Long-term unsupervised mobility assessment in movement disorders

Elke Warmerdam, Jeffrey M. Hausdorff, Arash Atrsaei, Yuhan Zhou, Anat Mirelman, Kamiar Aminian, Alberto J. Espay, Clint Hansen, Luc J.W. Evers, Andreas Keller, Claudine Lamoth, Andrea Pilotto, Lynn Rochester, Gerhard Schmidt, Bastiaan R. Bloem, Walter Maetzler

Department of Neurology, Kiel University, Kiel, Germany (E Warmerdam MSc, C Hansen PhD, Prof W Maetzler MD)

Faculty of Engineering, Kiel University, Kiel, Germany (E Warmerdam MSc, Prof G Schmidt PhD)

Center for the Study of Movement, Cognition, and Mobility, Tel Aviv Sourasky Medical Center; Department of Physical Therapy, Sackler Faculty of Medicine and Sagol School of Neuroscience, Tel Aviv University, Israel (Prof J M Hausdorff PhD, Prof A Mirelman PhD)

Rush Alzheimer's Disease Center and Department of Orthopaedic Surgery, Rush University Medical Center, Chicago, Illinois, USA (Prof J M Hausdorff PhD)

Laboratory of Movement Analysis and Measurement, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland (A Atrsaei MSc, Prof K Aminian PhD)

University of Groningen, University Medical Center Groningen, Center for Human Movement Sciences, Groningen, The Netherlands (Y Zhou MSc, C Lamoth PhD)

Radboud University Medical Center; Donders Institute for Brain, Cognition and Behaviour; Department of Neurology; Nijmegen, The Netherlands (LJW Evers BSc, Prof B R Bloem PhD)

Gardner Family Center for Parkinson's Disease and Movement Disorders, University of Cincinnati, Cincinnati, Ohio, USA (Prof A J Espay MD)

Center for Bioinformatics, Chair for Clinical Bioinformatics, Saarland University, Saarbrücken, Germany (Prof A Keller PhD)

Stanford University School of Medicine Office, Department of Neurology and Neurological Sciences (Prof A Keller PhD)

Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy (A Pilotto MD)

Parkinson's Disease Rehabilitation Centre, FERB ONLUS – S. Isidoro Hospital, Trescore Balneario (BG), Italy (A Pilotto MD)

Institute of Neuroscience, Newcastle University, Newcastle Upon Tyne, UK (Prof L Rochester PhD)

Newcastle upon Tyne Hospitals National Health Service Foundation Trust, Newcastle upon Tyne, UK (Prof L Rochester PhD)

Corresponding author: Prof. Jeffrey M. Hausdorff, Center for the Study of Movement, Cognition, and Mobility, Neurological Institute, Tel Aviv Sourasky Medical Center, 6 Weizmann Street, Tel Aviv, Israel, Tel: +972 3 697 3081, Fax: +972 3 697 4911, E-mail: jhausdor@tlvmc.gov.il

Glossary of unusual terms and terms used in the biomechanical field

Cadence	The rate at which a person steps (about 110-115 steps/min in healthy adults).
Chair rise peak power	The maximum power that is exerted to lift the body's centre of mass during a sit to stand movement.[1]
Daily-living	This term, also referred to as free-living, real-world, or community-living, is used to distinguish testing within the normal environment of a participant from testing in a standardized setting, such as in the clinic or laboratory setting.
Hawthorne effect	The change in behaviour of participants because of the awareness of being studied.[2]
Inertial measurement units	Sensors that measure acceleration and/or angular velocity. They can measure the quality and quantity of movement using specifically developed algorithms.
Median walking acceleration	The median of the magnitude of the acceleration during walking.
Mobile health technologies	Umbrella term for wearable or portable / domestic-integrated devices that can provide objective measures and that include digital applications, as well as body-worn (adhered to a body surface, mainly inertial measurement units) or frequently used patient-centred devices (e.g., smartphone, keyboard). Here we focus on technologies that can measure the frequency and quality of movement and mobility characteristics.[3]
Postural transitions	Sit to stand and stand to sit movements, or turning over in bed.
Reverse white coat effect	A change (typically an improvement) in a clinical parameter because it is measured in a clinical setting (it can be seen as the opposite of the white coat syndrome in hypertension).
Stance time	The time one leg is in contact with the surface during a step that is taken during walking.
Step time	The time it takes to perform one step (i.e., the time between initial contact of one foot and the initial contact of the contralateral foot).
Stride time	Also known as gait cycle time. This is the time to perform two steps (i.e., the time between initial contact of one foot and the next initial contact of the same foot).
Swing time	The time one leg is not in contact with the surface during a stride that is taken during walking. In healthy young adults, swing time is about 40% of the stride time. With aging and disease, the time spent in swing time often gets smaller.

