This is a repository copy of *Walking on common ground : a cross-disciplinary scoping review on the clinical utility of digital mobility outcomes*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/179422/

Version: Published Version

Article:

REVIEW ARTICLE   OPEN

Check for updates

Walking on common ground: a cross-disciplinary scoping review on the clinical utility of digital mobility outcomes

Ashley Polhemus [1 ✉], Laura Delgado Ortiz [2,3,4], Gavin Brittain [5], Nikolaos Chynkiamis [6], Francesca Salis [7], Heiko Gaßner [8], Michaela Gross [9], Cameron Kirk [10], Rachele Rossanigo [7], Kristin Taraldsen [11], Diletta Balta [12], Sofie Breuls [13,14], Sara Buttery [15], Gabriela Cardenas [2,3,4], Christoph Endress [9], Julia Gugenhan [9], Alison Keogh [16], Felix Kluge [17], Sarah Koch [2,3,4], M. Encarna Micó-Amigo [10], Corinna Nerz [9], Chloé Sieber [1], Parris Williams [15], Ronny Bergquist [11], Magda Bosch de Basea [2,3,4], Ellen Buckley [18], Clint Hansen [19], A. Stefanie Mikolaizak [9], Lars Schwickert [9], Kirsty Scott [18], Sabine Stallforth [8], Janet van Uem [19], Beatrix Vereijken [11], Andrea Cereatti [7,12], Heleen Demeyer [13,14,20], Nicholas Hopkinson [15], Walter Maetzler [19], Thierry Troosters [13,14], Ioannis Vogiatzis [6], Alison Yarnall [10], Clemens Becker [9], Judith Garcia-Aymerich [2,3,4], Letizia Leocani [21], Claudia Mazzà [18], Lynn Rochester [10], Basil Sharrack [5], Anja Frei [1], Milo Puhan [1] and Mobilise-D

Physical mobility is essential to health, and patients often rate it as a high-priority clinical outcome. Digital mobility outcomes (DMOs), such as real-world gait speed or step count, show promise as clinical measures in many medical conditions. However, current research is nascent and fragmented by discipline. This scoping review maps existing evidence on the clinical utility of DMOs, identifying commonalities across traditional disciplinary divides. In November 2019, 11 databases were searched for records investigating the validity and responsiveness of 34 DMOs in four diverse medical conditions (Parkinson's disease, multiple sclerosis, chronic obstructive pulmonary disease, hip fracture). Searches yielded 19,672 unique records. After screening, 855 records representing 775 studies were included and charted in systematic maps. Studies frequently investigated gait speed (70.4% of studies), step length (30.7%), cadence (21.4%), and daily step count (20.7%). They studied differences between healthy and pathological gait (36.4%), associations between DMOs and clinical measures (48.8%) or outcomes (4.3%), and responsiveness to interventions (26.8%). Gait speed, step length, cadence, step time and step count exhibited consistent evidence of validity and responsiveness in multiple conditions, although the evidence was inconsistent or lacking for other DMOs. If DMOs are to be adopted as mainstream tools, further work is needed to establish their predictive validity, responsiveness, and ecological validity. Cross-disciplinary efforts to align methodology and validate DMOs may facilitate their adoption into clinical practice.

npj Digital Medicine (2021)4:149 ; https://doi.org/10.1038/s41746-021-00513-5

INTRODUCTION

Physical mobility is an essential aspect of health. Mobility impairment is associated with reduced quality of life, falls, hospitalization, mortality, and other adverse events in many chronic conditions[1–7]. It is therefore unsurprising that people living with chronic conditions often rate physical mobility—and specifically walking ability—as one of the most important clinical outcomes[8–13].

Traditional mobility measures include patient-reported outcomes (how well an *individual thinks* they can walk), objective clinical assessments (an *individual's examined capacity* to walk), and subjective clinical assessments (how well a *clinician thinks* an individual can walk given a set of standard criteria). These measures can be subject to recall bias, Hawthorne effects, substantial training requirements, and ceiling or floor effects,

among other limitations[14–20]. They are acquired infrequently and often conducted in clinical settings that rarely reflect the complex environmental determinants of real-world function, raising questions of their ecological validity[14,21–23].

It is now technologically feasible to conduct quantitative mobility assessments during real-world walking, defined as unsupervised, unscripted walking behavior that occurs in non-simulated everyday situations[14,24,25]. Walking-related digital mobility outcomes (DMOs) including gait speed, step length, and step count are increasingly used to quantify gait quality in multiple medical conditions. Emerging evidence suggests that DMOs may be sensitive, ecologically valid markers of health status[14,21,23], but they are unvalidated and therefore not yet accepted as mainstream research and clinical assessment tools. This gap has sparked multidisciplinary calls to validate and qualify (i.e., seek
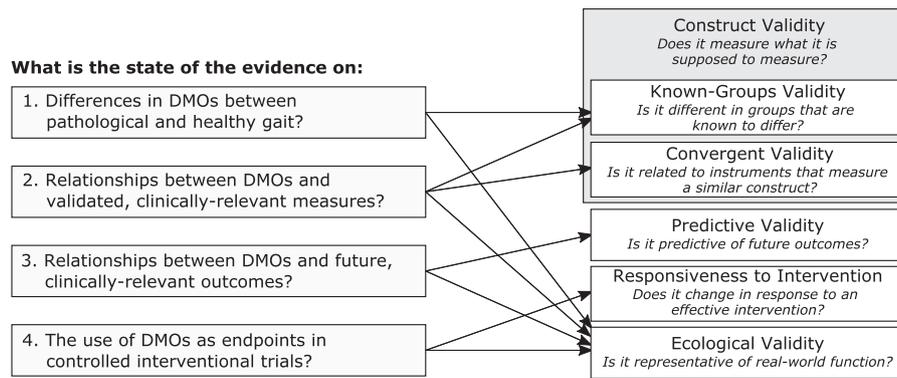
Fig. 1 Research questions (left) and psychometric properties (right) addressed by this review. DMO digital mobility outcome.

regulatory approval for) DMOs as clinical endpoint measures[26–31]. These calls suggest that collaboration across traditional clinical divides will accelerate the qualification process, which entails patient engagement, extensive technical validation, large clinical studies, and an intensive review by regulatory authorities[12,28,29,31]. This process must prove that DMOs are technically feasible to measure, relevant to patients, clinically meaningful, and cost-effective, among other considerations. In this context, clinical meaningfulness is judged by three psychometric properties: construct validity (i.e., they measure what they are supposed to measure), predictive validity (they are associated with important clinical outcomes such as mortality), and responsiveness (they change in response to effective interventions)[32–34].

Many DMOs have been investigated and proposed, but systematic evidence on their psychometric properties is often lacking. This is in part due to fragmentation of the literature by discipline, terminology, and methodology—both within and between clinical disciplines. Systematic evidence is beginning to accumulate, but is generally limited to clinical settings and specific medical conditions, DMOs, or psychometric properties[3,35–37]. These reviews, although foundational, provide a narrow and incomplete understanding of the research landscape. An over-arching view of existing evidence is needed to guide strategic priority setting, inform the design of validation efforts, and identify common research goals—and therefore opportunities for collaboration—which exist across traditional research domains.

Objective

The aim of this scoping review is to generate cross-disciplinary maps of existing evidence on the clinical meaningfulness of DMOs. We stratified our review by four research questions (Fig. 1) designed to map evidence pertaining to the known-groups validity, convergent validity, predictive validity, responsiveness, and ecological validity of a predefined set of DMOs. The resulting maps identify commonalities across disciplinary divides, suggest promising DMOs for further validation, and outline current research gaps. Although walking impairment is of interest in many medical conditions, it was impossible to map the entire research field in a single review. We selected four diverse medical conditions as exemplars, representing diverse etiologies and patterns of mobility impairment: Parkinson's disease (PD), multiple sclerosis (MS), chronic obstructive pulmonary disease (COPD), and proximal femoral fracture (PFF)[31]. They were selected due to their prevalence, impact on quality of life, economic burden, and evidence base[35,37–39]. Walking impairment is known to play a central role in the patient experience of each of these conditions[12,40–44]. These conditions are the focus of Mobilise-D, an Innovative Medicines Initiative 2 Joint Undertaking that aims to develop and validate DMOs for regulatory and clinical endorsement[45].

RESULTS

Characteristics of included studies

Searches yielded 19,672 unique records, of which 2903 were deemed eligible for full-text review. Of these, 855 records were eligible for inclusion (PD: $n = 307$; MS: $n = 270$; COPD: $n = 225$; PFF: $n = 53$), representing 5019 unique analyses from 775 studies (Fig. 2). The list of included records is available on our project repository[46]. Reviewer agreement was substantial at the abstract (16 raters, weighted Cohen's $\kappa = 0.77$, Fleiss' $\kappa = 0.56$) and full-text stages (22 raters, weighted Cohen's $\kappa = 0.75$). Gait speed was studied most frequently in all medical conditions except COPD, which favored daily step count (Supplementary Fig. 1). Characteristics of included studies and their populations are provided in Supplementary Table 8. Most studies were small (median [IQR]: 50 [30–94] participants) and included populations with moderate median disease severity. We observed substantial methodological heterogeneity both within and between the medical conditions, although the methods were often unclearly reported (Table 1).

Known-groups validity

Overall, 282 studies investigated differences in DMOs between healthy and pathological gait (Fig. 3) and 137 studies compared DMOs across disease severity strata (Supplementary Fig. 4). Several DMOs exhibited consistent evidence of known-groups validity in PD, MS, and COPD, although few investigated differences between known groups for any DMO in PFF. Gait speed, step/stride length, step/stride length variability, and measures describing the support phase of gait were consistently different between known groups, although the evidence is limited for disease severity strata in COPD. DMOs describing cadence, step/stride time, and daily step count were consistently different between known groups in MS and COPD, but less so in PD.

