Evaluating Motor Symptoms in Parkinson's Disease Through Wearable Sensors: A Systematic Review of Digital Biomarkers

Carlos Polvorinos-Fernández [1], Luis Sigcha [2], Luigi Borzì [3], Gabriella Olmo [3], César Asensio [4], Juan Manuel López [5], Guillermo de Arcas [1] and Ignacio Pavón [1,*]

[1] Instrumentation and Applied Acoustics Research Group (I2A2), Department of Mechanical Engineering, ETS. Ingenieros Industriales, Universidad Politécnica de Madrid, 28006 Madrid, Spain; c.polvorinos@upm.es (C.P.-F.); g.dearcas@upm.es (G.d.A.)

[2] Department of Physical Education and Sport Sciences (PESS), Data-Driven Computer Engineering Research Centre (D2iCE), and Health Research Institute (HRI), University of Limerick, V94 T9PX Limerick, Ireland; luis.sigcha@ul.ie

[3] Department of Control and Computer Engineering, Politecnico di Torino, 10129 Turin, Italy; luigi.borzi@polito.it (L.B.); gabriella.olmo@polito.it (G.O.)

[4] Instrumentation and Applied Acoustics Research Group (I2A2), Department of Audiovisual Engineering and Communications, ETS. de Ingeniería y Sistemas de Telecomunicación, Universidad Politécnica de Madrid, 28031 Madrid, Spain; c.asensio@upm.es

[5] Instrumentation and Applied Acoustics Research Group (I2A2), Department of Physical Electronics, Electrical Engineering and Applied Physics, ETS. de Ingeniería y Sistemas de Telecomunicación, Universidad Politécnica de Madrid, 28031 Madrid, Spain; juanmanuel.lopez@upm.es

* Correspondence: ignacio.pavon@upm.es

Abstract: Parkinson's disease (PD) is the second most common neurodegenerative disorder, entailing several motor-related symptoms that contribute to a reduced quality of life in affected subjects. Recent advances in wearable technologies and computing resources have shown great potential for the assessment of PD-related symptoms. However, the potential applications (e.g., early diagnosis, prognosis and monitoring) and key features of digital biomarkers for motor symptoms of PD (DB-MS-PD) have not been comprehensively studied. This study aims to provide a state-of-the-art review of current digital biomarker definitions for PD, focusing on the use of wearable devices. This review systematically examines research articles from 2012 to 2024, focusing on key features and recent technologies in PD research. A total of 22 studies were included and thoroughly analyzed. Results indicate that DB-MS-PD can accurately distinguish patients with PD (PwPD) from healthy controls (HC), assess disease severity or treatment response, and detect motor symptoms. Large sample sizes, proper validation, non-invasive devices, and ecological monitoring make DB-MS-PD promising for improving PD management. Challenges include sample and method heterogeneity and lack of public datasets. Future studies can leverage evidence of the current literature to provide more effective and ready-to-use digital tools for monitoring PD.

Keywords: Parkinson's disease; digital biomarkers; motor symptoms; wearables; body-worn sensors; machine learning

1. Introduction

Parkinson's disease (PD) is a chronic neurological disorder caused by the progressive loss of dopamine-producing cells. Dopamine is a neurotransmitter involved, among other things, in the organisation of movement and whose effect is to strenghten muscular activation [1,2]. PD is the second most common neurodegenerative disease after Alzheimer's dementia and mainly affects elderly people. The increasingly elderly population has led to an increase in the number of subjects affected by PD, reaching 8.5 million patients in 2019 [3]. Moreover, projections indicate a substantial increase, with the global patient population anticipated to 12 million in 2040 [4].

PD include motor symptoms such as tremor, bradykinesia, rigidity, or postural instability; and non-motor symptoms such as as loss of smell, constipation, fatigue, anxiety, depression, and REM sleep behavior disorder, which can occur years before the diagnosis [5–7]. These symptoms vary from patient to patient, as well as over the course of the disease [8], gradually reducing patients' quality of life [9].

The diagnosis of PD is complex, as there is no pathognomonic sign or biomarker to confirm the disease [2]. Instead, the diagnosis is based on a combination of medical history, clinical examination and the presence of specific motor and non-motor symptoms [10]. The diagnosis process involves an evaluation by neurologists, and may be difficult especially in the early stages of the disease, when symptoms may be mild or non-specific. Therefore, it is important to monitor symptoms regularly to appropriately adjust the treatment plan [11].

Currently, there is no cure that can stop or reverse the disease progression. Instead, several symptomatic therapies are available, which can guarantee a good control of the disease for several years. The most common options are drug treatment or, in severe cases, surgical interventions (HIFU, intracranial implants or duodenal pumps, among others) [12]. The most widely used drug is Levodopa, a metabolic precursor of dopamine [13], which is effective in controlling symptoms during the first years of treatment. However, after several years of Levodopa treatment, motor complications such as dyskinesias (i.e., involuntary movements) and motor fluctuations may occur, and the patient may switch between periods of adequate and poor control of the disease [14].

At present, the most commonly used scale for evaluating PD progression is the Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [15]. This scale is designed to evaluate both motor and non-motor symptoms during a clinical examination, using a series of questions and guided exercises performed by the patient.

The fluctuating nature of PD symptoms during the day [16] and the infrequent clinical consultations (6–9 months apart [17]) make it difficult to accurately assess the disease. Furthermore, the correct application and interpretation of this scale depend on the doctor's experience. The combination of these variables complicates the accuracy of therapeutic and pharmacological adjustments, often leading to complications derived from over- or under-medication [18].

In recent years, the emergence of new technologies, such as wearable sensors, has gained significant attention in the management of PD. These devices offer the potential to monitor various aspects of PD symptoms and motor function [**?**]. Wearable sensors enable continuous, objective, and long-term data collection and monitoring in a variety of contexts, including clinical and free-living environments [20].

Wearable devices can collect a large amount of data, which could give rise to digital biomarkers (i.e., quantitative and sensitive measurements of disease progression) for PD assessment, providing a wealth of data that traditional clinical assessments cannot capture. This extensive data collection can enable better understanding and management of the disease, facilitate personalized treatment, and support large-scale studies that contribute to the development of new therapies and interventions [21].

Review Objectives

This article aims to provide the reader with a summary of the trends and techniques used by the scientific community in the use of wearable technologies to acquire data and define digital biomarkers for motor symptoms, thus providing a review of the current status of the defined digital biomarkers.

Recent reviews have investigated the potential of digital biomarkers in PD, providing useful insights from different perspectives. A summary of biomarkers (speech, gait, handwriting) with potential clinical applications in PD was provided in [22], with a particular focus on validation, regulatory approval, ethical, legal, and social aspects. In [23], the potential of biomarkers extracted from facial expression analysis and eye tracking were investigated for early diagnosis and assessment of cognitive decline. Nevertheless, to the best

of our knowledge, no review studies have focused on digital biomarkers for motor symptoms of PD (DB-MS-PD) obtained through wearable devices, as well as their key features, encompassing sensor settings, experimental procedures, and data analysis methods.

This study reviews relevant research articles published between 2012 and 2024, focusing on digital biomarkers defined for PD through motion analysis using data collected through wearable devices. In specific, the objective of this study is to provide a comprehensive overview of sensors' settings, experimental procedures and data analysis methods that allow the extraction of relevant DB-MS-PD. Ultimately, comparative analysis of various studies can facilitate the identification of promising and minimally invasive systems (both hardware and software) for the supervised and unsupervised assessment of PD.

The remainder of the document is organized as follows. Section 2 presents an overview of the wearable devices and digital biomarkers. Section 3 describes the methodology used for the systematic review, including the search strategy used to find relevant articles. Section 4 presents the results of this literature review. Section 5 provides a discussion of the findings in the analyzed articles. Finally, the conclusions of this work are given in Section 6.

2. Background

2.1. Wearable Devices

The term wearable devices (*wearables*) refers to compact electronic devices as well as wireless-enabled computers that are seamlessly integrated into gadgets, accessories, or clothing designed to be worn on the human body. It also includes more invasive devices such as microchips or smart tattoos, or commercial and widespread devices including smart glasses, smartphones, smartwatches, smart clothing, and smart shoes, among others [24,25].

In recent years, wearable devices have shown the potential to overcome certain limitations of the traditional healthcare and medical assessments by harnessing digital and mobile health (m-health) technologies to progress towards efficient and personalized healthcare [26]. Moreover, these devices can facilitate long-term monitoring outside clinical settings, offering a discreet and comfortable solution [27].

Wearable devices have the ability to collect comprehensive health information while functioning as they were originally designed, like fashion or productivity devices. The collected data can be analysed using standard protocols by artificial intelligence, with the aim of identifying possible predictions of health problems [28], as well as recognising activities and detecting the context in which these activities are performed [29,30].

Since wearable sensors are going to be integrated into clothes, accessories or applied directly to the skin, they must be designed to be practical and accepted by people in their daily lives, so the materials used in their manufacture must be highly flexible, durable, anti-allergic and lightweight [31]. Advances in materials science are driving innovation in this field and researchers are exploring the use of nanomaterials, conductive polymers and elastic electronic components, among others [32]. In addition, these devices and sensors must be resistant to daily wear and tear if they are to be used for continuous remote monitoring, so sensors must be able to withstand stretching, bending, and exposure to sweat or moisture without loss of performance [33].

There are different types of sensors that can be integrated into a wearable device, from physical health sensors to obtain electrocardiogram, pulse or electromyogram, sensors related to human movement such as accelerometers, gyroscopes or magnetometers, to environmental sensors related to temperature, humidity or luminosity [34]. The geometry and type of wearable devices is defined in many cases by the type of sensor embedded in them, giving rise to a wide variety of options. There are very small sensors that can be included in devices as thin as a sheet of paper, flexible sensors to adapt to anybody surface or sensors that need a structure such as a watch to be used [35].

In the realm of monitoring and management of neurological disorders, wearable devices have emerged as a highly effective tool [36]. These devices allow for continuous and non-invasive monitoring of a range of physiological and behavioral parameters, both in medical consultation and in free-living conditions [37–39]. Wearables have the possibility to

provide an abundance of data that can support tasks such as early detection, optimization of treatment, and management of neurological disorders [40]. However, it is important to acknowledge that despite the potential advantages of this technology, their application should be complemented by professional medical supervision and should complement traditional clinical evaluations [41].

To carry out these tasks, wearable devices incorporate several core components that are common in most devices. The key hardware components include the display, which facilitates user interaction and presents information about activities to be performed, the battery, which provides the necessary power supply, and the processor, which handles computational tasks. In addition, sensors are used to gather physiological and environmental data. Storage is allocated for local data retention, and the housing encloses and protects these components. Arguably, the most critical components of wearable devices are sensors. In the field of human activity recognition, the accelerometer, gyroscope, magnetometer, and barometer are the most commonly used [42]. These sensors measure and collect data on motion, orientation, magnetic fields, and atmospheric pressure.

The comprehensive framework for wearable health monitoring systems is shown in Figure 1. The data collected from the sensors of wearable devices can be processed in two primary ways: Edge processing and Offline processing. In edge processing, data is analyzed in real-time directly on the device, enabling immediate responses, albeit with limited computational resources. In contrast, offline processing involves transmitting the data to external systems, such as servers or computers, where more extensive and resource-intensive analysis can be performed.
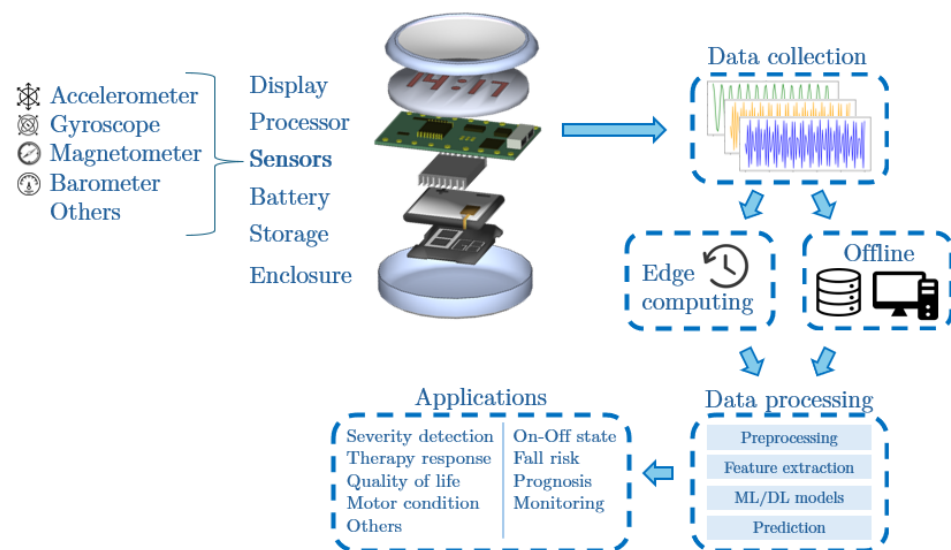

Figure 1. Comprehensive framework for wearable health monitoring systems.

Regardless of the processing method, the overall process follows a similar structure. It begins with a pre-processing stage, where raw sensor data is cleaned and transformed to ensure it is suitable for subsequent steps. Next, feature extraction is performed to identify key features or patterns within the data. These extracted features could then input into machine learning (ML) or deep learning (DL) models, which generate predictions, providing insights or diagnoses based on the analyzed data.