Supervised assessment	This refers to the traditional, conventional mode of assessing mobility in a lab or clinical setting. Typically, this is a qualitative or semi-quantitative “one-time snapshot” evaluation of mobility that is conducted by a trained healthcare professional.
Unsupervised assessment	This refers to the quantitative assessment of mobility in the home and daily-living environment that is conducted continuously with new, mainly mobile, health technologies over relatively long periods of time.
Wearables	Mobile devices worn on the body, e.g. inertial measurement units, smartwatches or Holter electrocardiogram monitors.
White coat effect	A change (typically worsening) in a parameter because it is measured in a clinical setting (a well-known example is the white coat effect that leads to an increase in blood pressure).

Abstract

Digital health technologies that quantify mobility in unsupervised, daily-living environments are emerging as a complementary evaluation approach in neurology. Data collected in these ecologically valid, patient-relevant settings can overcome significant limitations of conventional clinical assessments. Unsupervised assessments can capture fluctuating and rare events and have the promise of supporting clinical decision-making and serving as outcomes in clinical trials. However, studies that directly compared assessments made in unsupervised and supervised (i.e. in the lab or clinical) settings point to large disparities, even in the same parameters of mobility (up to 180% difference). These differences appear to be influenced by psychological, physiological, cognitive, environmental, and technical factors and by the specific aspect of mobility and diagnosis. To facilitate the successful adaptation of the unsupervised assessment of mobility in the clinic and in clinical trials, clinicians and future work should take into account these disparities and the multiple factors that contribute to them.

Introduction

Deficits in mobility are common among neurological patients and often impact daily-living activities, work, and socialization.[4] These deficits predict morbidity, cognitive decline and mortality[5–8] and negatively affect quality of life, especially in patients with neurological movement disorders.[9,10] For example, in patients with PD, health-related quality of life is strongly associated with the activities and participation components of the International Classification of Functioning, Disability and Health (ICF model).[11] It is, therefore, crucial for healthcare professionals to obtain a full and objective evaluation of a patient's mobility as a basis for individually tailored clinical decision-making and prognostication. Currently, mobility assessments are mainly performed under supervised conditions in a lab or hospital using standardized, mostly qualitative or semi-structured evaluations.[12–14] However, many patients perform paradoxically well when they know that they are being observed. Moreover, various clinically relevant events are difficult to capture during these "snapshot" observations, either because they take place over long periods of time (e.g., amount of physical activity), are rare (e.g., falls or freezing episodes[15]), occur at night (e.g., sleep disturbances), or have complex fluctuating patterns (e.g., the response to dopaminergic treatment in Parkinson's disease, PD). To reliably evaluate such events, it is necessary to measure patients unobtrusively and for much longer periods of time, while they move about freely and unsupervised in their daily-living environment.

Recent reviews describe the promise of unsupervised assessments of mobility using novel technologies.[3,16] Although very different from other daily-living acquired parameters already used in clinical routine, such as the Holter electrocardiogram[17,18] and blood glucose monitoring[19], these reviews suggest that the long-term evaluation of mobility will soon become increasingly relevant for personalised clinical decision-making in neurology. Unsupervised assessments may save time and costs by capturing health-related data largely independent of the availability of healthcare services. This is particularly important for patients living in rural areas or developing countries where the number of healthcare professionals is small relative to the population size.[16] Finally, unsupervised assessments offer patients the opportunity to become more actively involved, e.g., by using their own devices (such as smartphones) and receiving feedback about their own daily-living performance.[20]

Unsupervised assessments of mobility can provide additional and, at least partly, complementary information, compared to supervised assessments. Still, differences with respect to the conventional evaluation need to be considered. Here we summarize the existing findings on the poor, weak association between mobility assessed in the two settings and discuss potential reasons for the observed differences. We then present suggestions for clinical care and future research to help bring about the appropriate implementation of unsupervised mobility assessment.

