Gait Impairment Severity Estimation. *Lect Notes Comput Sc* **13438**, 130–139, doi:10.1007/978-3-031-16452-1_13 (2022).
- 19 Endo, M. *et al.* Data-driven discovery of movement-linked heterogeneity in neurodegenerative diseases. *Nat Mach Intell* **6**, doi:10.1038/s42256-024-00882-y (2024).
- 20 Naeini, S. A., Simmatis, L., Yunusova, Y. & Taati, B. Concurrent Validity of Automatic Speech and Pause Measures During Passage Reading in ALS. *2022 Ieee-Embs International Conference on Biomedical and Health Informatics (Bhi) Jointly Organised with the Ieee-Embs International Conference on Wearable and Implantable Body Sensor Networks (Bsn'22)*, doi:10.1109/Bhi56158.2022.9926862 (2022).
- 21 Zhao, X. H. *et al.* Validity, feasibility, and effectiveness of a voice-recognition based digital cognitive screener for dementia and mild cognitive impairment in community-dwelling older Chinese adults: A large-scale implementation study. *Alzheimers Dement* **20**, 2384–2396, doi:10.1002/alz.13668 (2024).
- 22 Staffaroni, A. M. *et al.* Reliability and Validity of Smartphone Cognitive Testing for Frontotemporal Lobar Degeneration. *Jama Network Open* **7**, doi:ARTN e244266 10.1001/jamanetworkopen.2024.4266 (2024).
- 23 Madrid-Navarro, C. J. *et al.* Validation of a Device for the Ambulatory Monitoring of Sleep Patterns: A Pilot Study on Parkinson's Disease. *Front Neurol* **10**, doi:ARTN 356 10.3389/fneur.2019.00356 (2019).
- 24 Coravos, A., Khozin, S. & Mandl, K. D. Developing and adopting safe and effective digital biomarkers to improve patient outcomes. *NPJ digital medicine* **2**, 14 (2019).
- 25 Goldhahn, J., Brasier, N. & Kehoe, L. Digitalizing health trials by the Clinical Trials Transformation Initiative. *Nature Reviews Bioengineering* **2**, 717–718 (2024).
- 26 Goldsack, J. C. *et al.* Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). *npj digital Medicine* **3**, 55 (2020).
- 27 Chen, C.-C., Lee, S.-H., Wang, W.-J., Lin, Y.-C. & Su, M.-C. EEG-based motor network biomarkers for identifying target patients with stroke for upper limb rehabilitation and its construct validity. *PLoS*

- One* **12**, e0178822 (2017).
- 28 Rutkowski, T. M., Abe, M. S., Tokunaga, S., Komendziński, T. & Otake-Matsuura, M. in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 4056–4059 (IEEE).
- 29 Marandi, R. Z. & Gazerani, P. Aging and eye tracking: in the quest for objective biomarkers. *Future Neurology* **14**, FNL33 (2019).
- 30 Diotaiuti, P. *et al.* Eye tracking in Parkinson's disease: a review of oculomotor markers and clinical applications. *Brain Sciences* **15**, 362 (2025).
- 31 Bueno, A. P. A., Sato, J. R. & Hornberger, M. Eye tracking - The overlooked method to measure cognition in neurodegeneration? *Neuropsychologia* **133**, doi:ARTN 107191 10.1016/j.neuropsychologia.2019.107191 (2019).
- 32 Kholghi, M., Szollosi, I., Hollamby, M., Bradford, D. & Zhang, Q. A validation study of a ballistocardiograph sleep tracker against polysomnography. *Journal of Clinical Sleep Medicine* **18**, 1203–1210 (2022).
- 33 O'Connor, Y., Rowan, W., Lynch, L. & Heavin, C. Privacy by design: informed consent and internet of things for smart health. *Procedia computer science* **113**, 653–658 (2017).
- 34 Khezzr, S., Moniruzzaman, M., Yassine, A. & Benlamri, R. Blockchain technology in healthcare: A comprehensive review and directions for future research. *Applied sciences* **9**, 1736 (2019).
- 35 Nebeker, C., Gholami, M., Kareem, D. & Kim, E. Applying a digital health checklist and readability tools to improve informed consent for digital health research. *Frontiers in Digital Health* **3**, 690901 (2021).