Convergent validity

We identified 378 studies that investigated associations between DMOs and validated measures of condition severity, lower-extremity function, health-related quality of life, and other constructs. Gait speed, step/stride length, cadence, and step/stride time exhibited consistent relationships with measures of condition severity (Fig. 4) and lower-extremity function (Supplementary Fig. 5). Mapped associations between DMOs and measures of balance, falls, and health-related quality of life are provided in Supplementary Figs. 6–9. Gait speed, daily step count, and daily walking time were consistently associated with health-related quality of life in all conditions. Gait speed, step/stride length variability, and step/stride time variability were most consistently related to balance and falls, although this primarily reflected studies in PD and MS.
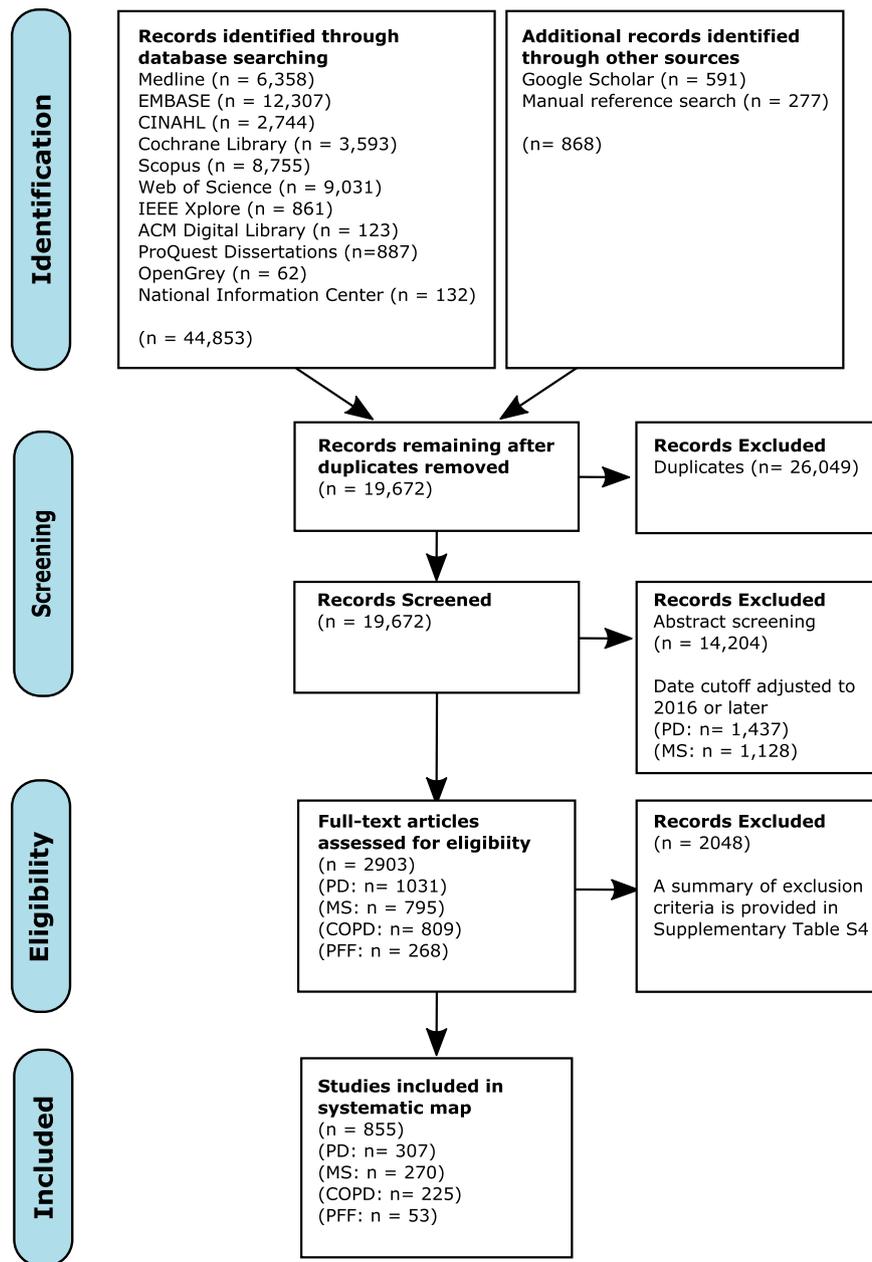
Fig. 2 PRISMA flow diagram. This diagram shows how records were screened for eligibility in this review.

Predictive validity

Only 33 studies investigated the predictive validity of DMOs (PD: $n = 10$; MS: $n = 7$; COPD: $n = 14$; PFF: $n = 2$). Most studies (22 of 33, 66.7%) adjusted at least some of their analyses for known predictors or common confounders. In PD, several DMOs were related to future disease progression, falls, physical function, and cognition[47-58] (Supplementary Fig. 10) In MS, gait speed was associated with disease progression[59,60], future falls[61,62], and functional status[63]. One study identified a relationship between daily step count and disease progression[64]. Another identified a relationship between stride time variability, but not stride speed variability, and falls[65]. In COPD, gait speed[66-70] and daily step count[71-75] demonstrated relationships with mortality. One study identified a relationship between step count and disease progression[72]. Evidence for relationships with exacerbations[76,77], activities of daily living[72], health-related quality of life[72], and healthcare utilization[69,74,78] was limited or inconsistent. In PFF, one

study found a relationship between gait speed and healthcare utilization[79], while the relationship between gait speed and activities of daily living was inconsistent[79,80].

Responsiveness to intervention

We identified 208 studies that used DMOs as outcome measures in controlled interventional trials. Of these, 140 (67.3%) reported using a DMO as a primary outcome and 79 (38.0%) reported using a DMO as a secondary outcome. However, many studies reported several "primary" outcomes and it was often unclear which outcomes, if any, were used in the power analysis. Studies were generally designed to evaluate the efficacy of interventions rather than the responsiveness of DMOs; therefore, evidence of DMOs' responsiveness could not be clearly disentangled from the efficacy of the various experimental interventions. Thus, we created two maps to estimate responsiveness. Figures 5 and 6 map analyses from all included studies and studies in which interventions were

Table 1. Walking conditions and measurement methods used in included studies.

	PD, $n = 265$	MS, $n = 250$	COPD, $n = 193$	PFF, $n = 48$
Measurement method				
Stopwatch	49 (18.5%)	158 (63.2%)	37 (19.2%)	34 (70.8%)
Video/optoelectronic system	65 (24.5%)	24 (9.6%)	4 (2.1%)	0 (0.0%)
Instrumented walkway	51 (19.2%)	34 (13.6%)	10 (5.2%)	8 (16.7%)
Instrumented treadmill	10 (3.8%)	6 (2.4%)	2 (1.0%)	0 (0.0%)
Instrumented environment	2 (0.8%)	1 (0.4%)	0 (0.0%)	0 (0.0%)
Wearable sensor[a] (hip/waist)	32 (12.1%)	16 (6.4%)	49 (25.4%)	0 (0.0%)
Wearable sensor (other/mixed locations)	58 (21.9%)	31 (12.4%)	74 (38.3%)	4 (8.3%)
Pedometer	1 (0.4%)	0 (0.0%)	31 (16.1%)	1 (2.1%)
Mobile phone	1 (0.4%)	1 (0.4%)	0 (0.0%)	0 (0.0%)
Video gaming system (e.g., Kinect)	6 (2.3%)	4 (1.6%)	0 (0.0%)	0 (0.0%)
Other	14 (5.3%)	3 (1.2%)	4 (2.1%)	4 (8.3%)
Measurement setting				
Clinic/lab	252 (95.1%)	240 (96.0%)	63 (32.6%)	41 (85.4%)
Home/real world	20 (7.5%)	25 (10.0%)	135 (69.9%)	5 (10.4%)
Walking bout length				
Short walk (≤1 min or <20 m)	204 (77.0%)	211 (84.4%)	35 (18.1%)	34 (70.8%)
Longer walk (>1 min or 20 m)	49 (18.5%)	54 (21.6%)	21 (10.9%)	13 (27.1%)
Real-world walking bouts	17 (6.4%)	24 (9.6%)	139 (72.0%)	4 (8.3%)
Unclear	13 (4.9%)	10 (4.0%)	3 (1.6%)	0 (0.0%)
Walking bout speed				
Habitual speed	205 (77.4%)	95 (38.0%)	35 (18.1%)	23 (47.9%)
Fast speed	34 (12.8%)	158 (63.2%)	22 (11.4%)	19 (39.6%)
Set speed (i.e., on a fixed-speed treadmill)	10 (3.8%)	5 (2.0%)	2 (1.0%)	0 (0.0%)
Averaged bouts of variable speeds	3 (1.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Real-world walking bouts	17 (6.4%)	23 (9.2%)	140 (72.5%)	3 (6.2%)
Unclear	32 (12.1%)	19 (7.6%)	6 (3.1%)	8 (16.7%)

Data are presented as n (%) of included studies. Multiple records were identified for several studies; thus, the total number of studies differs from the total number of records. The sum of percentages may exceed 100%, as studies often reported results for multiple measurement methods or walking conditions. Measurement method, measurement setting, walking bout length, and walking bout speed indicate the categories of walking conditions reported in included studies.

PD Parkinson's disease, *MS* multiple sclerosis, *COPD* chronic obstructive pulmonary disease, *PFF* proximal femoral fracture.

[a]Wearable sensors refer to any wearable data acquisition device other than pedometers, including accelerometers and inertial measurement units.

"effective," respectively. For our purposes, "effective" interventions are those which yield significant differences in any primary endpoint between experimental and control or comparator arms. The former is likely to underestimate responsiveness (it is confounded by the true efficacy of the experimental interventions), and the latter is likely to overestimate it (it is biased in favor of successful studies that used DMOs as primary outcomes). Gait speed, step/stride length, cadence, daily step count, and walking time often responded to "effective" interventions in all conditions, although outcomes were relatively inconsistent.

Ecological validity

Excluding measures of daily walking volume, only 17 studies measured spatiotemporal DMOs during real-world walking (PD: $n = 8$; MS: $n = 5$; COPD: $n = 2$; PFF: $n = 2$). Relationships exhibited by real-world and in-clinic DMOs are compared for PD[81–88] and MS[21,89–92] in Fig. 7 and Supplementary Fig. 12, respectively. In COPD, real-world walking cadence differed from healthy controls[93] and real-world gait speed was associated with disease severity[94]. In PFF, timed gait speed tests were conducted at home in two studies[95,96]. These tests were not responsive to intervention, although interventions in both studies were found to be ineffective. These relationships were similar to those observed in clinical settings, but such comparisons were qualitative.

Assessment of bias

Manual inspection revealed key differences in research strategy between the medical conditions. Records in PD were more likely to study specific subpopulations with gait impairments (e.g., fallers, individuals with freezing of gait). Several records in MS and PD, but not COPD or PFF, specifically studied populations with early-stage disease. The body of literature on PD, COPD, and PFF appeared to exhibit a survivorship bias (in this case, the tendency for healthier-than-average individuals with a given characteristic to be included in a study) with respect to age and condition severity. Only studies in MS, which represented a younger population, reflected the full range of disease severity and demonstrated the expected colinearity of age and disease severity (Supplementary Fig. 2). This is likely due to an association between age, condition severity, and comorbidities or cognitive impairment, which were often exclusion criteria in included studies.

Meta-regression showed that conference abstracts (adjusted odds ratio [95% confidence interval]: 2.44 [1.59–3.76], $p < 0.001$), studies with fast walking assessments (1.54 [1.10–2.17], $p = 0.02$), and studies on at-risk subgroups such as fallers (2.03 [1.47–2.80],
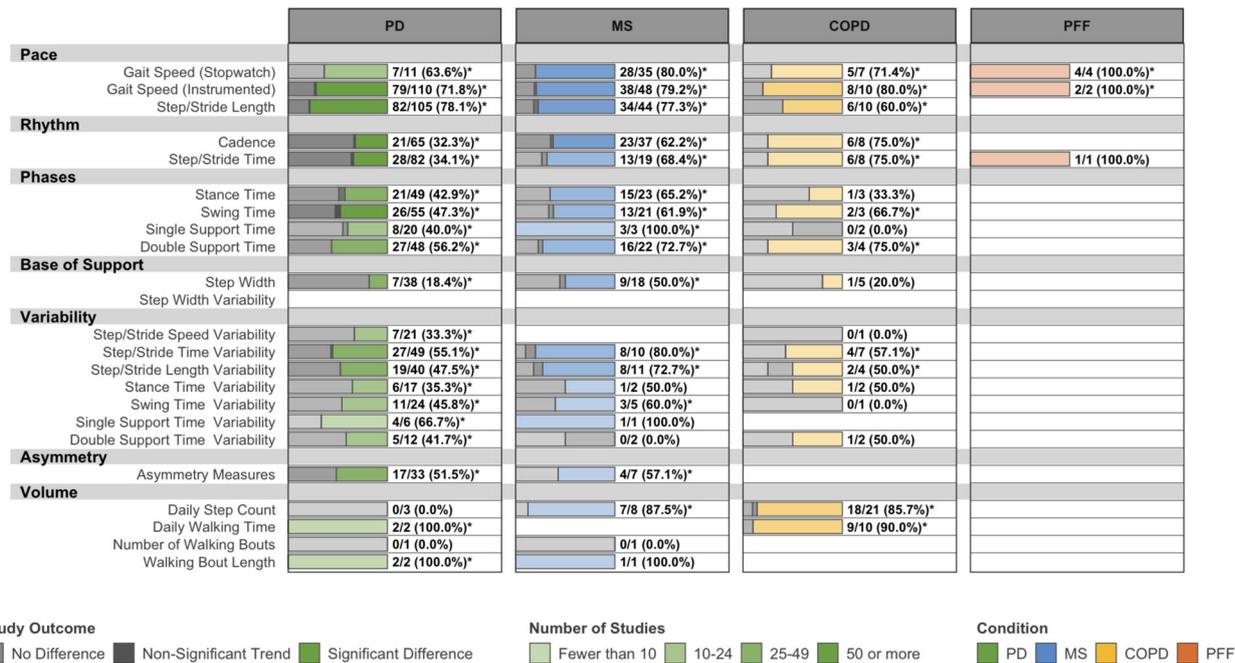
Fig. 3 Known-groups validity: number and outcome of eligible studies assessing differences in DMOs between patients and healthy controls. PD Parkinson's disease, MS multiple sclerosis, COPD chronic obstructive pulmonary disease, PFF proximal femoral fracture. Data are presented as: Number of studies with statistically significant differences between groups/Total studies (%). DMOs known to be highly intercorrelated were grouped (i.e., step length and stride length), and all DMOs were organized according to previously established domains of gait. *Proportion of studies exceeds the expected false-positive rate as determined by Bernoulli hypothesis testing and Benjamini–Hochberg adjustment.