In PD, wearable devices have been used to assess various symptoms. In some cases, symptom assessment has been done by time and frequency analysis, in other cases by extracting relevant features from the measured data, another strategy is the development of classifiers of disease severity [43,44], resting tremor [45], bradykinesia [46] or freezing of gait [47]. Several studies have presented devices specifically designed for PD patients such as STAT-ON® [48], KinetiGraphTM® [49] or PD Monitor® [50]. Figure 2 presents examples of wearable devices employed in the detection and monitoring of PD. These devices could

be smartwatches, inertial measurement units (IMUs), pressure insoles, sensor belts, or smartphones with body adaptors, among others.
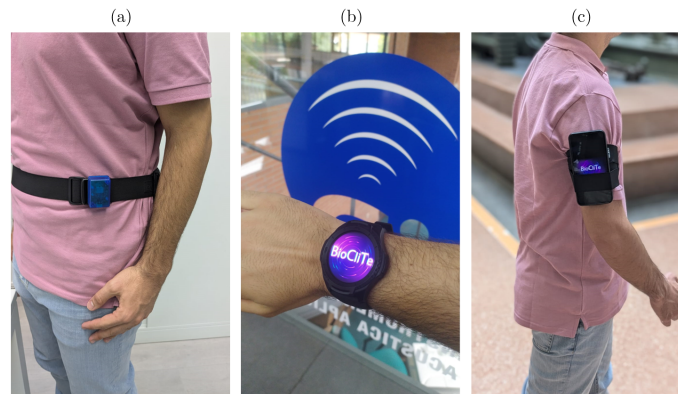

Figure 2. Examples of wearable devices for Parkinson's disease management: (**a**) IMU, (**b**) smartwatch, (**c**) smartphone.

In [51], commercial devices used for PD management were evaluated, considering their validation process and clinical applications along with the strengths and weaknesses of each. In [52] the authors explored the wearable devices used for PD in hospitals, concluding that the most common types of wearable devices include IMUs and smartwatches. In [53] a comprehensive review was performed on how wearable sensors can support tasks such as early diagnosis, human motion analysis, motor fluctuations, and home and long-term monitoring for PwPD. However, although wearable devices hold promise for improving PD management, their adoption in the clinical setting is limited by problems such as inadequate technical and clinical validation [54].

2.2. Digital Biomarkers

The term *biomarker* refers to objective medical signs that can be accurately and reproducibly measured outside the patient. In contrast, medical symptoms are health or disease indications perceived by patients [55]. Biomarkers can be defined as characteristics that are objectively measured and evaluated as indicators of normal biological processes, pathogenic processes, or pharmacological responses to therapeutic interventions [56]. Biomarkers can include any substance, structure, or process that can be measured from the body and can influence or predict the incidence of the outcome or disease. They can be diverse, from gait cadence and pulse measurements to complex laboratory tests of blood and other tissues [57]. However, the main objective of a biomarker remains to establish the connection between measurable parameters and relevant clinical endpoints [55].

The type of biomarkers studied in this work are known as *digital biomarkers*. According to the U.S. Food and Drug Administration (FDA), a digital biomarker is a feature or group of features obtained from digital health technologies that are measured to indicate normal biological processes, pathogenic processes, or reactions to exposure or intervention, such as therapeutic interventions [58,59]. Digital biomarkers are rapidly developing frontier enabled by the availability of sensors and personal devices that can assimilate information about an individual's psychological state, exercise level, cognitive abilities, eating patterns, movement, and tremor [60]. These data are largely derived from sources such as smartphones and portable electronic devices [61].

Although many additional studies are needed to link digital phenotypes and endpoints with traditional measurements, digital biomarkers have the potential to introduce novel measurements for phenomena that are already in use [62]. Depending on the applications of biomarkers, several types can be identified, such as diagnostic biomarkers (to confirm the presence of a certain disease), monitoring biomarkers (to determine the progression of a disease), pharmacodynamic/response biomarkers (to check the response to certain

therapies), predictive biomarkers (to know the response to certain medical products) and prognostic biomarkers (to identify the possibility of suffering a certain disease) [63].

Extracting digital biomarkers involves collecting and analyzing data from digital devices or platforms that can provide insights into physiological, behavioral and cognitive states, in a similar process as shown in Figure 1. These biomarkers can be derived from various sources such as smartphones, wearables, social media, and other digital interactions. The information defined in the biomarker must be meticulously processed and extracted, as raw data is influenced by numerous processes that obscure the underlying signals. To achieve this, various techniques in digital signal processing, statistical analysis, and artificial intelligence are employed to extract and refine the data into a coherent biomarker. Ultimately, these biomarkers undergo validation through clinical studies to ensure their accuracy and reliability in reflecting predicted physiological, behavioral, or cognitive states.

3. Methods

3.1. Research Questions

The aim of this review is to collect, analyze, and evaluate studies concerning the identification of DB-MS-PD. To achieve this objective, the following research questions were formulated:

RQ-1 What types of sensors and wearable devices are most commonly employed for monitoring and developing DB-MS-PD?
RQ-2 Are there specific digital biomarkers that are commonly measured or tracked using wearables in PD?
RQ-3 How reliable and accurate are the digital biomarkers captured by these wearables?
RQ-4 What are the main challenges or limitations associated with using wearables for capturing DB-MS-PD?

3.2. Search Strategy

On 14 February 2024, a literature search was conducted on the PubMed, Scopus, IEEE and Web of Science databases for all the returned results. The search string included keywords related to the disease under investigation, the type of biomarker searched and the devices used to collect the data. In more detail, the following Boolean search string was used:

((Parkinson) **OR** *(Motor symptoms))* **AND** *(Biomarker)* **AND** *(Digital)* **AND** *(Wearable)*.

No additional filters were applied in the literature search. All retrieved studies were systematically identified and screened, and the data were extracted for relevant information following the PRISMA guidelines [64].

3.3. Inclusion and Exclusion Criteria

The topic of this review concerns the definition of digital biomarkers of PD obtained from wearable devices. Journal articles published between January 2012 and February 2024 and written in English were included. Furthermore, the exclusion criteria were as follows:

1. Papers without peer review, books, book chapters, or published as "letter", "comments", "perspective" "case reports", "surveys" or "reviews".
2. Literature not written in English.
3. Studies related to diseases other than PD.
4. Studies that did not use any wearable devices or portable sensors for data acquisition.
5. Studies showing the results of a challenge, competition or programme.
6. Studies primarily focused on activities not related to motor symptoms in PD.
7. Studies that do not include humans.

3.4. Data Extraction

Four authors (C.PF, L.S., L.B., and I.P.) independently selected candidate studies by reviewing the title and abstract and repeated the process until they reached a consensus. The same procedure was performed for the selection based on the full-text evaluation.

Finally, candidate studies that met the eligibility criteria were selected for inclusion in the review. The following information was included in the data extraction procedure:

a. Identification of study data, including authors, title and citation.
b. Type of test performed.
c. Characteristics of the participants in the study.
d. Type, number, and location of the wearable sensors and devices used for data acquisition.
e. Objective of the study.
f. End points.

4. Results

4.1. Systematic Review

Based on the search criteria, 59 articles were retrieved from PubMed, 172 from Scopus, 16 from Web of Science and 8 from IEEE, for a total of 255 publications. After removing duplicates (57), 198 publications were examined for titles and abstracts. Then 143 articles were excluded according to the exclusion criteria. Of the 55 remaining records, 9 studies were excluded due to the unavailability of the full text. Consequently, a total of 46 full texts were screened for eligibility, and 24 records were excluded according to the exclusion criteria. Finally, 22 research studies were included and reviewed. The PRISMA flow chart used for the literature search and selection is shown in Figure 3.
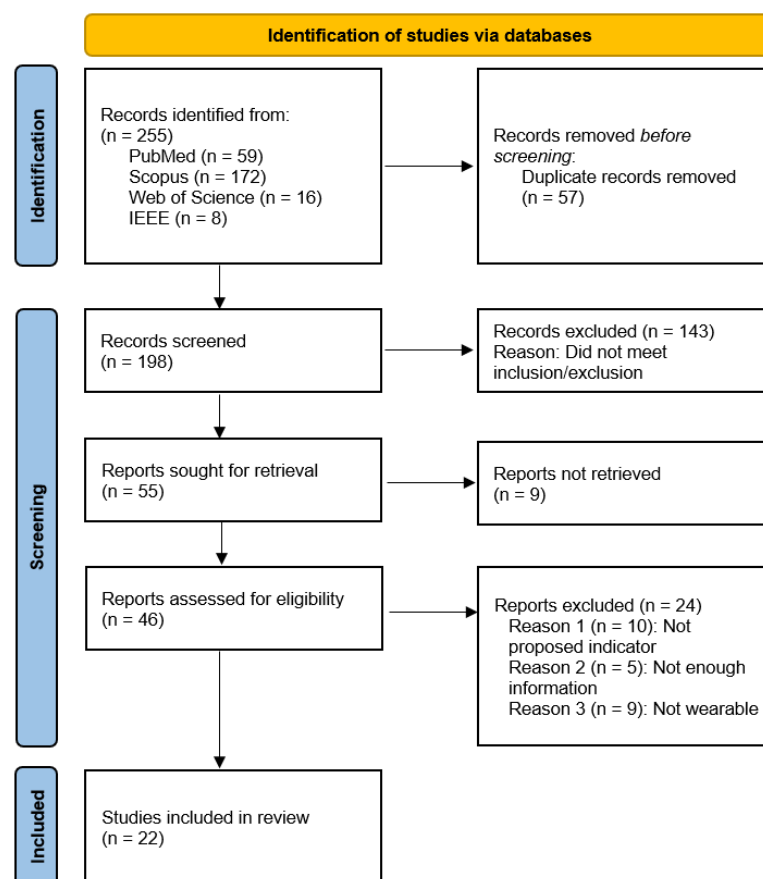

Figure 3. PRISMA flow diagram of literature search and selection process showing the number of studies identified, screened, and included in the review.

4.2. Study Characteristics

A summary of the findings of the 22 reviewed articles is given in Table 1. The information in each article was harmonized to facilitate comparability and analysis among studies.

Table 1. Summary of included papers. F: female; M: male; N: number; IMU: inertial measurement unit; PwPD: patients with Parkinson's Disease; HC: healthy control; AUC: area under the curve; VGRF: vertical ground reaction force; TUG: timed up and go; ML: machine learning; DL: deep learning

Reference	Study Design	Participants	Device, Sensors, Number (Device, Sensor), Body Location	Aim	Extraction Method	End Point
[65]	Gait (Physionet database)	PwPD: N = 90 (34 F; 56 M) HC: N = 62 (34 F; 28 M)	Pressure insoles VGRF sensors N = (1, 16) Foot (8 each)	Gait monitoring	Features combined with ML model (Random Forest)	Gait features that impact the predicted TUG scores are gait speed-based features (percentiles, mean, and kurtosis), with 84.8% accuracy.
[21]	Gait	PwPD: N = 29 (12 F; 17 M) HC: N = 27 (14 F; 13 M)	IMU (Opals by APDM®) 3-axial accelerometer, 3-axial gyroscope and 3-axial magnetometer N = (3, 3) Foot (1 each) and lower back	Classification PwPD-HC	Features combined with ML model (Logistic regression)	Turning and gait indicators discriminate PwPD from HC (Turn angle, swing time variability adn stride length with AUC = 0.87–0.89).
[66]	Gait	PwPD: N = 29 (12 F; 17 M) HC: N = 20 (8 F; 2 M)	IMU (Opals by APDM®) 3-axial accelerometer, 3-axial gyroscope and 3-axial magnetometer N = (3, 3) Foot (1 each) and lower back.	Classification PwPD-HC	Features combined with ML model (Polynomial regression)	Gait measures (gait speed, stride length) could be used to classify PwPD from HC, with AUC > 0.8.
[67]	Gait	PwPD: N = 5 HC: N = 5	IMU (Axivity AX3®) 3-axial accelerometer N = (1, 1) Lower back	Classification PwPD-HC	Features	The sample entropy of the gait signal of PwPD are higher than HC participants.
[68]	Gait Balance Task Finger Tapping (Mpower database)	PwPD: N = 1057 (359 F; 698 M) HC: N = 5343 (1014 F; 4329 M)	Smartphone 3-axial accelerometer (gait and balance) and pixel coordinates (tapping) N = (1, 2) Pocket (gait and balance) and front of participant (tapping)	Classification PwPD-HC	Raw data + DL model (Convolutional neural network)	Tapping positions (Centered tapping coordinates) are the most relevant data (AUC = 0.935) for PD detection.
[69]	Gait (Physionet database)	PwPD: N = 93 (35 F; 58 M) HC: N = 73 (33 F; 40 M)	Pressure insoles VGRF sensors N = (1, 16) Foot (8 each)	Gait monitoring and classification PwPD-HC	Features combined with ML model (Support vector machine)	Gait parameters (stride time, step time, stance time, swing time, cadence, step length, stride length, gait speed) differentiate PD severity and HC with 98.65% accuracy.