Unsupervised mobility assessment: State-of-the-art

Unsupervised assessments are usually done with mobile health technologies that can measure physical activity,[21–23] evaluate mobility or specific movements such as gait,[24–26] or detect specific symptoms in unsupervised environments.[27–29] The potential added value of unsupervised assessments in patients with mobility deficits has been demonstrated in several studies. For example, both predicting the risk of future falls and discriminating fallers from non-fallers in older adults[30–33] and stroke survivors[34] appears more accurate when using data collected in the unsupervised environment. Indeed, the relevance of unsupervised mobility parameters was recently acknowledged by the Food and Drug Administration (FDA)[35] and the European Medicines Agency (EMA).[36] Both regulatory agencies now encourage the inclusion of parameters from unsupervised mobility assessments as exploratory endpoints in clinical trials.

To compare the same features of mobility (i.e., gait, turns, postural transitions) in supervised and unsupervised assessments, a systematic search (Search strategy and selection criteria box) was performed. Twelve studies conducted in three different populations were identified (older adults,

PD, multiple sclerosis (MS); Supplementary Table 1). Strikingly, the same mobility parameters obtained in identical subjects differed by -40% to 180% (Figure 1). These differences are much larger than those seen after many interventions. Thus, small and even moderate treatment effects might be buried under the variations introduced merely by the measurement techniques themselves, if the differences between supervised and unsupervised assessments are not appropriately considered.

Why are supervised and unsupervised measurements different?

Several reasons could explain the substantial differences in mobility parameters when comparing supervised and unsupervised assessments (Table 1). First, unsupervised movements are typically self-initiated, embedded in a rich behavioural context, and goal-directed. In contrast, movements performed in a supervised setting are usually triggered by a command and performed in an isolated, standardized setting with limited ecological validity.[37] For example, self-initiated finger movements activate different brain structures compared to externally triggered movements.[38,39] Apparently, the brain generates supervised movements using networks that differ from those generating unsupervised movements. Moreover, with an external focus, attention is directed to the outcome of the action, e.g., leaving the room. With an internal focus, attention is directed to controlling the body parts when performing a certain movement.[40] An external focus of attention results, at least sometimes, in more fluent movements.[41]

Second, performance can be influenced by several psychological and physiological phenomena that may differ across settings. These factors include alertness, motivation, the white-coat-effect, the reverse white-coat-effect, the Hawthorne effect, as well as fatigue, pain, and stress. These effects may explain why rising from a chair was performed with lower peak power in unsupervised assessments, as compared to supervised assessments, even when these movements were performed in an identical environment and with the same equipment.[26] Similar disparities have been identified for several gait parameters.[42] Supervised assessments seemingly provide a measure of someone's best, rather than their usual performance; i.e., they capture "capacity" rather than "performance".[43,44]

Third, the environment is usually standardized in supervised conditions (e.g., walking in a clean and sterile environment, without distractions), but much more variable in unsupervised conditions (e.g., furniture, lighting, patterns, color of the environment, obstacles). This can induce large variability and asymmetry in mobility patterns, as shown by studies that assessed walking through busy corridors and through a city centre.[45,46] Different types of seats (e.g., firm chair, armchair, low couch) in unsupervised conditions can also partially explain the greater variability observed in postural transitions in daily-living.[26,32,47,48] Moreover, asymmetry can be introduced through a constrained environment that requires gait adaptation or turning in the same direction.