$p < 0.001$) were more likely to report significant results than their counterparts. Conversely, studies on populations with mild disease severity (0.46 [0.34–0.61], $p < 0.001$) were less likely to report significant findings than those with moderate severity. In studies comparing pathological to healthy gait, those that matched patients and controls for gait speed were less likely to report significant findings for any DMO (0.39 [0.18–0.83], $p = 0.014$). Contrary to our expectations, adjusted models were more likely to yield significant findings than univariate analyses in studies investigating the prognostic value of DMOs. This suggests that DMOs that did not reach significance in multivariate models were not consistently reported and that our maps may overestimate the true repeatability of these relationships. No other study characteristics were associated with study outcomes. Sensitivity analyses yielded similar estimates of all effects. It is important to note that these relationships are observational, and may not hold at the individual or study level[97]. They merely suggest that methodological and population heterogeneity contributed to the inconsistencies observed in our maps. Detailed results of these analyses are provided in the Supplementary materials (Supplementary Notes 3 and 5, Supplementary Fig. 3, and Supplementary Table 7).

Qualitative appraisal of existing evidence
We identified several notable evidence gaps. Few records studied the predictive validity of DMOs in any of the four conditions. Only gait speed and step count were regularly used as outcomes in interventional studies. Few DMOs were studied regularly in COPD and PFF.

Despite these gaps, evidence consistently supported the validity of gait speed, step/stride length, cadence, step/stride time, step/stride time variability, and daily step count whenever it was available (Table 2). In PD, more positive evidence was available for daily walking time than step count, but evidence supporting these measures was similar in the other conditions. Several DMOs exhibited evidence in PD and MS but lacked

evidence in the other two conditions. Additional detail is provided in Supplementary Tables 10–13.

DISCUSSION
Recent calls to validate real-world DMOs are based on three premises: that DMOs are clinically meaningful, that relationships observed in clinical settings translate to real-world walking, and that opportunities for collaboration across clinical disciplines exist[30,31]. This review conditionally supports these premises. Mobility indeed appears to be a concept of cross-disciplinary clinical interest. Multiple DMOs were regularly studied in the four included conditions and consistently exhibited evidence of construct validity, predictive validity, and responsiveness. Few studies measured real-world walking in this review, but those that did provide provisional evidence that relationships observed in clinical settings translate to real-world walking.

While condition- and context-specific validation studies are certainly required for the formal validation of DMOs[29,98], it appears that collaborative approaches to validation can speed this process[30,31]. The regulatory pathway for validating and qualifying DMOs—and digital outcomes in general—is taking shape due to the collaborative efforts of regulators, industry, academics, and precompetitive consortia[28,29,98–102]. The time is right for collaborative development of terminology, algorithms, methods, and evaluation frameworks for mutually interesting DMOs, which may streamline the validation of DMOs in PD, MS, COPD, PFF, and other medical conditions.

However, the volume of existing evidence varied across conditions and DMOs. Compared to PD and MS, evidence in PFF and COPD was sparse and concentrated on fewer DMOs. These differences, plus the differences in prevailing methodologies, suggest disparate research strategies between the conditions. DMOs appear to be more established in some conditions than others. Even in recent studies, uninstrumented gait speed tests
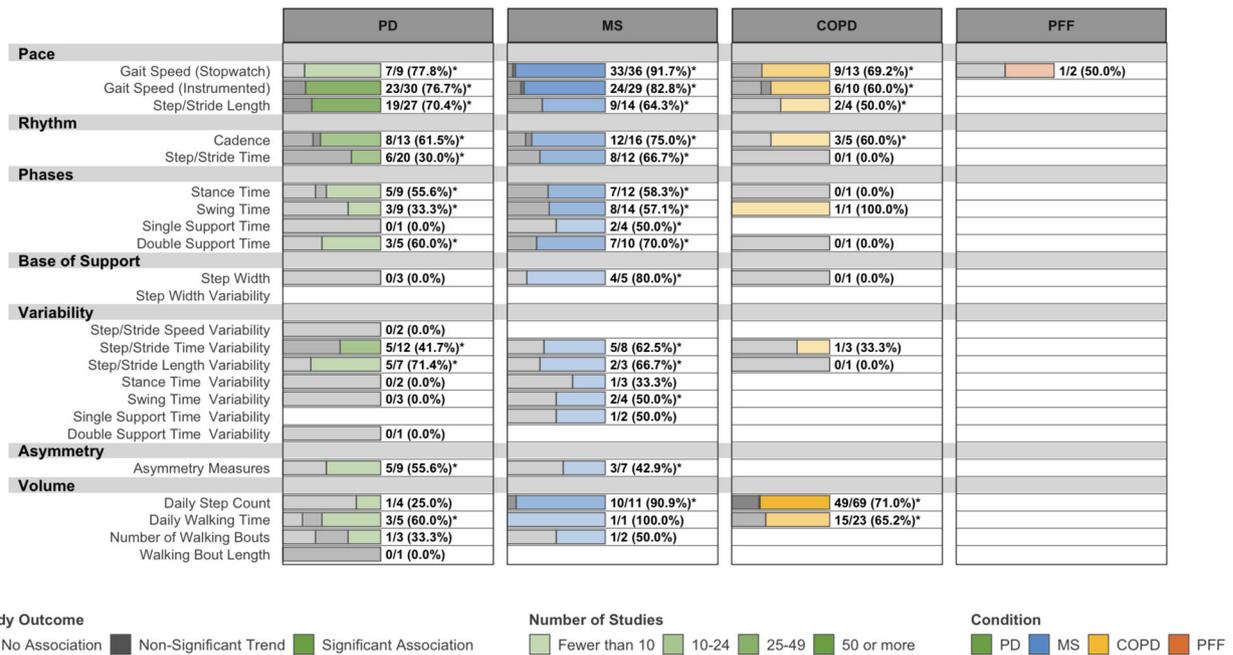
Domain / DMO	PD	MS	COPD	PFF
Pace				
Gait Speed (Stopwatch)	7/9 (77.8%)*	33/36 (91.7%)*	9/13 (69.2%)*	1/2 (50.0%)
Gait Speed (Instrumented)	23/30 (76.7%)*	24/29 (82.8%)*	6/10 (60.0%)*	
Step/Stride Length	19/27 (70.4%)*	9/14 (64.3%)*	2/4 (50.0%)*	
Rhythm				
Cadence	8/13 (61.5%)*	12/16 (75.0%)*	3/5 (60.0%)*	
Step/Stride Time	6/20 (30.0%)*	8/12 (66.7%)*	0/1 (0.0%)	
Phases				
Stance Time	5/9 (55.6%)*	7/12 (58.3%)*	0/1 (0.0%)	
Swing Time	3/9 (33.3%)*	8/14 (57.1%)*	1/1 (100.0%)	
Single Support Time	0/1 (0.0%)	2/4 (50.0%)*		
Double Support Time	3/5 (60.0%)*	7/10 (70.0%)*	0/1 (0.0%)	
Base of Support				
Step Width	0/3 (0.0%)	4/5 (80.0%)*	0/1 (0.0%)	
Step Width Variability				
Variability				
Step/Stride Speed Variability	0/2 (0.0%)			
Step/Stride Time Variability	5/12 (41.7%)*	5/8 (62.5%)*	1/3 (33.3%)	
Step/Stride Length Variability	5/7 (71.4%)*	2/3 (66.7%)*	0/1 (0.0%)	
Stance Time Variability	0/2 (0.0%)	1/3 (33.3%)		
Swing Time Variability	0/3 (0.0%)	2/4 (50.0%)		
Single Support Time Variability		1/2 (50.0%)		
Double Support Time Variability	0/1 (0.0%)			
Asymmetry				
Asymmetry Measures	5/9 (55.6%)*	3/7 (42.9%)*		
Volume				
Daily Step Count	1/4 (25.0%)	10/11 (90.9%)*	49/69 (71.0%)*	
Daily Walking Time	3/5 (60.0%)*	1/1 (100.0%)	15/23 (65.2%)*	
Number of Walking Bouts	1/3 (33.3%)	1/2 (50.0%)		
Walking Bout Length	0/1 (0.0%)			

Study Outcome: ■ No Association ■ Non-Significant Trend ■ Significant Association
Number of Studies: ■ Fewer than 10 ■ 10-24 ■ 25-49 ■ 50 or more
Condition: ■ PD ■ MS ■ COPD ■ PFF

Fig. 4 Convergent validity: associations between DMOs and disease severity measures. PD Parkinson's disease, MS multiple sclerosis, COPD: chronic obstructive pulmonary disease, PFF: proximal femoral fracture. Data are presented as: Number of studies with statistically significant associations between DMOs and measures of disease severity/Total studies (%). Disease severity measures include the UPDRS, UPDRS-III, and Hoehn & Yahr scale in PD, EDSS, and PDDS in MS, FEV_1 % predicted and GOLD stage in COPD, and patient- or physician-rated global measures of improvement in all four conditions. Most relevant measures in PFF fell under different categories, such as activities of daily living. DMOs known to be highly intercorrelated were grouped (i.e., step length and stride length), and all DMOs were organized according to previously established domains of gait. *Proportion of studies exceeds the expected false-positive rate as determined by Bernoulli hypothesis testing and Benjamini–Hochberg adjustment.

were used more frequently than other spatiotemporal DMOs in MS, COPD, and PFF. These tests are widely used, well-established, inexpensive, and simple to implement. However, they are limited by sensitivity to methodology, Hawthorne effects, and other shortcomings[103,104]. Current methods to measure other DMOs are newer, less mature, and more expensive, requiring significant infrastructure and technical expertise[105]. As these factors likely influence the adoption and study of DMOs, we do not consider lack of evidence to constitute negative evidence or inferiority in this review. These gaps are merely areas in which more evidence must be established.

The purpose of scoping reviews is to map research fields and set an agenda for future research[106,107]. The relationships observed here provide clues on the contexts in which DMOs might be useful as outcome measures and suggest gaps that should be addressed to inform DMO validation.

Predictive Validity

Endpoint qualification requires evidence that DMOs are associated with "hard" clinical outcomes such as falls, hospitalization, and mortality. Relationships between these outcomes and in-clinic gait speed are established in many conditions[108,109], including those studied here. However, with few exceptions, evidence on the predictive validity of other DMOs is sparse. Despite the purported potential of spatiotemporal parameters and real-world DMOs as clinical measures, further work is needed to confirm their predictive validity before they can be considered for regulatory qualification.