Table 1. *Cont.*

Reference	Study Design	Participants	Device, Sensors, Number (Device, Sensor), Body Location	Aim	Extraction Method	End Point
[70]	Gait (Physionet database)	PwPD: N = 93 (35 F; 58 M) HC: N = 72 (32 F; 40 M)	Pressure insoles VGRF sensors N = (1, 16) Foot (8 each)	Gait monitoring and classification PwPD-HC	Features combined with ML model (Decission Tree)	Gait parameters (step length, force variations at heel strike, centre of pressure variability, swing stance ratio, and double support phase) are able to detect PwPD with 99.9% accuracy and its severity shows $R^2 = 98.7\%$.
[71]	Gait Activities of daily living	PwPD: N = 27 (11 F; 16 M)	IMU (RehaGait®) (clinical assessment) and IMU (Physilog 5®) (home assessment) 3-axial accelerometer and 3-axial gyroscope (clinical assessment), and 3-axial accelerometer, 3-axial gyroscope, and barometrer (home assessment) N = (3, 3) Foot (1 each in clinical assessment) (only 1 in home assessment)	Gait monitoring and treatment detection	Features	Gait speed could be used to control of medication intake in PD.
[72]	Finger Tapping Pronation-supination	PwPD: N = 11 (3 F; 8 M) HC: N = 11 (6 F; 5 M)	Smartphone Pixel coordinates N = (1, 1) Front of participant	Classification PwPD-HC and ON-OFF states monitoring	Features combined with ML model (Logistic regression)	Tapping features (total taps, tap interval, and tap accuracy) can detect PwPD with $p < 0.0005$ and detect ON/OFF state with AUC 0.82.
[73]	Pronation-supination Leg Agility Toe Tapping TUG test Postural stability Postural Tremor Rest Tremor	PwPD: N = 36 (9 F; 27 M)	IMU (Movit G1®) 3-axial accelerometer and 3-axial gyroscope N = (14, 2) Lower back, upper back, forearm (1 each), arm (1 each), upper leg (1 each), lower leg (1 each), hand (1 each), foot (1 each)	Prognosis (motor symptoms) and therapy response monitoring	Features	A correlation was found between motor symptoms progression and some features (toe tapping amplitude decrement, velocity of arms and legs, sit-to-stand time, $p < 0.01$).

Reference	Study Design	Participants	Device, Sensors, Number (Device, Sensor), Body Location	Aim	Extraction Method	End Point
[74]	Gait TUG test Sit-to-tand test	PwPD: N = 10 (4 F; 6 M) PSP (Progressive Supranuclear Palsy): N = 10 (4 F; 6 M)	IMU (LEGSys®) 3-axial accelerometer, 3-axial gyroscope, 3-axial magnetometer N = (3, 3) Shin (1 each) and lower Back	Classification PwPD-PSP	Features	Gait speed was significantly slower in PSP ($p < 0.001$).
[75]	Balance Task Gait Finger tapping Reaction time Rest tremor Postural tremor	PwPD: N = 334 (125 F; 209 M) HC: N = 84 (17 F; 67 M) iRBD (idiopathic REM sleep behavior disorder): N = 104 (88 F; 16 M)	Smartphone 3-axial accelerometer (Balance, gait, rest tremor and postural tremor) and pixel coordinates (Tapping and reaction time) N = (1, 2) Pocket (balance and gait), front of participant (tapping and reaction time) and hand (postural and rest tremor)	Features combined with ML model (Random Forest)	Clasification PwPD-HC and clasification PwPD-iRBD	Postural tremor (mean squared energy, azimuth, 25th quartile, mode, radius) and rest tremor (entropy, root mean square) were the most discriminatory task between PD-HC-iRBD, with 85–88% of sensitivity.
[76]	Finger Tapping - Two-target finger tapping test - Reaction time - Pronation-supination	PwPD: N = 19 HC: N = 17	Tablet Pixel coordinates N = (1, 1) Front of participant	Classification of PwPD-HC and ON-OFF states monitoring	Features combined with ML model (Artificial neural network)	All test combined classify PwPD-HC with 93.11% accuracy. Most differentiating test is reaction time (inter-tap interval, tap accuracy) with 83.9% accuracy while ON-OFF state classifies with 76.5% accuracy.
[77]	Gait	PwPD: N = 81 (28 F; 53 M) HC: N = 61 (27 F; 34 M)	IMU (Axivity AX3®) 3-axial accelerometer N = (1, 1) Lower Back	Clasification PwPD-HC	Features combined with ML model (Discriminant analysis)	Gait features (root mean square values, power spectral density, gait speed velocity, step length, step time and age) classify PwPD with AUC = 0.94.
[78]	Gait	PwPD: N = 40 (19 F; 21 M)	IMU (+sMotion®) 3-axial accelerometer and 3-axial gyroscope N = (1, 2) Lower back	Classification motor condition and Quality of Life.	Features combined with ML model (Logistic regression)	Gait Features (velocity pace, SD swing time variability, Antero-posterior center of mass angle of postural control) classify UPDRS-III severity with AUC 0.89. Gait Features (gait speed, step time rhythm, stance time, step length) correlated with PDQ39 with AUC 0.95.

Table 1. *Cont.*

Reference	Study Design	Participants	Device, Sensors, Number (Device, Sensor), Body Location	Aim	Extraction Method	End Point
[79]	Gait Balance Task Finger Tapping (Mpower database)	PwPD: N = 610 (211 F; 399 M) (gait), 612 (211 F; 401 M) (balance), 970 (340 F; 630 M) (tapping) HC: N = 807 (152 F; 655 M) (gait), 823 (155 F; 668 M) (balance), 1674 (304 F; 1370 M) (tapping)	Smartphone 3-axial accelerometer (gait and balance) and pixel coordinates (tapping) N = (1, 2) Pocket (gait and balance) and front of participant (tapping)	Classification PwPD-HC and therapy response monitoring	Features	Tapping features (total taps, inter-tap intervals, median/standard deviation absolute deviations, correlation X-Y tap) displayed the best performance in classify PwPD-HC ($p < 0.05$).
[80]	Activities of Daily Living Rest tremor Postural tremor Finger tapping Balance task Gait	PwPD: N = 43 (8 F; 35 M) HC: N = 35 (8 F; 27 M)	Smartphone 3-axial accelerometer, 3-axial gyroscope and 3-axial magnetometer N = (1, 3) Waist (balance and gait), hand (tremor) and front of participant (tapping)	Classification PwPD-HC and Motor symptoms monitoring	Features	Tapping (inter-tap variability), rest tremor (acceleration skewness), postural tremor (total power of accelerometer), balance (mean velocity), gait (turn speed) differentiated HC from PwPD and PD abnormalities ($p < 0.005$).
[81]	Activities of daily living MDS-UPDRS task	PwPD: N = 31 (11 F; 20 M) HC: N = 50 (27 F; 23 M)	IMU (Opals by APDM®) 3-axial accelerometer, 3-axial gyroscope and 3-axial magnetometer N = (1, 3) Wrist	Motor symptoms monitoring; Therapy-response monitoring	Features	RMS (amplitude) of the magnitude vector for resting tremor ($p < 0.0004$) and RMS (amplitude) and jerk (smoothness) of the magnitude vector forbradykinesia ($p < 0.0001$) achieve agreement with clinical assessment of symptom severity and treatment-related changes in motor states.

Table 1. *Cont.*

Reference	Study Design	Participants	Device, Sensors, Number (Device, Sensor), Body Location	Aim	Extraction Method	End Point
[82]	Activities of daily living TUG test Abnormal Involuntary Movement Scale MDS-UPDRS task Gait	PwPD: N = 18 (7 F; 11 M) HC: N = 24 (11 F; 13 M)	IMU (Physilog 4®), Android smartwatch, Android smartphone, Empatica E4® smartwatch 3-axial accelerometer, 3-axial gyroscope, 3-axial magnetometer, and barometer (IMU), 3-axial accelerometer, 3-axial gyroscope, barometer, and light (Android smartwatch), 3-axial accelerometer, 3-axial magnetometer, light, proximity, GPS, WiFi, and cellular networks (Android smartphone), and Galvanic skin response, photoplethysmogram, skin temperature, 3-axial accelerometer (Empatica) N = (8, 12) Ankles (1 each), wrist (1 each), lower back (IMU), wrist (Android smartwatch), pocket (Android smartphone), and wrist (Empatica)	Classification of PwPD-HC; ON-OFF states monitoring	Features combined with ML model (Logistic regression)	The total power in the 0.5- to 10-Hz band was most discriminate feature to classify PwPD-HC (AUC = 0.76) and ON-OFF detection (AUC = 0.84).
[83]	Finger Tapping - Index and middle finger tapping (IMFT) - Alternate index finger tapping (IFT) - Thumb index finger tapping (TIFT)	PwPD: N = 20 (6 F; 14 M)	Tablet (IMFT and IFT) and Biometrics® (TIFT) Pixel Coordinates (IMFT and IFT) and Goniometer (TIFT) N = (2, 2) Front of participant (IMFT and IFT) and hand (TIFT)	Therapy response monitoring and Classification of subjects with therapy and placebo	Features combined with ML model (Discriminant analysis)	The IFT features (total taps, bivariate contour ellipse area, spatial error, velocity changes, intertap intervals) provides the best performance in estimating MDS-UPDRS III, with $p < 0.001$ and accuracy of 84% in classification of subjects.

Table 1. *Cont.*

Reference	Study Design	Participants	Device, Sensors, Number (Device, Sensor), Body Location	Aim	Extraction Method	End Point
[84]	Gait Balance Task Finger Tapping (Mpower database)	PwPD: N = 610 (211 F; 399 M) (gait), 612 (211 F; 401 M) (balance), 970 (340 F; 630 M) (tapping) HC: N = 787 (147 F; 640 M) (gait), 803 (150 F; 653 M) (balance), 1257 (239 F; 1018 M) (tapping)	Smartphone 3-axial accelerometer (gait and balance) and pixel coordinates (tapping) N = (1, 2) Pocket (gait and balance) and front of participant (tapping)	Classification PwPD-HC	Features combined with ML model (Support vector machine)	Tapping features (inter-tap interval (range, maximum value and Teager-Kaiser energy operator) detect PwPD with AUC = 0.74.
[85]	TUG test	PwPD dataset 1: N = 15 (5 F; 10 M) PwPD dataset 2: N = 27 (9 F; 17 M) HC: N = 1015 (671 F; 344 M)	IMU (Kinesis QTUG®) 3-axial accelerometer and 3-axial gyroscope N = (1, 2) Shin	Fall risk, prognosis and gait monitoring	Features combined with ML model (Logistic regression)	The mobility parameters (speed, turn, transfers, symmetry, variability) could be used to predict number of fall counts of PwPD (R^2 = 43%).

4.3. Study Design

The studies under investigation analyzed different sets of activities and tasks. These can be classified into 11 different categories: activities of daily living, balance task, finger tapping, MDS-UPDRS task, postural tremor, pronation-supination movements of the hands, reaction time, rest tremor, timed up and go (TUG) test, gait task, and others.

Due to the diversity of exercises performed within a category, similar activities were grouped in a more general category. Specifically, the finger tapping activity includes four types of sub-activities: index and middle finger tapping, alternate index finger tapping, thumb-index finger tapping, and two-target finger tapping test. Similarly, the gait includes: 2-min walk at convenient speed, 20-m straight walk at convenient speed, 20-m straight walk at fast speed, 10-m at convenient speed, 20-m circular walk at fast speed, 20-steps straight walk, normal walking, rhythmic auditory cued walking, treadmill walking, and 2-min circular walk at convenient speed.

Figure 4 shows the distribution of the different activities evaluated in the selected studies.
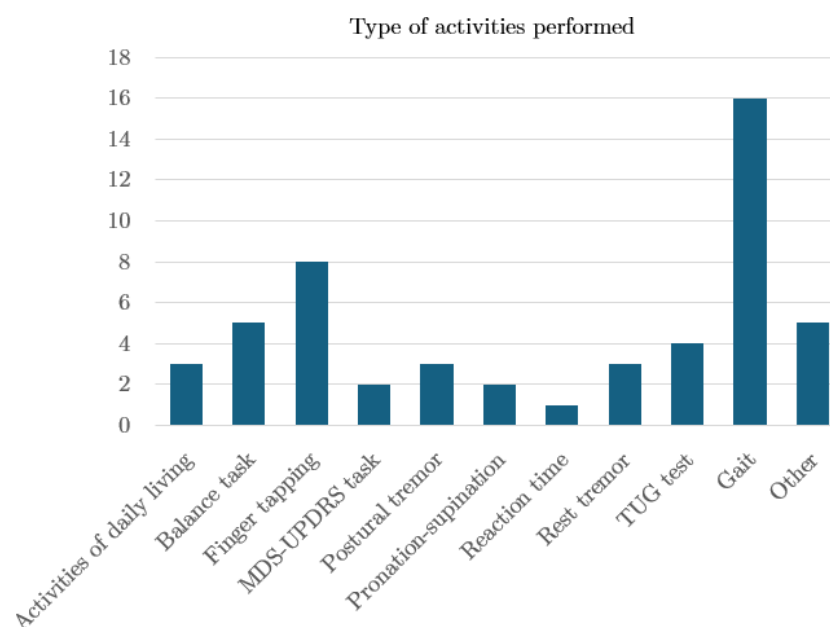

Figure 4. Distribution of the activities performed in the different studies.

A total of 52 different activities were investigated in the 22 articles analyzed. The most frequent activity is gait (N = 16, 31%), followed by finger tapping (N = 8, 15%), balance task (N = 5, 10%), and TUG test (N = 4, 8%). Other activities that appear less frequently include postural tremor, pronation-supination, rest tremor, MDS-UPDRS tasks and reaction time.

Less than one third of the studies (N = 6, 27%) used public datasets. In more detail, 3 studies [68,79,84] used the mPower dataset [86] and 3 studies [65,69,70] used data from the Physionet vertical ground reaction force (VGRF) dataset [87]. The Physionet database contains measures of gait from 93 PwPD and 73 HC. The database includes the VGRF records of subjects as they walked at their usual, self-selected pace for approximately 2 min on level ground. The mPower database includes data from more than 1000 PwPD and 4000 HC. Different activities (i.e., memory, tapping, voice and walking) were recorded with the built-in smartphone sensors (i.e., triaxial accelerometer, microphone and touchscreen) in unsupervised environments.