Fourth, multi-tasking situations are common in unsupervised environments, e.g. walking and talking, but uncommon in supervised assessments. This may further contribute to the observed differences. Even during supervised dual-task walking, the gait quality was usually "better" compared to unsupervised walking.[49]

Fifth, the presence of a partner/caregiver can also affect mobility in unsupervised conditions. Indeed, social interactions are common during every-day walking, e.g., a spouse who acts as a type of external cue to improve walking in PD patients or to relieve anxiety in someone with a cautious gait disorder.

Sixth, technical limitations may add to the differences observed. Most currently available algorithms were developed and validated in supervised environments. As the amount and variability of activities and mobility are much larger in unsupervised environments, algorithms may have difficulties with differentiating similar movements (such as picking something up from the floor and sit-to-stand movements) that were not evaluated in the supervised assessment.[50,51] Of note, only one study found in our systematic search used algorithms that were explicitly validated in both standardized

and non-standardized settings.[52] A further bias may be introduced by the use of different device locations (e.g. waist or ankle). The use of different mobile health technologies (e.g., hardware, algorithms)[47,53] could play a role but we consider this as relatively limited (Supplementary Figure 1). The validation of algorithms for unsupervised daily-living assessments brings new challenges as gold-standard references are lacking.[54,55]

Finally, the current statistical approaches for analysing supervised assessments, e.g., conventional means and standard deviations, may not be optimal for characterising complex data coming from the unsupervised arena. The supervised assessment typically involves one test, a single snapshot, whereas the unsupervised evaluation may include thousands of bouts of walking. It remains to be seen how to best compare a single value with values obtained from a distribution (or histogram). (Figure 2 and Supplementary Table). Indeed, the tails of an individual's distribution may correspond better to clinical endpoints such as fall risk, mobility, limitation in activities, frailty and supervised gait speed, than do mean and median values.[26,37,56]

Does the type of movement and disease matter?

Some types of mobility, such as postural transitions, show seemingly larger differences than others (e.g., walking) when comparing supervised to unsupervised conditions (Figure 1). This difference may even depend on specific parameters. In a study of PD patients, the velocity at the beginning of the turn was similar in unsupervised and supervised conditions but was lower at the middle, and substantially higher at the end of turns in the unsupervised condition.[53]

Interestingly, the type and severity of a disease may also influence the differences between supervised and unsupervised assessments (Figure 1).[25,57] The differences in stand-to-sit duration between both settings were smaller in older adults than PD patients.[47] Patients with MS showed a different pattern. Their gait speed was similar under supervised and unsupervised assessments,[58] while the “opposite” behaviour compared to PD and older adults was seen for stance, step, and swing time.[52] The reasons for these observations are not yet clear, but differences in physical, attentional and cognitive capabilities may contribute.[59] These differences between supervised and unsupervised performance may even be relevant at a subgroup level. The above-reported changes in turning parameters in PD patients[53] differed substantially between fallers and non-fallers, with or without fear of falling. Remarkably, fallers with fear of falling showed slower turns in the supervised assessment but *faster* turns in the unsupervised assessment, compared to other PD subgroups.[53]

How to implement in clinical routine and future research?

Unsupervised and supervised measurements of mobility often strikingly differ. As we anticipate that unsupervised assessments will become a prerequisite for future clinical decision-making and clinical trials, here we provide directions to help move this emerging field forward (see Table 2).

First, we should *acknowledge*, as neurologists and other healthcare professionals, patients, and researchers, the *limited understanding of* the association between supervised and unsupervised mobility when interpreting data obtained from *unsupervised environments*. The published literature, although not yet fully mature, strongly suggests that any extrapolation of unsupervised mobility based on our knowledge obtained from supervised mobility may be substantially influenced by the type and subtype of disease as well as the disease stage and may affect all or only parts of a patient’s mobility.