Responsiveness to Intervention

This review identified preliminary evidence for the responsiveness of common DMOs. However, included studies were not specifically designed to assess the responsiveness of DMOs; they were

designed to test the efficacy of interventions. It is not yet clear which DMOs are responsive to which types of interventions, nor is it clear what constitutes clinically meaningful changes in these DMOs. The context-dependency and relative magnitude of DMOs' responsiveness should be confirmed against "gold standard" outcomes through adequately powered interventional studies and meta-analyses. Any future work should report measures of effect size to quantify the responsiveness of DMOs.

Ecological validity

Scripted walking assessments, which test functional capacity at a single timepoint, are not necessarily representative of habitual or spontaneous walking behavior[21,110,111]. If DMOs are to be used as real-world measures or interpreted as "ecologically valid," the psychometric properties of DMOs measured during real-world walking must be established. At the time of our search, records measuring real-world walking were relatively rare. Real-world walking assessment remains technically and logistically challenging[105]. The performance of existing algorithms, which are usually validated under controlled clinical conditions, varies with changing environment, activities, and walking speed[112]. The effect of this variation on DMOs' clinical utility is unclear. While research on real-world walking has recently accelerated in PD and MS, future work should enrich this evidence. As a priority, this work should supplement the pioneering studies that conduct head-to-head comparisons of DMOs measured in the clinic and in the real world[84,110,111]. In the near future, literature on real-world DMOs should be systematically reviewed to establish similarities and differences between real-world and in-clinic walking assessments.

Importance to patients

The clinical perspectives and psychometric properties discussed here, while necessary, are ultimately insufficient to guide DMO
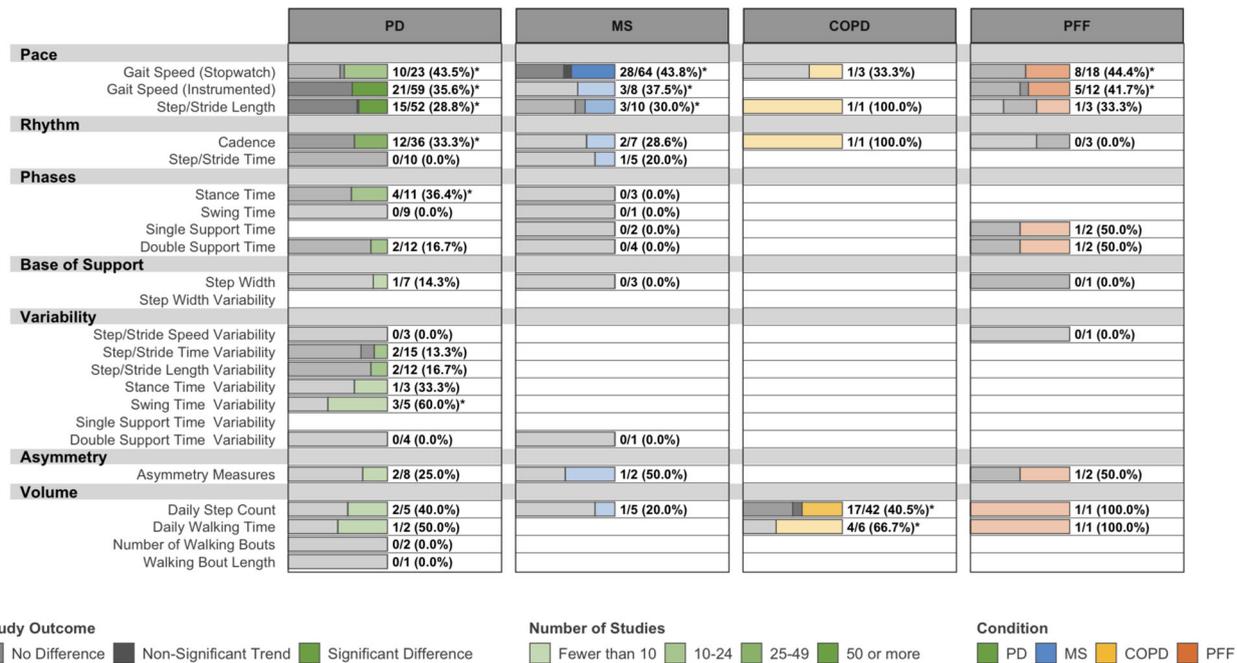
Fig. 5 Responsiveness of DMOs used as primary or secondary endpoints in all eligible interventional studies. PD Parkinson's disease, MS multiple sclerosis, COPD chronic obstructive pulmonary disease, PFF proximal femoral fracture. Data are presented as: Number of studies with statistically significant differences between groups/Total studies (%). Interventions in eligible studies were not necessarily effective, and this map may underestimate the responsiveness of DMOs. DMOs known to be highly intercorrelated were grouped (i.e., step length and stride length), and all DMOs were organized according to previously established domains of gait. *Proportion of studies exceeds the expected false positive rate as determined by Bernoulli hypothesis testing and Benjamini–Hochberg adjustment.

selection and validation. This review mapped hundreds of promising relationships from studies conducted in various contexts, populations, and settings, begging the question, "Of the DMOs and relationships we *can* validate, which *should we* validate?"

Regulatory bodies such as the Food and Drug Administration, the European Medicines Agency, and local Health Technology Assessment bodies expect this question to be addressed from the patient perspective[113–115]. The onus is on researchers to prove that new digital outcomes are important and meaningful within the context of patients' daily lives[114,116]. This refers not only to the construct the measure assesses but also to the level of change the measure can detect. However, relationships between DMOs and the constructs that matter to patients are not always direct. While some DMOs (i.e., gait speed, daily step count) are readily interpretable, others may have little intuitive or practical meaning to anyone other than a gait specialist (i.e, stance time variability). The relationships between DMOs and meaningful constructs must be established both statistically and through early patient engagement, journey mapping, and formal qualitative research, and collaborative agenda setting. Existing guidance[113,115,117,118] and worked examples from past projects[12,102] can shape this interaction. The maps generated in this study can be used to match candidate DMOs with prioritized walking-related constructs and experiences. These candidate DMOs should then be considered specifically when addressing the evidence gaps described here.

Generalizability and context

Relationships between DMOs, condition severity, and physical function in all four conditions suggest that DMOs may be useful to monitor disease progression or changes in mobility status over time. Similarly, cross-sectional and longitudinal associations between DMOs and falls suggest that DMOs may be useful to quantify fall risk in PD and MS. However, the maps presented here

should be treated as directional. The included conditions are highly heterogeneous, representing an array of symptoms under single diagnostic umbrellas. It is entirely possible that the utility of any given DMO is context-dependent, differing between environments, early and late-stage disease, during an acute health event or exacerbation, or between disease subtypes. Examples include individuals with relapsing vs. progressive courses[119] or ataxic vs. paretic gait[120] in MS, freezing of gait[121], orthostatic hypotension[122], tremor-dominant vs. postural instability gait disorder subtypes[123] in PD, oxygen users in COPD[124], or those with different fracture and surgery types in PFF[125]. For many subpopulations, additional original research may be required. When evidence exists, nuanced perspectives on DMOs' clinical utility can be explored through a systematic review.

Strengths and limitations

Our maps aggregated a large, heterogeneous body of literature to identify overarching trends, inform future research, and identify opportunities for cross-disciplinary collaboration. Technical and clinical subject matter experts took part in multidisciplinary review teams, guiding the design of the review and interpretation of the results. Despite its rigor, this review has several notable limitations. For feasibility, we limited included records on PD and MS to those published in 2016 or later. Thus, findings for these two diseases should be interpreted as trends in the literature, rather than an exhaustive tabulation of existing evidence. However, methods remained systematic and data saturation was generally observed. Inconsistent reporting necessitated the use of a relationship's statistical significance, rather than its effect size, in our maps and analyses. Therefore, trends should be interpreted as the repeatability, rather than strength, of observed relationships. Additional systematic reviews and meta-analyses are needed to estimate the strength of key relationships and assess the quality of existing evidence. Despite the breadth of mobility symptoms and disease trajectories covered in the four included medical conditions, the
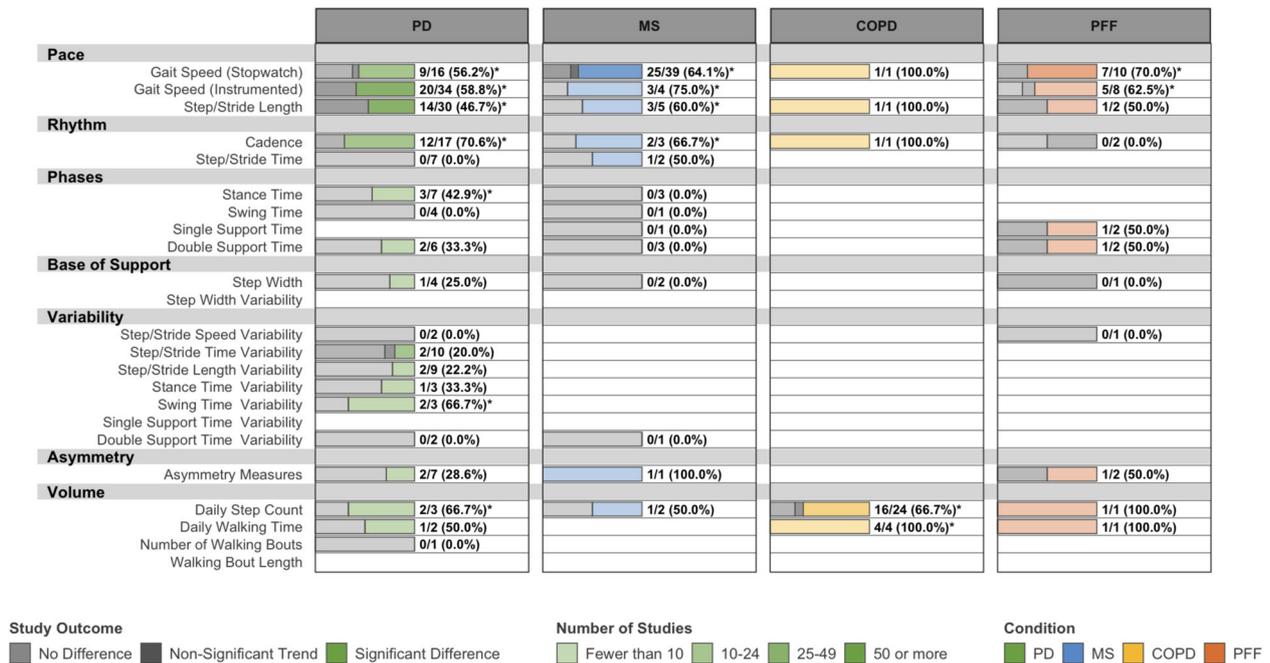
Fig. 6 Responsiveness of DMOs used as primary or secondary endpoints when a studied intervention was effective. PD Parkinson's disease, MS multiple Sclerosis, COPD chronic obstructive pulmonary disease, PFF proximal femoral fracture. Data are presented as: Number of studies with statistically significant differences between groups/Total studies (%). This map may overestimate the responsiveness of DMOs, which were occasionally used as sole primary outcomes (i.e., gait speed and step count), since negative results could be due either to the DMO's responsiveness or to the intervention's efficacy. DMOs known to be highly intercorrelated were grouped (i.e., step length and stride length), and all DMOs were organized according to previously established domains of gait. *Proportion of studies exceeds the expected false-positive rate as determined by Bernoulli hypothesis testing and Benjamini–Hochberg adjustment.

relationships observed here may differ in other conditions, or even within specific subpopulations of the included conditions. Although many DMOs were included, evidence is emerging for other DMOs such as sample entropy, Lyapunov exponents and detrended fluctuation analysis, which are of special interest in real-world assessments[126–131]. As these measures mature, this review should be updated and expanded to include additional DMOs. Finally, we present our results at a high level and many interesting subanalyses were not conducted. We hope that this work will inspire and enable a deeper investigation into the topics discussed here.