4.4. Participant Characteristics

Of the 22 studies reviewed, 5 of them evaluated only PwPD [71,73,74,78,83], while the other 17 papers included PwPD and HC in their study. A single study [75] included people with idiopathic REM sleep behavior disorder, while [74] also evaluated subjects

with progressive supranuclear palsy. Furthermore, in most studies (53%) the number of PwPD was higher than the number of HC.

The distribution of the number of articles according to the number of participants is shown in Figure 5. In terms of the number of PwPD, the mean and median values are 44 and 33 subjects, respectively. However, it should be noted that there is a high variability, with a minimum of 5 and a maximum of more than 1000 PwPD. More than half of the articles (N = 14, 63%) presented a sample size of 1 to 50 patients, while 14% (N = 3) enrolled more than 1000 patients.
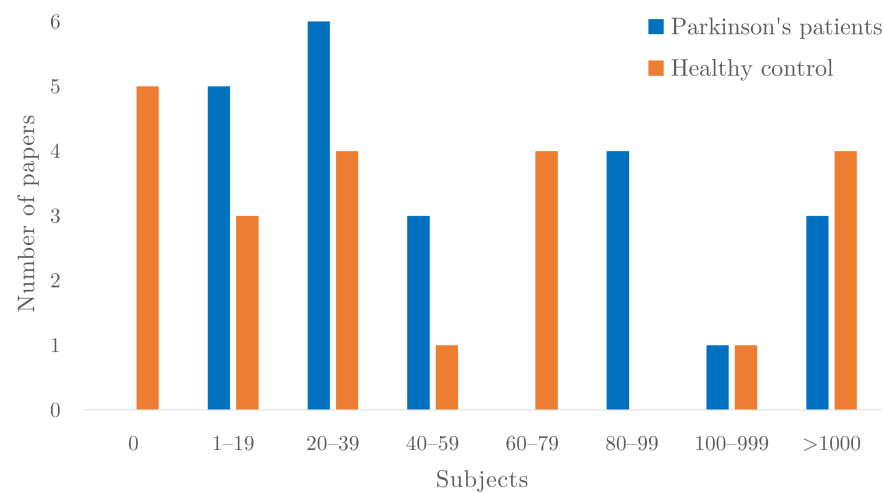
Figure 5. Distribution of the number of articles according to the number of participants.

This variability can also be observed with the distribution of HC, which presents mean and median values of 57 and 50 participants, respectively. In this case, 59% (N = 13) of the studies enrolled less than 50 controls while 18% (N = 4) use a sample size of more than 1000 controls.

Regarding the gender distribution (Figure 6), 65% PwPD were male while 35% were female, reflecting the actual gender distribution in the PD population. This difference increases in HC, where 71% of the participants were male and 29% were female.
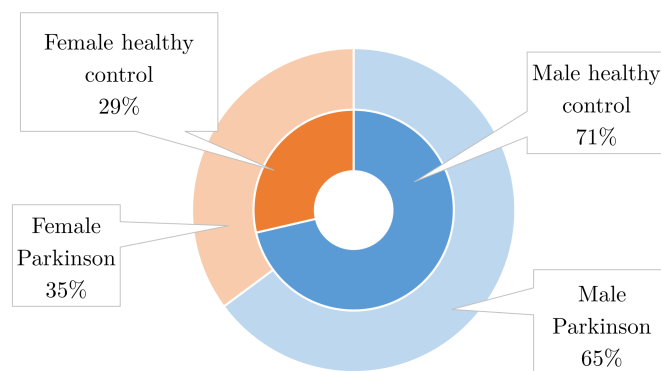

Figure 6. Gender distribution in the selected papers.

4.5. Device, Sensor and Body Location

Regarding the type of devices used, 20 studies (91%) used a single type of device for data recording, while [83] employed two different devices and [82] used three types of devices. In total, 25 devices were used in the 22 studies. As shown in Figure 7a, the most commonly used device is the IMU (N = 11 of the 25 devices, 44%) [21,66,67,71,73,74,77,78,81,82,85], followed by the smartphone (N = 7, 28%). [68,72,75,79,80,82,84], pressure insoles were used

in 3 cases (12%) [65,69,70], while other devices such as tablets were used in 4 studies (16%) [76,82,83].
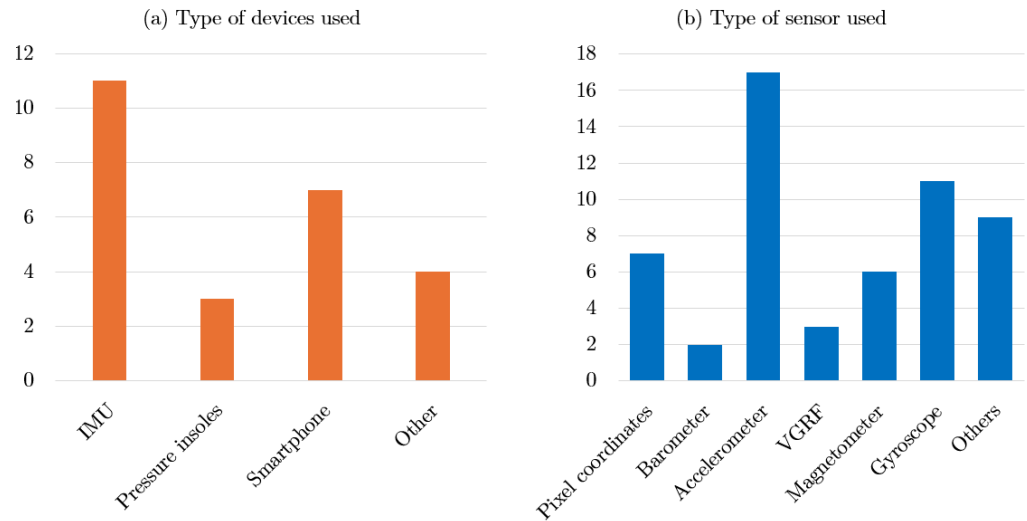

Figure 7. Distribution of types of (**a**) devices and (**b**) sensors used.

In terms of sensor types (Figure 7b), of the 55 total sensors used, the most used sensor is the triaxial accelerometer (N = 17, 31%), followed by the triaxial gyroscope (N = 11, 20%), the touch screen (N = 7, 13%), triaxial magnetometer (N = 6, 11%), VGRF sensors (N = 3, 5%), and barometer (N = 2, 4%). Other sensors (N = 9, 16%) include the goniometer, GPS or galvanic skin response.

Figure 8a presents the distribution of the number of devices used. The most common option involves a single device (N = 14, 64%). The remaining studies used multiple sensors. Specifically, 6 papers (27%) [21,66,71,74,83,85] used between 2 and 3 devices, while the 2 remaining papers (9%) [73,82] used equal or more than 8 devices. Figure 8b presents the number of sensors used. The most popular choice is to use 2 sensors (N = 8, 36%). A single sensor was used in 3 articles [67,76,77] (14%), 3 sensors in 6 articles [21,66,67,74,80,81] (27%), while 5 articles (23%) use more than 10 sensors.
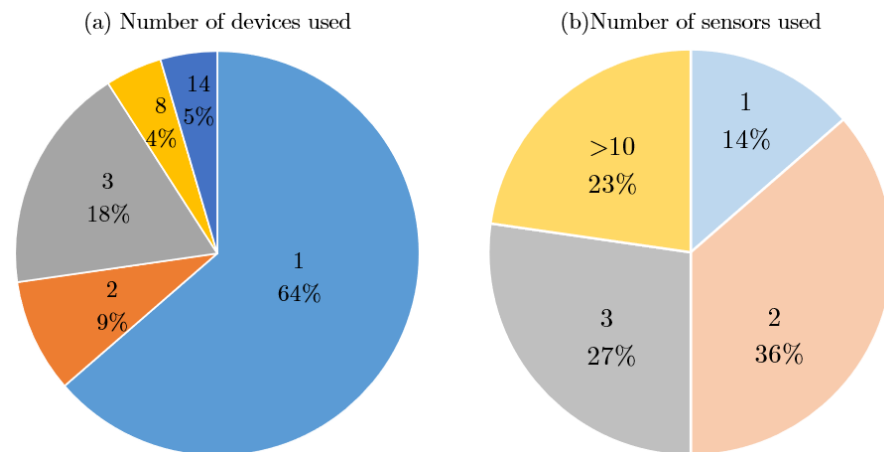

Figure 8. Number of (**a**) devices and (**b**) sensors used.

Figure 9 shows common positions of the sensor on the human body. The device was placed in front of the person (e.g., smartphone, tablet) or on the foot or in the lower back in 8 articles each (18%). In 5 cases (11%), the wearable device was placed in the pocket, in the hand in 4 papers (9%) and the wrist in 2 cases (5%). The rest of the studies positioned the sensors in other locations like the ankle or the waist. Half of the included articles used a single location, while the other half used more than one body position.
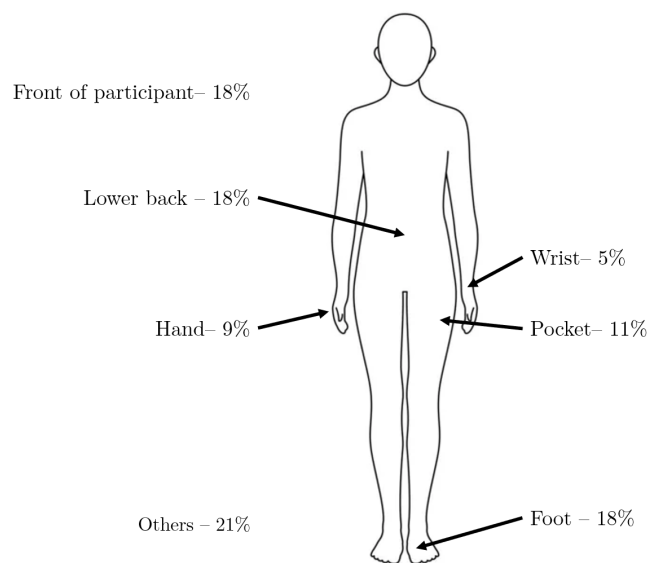
Figure 9. Distribution of the location of the sensor on the human body.

4.6. Aim

In terms of the study aims, the reviewed articles were categorized into the following primary areas: diagnosis (i.e., classification between PwPD and HC), monitoring, and prognosis, as outlined in [20]. Figure 10 presents a summary of the principal aims of the selected studies.
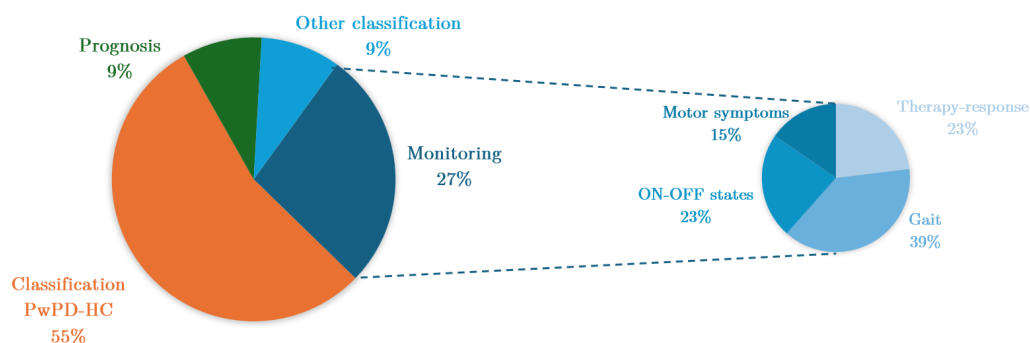

Figure 10. Principal aims of the selected studies.

Most of these studies focused on the classification between PwPD and HC (N = 12, 55%), followed by monitoring (N = 6, 27%), prognosis (N = 2, 9%) and other type of classification (N = 2, 9%). The six studies on PD monitoring included the assessment of gait impairment (N = 5, 39%), therapy response (N = 3, 23%), ON-OFF state (N = 3, 23%) estimation, and automatic detection of motor symptoms (bradykinesia, tremor, and postural instability) (N = 2, 15%). Regarding prognosis, the studies aimed to predict the risk of falls (number of falls) and the evolution of motor symptoms such as abnormalities in movement of the lower-limbs.

Figure 11 shows the distribution of the aims addressed in the selected studies.

Specifically, 59% (N = 13) of the studies addressed more than one aim, while the remaining 41% (N = 9) addressed a single primary area (i.e., diagnosis, monitoring, and prognosis). Three studies (14%) [72,76,82] combined ON-OFF states monitoring with PwPD and HC classification, while three studies (14%) [73,79,81] combined therapy response evaluation with PwPD/HC classification and prognosis. Two studies (9%) [80,85] combined motor symptoms monitoring with PwPD and HC and prognosis, respectively. Two studies (9%) [69,70] combined classification PwPD/HC with gait monitoring. Three studies (14%) combined other aims, such as classification of therapy/placebo subjects and therapy response monitoring [83], the classification of PD/REM sleep behavior disorder with the

classification of PwPD/HC [75] or the classification of motor condition was addressed, together with the assessment of the quality of life of PwPD [78].
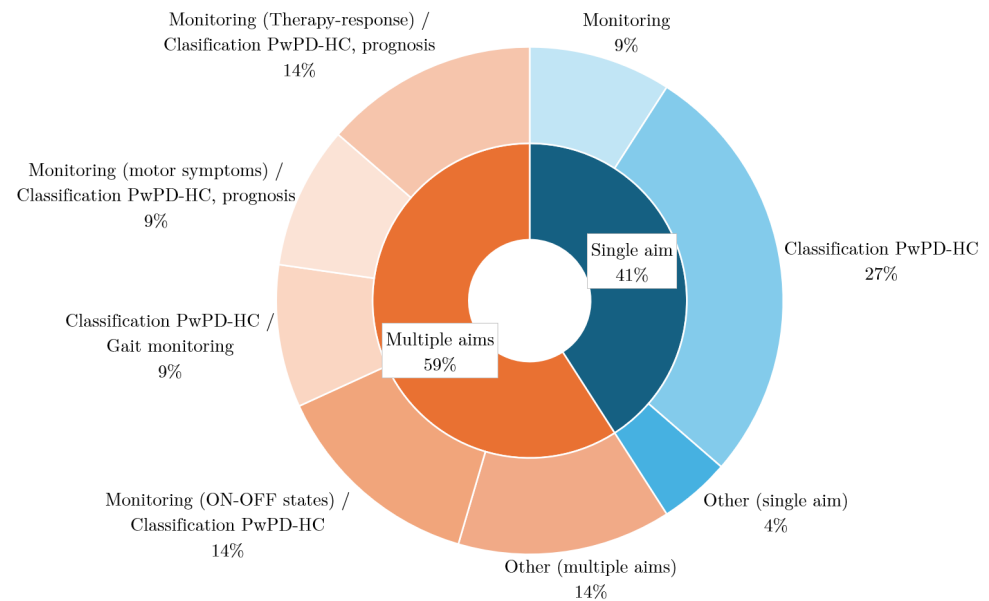

Figure 11. Distribution of the aims addressed in the selected studies.