- 36 Manly, J. J., Gilmore-Bykovskyi, A. & Deters, K. D. Inclusion of Underrepresented Groups in Preclinical Alzheimer Disease Trials-Opportunities Abound. *Jama Network Open* **4**, doi:ARTN e2114606 10.1001/jamanetworkopen.2021.14606 (2021).
- 37 Wise-Brown, A. *et al.* Promoting diversity in clinical trials: insights from planning the ALUMNI AD study in historically underrepresented US populations with early symptomatic Alzheimer ' s disease. *Eclinicalmedicine* **73**, doi:ARTN 102693 10.1016/j.eclinm.2024.102693 (2024).
- 38 Brewer, L. C. *et al.* Back to the future: achieving health equity through health informatics and digital health. *JMIR mHealth and uHealth* **8**, e14512 (2020).
- 39 Sieck, C. J. *et al.* Digital inclusion as a social determinant of health. *NPJ digital medicine* **4**, 52 (2021).
- 40 Lyles, C. R., Wachter, R. M. & Sarkar, U. Focusing on digital health equity. *Jama* **326**, 1795–1796 (2021).
- 41 Miller, I. N. & Cronin-Golomb, A. Gender differences in Parkinson's disease: clinical characteristics and cognition. *Movement disorders* **25**, 2695–2703 (2010).
- 42 Schaeckermann, M. *et al.* Health equity assessment of machine learning performance (HEAL): a framework and dermatology AI model case study. *EclinicalMedicine* **70** (2024).
- 43 Alegria, M., Sud, S., Steinberg, B. E., Gai, N. & Siddiqui, A. Reporting of participant race, sex, and socioeconomic status in randomized clinical trials in general medical journals, 2015 vs 2019. *JAMA Network Open* **4**, e2111516 (2021).
- 44 Bòyd, A. D. *et al.* Potential bias and lack of generalizability in electronic health record data: reflections on health equity from the National Institutes of Health Pragmatic Trials Collaboratory.

- Journal of the American Medical Informatics Association* **30**, 1561–1566, doi:10.1093/jamia/ocad115 (2023).
- 45 Nouri, S. S. *et al.* Patient characteristics associated with objective measures of digital health tool use in the United States: A literature review. *Journal of the American Medical Informatics Association* **27**, 834–841, doi:10.1093/jamia/ocaa024 (2020).
- 46 Liu, X. *et al.* Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *The Lancet Digital Health* **2**, e537–e548 (2020).
- 47 Rivera, S. C. *et al.* Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *The Lancet Digital Health* **2**, e549–e560 (2020).
- 48 Norgeot, B. *et al.* Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. *Nat Med* **26**, 1320–1324, doi:10.1038/s41591-020-1041-y (2020).
- 49 Collins, G. S. *et al.* TRIPOD plus AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *Bmj-Brit Med J* **385**, doi:ARTN e078378 10.1136/bmj-2023-078378 (2024).
- 50 Whitehead, D. & Conley, J. The Next Frontier of Remote Patient Monitoring: Hospital at Home. *J Med Internet Res* **25**, doi:ARTN e42335 10.2196/42335 (2023).
- 51 Rakers, M. M. *et al.* Perceived barriers and facilitators of structural reimbursement for remote patient monitoring, an exploratory qualitative study. *Health Policy Techn* **12**, doi:ARTN 100718 10.1016/j.hlpt.2022.100718 (2023).
- 52 Aguilar, M. What Kaiser Permanente learned from using Apple Watch tech in Parkinson's care. *STAT+ healthtech* (2024). <<https://www.statnews.com/2024/10/03/kaiser-permanente-apple-watch-rune-strivepd-parkinsons/>>.
- 53 Owen, M. Apple Watch 'black box' algorithms unreliable for medical research [u]. *Apple Insider* (2021). <<https://appleinsider.com/articles/21/07/27/apple-watch-black-box-algorithms-unreliable-for-medical-research>>.
- 54 Sieberts, S. K. *et al.* Crowdsourcing digital health measures to predict Parkinson's disease severity: the. *Npj Digital Medicine* **4**, doi:ARTN 53 10.1038/s41746-021-00414-7 (2021).
- 55 Ayaz, M., Pasha, M. F., Alzahrani, M. Y., Budiarto, R. & Stiawan, D. The Fast Health Interoperability Resources (FHIR) Standard: Systematic Literature Review of Implementations, Applications, Challenges and Opportunities (vol 9, e21929, 2021). *Jmir Med Inf* **9**, doi:ARTN e32869 10.2196/32869 (2021).