Conclusions

Existing evidence supports cross-disciplinary validation efforts for gait speed, step and stride length, cadence, and step count, but is inconsistent or lacking for other DMOs. The relationships exhibited by these DMOs were similar across conditions, signaling potential opportunities for cross-disciplinary collaboration. Future work should include further epidemiological studies, systematic reviews, and meta-analyses to confirm and quantify the relationships observed in this scoping review.

METHODS

Review methodology

We followed the scoping review framework developed by Arksey and O'Malley and advanced by Levac et al.[106,107]. This framework consists of six stages: (1) identifying the research question, (2) identifying relevant studies, (3) selecting studies, (4) charting the data, (5) collating, summarizing, and reporting results, and (6) consulting with relevant stakeholders. Study conduct and reporting adhered to the PRISMA-ScR (PRISMA Extension for Scoping Reviews) guidelines for scoping reviews[132]. A detailed review

protocol was designed and published a priori[133], and is summarized here.

Identifying relevant studies

Search strategies were iteratively developed and tested in exchange with a research librarian and subject matter experts. In November 2019, the librarian searched 11 databases for scientific and gray literature (MEDLINE, EMBASE, CINAHL, Cochrane Library, Scopus, Web of Science, IEEE Xplore, ACM Digital Library, ProQuest Dissertations, OpenGrey, National Information Center's Projects in Progress Database). Final searches with structure (mobility terms) AND (population terms) identified all English-language abstracts published between January 1999 and November 2019. Similar searches in Google Scholar and manual collation of references supplemented this corpus. The search strategy for MEDLINE is provided in Supplementary Table 1 and all search strategies are provided on the project repository[46].

Selecting studies and charting the data

All relevant definitions, eligibility criteria, reference sheets, and data extraction forms are provided in Supplementary Note 1, Supplementary Table 2, or the project protocol[133]. To be eligible, a record must have reported an original analysis that addressed at least one of our research questions with respect to an included DMO in an included population. For the sake of feasibility, we prespecified a list of included DMOs (Supplementary Table 3), limited assessments of construct validity and predictive validity to predefined lists of validated measures, and set a lower limit of ten patients per analysis (or study arm, in the case of interventional trials). We did not otherwise exclude based on methodology. Predefined lists were developed by internal panels of clinical, technical, and research experts. Texts published in any language spoken within our research group (English, German, Spanish,
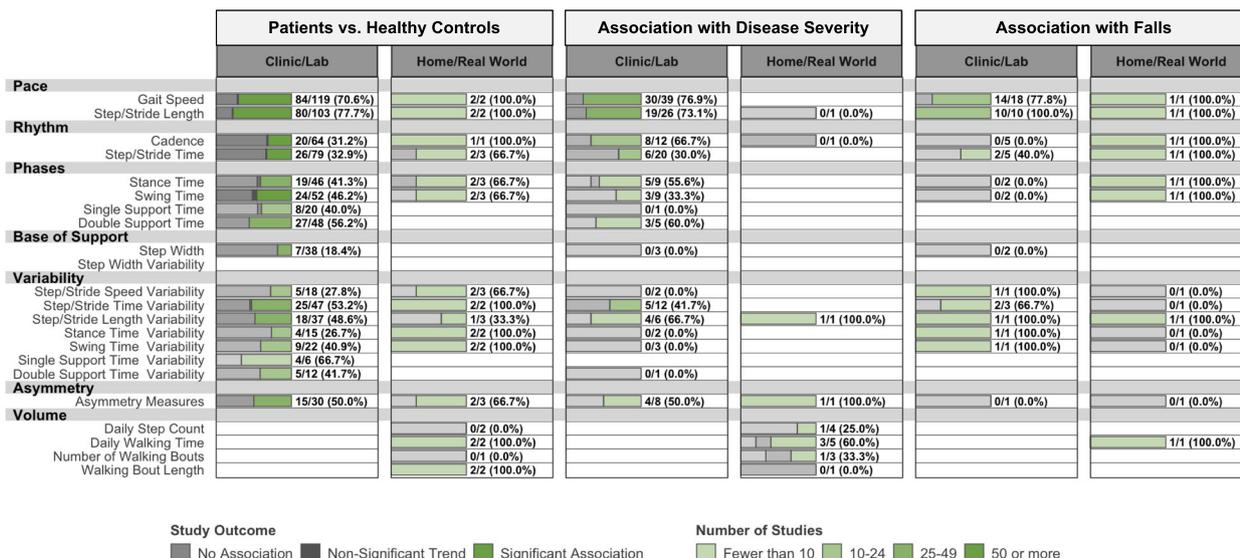
Fig. 7 Ecological validity of DMOs in Parkinson's disease: DMOs collected in clinical vs real-world environments.

Gait domain / DMO	PvHC Clinic/Lab	PvHC Home/Real World	Severity Clinic/Lab	Severity Home/Real World	Falls Clinic/Lab	Falls Home/Real World
Pace						
Gait Speed	84/119 (70.6%)	2/2 (100.0%)	30/39 (76.9%)		14/18 (77.8%)	1/1 (100.0%)
Step/Stride Length	80/103 (77.7%)	2/2 (100.0%)	19/26 (73.1%)	0/1 (0.0%)	10/10 (100.0%)	1/1 (100.0%)
Rhythm						
Cadence	20/64 (31.2%)	1/1 (100.0%)	8/12 (66.7%)	0/1 (0.0%)	0/5 (0.0%)	1/1 (100.0%)
Step/Stride Time	26/79 (32.9%)	2/3 (66.7%)	6/20 (30.0%)		2/5 (40.0%)	1/1 (100.0%)
Phases						
Stance Time	19/46 (41.3%)	2/3 (66.7%)	5/9 (55.6%)		0/2 (0.0%)	1/1 (100.0%)
Swing Time	24/52 (46.2%)	2/3 (66.7%)	3/9 (33.3%)		0/2 (0.0%)	1/1 (100.0%)
Single Support Time	8/20 (40.0%)		0/1 (0.0%)			
Double Support Time	27/48 (56.2%)		3/5 (60.0%)			
Base of Support						
Step Width	7/38 (18.4%)		0/3 (0.0%)		0/2 (0.0%)	
Step Width Variability						
Variability						
Step/Stride Speed Variability	5/18 (27.8%)	2/3 (66.7%)	0/2 (0.0%)		1/1 (100.0%)	0/1 (0.0%)
Step/Stride Time Variability	25/47 (53.2%)	2/2 (100.0%)	5/12 (41.7%)		2/3 (66.7%)	0/1 (0.0%)
Step/Stride Length Variability	18/37 (48.6%)	1/3 (33.3%)	4/6 (66.7%)	1/1 (100.0%)	1/1 (100.0%)	1/1 (100.0%)
Stance Time Variability	4/15 (26.7%)	2/2 (100.0%)	0/2 (0.0%)		1/1 (100.0%)	0/1 (0.0%)
Swing Time Variability	9/22 (40.9%)	2/2 (100.0%)	0/3 (0.0%)		1/1 (100.0%)	0/1 (0.0%)
Single Support Time Variability	4/6 (66.7%)					
Double Support Time Variability	5/12 (41.7%)		0/1 (0.0%)			
Asymmetry						
Asymmetry Measures	15/30 (50.0%)	2/3 (66.7%)	4/8 (50.0%)	1/1 (100.0%)	0/1 (0.0%)	0/1 (0.0%)
Volume						
Daily Step Count		0/2 (0.0%)		1/4 (25.0%)		
Daily Walking Time		2/2 (100.0%)		3/5 (60.0%)		1/1 (100.0%)
Number of Walking Bouts		0/1 (0.0%)		1/3 (33.3%)		
Walking Bout Length		2/2 (100.0%)		0/1 (0.0%)		

Study Outcome: ▪ No Association ▪ Non-Significant Trend ▪ Significant Association
Number of Studies: □ Fewer than 10 ▫ 10–24 ▫ 25–49 ▪ 50 or more

Data are presented as: Number of studies with statistically significant associations between DMOs and measures of lower-extremity function/Total studies (%). DMOs known to be highly intercorrelated were grouped (i.e., step length and stride length), and all DMOs were organized according to previously established domains of gait. *Proportion of studies exceeds the expected false-positive rate as determined by Bernoulli hypothesis testing and Benjamini–Hochberg adjustment.

Table 2. Qualitative appraisal of existing evidence.

Gait domain	Digital mobility outcome	PD	MS	COPD	PFF
Pace	Gait speed	++	++	++	++
	Step/stride length	++	+	+	?
Rhythm	Cadence	++	++	+	?
	Step/stride time	+	+	?	?
Phase	Stance time	++	+	?	?
	Swing time	+	+	?	?
	Single support time	+	+	?	?
	Double support time	+	+	?	?
Base of support	Step width	−	+	?	?
	Step width variability	−	−	?	?
Variability	Step/stride speed variability	+	?	?	?
	Step/stride length variability	++	+	?	?
	Step/stride time variability	+	+	?	?
	Stance time variability	+	?	?	?
	Swing time variability	+	+	?	?
	Single support time variability	?	+	?	?
	Double support time variability	?	−	?	?
Asymmetry	All asymmetry measures	++	+	?	?
Volume	Daily step count	+	+	++	?
	Daily walking time	+	?	+	?
	Number of walking bouts	−	?	?	?
	Walking bout length	?	+	?	?

PD Parkinson's disease, *MS* multiple sclerosis, *COPD* chronic obstructive pulmonary disease, *PFF* proximal femoral fracture.

French, Italian, Portuguese, Danish, Norwegian, Swedish, Hebrew, Dutch, Catalan, Russian) were eligible.

We assessed eligibility through abstract and full-text screening. All reviewers were trained, piloted study materials, and completed consistency checks before each screening phase. Records were included in full-text screening if a single reviewer deemed an abstract eligible, while rejection by two reviewers was required to exclude. Full-text screening was conducted by those with relevant clinical and technical expertise. One reviewer screened each full-text and, if eligible, extracted data. One of three senior reviewers (A.P., N.C., and H.G.) then checked each review for accuracy. Disagreements were resolved through discussion or, when necessary, a third review. Records stemming from the same study were identified through keyword and author searches and confirmed via manual review. These records were linked and duplicate analyses were removed. The net agreement was assessed via Fleiss' κ[134] and individual agreement between each reviewer and the primary reviewer (A.P.) was monitored via Cohen's κ[135]. Record screening and data management were conducted in DistillerSR (Evidence Partners, Ottawa, Canada).

Scoping reviews map broad, previously uncharted bodies of literature; thus, Arksey and O'Malley's framework allows for the reflexive adaptation of eligibility criteria to ensure scope remains manageable[106]. We made three such adaptations according to a predefined process. First, we added a second, condition-specific abstract screening phase because limited disease-area knowledge led to the overinclusion of ineligible records. Exclusion during this phase was restricted to criteria associated with disease-specific knowledge. Due to the volume of relevant literature, we limited full-text review in PD and MS to literature published during or after 2016. Therefore, maps of PD and MS must be interpreted as trends in recent research, rather than an exhaustive tabulation of evidence. These maps were monitored for data saturation, defined as "the point in the research process when no new information is discovered in data analysis… [and the] researcher can be reasonably assured that further data collection would yield similar results."[136] Maps of COPD and PFF remained exhaustive. Finally, we shifted from a parallel full-text review paradigm, in which all records are reviewed independently in duplicate, to the review/quality-control paradigm described above. All changes were made between the abstract and full-text review stages, approved by the study team, and applied to all records.

Table 3. Psychometric properties mapped in this review.