4.7. Extraction Methods

DB-MS-PD were extracted by different methods (Figure 12). 7 studies (32%) assessed the potential of single features as biomarkers. Most studies (N = 14, 64%) combined multiple features and used ML models to provide a robust result. Logistic regression (N = 5, 33%) and random forest (3, 20%) was the most common between the ML models. Finally, only one study used raw data as input for a DL model (N = 1, 4%).
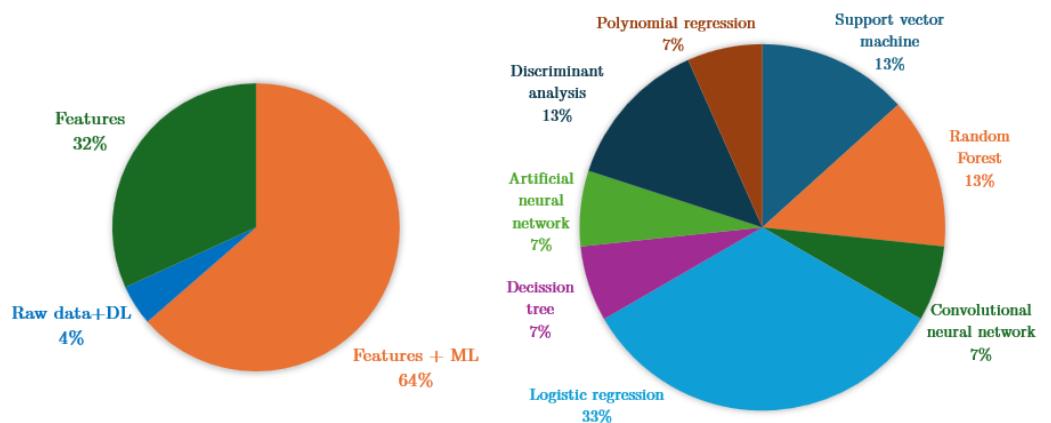

Figure 12. Biomarker extraction methods. ML and DL models used.

4.8. Endpoints

In 11 studies (50%), different motor tasks were evaluated. The distribution of the most relevant tasks in studies addressing multiple tasks is shown in Figure 13.
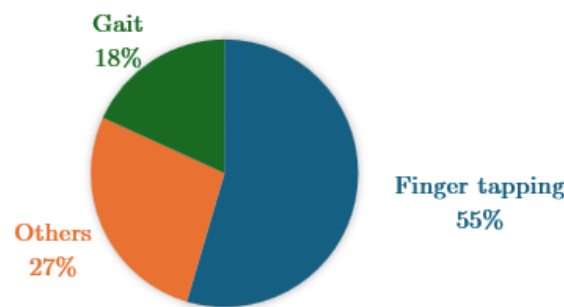
Figure 13. Most relevant tasks in studies addressing multiple tasks.

In 7 studies (32%), finger tapping was associated to different tasks, including balance, gait, rest tremor, postural tremor, pronation-supination, leg agility and reaction time. In 6 of these studies, finger tapping performed better than the other tasks for PD diagnosis, ON/OFF detection and PD progression. Gait gave better results than other activities in 2 cases (18%) for PD detection and progression. The 3 remaining articles (27%) studies another tasks. In a single study [75], resting and postural tremor provided better results than the other tasks (including finger tapping) for the diagnosis of PD. A single study [73] (which included neither finger tapping nor gait) found that toe tapping and TUG tasks give better results than leg agility and hand movements for PD progression estimation. Finally, a single study [81] concluded that resting tremor and bradykinesia achieve agreement with clinical assessment of symptom severity and treatment-related changes in motor states.

In half of the studies, DB-MS-PD were extracted in the laboratory/clinic, whereas in the other half, remote home monitoring was used. In the laboratory-based studies, mostly supervised active tasks (participants were asked to perform different tasks) were used, whereas in one study both active and passive monitoring were used. When data were recorded at home (11 studies), active tasks were used in 6 and passive monitoring (extraction of DB-MS-PD during daily life) in the other 5.

Focusing on the tapping task (Figure 14), of the 7 studies in which tapping was the most relevant task, most found that, in 5 cases each (71% of the cases), statistical metrics calculated on the intervals between taps (i.e., mean value, variability, percentiles) or on the spatial accuracy (i.e., two-dimensional distance between tap position and target point) were the most significant DB-MS-PD. This is followed by total taps (N = 4, 57%) and the number of on-target taps (N = 2, 29%). Other characteristics include tap duration and fatigue. The latter was calculated as the difference in tap speed between the first and last tap. These features were highly relevant for the diagnosis of PD, the estimation of motor fluctuations, the assessment of disease progression and the prediction of clinical motor scores.

Regarding to the gait task (Figure 15), of the 11 studies where gait was the primary task, most (N = 7, 63%) found gait speed and its variability to be the one of most significant DB-MS-PD in both supervised and unsupervised contexts. This is followed by average step length and their variability (N = 4, 36%), stride length (N = 3, 27%), step time (N = 3, 27%), swing time (N = 2, 18%), stance time (N = 2, 18%), and cadence (N = 2, 18%). Other characteristics include total power, center of pressure, force variations, power spectrals and entropy. These features were found to be highly relevant for the diagnosis of PD, the estimation of motor fluctuations and the assessment of disease progression.
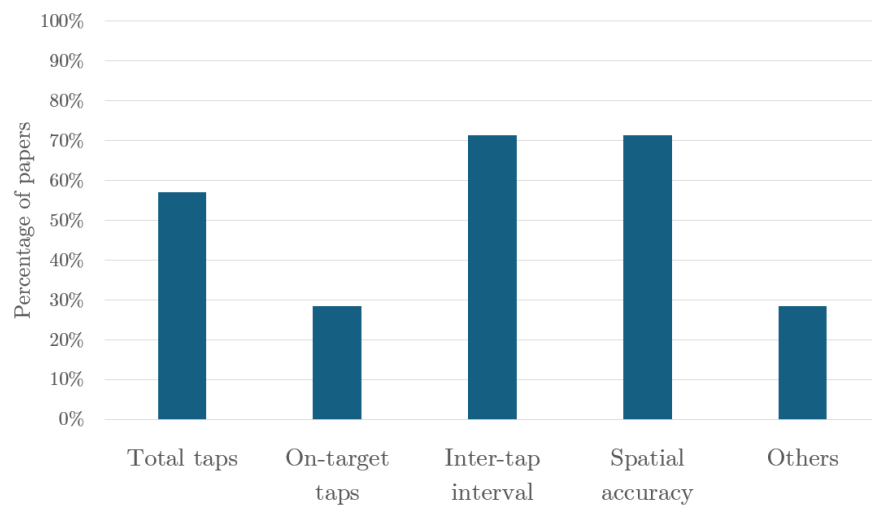
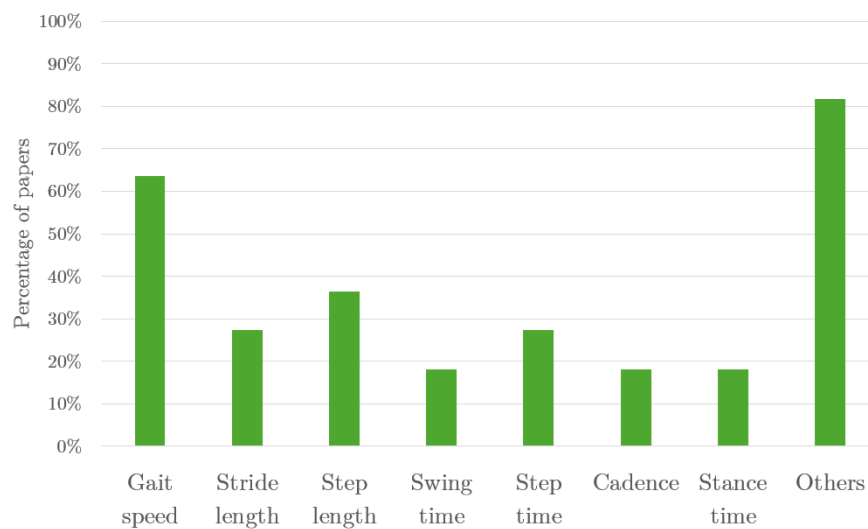**Figure 14.** Most relevant features in the finger tapping task.

Figure 15. Most relevant features in the gait task.

Overall, the supervised and unsupervised finger tapping task (assessed through position and acceleration analysis on a smartphone or tablet) and the supervised and unsupervised gait task (assessed through inertial or force sensors) represented the most explored tasks. When comparing the results obtained from different studies, it is evident that these two tasks provide similar performance for PD diagnosis and treatment response assessment, as shown in Table 2. On the other hand, gait was also explored for severity estimation, while the finger tapping task for predicting the clinical motor score. The classification and regression results are reported in the Table 2, where the results of the different studies are summarized using the performance range.

Table 2. Performance comparison of tasks for disease diagnosis, prognosis, and monitoring. Acc: accuracy; r: Pearson's correlation coefficient; MAE: mean absolute value.

Task	Diagnosis	Treatment	Severity	UPDRS–III
Finger tapping	AUC 0.74–0.95	Acc 0.75–0.84	-	$r = 0.51$–0.69, MAE = 8
Gait	AUC 0.76–0.98	AUC 0.82	AUC 0.85–0.98	-

It is worth noting that all finger tapping tasks represent active tests, in which subjects have to actively participate in the activity. On the other hand, tapping on touch screens

represents an easy and effective activity that can be performed at home. Very good performance was obtained both in supervised laboratory settings and in unsupervised remote settings (i.e., at home), demonstrating that finger tapping represents an easy and effective task for the extraction of DB-MS-PD. With regard to gait, data recorded in supervised laboratory environments performed better for PD diagnosis (accuracy 0.94–0.98) and severity estimation (accuracy 0.85–0.98), compared to unsupervised data recorded continuously in the home environment, with an area under the curve (AUC) between 0.76–0.95.

Analyzing tremor, a sensitivity of 0.85 and a specificity of 0.88 were obtained in distinguishing PwPD from HC; A sensitivity of 0.88 and specificity of 0.90 were obtained for the classification of PD versus idiopatic sleep behavior disorder (iRBD) [75]. The performance of tremor detection achieved an accuracy of 0.83 and a correlation of r = 0.97 ($p < 0.001$) with the UPDRS tremor score (item 3.18). A correlation of r = 0.67 ($p < 0.001$, RMSE = 2) was obtained between the sensor measurements (amplitude of hand movement) and the clinical bradykinesia score (sum of items 3.4–3.6) [81]. For toe tapping, a correlation of r = 0.74 ($p < 0.001$) was found in predicting motor progression at 30 months [73]. Finally, the TUG test analysis provided an r = 0.55 and an RMSE of 0.33 in predicting the number of falls in PD [85].

5. Discussion

This article provides an updated review of the existing literature on digital biomarkers for motor symptoms assessment in PD. This disorder represents the second most common neurodegenerative disease in the world [3], and new digital technologies promise to significantly support the approach to the diagnosis, prognosis and monitoring of PD.

In this review, a total of 22 articles were selected and thoroughly analyzed to provide an summary of current DB-MS-PD in PD. As the term biomarker is very broad and the use of wearable devices can be extended to different locations, a large heterogeneity was found among the studies examined. Nevertheless, it was possible to identify trends and patterns in the definition of the experimental protocols, the number of participants, the number and type of devices and sensors used, the location of the devices on the human body, the objectives pursued and the types of DB-MS-PD proposed.

Overall, the results indicate that wearable devices have the potential to be used to define DB-MS-PD. These can provide measurable and objective evaluations in clinical or hospital settings. Furthermore, the indicators collected through wearable devices could contribute to the development of remote and continuous patient monitoring systems to follow the evolution of different symptoms, especially in unsupervised settings such as patient's home.