Property	Maps generated in this review
Known-groups validity	Number and proportion of analyses per DMO and medical condition, which found a statistically significant difference (1) between pathological and healthy gait, or (2) between disease severity strata
Convergent validity	Number and proportion of analyses per DMO and medical condition, which found a statistically significant, cross-sectional association between a DMO and validated measures of relevant constructs (e.g., disease severity, physical function, health-related quality of life, etc.)
Predictive validity	Number and proportion of analyses per DMO and medical condition, which found a statistically significant association between a DMO measured at baseline and a clinically relevant outcome at follow-up (i.e., mortality, physical function, healthcare utilization, etc.)
Responsiveness to intervention	Number and proportion of analyses per DMO and medical condition, which found a significant difference between experimental and control groups in an interventional study
Ecological validity	DMOs measured in clinical and real-world settings were mapped separately and trends were compared qualitatively

Collating, summarizing, and reporting results

We systematically mapped the results of eligible records through frequency analysis, which was stratified by medical condition and DMO. Each map reflects the volume and outcome of existing evidence on a specific association, relationship, or characteristic of DMOs, as described in Table 3. For our purposes, "volume" of evidence refers to the number of unique analyses reported in the corpus, and "outcome" refers to the proportion of these analyses which yielded statistically significant results according to the authors' original analyses. It is important to note that this approach describes an association's repeatability, rather than its strength. The latter would be more appropriately described by measures of effect size, but because effect sizes were not consistently reported or interpreted in included studies, they were not feasible to map. This issue should be the focus of future systematic reviews and meta-analyses.

For ease of interpretation, DMOs were organized into the previously established domains Pace, Rhythm, Phase, Base of Support, Variability, and Symmetry[137–141]. DMOs within each of these domains are known to exhibit similar characteristics and are highly inter-correlated. Step count, walking time, walking bout length, or duration were categorized as "Volume of Walking." Operational definitions of these domains and their associated DMOs are provided in Supplementary Table 3.

Assessing risk of bias

Critical appraisal of individual studies is not required for scoping reviews[132], and was not conducted here. However, given the heterogeneity of included records, it was necessary to identify sources of bias and assess sensitivity to study design before we could confidently interpret our maps.

Because our data were comprised of statistical tests reported by the authors themselves, it was important to consider the potential influence of statistical heterogeneity, multiple testing, and false positives. Inclusion of underpowered analyses or those unadjusted for multiple testing could inflate the observed repeatability of mapped relationships. It was, therefore, necessary to confirm whether observed proportions exceeded type 1 error rates (i.e., false positives) expected under conservative conventional assumptions. Observed distributions were subjected to single-population Bernoulli hypothesis tests with a mean of zero and an expected false-positive rate of 5% (assuming the conventional $\alpha = 0.05$). These hypothesis tests were then adjusted for multiple testing through a Benjamini–Hochberg procedure[142]. When the proportion of studies reporting statistically significant results exceeded expected false-positive rates, this was indicated on the maps. The proportions themselves were not adjusted and should

be interpreted with the potential impact of statistical heterogeneity in mind.

We also assessed clinically plausible sources of bias and effect modification in the entire corpus through manual inspection and random-effects meta-regression[97,143,144]. Potential effect modifiers included the speed and length of walking bouts, statistical analysis methods, and the size, median age, and disease severity of study populations. Supplementary methods and variable definitions are provided in Supplementary Note 3 and Supplementary Table 4. Associations between study outcomes and potential effect modifiers were modeled on the entire corpus through univariate logistic regression assuming random-effects per study. Models were subsequently adjusted for medical condition, research question, and DMO domain and significance tests were adjusted through a Benjamini–Hochberg procedure[142]. In a sensitivity analysis, unreported study characteristics were treated as missing and multiple imputed using the method of chained equations and assuming the missing-at-random hypothesis[97,145]. Data analysis was conducted in R (version 3.6.1)[146].

Qualitative appraisal of the evidence

We defined a systematic qualitative appraisal protocol to synthesize and interpret our maps. For each medical condition, we compiled and appraised evidence related to DMOs' known-groups validity, convergent validity, predictive validity, responsiveness, and ecological validity. Each DMO's overall rating in each condition describes its potential for further validation as a clinical endpoint according to current evidence. The appraisal protocol is provided in Supplementary Tables 5 and 6.

Consulting with relevant stakeholders

Levac et al. recommend that research teams involve stakeholders throughout the review process, as stakeholders can provide nuanced insights beyond those reported in the literature[107]. We regularly discussed review design and data interpretation with clinical, technical, epidemiological, regulatory, academic, and industry subject matter experts. Patients were not directly involved in this review. However, we plan to use these results in future priority-setting exercises with patient representatives.

DATA AVAILABILITY

The datasets generated during and/or analyzed during the current study are available in the OSF project repository (https://osf.io/k7395).

REFERENCES

1. Kohn, C. G., Baker, W. L., Sidovar, M. F. & Coleman, C. I. Walking speed and health-related quality of life in multiple sclerosis. *Patient* **7**, 55–61 (2014).
2. Greene, B. R., McManus, K., Redmond, S. J., Caulfield, B. & Quinn, C. C. Digital assessment of falls risk, frailty, and mobility impairment using wearable sensors. *npj Digit. Med.* **2**, 1–7 (2019).
3. Creaby, M. W. & Cole, M. H. Gait characteristics and falls in Parkinson's disease: a systematic review and meta-analysis. *Parkinsonism Relat. Disord.* **57**, 1–8 (2018).
4. Duan-Porter, W. et al. Hospitalization-associated change in gait speed and risk of functional limitations for older adults. *J. Gerontol. Ser. A* **74**, 1657–1663 (2019).
5. Hardy, S. E., Perera, S., Roumani, Y. F., Chandler, J. M. & Studenski, S. A. Improvement in usual gait speed predicts better survival in older adults. *J. Am. Geriatr. Soc.* **55**, 1727–1734 (2007).
6. Pulignano, G. et al. Incremental value of gait speed in predicting prognosis of older adults with heart failure: insights from the IMAGE-HF study. *JACC Heart Fail.* **4**, 289–298 (2016).
7. Cesari, M. et al. Prognostic value of usual gait speed in well-functioning older people;results from the health, aging and body composition study. *J. Am. Geriatr. Soc.* **53**, 1675–1680 (2005).
8. Heesen, C. et al. Patient perception of bodily functions in multiple sclerosis: gait and visual function are the most valuable. *Mult. Scler.* **14**, 988–991 (2008).
9. Day, G. S. et al. Identifying priority outcomes that influence selection of disease-modifying therapies in MS. *Neurol. Clin. Pract.* **8**, 179–185 (2018).
10. Deane, K. H. O. et al. Priority setting partnership to identify the top 10 research priorities for the management of parkinson's disease. *BMJ Open* **4**, e006434 (2015).
11. Hammarlund, C. S., Nilsson, M. H. & Hagell, P. Measuring outcomes in Parkinson's disease: a multi-perspective concept mapping study. *Qual. Life. Res.* **21**, 453–463 (2012).
12. Dobbels, F. et al. The PROactive innovative conceptual framework on physical activity. *Eur. Respir. J.* **44**, 1223–1233 (2014).
13. Port, R. J. et al. People with Parkinson's disease: What symptoms do they most want to improve and how does this change with disease duration? *J. Parkinsons. Dis.* **11**, 715–724 (2021).
14. Del Din, S., Godfrey, A., Mazzà, C., Lord, S. & Rochester, L. Free-living monitoring of Parkinson's disease: Lessons from the field. *Mov. Disord.* **31**, 1293–1313 (2016).
15. Goldhahn, J. Need for digital biomarkers in musculoskeletal trials. *Digit. Biomarkers* **1**, 82–86 (2017).
16. Stull, D. E., Leidy, N. K., Parasuraman, B. & Chassany, O. Optimal recall periods for patient-reported outcomes: challenges and potential solutions. *Curr. Med. Res. Opin.* **25**, 929–942 (2009).
17. Eastlack, M. E., Arvidson, J., Snyder-Mackler, L., Danoff, J. V. & McGarvey, C. L. Interrater reliability of videotaped observational gait-analysis assessments. *Phys. Ther.* **71**, 465–472 (1991).
18. Wrisley, D. M., Walker, M. L., Echternach, J. L. & Strasnick, B. Reliability of the dynamic gait index in people with vestibular disorders. *Arch. Phys. Med. Rehabil.* **84**, 1528–1533 (2003).
19. Zhang, J., Waubant, E., Cutter, G., Wolinsky, J. S. & Glanzman, R. EDSS variability before randomization may limit treatment discovery in primary progressive MS. *Mult. Scler. J.* **19**, 775–781 (2013).
20. McCambridge, J., Witton, J. & Elbourne, D. R. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *J. Clin. Epidemiol.* **67**, 267–277 (2014).
21. Storm, F. A., Nair, K. P. S., Clarke, A. J., Van der Meulen, J. M. & Mazzà, C. Free-living and laboratory gait characteristics in patients with multiple sclerosis. *PLoS ONE* **13**, e0196463 (2018).
22. Dorsey, E. R., Papapetropoulos, S., Xiong, M. & Kieburtz, K. The first frontier: digital biomarkers for neurodegenerative disorders. *Digit. Biomarkers* **1**, 6–13 (2017).
23. Shema-Shiratzky, S. et al. Deterioration of specific aspects of gait during the instrumented 6-min walk test among people with multiple sclerosis. *J. Neurol.* **266**, 3022–3030 (2019).
24. Iluz, T. et al. Automated detection of missteps during community ambulation in patients with Parkinson's disease: a new approach for quantifying fall risk in the community setting. *J. Neuroeng. Rehabil.* **11**, 48 (2014).
25. Kluge, F. et al. Consensus based framework for digital mobility monitoring. *PLoS ONE* **16**, e0256541. https://doi.org/10.1371/journal.pone.0256541. PMID: 34415959; PMCID: PMC8378707 (2021)..
26. Lord, S., Galna, B. & Rochester, L. Moving forward on gait measurement: toward a more refined approach. *Mov. Disord* **28**, 1534–1543 (2013).
27. Coravos, A., Khozin, S. & Mandl, K. D. Developing and adopting safe and effective digital biomarkers to improve patient outcomes. *npj Digit. Med.* **2**, 1–5 (2019).
28. Viceconti, M. et al. Toward a regulatory qualification of real-world mobility performance biomarkers in parkinson's patients using digital mobility outcomes. *Sensors* **20**, 1–13 (2020).
29. Cerreta, F. et al. Digital technologies for medicines: shaping a framework for success. *Nat. Rev. Drug Discov* **19**, 573–574 (2020).
30. Stephenson, D. et al. Precompetitive consensus building to facilitate the use of digital health technologies to support parkinson disease drug development through regulatory science. *Digit. Biomarkers* **4**, 28–49 (2020).
31. Rochester, L. et al. A roadmap to inform development, validation and approval of digital mobility outcomes: the Mobilise-D approach. *Digit. Biomarkers* **4**, 13–27 (2020).
32. Kraus, V. B. Biomarkers as drug development tools: discovery, validation, qualification and use. *Nat. Rev. Rheumatol.* **14**, 354–362 (2018).
33. Lassere, M. N. A users guide to measurement in medicine. *Osteoarthr. Cartil.* **14**, 10–13 (2006).
34. Katz, R. Biomarkers and surrogate markers: an FDA perspective. *NeuroRx* **1**, 189–195 (2004).
35. Gore, S., Blackwood, J., Guyette, M. & Alsalaheen, B. Validity and reliability of accelerometers in patients With COPD. *J. Cardiopulm. Rehabil. Prev.* **38**, 147–158 (2018).
36. Zanardi, A. P. J. et al. Gait parameters of Parkinson's disease compared with healthy controls: a systematic review and meta-analysis. *Sci. Rep.* **11**, 1–13 (2021).
37. Zago, M., Sforza, C., Bonardi, D. R., Guffanti, E. E. & Galli, M. Gait analysis in patients with chronic obstructive pulmonary disease: a systematic review. *Gait Posture* **61**, 408–415 (2018).
38. König, N., Singh, N. B., Baumann, C. R. & Taylor, W. R. Can gait signatures provide quantitative measures for aiding clinical decision-making? A systematic meta-analysis of gait variability behavior in patients with parkinson's disease. *Front. Hum. Neurosci.* **10**, 319 (2016).
39. Fox, K. M. et al. Mobility after hip fracture predicts health outcomes. *J. Am. Geriatr. Soc.* **46**, 169–173 (1998).
40. Griffiths, F. et al. Evaluating recovery following hip fracture: a qualitative interview study of what is important to patients. *BMJ Open* **5**, e005406 (2015).
41. Knox, K. B., Clay, L., Stuart-Kobitz, K. & Nickel, D. Perspectives on walking from people with multiple sclerosis and reactions to video self-observation. *Disabil. Rehabil.* **42**, 211–218 (2018).
42. Thinuan, P. et al. Narratives of older persons' frailty and physical activity in relation to environmental landscapes and time. *Int. J. Older People Nurs.* **15**, e12298 (2020).
43. Parry, R. et al. "The whole perimeter is difficult": Parkinson's disease and the conscious experience of walking in everyday environments. *Disabil. Rehabil.* **41**, 2784–2791 (2019).
44. Hammarlund, C. S., Andersson, K., Andersson, M., Nilsson, M. H. & Hagell, P. The significance of walking from the perspective of people with Parkinson's disease. *J. Parkinsons. Dis.* **4**, 657–663 (2014).
45. Mobilise-D. Mobilise-D Homepage. https://www.mobilise-d.eu/ (2021).
46. Polhemus, A. M. Mobilise-D scoping review. *OSF* https://osf.io/k7395 (2021).
47. Amboni, M. et al. Step length predicts executive dysfunction in Parkinson's disease: a 3-year prospective study. *J. Neurol.* **265**, 2211–2220 (2018).
48. Morris, R. et al. Gait rather than cognition predicts decline in specific cognitive domains in early pParkinson's disease. *J. Gerontol. Ser. A* **72**, 1656–1662 (2017).
49. Herman, T., Shema-Shiratzky, S., Arie, L., Giladi, N. & Hausdorff, J. M. Depressive symptoms may increase the risk of the future development of freezing of gait in patients with Parkinson's disease: findings from a 5-year prospective study. *Parkinsonism Relat. Disord.* **60**, 98–104 (2019).
50. Mico-Amigo, M. E. et al. Potential markers of progression in idiopathic parkinson's disease derived from assessment of circular gait with a single body-fixed-sensor: a 5 year longitudinal study. *Front. Hum. Neurosci.* **13**, 59 (2019).
51. Combs-Miller, S. A. & Moore, E. S. Predictors of outcomes in exercisers with Parkinson disease: a two-year longitudinal cohort study. *NeuroRehabilitation* **44**, 425–432 (2019).