The vast majority of studies included in this review employed commercial IMUs, mostly embedding at least accelerometer and gyroscope, or the touchscreen of smartphones and tablets. These sensors allowed assessment and monitoring of a wide spectrum of motor symptoms in PD for diagnosis, prognosis, and monitoring purposes. Importantly, more than half studies used a single device embedding multiple sensors. Among these, IMUs seem to represent the most promising solution (11 of 22 studies). IMUs are widespread and used on various applications due to their versatility, accuracy, and relatively compact form factor, thus can be used for data collection and analysis without requiring additional dedicated hardware. They feature built-in sensors like an accelerometer, gyroscope, and magnetometer for comprehensive motor symptom evaluation. The results are promising, with high diagnostic accuracy in discriminating PwPD from HC and high correlation with clinical scores. However, smartphones (7 of 22 studies) are used on a daily basis, and are equipped with a large number of built-in sensors such as accelerometer, gyroscope, magnetometer, touchscreen, camera, and microphone. Smartphones were mostly used in the execution of scripted active tasks (e.g., finger tapping, memory, walking). None of the studies evaluated the potential of smartphones for long term, continuous monitoring in unsupervised environments. While representing technology which subjects are familiar with, the dimension and weight of smartphones are not comparable to the small and tiny

inertial modules that can be even embedded in smart-clothes or devices. Passive continuous monitoring through IMUs or smartphones (i.e., in the pocket) was not addressed and needs further evaluation.

Gait and finger tapping represent the most frequent activities in the studies analyzed. Other tasks and activities such as TUG test, balance task, and reaction times were poorly addressed and thus do not allow to obtain robust conclusions. As previously discussed, gait and finger tapping provided similar results in terms of discrimination capability of PwPD from HC and for therapy condition assessment. These tasks were analysed both in supervised clinical settings and in unsupervised home environments. Despite similar results, gait has the potential to provide a continuous passive evaluation of disease severity, the presence of motor symptoms, and the effect of therapy. On the other hand, finger tapping represents an active task and should be performed several times a day to accurately estimate motor fluctuations.

In general, this review emphasizes the potential of DB-MS-PD for PD diagnosis, prognosis, and monitoring. Data can be collected in unsupervised environments for long periods of time using widespread commercial devices such as smartphones. Individual features or a combination of multiple features and ML models can be used to detect symptoms, predict severity, and evaluate therapy response. The evidences are convincing, as large sample sizes, correct validation procedures, and robust methods were used. Ultimately, this review article is intended to provide the reader with a comprehensive set of information that demonstrates the potential of DB-MS-PD in PD, critically discussing current limitations and providing recommendations for future work.

5.1. Challenges

The careful analysis of the articles included in the review highlights several challenges, which are reported below. The identification of these challenges can help design and conduct future studies related to digital biomarkers for motor symptoms assessment in PwPD.

Participant Selection: Most studies focused on the diagnosis of PD, alone or in combination with other objectives (e.g., estimation of disease severity, treatment conditions). The proper evaluation of digital tools for computerized early diagnosis should be conducted on newly diagnosed PwPD. However, some studies have not reported any measurement of duration and/or severity of disease, which makes it complicated to assess the potential of the proposed solution. When reported, the average duration of the disease was in the range of 3.5–13 years and the H&Y stage was mostly equal to or greater than 2. This means that the recruited PwPD were mostly at an advanced stage of the disease, when motor symptoms were fully visible. This raises doubts about the usefulness of the solutions developed and their real potential for early diagnosis. It is essential to define appropriate standardized criteria that provide guidelines for the recruitment process and the reporting of participant characteristics. For example, in order to enable a proper evaluation of the PD, at least disease duration, H&Y and MDS-UPDRS scores must be reported. Strict inclusion criteria with regard to disease duration (e.g., less than 2) and H&Y stage (e.g., less than or equal to 1) must be adopted when evaluating digital systems for early diagnosis.

Additionally, regarding the number of participants, the sample size significantly varies. Two clear trends are observed, 14 studies with a number of PwPD between 5 and 42, and 8 works with more than 80 participants (3 studies with more than 1000). Furthermore, studies tend to have a higher number of HC than PwPD (of the articles that are clearly specified, 77% of the total participants are HC). It would be interesting for a correct definition of DB-MS-PD that these are defined from a significant number of PwPD registries, close to 50% of the total. In addition, another factor that should be taken into account is the gender balance of the participants, as, for HC and PwPD, the percentage of male participants is high compared to female participants (71% HC, 65% PD). It could be interesting to try to define a minimum number of patients or healthy participants and age or gender distribution for a biomarker to be proposed as such, as it would allow us to know that it has been tested with a relevant number of people.

Reporting of methods: Related to the activities and tasks performed, in some cases, the activity carried out is very detailed, while in others it is described in a general way. For example, distance travelled, gait trajectories to be followed, and speeds are indicated in a limited number of studies. Similarly, the addressed tasks are sometimes referred as MDS-UPDRS tasks, without precise indication of the specific UPDRS item. A similar phenomenon is observed in the description of the positioning of wearable devices on the body. Some studies explicitly describe the position of the device through photographs or diagrams, while others merely refer to general anatomical regions, such as the upper limb or lower limb. A possible option would be to use a diagram in the shape of a human body to identify the location of each of the devices on the body. Several articles [67,73,74,77] have already used this approach and it is an effective way to identify the positions discussed in detail.

Tasks performed: Gait and finger tapping have been found to be the most studied activities for monitoring and understanding the evolution of PD severity and motor symptoms. In the reviewed studies, most PD participants were at the same stage of the disease and performing the movements in a fairly similar way, which allows for a classification of each stage easily. Gait is the task performed in the studied papers that is the easiest to apply in free-living conditions. However, the use of other activities such as reaction times, pronation-supination movements or TUG test can help to get a broader and more detailed overview of the disease for each PwPD. In addition, analysis of these less studied activities, or new ones yet to be discovered, could provide valuable information on more subtle aspects of disease progression. For example, reaction times may reveal alterations in processing speed in patients PwPD, while pronation-supination movements could provide clues about coordination and fine motor control. Therefore, all possible activities that have significant results for the follow-up and monitoring of PwPD should be further studied and promoted.

Monitoring settings: Regarding the evaluation of therapy condition (ON/OFF), half of the studies were conducted in clinical, supervised settings. This raises questions about the applicability of the developed digital tools to unsupervised remote environments. When performing at-home monitoring, passive monitoring during activities of daily living lasted 1 h to 1 day [71,82]. This represents rather a short period of time for fully evaluating the performance of automatic motor condition assessment. On the other hand, long-term monitoring of 2 weeks to 6 months [79,80] provide robust performance estimate. However, this was achieved using active tests, involving patients to complete a scripted set of tasks. Ideally, passive monitoring over a long period of time would be desirable, so the evolution of biomarkers over time can be monitored without forcing patients to engage in sustained activities.

Report of the results: A comprehensive reporting of performance metrics is essential to fully evaluate the potential of the prediction system and to fairly compare similar studies. Focusing on diagnosis, almost half of the studies reported classification performance in terms of AUC. This allows the diagnostic ability of DB-MS-PD to be assessed regardless of the selected classification threshold and serves as a summary of the overall model performance. However, AUC was not reported in the other half of the studies, where accuracy or the combination of sensitivity and specificity were preferred. With regard to regression metrics, the performance evaluation is rather heterogeneous across studies. A substantial number of studies reported the correlation coefficient alone or a single error measure (i.e., MAE, RMSE). Furthermore, test-retest reliability was assessed in a single study using the intra-class correlation coefficient. Finally, inter-rater variability was not studied. Establishing common metrics across all proposals would facilitate a meaningful comparison of their respective performances.

Overall, the heterogeneity of the performance metrics hinders a fair comparison with similar work and does not allow a comprehensive evaluation of performance. Again, some guidelines are needed to suggest the minimum set of performance metrics to be reported. These may include at least the receiver operating characteristic-ROC curve and

AUC value in classification problems and the correlation coefficient, the associated p-value and a measure of error (i.e., MAE, RMSE or MSE) for regression tasks. In the latter case, test-retest reliability and inter-rater variability are essential to assess the consistency of multiple measurements and multiple clinical raters, respectively [88,89].

Proposed Biomarkers: Above all, the definition of biomarkers is heterogeneous in the evaluated studies, with no common report structure. Some studies used single features as DB-MS-PD and assessed their classification or regression performance. Other studies extracted multiple features from the data and combined them using an ML model. In this case, if adequately described, the aggregation of features and the ML model can represent a DB-MS-PD. In most of these studies, the contribution of each feature to final performance was assessed using regression metrics or statistical tests between groups of subjects, conditions, or treatments. This is useful as it is possible to identify the most important feature for the specific purpose. Therefore, it would be better if the proposal studies use a standardized reporting structure or presentation of results that reflects certain key aspects and allows other authors to easily identify what the proposal is.

The reliability and accuracy of a biomarker depend on their development context, such as patient characteristics (e.g., Hoehn and Yahr stage, MDS-UPDRS), tasks performed, exact placement of sensors and metrics calculated. To assess whether biomarkers have been calculated correctly, appropriate reporting of key aspects is essential. To ensure reliability, biomarkers must be reproduced in multiple studies around the world, so reproducibility is very important, because if a biomarker is consistently reproduced and applied to many patients, it can be considered reliable and accurate for the study of the disease.

Reproducibility is a key aspect to allow for repeatable experiments ready-to-use solutions. However, methods and implementation details were not always exhaustively reported. Relevant information such as sensor characteristics and specific positioning on the body, signal pre-processing and conditioning steps, and ML models' parameters should be carefully described. In addition, most of the evaluated studies used proprietary datasets that were not made publicly available. Although the diversity between the different datasets provides new perspectives and enrich the current body of knowledge, proprietary datasets hinder reproducibility of the experiments and results, and limits the advance in scientific research.

The recent significant technological advances in wearable technology allows for a wide range of data measurement modality. This is obviously a great advantage, as a large spectrum of physical and physiological parameters can be easily extracted from on-body sensors. However, it is worth considering that the number of devices, their wearability and comfort are of utmost importance when designing systems for long-term unsupervised monitoring. PwPD often suffer from non-motor symptoms such as sleep disorders, anxiety, and depression. Thus, cognitive load and patient compliance are essential aspects to consider when developing digital solutions.

5.2. Limitations of This Study

This review has some limitations. First, despite extensive search on several digital journals, only 22 studies were finally included and fully evaluated. Considering the heterogeneity in the sample size, objective, methods and results, a direct and comprehensive comparison of similar studies is not always possible. This is particularly evident for studies focusing on activities different from gait and finger tapping, for which it is not possible to provide robust conclusions. In addition, the distribution of study objectives is very unbalanced, with most studies focusing of diagnosis and therapy condition estimation, while the severity of motor symptoms and overall motor condition were poorly addressed. Finally, this review focused specifically on DB-MS-PD. This led to consider motor symptoms such as gait impairment, postural instability, tremor, bradykinesia, and dyskinesia. Other motor symptoms such as rigidity, fatigue, and hypomimia were not addressed in the reviewed studies.

Another limitation that may be encountered is the replicability, as each study uses a different wearable device with its own sampling frequency and filtering techniques (not all of which are indicated), using a wide variety of devices.

A high percentage of the assessed studies (84.8%, n = 168 of the 198 articles screened) are from 2020 or later. Furthermore, the observable trend in recent years has been that more articles are being produced each year, so it can be deduced that the definition of biomarkers using wearable technology is a subject on which numerous research projects are being carried out and that new results and DB-MS-PD will be published in the coming years. Therefore, this work can be a reference to know the current trends and future literature reviews on this topic can expand the information shown and know in further detail the way to the definition of DB-MS-PD.

In addition, 6 of the 22 articles use public databases, therefore, the variability in the results is conditioned to a certain extent, since the results of these databases are similar. It would be interesting to extend the results obtained by using other different databases or by proposing new activities or measures.

6. Conclusions

This article explores the use of wearable sensors to extract digital biomarkers to assess motor symptoms of Parkinson's disease (PD). Findings show these digital biomarkers can accurately track severity symptoms, monitor the response to treatment, and distinguish between patients with PD (PwPD) and healthy controls. Although these tools can provide objective and continuous information for improved monitoring, findings highlight the importance of standardized data collection and analysis methods to ensure reliable integration into clinical practice. Among the reviewed studies, a total of 25 different devices (20 out of 22 articles used a single type of device) were used for developing digital biomarkers for motor symptoms of PD (DB-MS-PD), with the inertial measurement units (IMU) being the most commonly used, followed by smartphones. In terms of sensor types, the most frequently used sensors were triaxial accelerometers and gyroscopes.

DB-MS-PD were extracted primarily through three approaches: combining multiple features with machine learning (ML), deep learning (DL) algorithms, or evaluating single features using statistical methods. To extract these digital biomarkers, different motor tasks were assessed, with finger tapping and gait being the most frequent motor activities.

Finger tapping, particularly in terms of tap intervals, spatial accuracy, and total taps, consistently stands out as a significant indicator for the diagnosis and monitoring of PD progression. Performance metrics indicate that finger tapping tasks, which involve active participation, are effective for both laboratory and home environments, with accuracies ranging from 0.75 to 0.84 and area under the curve (AUC) values between 0.74 and 0.95.

Gait analysis, on the other hand, using measurements such as speed, variability, and stride length was also important to assess motor function in both supervised and unsupervised contexts. Furthermore, the results indicate that gait tasks are similarly effective in the diagnosis of PD and in evaluating treatment responses. AUC values for gait analysis range from 0.76 to 0.98, with the highest accuracy (0.94 to 0.98) observed in controlled laboratory settings.

The use of wearable devices can provide many advantages for traditional clinical monitoring and assessment, as they can generate a high volume of data in a relatively short time, provide objective opinions on the patient's actual condition, or be able to assess or account for aspects that traditionally have not been possible, such as falls or freezing. In addition, they introduce an unexploited aspect, such as the possibility of monitoring a patient in a home environment without the pressure of feeling observed by the neurologist.

However, in order to develop remote patient monitoring systems, several issues identified in the review need to be addressed. One of the most relevant challenges is the standardization of data collection, analysis, processing, evaluation, and reporting of DB-MS-PD in order to facilitate comparability and replicability of results. In the absence of general guidelines for experimental development, each trial can be conducted in a different

way and managed differently. This idea of standardization has been pursued in other fields such as neuroimaging in Alzheimer [90] or in the use of public tools for database publication [91].

The review highlighted multiple issues and constraints within the included studies, offering suggestions for future research to overcome these limitations and improve PD assessment. Rapid progress in sensing and data analysis technologies is expected to significantly expedite the integration of wearable devices in this field. These DB-MS-PD can be used to obtain objective and measurable information on the status of certain disease symptoms and could be used in continuous patient monitoring systems.

Author Contributions: Conceptualization: C.P.F., L.S. and I.P.; Data curation: C.P.F.; Formal analysis: C.P.F., L.S., L.B. and I.P.; Funding acquisition: C.A., J.M.L., G.d.A. and I.P.; Investigation: C.P.F., L.S., L.B. and I.P.; Methodology: C.P.F.; Project administration: I.P.; Resources: C.A., J.M.L., G.d.A. and I.P.; Supervision: G.O. and I.P.; Validation: C.P.F.; Visualization: C.P.F., L.S., L.B. and I.P.; Writing—original draft: C.P.F., L.S., L.B. and I.P.; Writing—review & editing: C.P.F., L.S., L.B., G.O., C.A., J.M.L., G.d.A. and I.P. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Abbreviations

The following abbreviations are used in this manuscript:

DB-MS-PD	Digital Biomarkers for Motor Symptoms of Parkinson's Disease
PD	Parkinson's Disease
PwPD	Patients with Parkinson's Disease
HC	Healthy Control
ML	Machine Learning
DL	Deep Learning
AUC	Area Under the Curve
MAE	Mean Absolute Value
IMU	Inertial Measurement Unit
VGRF	Vertical Ground Reaction Force
TUG	Timed Up and Go
MDS-UPDRS	Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale

References

1. Goetz, C.G. The history of Parkinson's disease: Early clinical descriptions and neurological therapies. *Cold Spring Harb. Perspect. Med.* **2011**, *1*, a008862. [CrossRef] [PubMed]
2. Armstrong, M.J.; Okun, M.S. Diagnosis and treatment of Parkinson disease: A review. *JAMA* **2020**, *323*, 548. [CrossRef] [PubMed]
3. World Health Organization. Parkinson Disease. 2022. Available online: https://www.who.int/news-room/fact-sheets/detail/parkinson-disease (accessed on 18 March 2024).
4. Dorsey, E.R.; Sherer, T.; Okun, M.S.; Bloem, B.R. The emerging evidence of the Parkinson pandemic. *J. Parkinsons Dis.* **2018**, *8*, S3–S8. [CrossRef] [PubMed]
5. Sveinbjornsdottir, S. The clinical symptoms of Parkinson's disease. *J. Neurochem.* **2016**, *139*, 318–324. [CrossRef] [PubMed]
6. Chaudhuri, K.R.; Healy, D.G.; Schapira, A.H. Non-motor symptoms of Parkinson's disease: Diagnosis and management. *Lancet Neurol.* **2006**, *5*, 235–245. [CrossRef]

7.   Tolosa, E.; Wenning, G.; Poewe, W. The diagnosis of Parkinson's disease. *Lancet Neurol.* **2006**, *5*, 75–86. [CrossRef]

8.   Xia, R.; Mao, Z.H. Progression of motor symptoms in Parkinson's disease. *Neurosci. Bull.* **2012**, *28*, 39–48. [CrossRef] [PubMed]

9.   Zhao, N.; Yang, Y.; Zhang, L.; Zhang, Q.; Balbuena, L.; Ungvari, G.S.; Zang, Y.; Xiang, Y. Quality of life in Parkinson's disease: A systematic review and meta-analysis of comparative studies. *CNS Neurosci. Ther.* **2020**, *27*, 270–279. [CrossRef] [PubMed]

10.   Jankovic, J. Parkinson's disease: Clinical features and diagnosis. *J. Neurol. Neurosurg. Psychiatry* **2008**, *79*, 368–376. [CrossRef]

11.   Jankovic, J. The evolution of diagnosis in early Parkinson disease. *Arch. Neurol.* **2000**, *57*, 369. [CrossRef] [PubMed]

12.   Singh, N.; Pillay, V.; Choonara, Y.E. Advances in the treatment of Parkinson's disease. *Prog. Neurobiol.* **2007**, *81*, 29–44. [CrossRef]

13.   The Parkinson Study Group. Levodopa and the progression of Parkinson's disease. *N. Engl. J. Med.* **2004**, *351*, 2498–2508. [CrossRef] [PubMed]

14.   Jankovic, J. Motor fluctuations and dyskinesias in Parkinson's disease: Clinical manifestations. *Mov. Disord.* **2005**, *20*, S11–S16. [CrossRef]

15.   Goetz, C.G.; Tilley, B.C.; Shaftman, S.R.; Stebbins, G.T.; Fahn, S.; Martinez-Martin, P.; Poewe, W.; Sampaio, C.; Stern, M.B.; Dodel, R.; et al. Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov. Disord.* **2008**, *23*, 2129–2170. [CrossRef]

16.   Schrag, A. How valid is the clinical diagnosis of Parkinson's disease in the community? *J. Neurol. Neurosurg. Psychiatry* **2002**, *73*, 529–534. [CrossRef] [PubMed]

17.   Albanese, A. Standard strategies for diagnosis and treatment of patients with newly diagnosed Parkinson's disease. *Neurol. Clin. Pract.* **2013**, *3*, 476–477. [CrossRef] [PubMed]

18.   Davidson, M.B.; McGhee, D.J.; Counsell, C.E. Comparison of patient rated treatment response with measured improvement in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **2012**, *83*, 1001–1005. [CrossRef] [PubMed]

19.   Luis-Martínez, R.; Monje, M.H.G.; Antonini, A.; Sánchez-Ferro, Á.; Mestre, T.A. Technology-enabled care: Integrating multidisciplinary care in Parkinson's disease through digital technology. *Front. Neurol.* **2020**, *11*, 575975. [CrossRef] [PubMed]

20.   Sigcha, L.; Borzì, L.; Amato, F.; Rechichi, I.; Ramos-Romero, C.; Cárdenas, A.; Gascó, L.; Olmo, G. Deep learning and wearable sensors for the diagnosis and monitoring of Parkinson's disease: A systematic review. *Expert Syst. Appl.* **2023**, *229*, 120541. [CrossRef]

21.   Shah, V.V.; McNames, J.; Mancini, M.; Carlson-Kuhta, P.; Nutt, J.G.; El-Gohary, M.; Lapidus, J.A.; Horak, F.B.; Curtze, C. Digital biomarkers of mobility in Parkinson's disease during daily living. *J. Parkinsons Dis.* **2020**, *10*, 1099–1111. [CrossRef] [PubMed]

22.   Fröhlich, H.; Bontridder, N.; Petrovska-Delacréta, D.; Glaab, E.; Kluge, F.; Yacoubi, M.E.; Marín Valero, M.; Corvol, J.C.; Eskofier, B.; Van Gyseghem, J.M.; et al. Leveraging the potential of digital technology for better individualized treatment of Parkinson's disease. *Front. Neurol.* **2022**, *13*, 788427. [CrossRef] [PubMed]

23.   Chudzik, A.; Śledzianowski, A.; Przybyszewski, A.W. Machine learning and digital biomarkers can detect early stages of neurodegenerative diseases. *Sensors* **2024**, *24*, 1572. [CrossRef] [PubMed]

24.   Bonato, P. Wearable sensors and systems. *IEEE Eng. Med. Biol. Mag.* **2010**, *29*, 25–36. [CrossRef] [PubMed]

25.   Ometov, A.; Shubina, V.; Klus, L.; Skibińska, J.; Saafi, S.; Pascacio, P.; Flueratoru, L.; Gaibor, D.Q.; Chukhno, N.; Chukhno, O.; et al. A survey on wearable technology: History, state-of-the-art and current challenges. *Comput. Netw.* **2021**, *193*, 108074. [CrossRef]

26.   Iqbal, S.M.A.; Mahgoub, I.; Du, E.; Leavitt, M.A.; Asghar, W. Advances in healthcare wearable devices. *npj Flex. Electron.* **2021**, *5*, 9. [CrossRef]

27.   Dunn, J.; Runge, R.; Snyder, M. Wearables and the medical revolution. *Pers. Med.* **2018**, *15*, 429–448. [CrossRef]

28.   Martin, T.; Healey, J. 2006's wearable computing advances and fashions. *IEEE Pervasive Comput.* **2007**, *6*, 14–16. [CrossRef]

29.   Jiang, W.; Yin, Z. Human activity recognition using wearable sensors by deep convolutional neural networks. In Proceedings of the 23rd ACM International Conference on Multimedia, MM '15, Brisbane, Australia, 26–30 October 2015; ACM: New York, NY, USA, 2015. [CrossRef]

30.   Borzì, L.; Sigcha, L.; Olmo, G. Context recognition algorithms for energy-efficient freezing-of-gait detection in Parkinsons disease. *Sensors* **2023**, *23*, 4426. [CrossRef]

31.   Li, B.; Yao, Z.; Wang, J.; Wang, S.; Yang, X.; Sun, Y. Improved deep learning technique to detect freezing of gait in Parkinson's disease based on wearable sensors. *Electronics* **2020**, *9*, 1919. [CrossRef]

32.   Yao, S.; Swetha, P.; Zhu, Y. Nanomaterial-enabled wearable sensors for healthcare. *Adv. Healthc. Mater.* **2017**, *7*, 1700889. [CrossRef] [PubMed]

33.   Ates, H.C.; Nguyen, P.Q.; Gonzalez-Macia, L.; Morales-Narváez, E.; Güder, F.; Collins, J.J.; Dincer, C. End-to-end design of wearable sensors. *Nat. Rev. Mater.* **2022**, *7*, 887–907. [CrossRef] [PubMed]

34.   Nasiri, S.; Khosravani, M.R. Progress and challenges in fabrication of wearable sensors for health monitoring. *Sens. Actuators A Phys.* **2020**, *312*, 112105. [CrossRef]

35.   Stefana, E.; Marciano, F.; Rossi, D.; Cocca, P.; Tomasoni, G. Wearable devices for ergonomics: A systematic literature review. *Sensors* **2021**, *21*, 777. [CrossRef] [PubMed]

36.   Son, D.; Lee, J.; Qiao, S.; Ghaffari, R.; Kim, J.; Lee, J.E.; Song, C.; Kim, S.J.; Lee, D.J.; Jun, S.W.; et al. Multifunctional wearable devices for diagnosis and therapy of movement disorders. *Nat. Nanotechnol.* **2014**, *9*, 397–404. [CrossRef] [PubMed]

37.   Del Din, S.; Godfrey, A.; Mazzà, C.; Lord, S.; Rochester, L. Free-living monitoring of Parkinson's disease: Lessons from the field. *Mov. Disord.* **2016**, *31*, 1293–1313. [CrossRef] [PubMed]

38. Wen, D.; Zhang, X.; Liu, X.; Lei, J. Evaluating the consistency of current mainstream wearable devices in health monitoring: A comparison under free-living conditions. *J. Med. Internet Res.* **2017**, *19*, e68. [CrossRef]

39. Guk, K.; Han, G.; Lim, J.; Jeong, K.; Kang, T.; Lim, E.K.; Jung, J. Evolution of wearable devices with real-time disease monitoring for personalized healthcare. *Nanomaterials* **2019**, *9*, 813. [CrossRef]

40. Erdmier, C.; Hatcher, J.; Lee, M. Wearable device implications in the healthcare industry. *J. Med. Eng. Technol.* **2016**, *40*, 141–148. [CrossRef] [PubMed]

41. Patel, M.S.; Asch, D.A.; Volpp, K.G. Wearable devices as facilitators, not drivers, of health behavior change. *JAMA* **2015**, *313*, 459. [CrossRef]

42. Aroganam, G.; Manivannan, N.; Harrison, D. Review on wearable technology sensors used in consumer sport applications. *Sensors* **2019**, *19*, 1983. [CrossRef] [PubMed]

43. Aşuroğlu, T.; Oğul, H. A deep learning approach for Parkinson's disease severity assessment. *Health Technol.* **2022**, *12*, 943–953. [CrossRef]

44. Sigcha, L.; Polvorinos-Fernández, C.; Costa, N.; Costa, S.; Arezes, P.; Gago, M.; Lee, C.; López, J.M.; de Arcas, G.; Pavón, I. Monipar: Movement data collection tool to monitor motor symptoms in Parkinson's disease using smartwatches and smartphones. *Front. Neurol.* **2023**, *14*, 1326640. [CrossRef] [PubMed]

45. Sigcha, L.; Pavón, I.; Costa, N.; Costa, S.; Gago, M.; Arezes, P.; López, J.M.; De Arcas, G. Automatic resting tremor assessment in Parkinson's disease using smartwatches and multitask convolutional neural networks. *Sensors* **2021**, *21*, 291. [CrossRef] [PubMed]

46. Sigcha, L.; Domínguez, B.; Borzì, L.; Costa, N.; Costa, S.; Arezes, P.; López, J.M.; De Arcas, G.; Pavón, I. Bradykinesia detection in Parkinson's disease using smartwatches' inertial sensors and deep learning methods. *Electronics* **2022**, *11*, 3879. [CrossRef]

47. Sigcha, L.; Borzì, L.; Pavón, I.; Costa, N.; Costa, S.; Arezes, P.; López, J.M.; De Arcas, G. Improvement of performance in freezing of gait detection in Parkinson's disease using transformer networks and a single waist-worn triaxial accelerometer. *Eng. Appl. Artif. Intell.* **2022**, *116*, 105482. [CrossRef]