52. Lindholm, B., Nilsson, M. H., Hansson, O. & Hagell, P. The clinical significance of 10-m walk test standardizations in Parkinson's disease. *J. Neurol.* **265**, 1829–1835 (2018).

53. Lord, S. et al. Natural history of falls in an incident cohort of Parkinson's disease: early evolution, risk and protective features. *J. Neurol.* **264**, 2268–2276 (2017).

54. Lord, S. et al. Predicting first fall in newly diagnosed Parkinson's disease: Insights from a fall-naive cohort. *Mov. Disord.* **31**, 1829–1836 (2016).

55. Vitório, R. et al. The combination of clinical scales and walking measures to predict falls in Parkinson's disease: Does the length of the prospective follow-up period matter? *Mov. Disord.* **32**, 857–859 (2017).

56. Ma, L. & Chan, P. Fall history and gait variability can predict falls in individuals with Parkinson's disease [abstract]. *Mov. Disord.* **34**, suppl 2 (2019) https://www.mdsabstracts.org/abstract/fall-history-and-gait-variability-can-predictfalls-in-individuals-with-parkinsons-disease/. Accessed 20 Sept 2021.

57. Srulijes, K. et al. Fall risk in relation to individual physical activity exposure in patients with different neurodegenerative diseases: a pilot study. *Cerebellum* **18**, 340–348 (2019).

58. Adams, E. V. et al. Predictors of functional improvements after therapeutic yoga intervention for people with Parkinson's disease. *Int. J. Yoga Ther.* **26**, 26 (2019).

59. Koch, M. W., Mostert, J., Uitdehaag, B. & Cutter, G. Clinical outcome measures in SPMS trials: an analysis of the IMPACT and ASCEND original trial data sets. *Mult. Scler. J.* **26**, 1540–1549 (2019).

60. Cadavid, D. et al. The EDSS-Plus, an improved endpoint for disability progression in secondary progressive multiple sclerosis. *Mult. Scler.* **23**, 94–105 (2017).

61. Grobelny, A. et al. Predicting falls in multiple sclerosis: utility of visual perceptive computing (P2.131). *Neurology* **86**, 16 Supplement (2016).

62. Tajali, S. et al. Predicting falls among patients with multiple sclerosis: comparison of patient-reported outcomes and performance-based measures of lower extremity functions. *Mult. Scler. Relat. Disord.* **17**, 69–74 (2017).

63. Johansson, S. et al. Participation in social/lifestyle activities in people with multiple sclerosis: changes across 10 years and predictors of sustained participation. *Mult. Scler.* **26**, 1775–1784 (2020).

64. Block, V. J. et al. Association of continuous assessment of step count by remote monitoring with disability progression among adults with multiple sclerosis. *JAMA Netw. Open* **2**, e190570 (2019).

65. Allali, G. et al. Gait variability in multiple sclerosis: a better falls predictor than EDSS in patients with low disability. *J. Neural Transm.* **123**, 447–450 (2016).

66. Jones, M. A. et al. P262 Can we improve upon clinician prediction of survival in advanced COPD using clinically measurable prognostic factors? *Thorax* **74**, A233 (2019).

67. Benzo, R. et al. Factors to inform clinicians about the end of life in severe chronic obstructive pulmonary disease. *J. Pain Symptom Manag.* **46**, 491–499.e4 (2013).

68. Pilsworth, S., Crane, J., Wat, D., Sibley, S. & Shaw, M. A service evaluation to assess the accuracy of the gold standard framework proactive indicator guidance (GSF PIG) in predicting 12 month mortality in patients with a diagnosis with chronic obstructive pulmonary disease. *Thorax* **72**, A98 (2017).

69. Kon, S. S. C. et al. Gait speed is a predictor of mortality following hospitalisation for acute exacerbations of COPD. *Thorax* **69**, A45 (2014).

70. Kon, S. et al. Gait Speed as a predictor of mortality in COPD. *Eur. Respir. J.* **46**, 123–132 (2015).

71. Kang, Y., Steele, B. G., Burr, R. L. & Dougherty, C. M. Mortality in advanced chronic obstructive pulmonary disease and heart failure following cardiopulmonary rehabilitation. *Biol. Res. Nurs.* **20**, 429–439 (2018).

72. Demeyer, H. et al. Physical activity is associated with attenuated disease progression in chronic obstructive pulmonary disease. *Med. Sci. Sports Exerc.* **51**, 833–840 (2019).

73. Neumannova, K. et al. Association between number of steps and mortality in patients with severe chronic obstructive pulmonary disease. *Eur. Respir. J.* **50**, PA4698 (2017).

74. Durheim, M. T. et al. Physical function as measured by 6-minute walk distance or accelerometry predicts clinical outcomes in COPD patients independent of gold 2011. *Am. J. Respir. Crit. Care Med.* **189**, A6679 (2014).

75. Waschki, B. et al. Physical activity is the strongest predictor of all-cause mortality in patients with COPD: a prospective cohort study. *Chest* **140**, 331–342 (2011).

76. Cushen, B. et al. A pilot study to monitor changes in spirometry and lung volume, following an exacerbation of chronic obstructive pulmonary disease (COPD), as part of a supported discharge program. *Respir. Med.* **119**, 55–62 (2016).

77. Sievi, N. A. et al. No impact of exacerbation frequency and severity on the physical activity decline in COPD: a long-term observation. *Int. J. Chron. Obstruct. Pulmon. Dis.* **14**, 431–437 (2019).

78. Spacht, A. et al. Grip weakness increases 30-day readmission for patients with acute exacerbation of copd. *J. Gen. Intern. Med.* **33**, 211–212 (2018).

79. Hershkovitz, A., Beloosesky, Y. & Brill, S. Mobility assessment of hip fracture patients during a post-acute rehabilitation program. *Arch. Gerontol. Geriatr.* **55**, 35–41 (2012).

80. Gherardini, S. et al. Prognostic implications of predischarge assessment of gait speed after hip fracture surgery. *J. Geriatr. Phys. Ther.* **42**, 148–152 (2019).

81. Cheng, F. Y., Yang, Y. R., Wu, Y. R., Cheng, S. J. & Wang, R. Y. Effects of curved-walking training on curved-walking performance and freezing of gait in individuals with Parkinson's disease: a randomized controlled trial. *Parkinsonism Relat. Disord.* **43**, 20–26 (2017).

82. Cheng, W. Y. et al. Smartphone-based continuous mobility monitoring of Parkinsons disease patients reveals impacts of ambulatory bout length on gait features. In *2017 IEEE Life Sciences Conference* 166–169 (IEEE, 2017).

83. Del Din, S. et al. Analysis of free-living gait in older adults with and without parkinson's disease and with and without a history of falls: identifying generic and disease-specific characteristics. *J. Gerontol. Ser. A* **74**, 500–506 (2019).

84. Del Din, S., Godfrey, A., Galna, B., Lord, S. & Rochester, L. Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length. *J. Neuroeng. Rehabil.* **13**, 46 (2016).

85. Galperin, I. et al. Associations between daily-living physical activity and laboratory-based assessments of motor severity in patients with falls and Parkinson's disease. *Parkinsonism Relat. Disord.* **62**, 85–90 (2019).

86. Klenk, J. et al. Ambulatory activity components deteriorate differently across neurodegenerative diseases: a cross-sectional sensor-based study. *Neurodegener. Dis.* **16**, 317–323 (2016).

87. Landers, M. R., Hatlevig, R. M., Davis, A. D., Richards, A. R. & Rosenlof, L. E. Does attentional focus during balance training in people with Parkinson's disease affect outcome? A randomised controlled clinical trial. *Clin. Rehabil.* **30**, 53–63 (2016).

88. Ardle, R. M., Galna, B., Del Din, S., Thomas, A. J. & Rochester, L. Gait impairments in dementia subtypes: considering the impact of environmental context. *Alzheimer's Dement.* **15**, P661–P662 (2019).

89. Engelhard, M. M., Patek, S. D., Lach, J. C. & Goldman, M. D. Real-world walking in multiple sclerosis: separating capacity from behavior. *Gait Posture* **59**, 211–216 (2018).

90. Brown, T. R. & Simnad, V. I. A randomized crossover trial of dalfampridine extended release for effect on ambulatory activity in people with multiple sclerosis. *Int. J. MS Care* **18**, 170–176 (2016).

91. Stellmann, J. P. et al. Fampridine and real-life walking in multiple sclerosis: low predictive value of clinical test for habitual short-term changes. *J. Neurol. Sci.* **368**, 318–325 (2016).

92. Chitnis, T. & Glanz, B. I. Quantifying neurologic disease using biosensor measurements in-clinic and in free-living settings in multiple sclerosis. *npj Dig. Med.* **2**, 123 (2019).