48. Rodríguez-Martín, D.; Cabestany, J.; Pérez-López, C.; Pie, M.; Calvet, J.; Samà, A.; Capra, C.; Català, A.; Rodríguez-Molinero, A. A new paradigm in Parkinson's disease evaluation with wearable medical devices: A review of STAT-ON. *Front. Neurol.* **2022**, *13*, 912343. [CrossRef] [PubMed]

49. Santiago, A.; Langston, J.W.; Gandhy, R.; Dhall, R.; Brillman, S.; Rees, L.; Barlow, C. Qualitative evaluation of the Personal KinetiGraph™ movement recording system in a Parkinson's clinic. *J. Parkinsons Dis.* **2019**, *9*, 207–219. [CrossRef] [PubMed]

50. Antonini, A.; Reichmann, H.; Gentile, G.; Garon, M.; Tedesco, C.; Frank, A.; Falkenburger, B.; Konitsiotis, S.; Tsamis, K.; Rigas, G.; et al. Toward objective monitoring of Parkinson's disease motor symptoms using a wearable device: Wearability and performance evaluation of PDMonitor®. *Front. Neurol.* **2023**, *14*, 1080752. [CrossRef] [PubMed]

51. Moreau, C.; Rouaud, T.; Grabli, D.; Benatru, I.; Remy, P.; Marques, A.R.; Drapier, S.; Mariani, L.L.; Roze, E.; Devos, D.; et al. Overview on wearable sensors for the management of Parkinson's disease. *npj Parkinsons Dis.* **2023**, *9*, 153. [CrossRef]

52. Tam, W.; Alajlani, M.; Abd-alrazaq, A. An exploration of wearable device features used in UK hospital Parkinson disease care: Scoping review. *J. Med. Internet Res.* **2023**, *25*, e42950. [CrossRef]

53. Rovini, E.; Maremmani, C.; Cavallo, F. How wearable sensors can support Parkinson's disease diagnosis and treatment: A systematic review. *Front. Neurosci.* **2017**, *11*, 555. [CrossRef]

54. Del Din, S.; Kirk, C.; Yarnall, A.J.; Rochester, L.; Hausdorff, J.M. Body-worn sensors for remote monitoring of Parkinson's disease motor symptoms: Vision, state of the art, and challenges ahead. *J. Parkinsons Dis.* **2021**, *11*, S35–S47. [CrossRef] [PubMed]

55. Strimbu, K.; Tavel, J.A. What are biomarkers? *Curr. Opin. HIV AIDS* **2010**, *5*, 463–466. [CrossRef] [PubMed]

56. Atkinson, A.J., Jr.; Colburn, W.A.; DeGruttola, V.G.; DeMets, D.L.; Downing, G.J.; Hoth, D.F.; Oates, J.A.; Peck, C.C.; Schooley, R.T.; et al. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clin. Pharmacol. Ther.* **2001**, *69*, 89–95. [CrossRef]

57. Park, J.E.; Gunasekaran, T.I.; Cho, Y.H.; Choi, S.M.; Song, M.K.; Cho, S.H.; Kim, J.; Song, H.C.; Choi, K.Y.; Lee, J.J.; et al. Diagnostic blood biomarkers in Alzheimer's disease. *Biomedicines* **2022**, *10*, 169. [CrossRef]

58. US Food and Drug Administration. Patient-Focused Drug Development: Collecting Comprehensive and Representative Input. Available online: https://www.federalregister.gov/d/2020-13046 (accessed on 18 March 2024).

59. Vasudevan, S.; Saha, A.; Tarver, M.E.; Patel, B. Digital biomarkers: Convergence of digital health technologies and biomarkers. *npg Digit. Med.* **2022**, *5*, 36. [CrossRef]

60. Manta, C.; Patrick-Lake, B.; Goldsack, J.C. Digital measures that matter to patients: A framework to guide the selection and development of digital measures of health. *Digit. Biomark.* **2020**, *4*, 69–77. [CrossRef]

61. Insel, T.R. Digital phenotyping: Technology for a new science of behavior. *JAMA* **2017**, *318*, 1215. [CrossRef]

62. Babrak, L.M.; Menetski, J.; Rebhan, M.; Nisato, G.; Zinggeler, M.; Brasier, N.; Baerenfaller, K.; Brenzikofer, T.; Baltzer, L.; Vogler, C.; et al. Traditional and digital biomarkers: Two worlds apart? *Digit. Biomark.* **2019**, *3*, 92–102. [CrossRef]

63. Califf, R.M. Biomarker definitions and their applications. *Exp. Biol. Med.* **2018**, *243*, 213–221. [CrossRef]

64. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*, e1000097. [CrossRef] [PubMed]

65. Hao, T.; Yamada, Y.; Rogers, J.L.; Shinakwa, K.; Nemoto, M.; Nemoto, K.; Arai, T. An automated digital biomarker of mobility. In Proceedings of the 2023 IEEE International Conference on Digital Health (ICDH), Chicago, IL, USA, 2–8 July 2023; IEEE: Piscataway, NJ, USA, 2023. [CrossRef]

66. Shah, V.V.; McNames, J.; Harker, G.; Mancini, M.; Carlson-Kuhta, P.; Nutt, J.G.; El-Gohary, M.; Curtze, C.; Horak, F.B. Effect of bout length on gait measures in people with and without Parkinson's disease during daily life. *Sensors* **2020**, *20*, 5769. [CrossRef] [PubMed]

67. Coates, L.; Shi, J.; Rochester, L.; Del Din, S.; Pantall, A. Entropy of real-world gait in Parkinson's disease determined from wearable sensors as a digital marker of altered ambulatory behavior. *Sensors* **2020**, *20*, 2631. [CrossRef] [PubMed]

68. Deng, K.; Li, Y.; Zhang, H.; Wang, J.; Albin, R.L.; Guan, Y. Heterogeneous digital biomarker integration out-performs patient self-reports in predicting Parkinson's disease. *Commun. Biol.* **2022**, *5*, 58. [CrossRef]

69. Vidya, B.; Sasikumar, P. Gait based Parkinson's disease diagnosis and severity rating using multi-class support vector machine. *Appl. Soft Comput.* **2021**, *113*, 107939. [CrossRef]

70. Khera, P.; Kumar, N. Novel machine learning-based hybrid strategy for severity assessment of Parkinson's disorders. *Med. Biol. Eng. Comput.* **2022**, *60*, 811–828. [CrossRef] [PubMed]

71. Atrsaei, A.; Corrà, M.F.; Dadashi, F.; Vila-Chã, N.; Maia, L.; Mariani, B.; Maetzler, W.; Aminian, K. Gait speed in clinical and daily living assessments in Parkinson's disease patients: Performance versus capacity. *npj Parkinsons Dis.* **2021**, *7*, 24. [CrossRef] [PubMed]

72. Wissel, B.D.; Mitsi, G.; Dwivedi, A.K.; Papapetropoulos, S.; Larkin, S.; López Castellanos, J.R.; Shanks, E.; Duker, A.P.; Rodriguez-Porcel, F.; Vaughan, J.E.; et al. Tablet-based application for objective measurement of motor fluctuations in Parkinson disease. *Digit. Biomark.* **2018**, *1*, 126–135. [CrossRef]

73. Di Lazzaro, G.; Ricci, M.; Saggio, G.; Costantini, G.; Schirinzi, T.; Alwardat, M.; Pietrosanti, L.; Patera, M.; Scalise, S.; Giannini, F.; et al. Technology-based therapy-response and prognostic biomarkers in a prospective study of a de novo Parkinson's disease cohort. *npj Parkinsons Dis.* **2021**, *7*, 82. [CrossRef]

74. Sharma, M.; Mishra, R.k.; Hall, A.J.; Casado, J.; Cole, R.; Nunes, A.S.; Barchard, G.; Vaziri, A.; Pantelyat, A.; Wills, A.M. Remote at-home wearable-based gait assessments in progressive supranuclear palsy compared to Parkinson's disease. *BMC Neurol.* **2023**, *23*, 434. [CrossRef]

75. Arora, S.; Baig, F.; Lo, C.; Barber, T.R.; Lawton, M.A.; Zhan, A.; Rolinski, M.; Ruffmann, C.; Klein, J.C.; Rumbold, J.; et al. Smartphone motor testing to distinguish idiopathic REM sleep behavior disorder, controls, and PD. *Neurology* **2018**, *91*, e1528–e1538. [CrossRef] [PubMed]

76. Tsoulos, I.G.; Mitsi, G.; Stavrakoudis, A.; Papapetropoulos, S. Application of machine learning in a Parkinson's disease digital biomarker dataset using Neural Network Construction (NNC) methodology discriminates patient motor status. *Front. ICT* **2019**, *6*, 10. [CrossRef]

77. Rehman, R.Z.U.; Buckley, C.; Mico-Amigo, M.E.; Kirk, C.; Dunne-Willows, M.; Mazza, C.; Shi, J.Q.; Alcock, L.; Rochester, L.; Del Din, S. Accelerometry-based digital gait characteristics for classification of Parkinson's disease: What counts? *IEEE Open J. Eng. Med. Biol.* **2020**, *1*, 65–73. [CrossRef]

78. Gonçalves, H.R.; Branquinho, A.; Pinto, J.; Rodrigues, A.M.; Santos, C.P. Digital biomarkers of mobility and quality of life in Parkinson's disease based on a wearable motion analysis LAB. *Comput. Methods Programs Biomed.* **2024**, *244*, 107967. [CrossRef]

79. Sahandi Far, M.; Eickhoff, S.B.; Goni, M.; Dukart, J. Exploring test-retest reliability and longitudinal stability of digital biomarkers for Parkinson disease in the m-Power data set: Cohort study. *J. Med. Internet Res.* **2021**, *23*, e26608. [CrossRef] [PubMed]

80. Lipsmeier, F.; Taylor, K.I.; Kilchenmann, T.; Wolf, D.; Scotland, A.; Schjodt-Eriksen, J.; Cheng, W.; Fernandez-Garcia, I.; Siebourg-Polster, J.; Jin, L.; et al. Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. *Mov. Disord.* **2018**, *33*, 1287–1297. [CrossRef]

81. Mahadevan, N.; Demanuele, C.; Zhang, H.; Volfson, D.; Ho, B.; Erb, M.K.; Patel, S. Development of digital biomarkers for resting tremor and bradykinesia using a wrist-worn wearable device. *npg Digit. Med.* **2020**, *3*, 5. [CrossRef]

82. Evers, L.J.; Raykov, Y.P.; Krijthe, J.H.; Silva de Lima, A.L.; Badawy, R.; Claes, K.; Heskes, T.M.; Little, M.A.; Meinders, M.J.; Bloem, B.R. Real-life gait performance as a digital biomarker for motor fluctuations: The Parkinson@Home validation study. *J. Med. Internet Res.* **2020**, *22*, e19068. [CrossRef]

83. ZhuParris, A.; Thijssen, E.; Elzinga, W.O.; Makai-Bölöni, S.; Kraaij, W.; Groeneveld, G.J.; Doll, R.J. Treatment detection and movement disorder society-unified Parkinson's disease rating scale, part III estimation using finger tapping tasks. *Mov. Disord.* **2023**, *38*, 1795–1805. [CrossRef] [PubMed]

84. Goni, M.; Eickhoff, S.B.; Far, M.S.; Patil, K.R.; Dukart, J. Smartphone-based digital biomarkers for Parkinson's disease in a remotely-administered setting. *IEEE Access* **2022**, *10*, 28361–28384. [CrossRef]

85. Greene, B.R.; Premoli, I.; McManus, K.; McGrath, D.; Caulfield, B. Predicting fall counts using wearable sensors: A novel digital biomarker for Parkinson's disease. *Sensors* **2021**, *22*, 54. [CrossRef] [PubMed]

86. mPower Public Researcher Portal. Mobile Parkinson's Disease Study. Available online: https://doi.org/10.7303/syn4993293 (accessed on 20 March 2024). [CrossRef]

87. Hausdorff, J.M. Gait in Parkinson's Disease. Available online: https://physionet.org/content/gaitpdb/ (accessed on 18 March 2024).

88. Adams, J.L.; Kangarloo, T.; Gong, Y.; Khachadourian, V.; Tracey, B.; Volfson, D.; Latzman, R.D.; Cosman, J.; Edgerton, J.; Anderson, D.; et al. Using a smartwatch and smartphone to assess early Parkinson's disease in the WATCH-PD study over 12 months. *npj Parkinsons Dis.* **2024**, *10*, 112. [CrossRef] [PubMed]

89. Borzì, L.; Varrecchia, M.; Sibille, S.; Olmo, G.; Artusi, C.A.; Fabbri, M.; Rizzone, M.G.; Romagnolo, A.; Zibetti, M.; Lopiano, L. Smartphone-based estimation of item 3.8 of the MDS-UPDRS-III for assessing leg agility in people with Parkinson's disease. *IEEE Open J. Eng. Med. Biol.* **2020**, *1*, 140–147. [CrossRef] [PubMed]

90. Jack, C.R.; Bernstein, M.A.; Fox, N.C.; Thompson, P.; Alexander, G.; Harvey, D.; Borowski, B.; Britson, P.J.; L. Whitwell, J.; Ward, C.; et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* **2008**, *27*, 685–691. [CrossRef] [PubMed]

91. Siirtola, P.; Koskimäki, H.; Röning, J. OpenHAR: A MATLAB toolbox for easy access to publicly open human activity data sets. In Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, UbiComp '18, Singapore , 8–12 October 2018; ACM: New York, NY, USA, 2018. [CrossRef]