93. Yilmaz, D., Dodd, J. W., Shrikrishna, D., Hopkinson, N. & Jones, P. W. Walking behaviours of COPD patients following hospitalisation with acute exacerbation vs. Stable COPD and age matched control. *Am. J. Respir. Crit. Care Med.* **187**, 1418–1422 (2013).

94. Cohen, M. D. & Cutaia, M. A novel approach to measuring activity in chronic obstructive pulmonary disease: using 2 activity monitors to classify daily activity. *J. Cardiopulm. Rehabil. Prev.* **30**, 186–194 (2010).

95. Sherrington, C., Lord, S. R. & Herbert, R. D. A randomized controlled trial of weight-bearing versus non-weight-bearing exercise for improving physical ability after usual care for hip fracture. *Arch. Phys. Med. Rehabil.* **85**, 710–716 (2004).

96. Tsauo, J. Y., Leu, W. S., Chen, Y. T. & Yang, R. S. Effects on function and quality of life of postoperative home-based physical therapy for patients with hip fracture. *Arch. Phys. Med. Rehabil.* **86**, 1953–1957 (2005).

97. Deeks, J. J., Higgins, J. P.T., Altman, D. G. in *Cochrane Handbook for Systematic Reviews of Interventions Version 6.2* (eds. Higgins, J. et al.). www.training.cochrane.org/handbook (2021).

98. Mantua, V., Arango, C., Balabanov, P. & Butlen-Ducuing, F. Digital health technologies in clinical trials for central nervous system drugs: an EU regulatory perspective. *Nat. Rev. Drug Discov.* **20**, 83–84 (2021).

99. Izmailova, E. S. et al. Remote digital monitoring for medical product development. *Clin. Transl. Sci.* **14**, 94–101 (2021).

100. Goldsack, J. C., Dowling, A. V., Samuelson, D., Patrick-Lake, B. & Clay, I. Evaluation, acceptance, and qualification of digital measures: from proof of concept to endpoint. *Digit. Biomarkers* **5**, 53–64 (2021).

101. Goldsack, J. C. et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). *npj Digit. Med.* **3**, 1–15 (2020).

102. Clinical Trials Transformation Initiative. Developing novel endpoints generated by mobile technology for use in clinical trials. https://www.ctti-clinicaltrials.org/

briefing-room/recommendations/developing-novel-endpoints-generated-mobile-technology-use-clinical (2021).

103. Stuck, A. K., Bachmann, M., Füllemann, P., Josephson, K. R. & Stuck, A. E. Effect of testing procedures on gait speed measurement: a systematic review. *PLoS ONE* **15**, e0234200 (2020).

104. Friesen, K. B., Zhang, Z., Monaghan, P. G., Oliver, G. D. & Roper, J. A. All eyes on you: how researcher presence changes the way you walk. *Sci. Rep.* **10**, 1–8 (2020).

105. 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.* https://doi.org/10.3233/jpd-202471, 1–13 (2021).

106. Arksey, H. & O'Malley, L. Scoping studies: towards a methodological framework. *Int. J. Soc. Res. Methodol. Theory Pract.* **8**, 19–32 (2005).

107. Levac, D., Colquhoun, H. & O'Brien, K. K. Scoping studies: advancing the methodology. *Implement. Sci.* **5**, 69 (2010).

108. Middleton, A., Fritz, S. L. & Lusardi, M. Walking speed: the functional vital sign. *J. Aging Phys. Act.* **23**, 314–322 (2015).

109. Fritz, S. & Lusardi, M. White paper: walking speed: the sixth vital sign. *J. Geriatr. Phys. Ther.* **32**, 46–49 (2009).

110. Atrsaei, A. et al. Gait speed in clinical and daily living assessments in Parkinson's disease patients: performance versus capacity. *npj Park. Dis.* **7**, 24 (2021).

111. Shah, V. V. et al. Laboratory versus daily life gait characteristics in patients with multiple sclerosis, Parkinson's disease, and matched controls. *J. Neuroeng. Rehabil.* **17**, 159 (2020).

112. Khandelwal, S. & Wickström, N. Evaluation of the performance of accelerometer-based gait event detection algorithms in different real-world scenarios using the MAREA gait database. *Gait Posture* **51**, 84–90 (2017).

113. European Medicines Agency. *Revised Framework for Interaction Between the European Medicines Agency and Patients and Consumers and Their Organisations* (European Medicines Agency, 2014).

114. Geissler, J., Ryll, B., di Priolo, S. L. & Uhlenhopp, M. Improving patient involvement in Medicines Research and Development:: A Practical Roadmap. *Ther. Innov. Regul. Sci.* **51**, 612–619 (2017).

115. US Food and Drug Administration. CDRH patient engagement. *CDRH Patient Science and Engagement Program* https://www.fda.gov/about-fda/cdrh-patient-science-and-engagement-program/cdrh-patient-engagement (2021).

116. Haerry, D. et al. EUPATI and patients in medicines research and development: guidance for patient involvement in regulatory processes. *Front. Med.* **5**, 230 (2018).

117. Guidance. EUPATI Toolbox. https://toolbox.eupati.eu/guidance/ (2021).

118. Clinical Trials Transformation Initiative. Recommendations on effective engagement with patient groups around clinical trials. https://www.ctti-clinicaltrials.org/briefing-room/recommendations/ctti-recommendations-effective-engagement-patient-groups-around (2021).

119. Dujmovic, I. et al. Gait pattern in patients with different multiple sclerosis phenotypes. *Mult. Scler. Relat. Disord.* **13**, 13–20 (2017).

120. Filli, L. et al. Profiling walking dysfunction in multiple sclerosis: characterisation, classification and progression over time. *Sci. Rep.* **8**, 1–13 (2018).

121. Vervoort, G. et al. Progression of postural control and gait deficits in Parkinson's disease and freezing of gait: A longitudinal study. *Park. Relat. Disord.* **28**, 73–79 (2016).

122. Sturchio, A. et al. Kinematic but not clinical measures predict falls in Parkinson-related orthostatic hypotension. *J. Neurol.* **268**, 1006–1015 (2021).

123. Galna, B., Lord, S., Burn, D. J. & Rochester, L. Progression of gait dysfunction in incident Parkinson's disease: Impact of medication and phenotype. *Mov. Disord.* **30**, 359–367 (2015).

124. Kantorowski, A. et al. Determinants and outcomes of change in physical activity in COPD. *ERJ Open Res.* **4**, 00054–02018 (2018).

125. Pfeufer, D. et al. Weight bearing in patients with femoral neck fractures compared to pertrochanteric fractures: a postoperative gait analysis. *Injury* **50**, 1324–1328 (2019).

126. 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* **20**, 2631 (2020).

127. Peebles, A. T., Reinholdt, A., Bruetsch, A. P., Lynch, S. G. & Huisinga, J. M. Dynamic margin of stability during gait is altered in persons with multiple sclerosis. *J. Biomech.* **49**, 3949–3955 (2016).

128. Peebles, A. T., Bruetsch, A. P., Lynch, S. G. & Huisinga, J. M. Dynamic balance in persons with multiple sclerosis who have a falls history is altered compared to non-fallers and to healthy controls. *J. Biomech.* **63**, 158–163 (2017).

129. Ihlen, E. A. F. Improved prediction of falls in community-dwelling older adults through phase-dependent entropy of daily-life walking. *Front. Aging Neurosci.* **10**, 44 (2018).

130. Ihlen, E. A. F., Weiss, A., Bourke, A., Helbostad, J. L. & Hausdorff, J. M. The complexity of daily life walking in older adult community-dwelling fallers and non-fallers. *J. Biomech.* **49**, 1420–1428 (2016).

131. Toebes, M. J. P., Hoozemans, M. J. M., Furrer, R., Dekker, J. & Van Dieën, J. H. Local dynamic stability and variability of gait are associated with fall history in elderly subjects. *Gait Posture* **36**, 527–531 (2012).

132. Tricco, A. C. et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. *Ann. Intern. Med.* **169**, 467 (2018).

133. Polhemus, A. M. et al. Walking-related digital mobility outcomes as clinical trial endpoint measures: protocol for a scoping review. *BMJ Open* **10**, e038704 (2020).

134. Fleiss, J. L. Measuring nominal scale agreement among many raters. *Psychol. Bull.* **76**, 378–382 (1971).

135. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **20**, 37–46 (1960).

136. Faulkner, S. L. & Trotter, S. P. Data Saturation. in *The International Encyclopedia of Communication Research Methods* (eds Matthes, J., Davis, C. S. & Potter, R. F.) 1–2 (Wiley, 2017).

137. Thingstad, P. et al. Identification of gait domains and key gait variables following hip fracture. *BMC Geriatr.* **15**, 150 (2015).

138. Morris, R. et al. A model of free-living gait: a factor analysis in Parkinson's disease. *Gait Posture* **52**, 68–71 (2017).

139. Lahousse, L. et al. Gait patterns in COPD: the Rotterdam Study. *Eur. Respir. J.* **46**, 88–95 (2015).

140. Hollman, J. H., McDade, E. M. & Petersen, R. C. Normative spatiotemporal gait parameters in older adults. *Gait Posture* **34**, 111–118 (2011).

141. Lord, S. et al. Independent domains of gait in older adults and associated motor and nonmotor attributes: validation of a factor analysis approach. *J. Gerontol. Ser. A* **68**, 820–827 (2013).

142. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).

143. Boutron, I. et al. in *Cochrane Handbook for Systematic Reviews of Interventions Version 6.2* (eds Higgins, J. et al). www.training.cochrane.org/handbook (2021).

144. Thompson, S. G. & Sharp, S. J. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat. Med.* **18**, 2693–2708 (1999).

145. Donders, A. R. T., van der Heijden, G. J. M. G., Stijnen, T. & Moons, K. G. M. Review: a gentle introduction to imputation of missing values. *J. Clin. Epidemiol.* **59**, 1087–1091 (2006).

146. R Core Team. *R: A Language and Environment for Statistical Computing* (R Core Team, 2020).

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

Author A.P. led and coordinated the project. A.P., L.D.O., G.B., H.G., K.T., M.B.d.B., A.M., L.S., B.V., H.D., N.H., W.M., T.T., I.V., A.Y., C.B., J.G.-A., L.L., C.M., L.R., B.S., A.F., and M.P. contributed to the conceptualization and design of the review. G.B., K.T., A.S.M., L.S., B.V., H.D., N.H., W.M., T.T,. I.V., A.Y., C.B., J.G.-A, L.L., C.M., L.R., B.S., A.F., and M.P. contributed as subject matter experts and consulted on study scope, search strategies, and reference material design. A.P., L.D.O., G.B., N.C., H.G., C.K., K.T., A.S.M., L.S., and B.V. contributed to search strategies, form design, and reference material design. A.P., L.D.O, G.B., N.C., F.S., H.G., C.K., R.R., D.B., G.C., F.K., S.K., M.E.M.-A., C.N., R.B., M.B.d.B., E.B., C.H., K.S. contributed to abstract screening and A.P., L.D.O., G.B., N.C., F.S., H.G., M.G., C.K., R.R., K.T., D.B., S.Breuls, S.Buttery, C.E., J.G., A.K., S.K., M.E.M.-A., C.S., P.W., M.B.d.B., S.S., J.v.U., and A.C. contributed to full-text review. A.P. and L.D.O. contributed to data analysis. M.P. acted as the primary supervisor, and A.C., H.D., N.H., W.M., T.T., I.V., A.Y., C.B., J.G.-A., L.L., C.M., L.R., B.S., and A.F. acted as supervisors for the condition-specific subgroups. A.P. wrote the manuscript, and all authors reviewed and approved the final version.

COMPETING INTERESTS

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-021-00513-5.

Correspondence and requests for materials should be addressed to Ashley Polhemus.

Reprints and permission information is available at http://www.nature.com/reprints

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