Technical limitations should be addressed, e.g., by using the same mobile health technologies with the same body location for both supervised and unsupervised measurements. The algorithms used to calculate mobility parameters should be validated, to the degree that is possible, in both settings. Another requirement to increase the utility of unsupervised measures is a harmonized reporting of parameters, for example as a core dataset, across studies. This should include the reporting of metadata, i.e., data that accompany and describe the primary data. The duration of the assessments

should be harmonized and the type of movement should be reported in detail.[60,61] We must also think of validation approaches that go beyond clinical observation and lab analysis tools. Moreover, algorithms for mobility assessment should be validated separately for each type of neurological movement disorder, as they may be associated with distinct movements patterns.[25,62] Even healthy people move differently in different ages[63,64] and with different fitness levels.[65]

Special emphasis must be placed on a more *sophisticated analysis of unsupervised data*. A promising approach is to consider and leverage specific episodes of mobility (e.g., turning, sit-to-stand and stand-to-sit movements and other movements performed regularly during the day) and novel parameters, such as the distribution and extreme values of mobility parameters (Figure 2).[26,37,56,66] These analyses have only been performed for healthy (older) adults and not yet for neurological patients. An example could be the evaluation of the effects of an experimental therapy. This might be measured best in its optimum state (visible by improvement in the supervised assessment and, e.g., the best 10% of an individual's distribution during unsupervised assessment), while the typical and worse values may also be informative of changes throughout the day. Phase 3 trials could use this information as outcomes.

Variability measures can serve as a useful example for how important it is for clinicians and principal investigators to have a profound understanding of how their treatment and compounds influence mobility in daily-life. Some variability measures are highly influenced by the environment and might therefore be measured best in a supervised setting, where it better reflects someone's capacity.[67] In the home environment, decreased variability with similar mean values might be a positive outcome if the goal of an intervention is to reduce motor response fluctuations in patients with PD. In a trial investigating generally undertreated patients, a decrease in variability associated with an improvement of mean values can indicate more consistent good performance during the day. In trials focusing on behavioural symptoms, increased variability might indicate better adaptability, more variable and enriched physical activity, and social interactions. The context is critical.

More generally, we must evaluate whether data obtained from unsupervised environments provide relevant progression and treatment response information, rather than acting as markers of routine, fixed behaviours or trait markers. Trait markers could still be good measures of progression, but appropriate interpretation is key to utility. For example, the actions performed during daily-living are very different per individual but show a surprisingly comparable pattern within an individual.[68]

Moreover, statistical approaches should take advantage of the high number of repeated, specific movements occurring during long-term observation periods in unsupervised environments (Figure 2).[26,53] Here, also deep learning, machine learning, and artificial intelligence approaches could be applied. Such algorithms that "learn" from data have shown remarkable success in making accurate predictions in complex problems that previously depended on human skills. Existing examples include referral for eye diseases[69] and detection of PD motor fluctuations.[70]

To better understand the *clinical relevance of novel outcomes in the unsupervised settings*, future work should further explore the relationship between these objective digital measures, on the one hand, with conventional measures of mobility, as well as with patient-reported outcome measures (PROMs) and caregiver reported outcomes, on the other hand. Both PROMs (in this case subjectively) and mobile health technologies (in this case objectively) offer remote measurements in the unsupervised setting, and both approaches are potentially more ecologically valid and more meaningful to patients and their caregivers than data acquired in the traditional clinical setting. Among the studies that we identified, only four assessed correlations with PROMs related to mobility, with contrasting findings (Supplementary Table).

We must also keep in mind that mobile health technologies may cause behavioural changes by themselves, even when no feedback is provided (e.g., Hawthorne effect) but especially if feedback is provided (e.g., to induce compliance). Studies are needed to investigate: (i) if and when the

performance of the user in the unsupervised setting becomes more like that in the supervised setting and (ii) whether the induced behavioural changes themselves may have therapeutic effects that could interfere with the evaluation. For example, subjects who know that they are equipped with mobile health technologies may increase their level of physical activitity, particularly when feedback about their own performance is provided.

Eventually, healthcare professionals should *interpret* their *supervised assessments cautiously*, as these could have limited value when transferred into daily-life. To improve their value, we recommend providing natural, everyday life-like situations and instructions when performing supervised assessments. Explicit goals should be given, forcing patients to focus on the goal instead of on the actual movements that must be performed to reach the goal.[60] For example, instructing a sitting person to walk allows for a more naturalistic observation of the sit-to-stand performance, because the person focuses more on the walking task rather than the necessary transition from sit-to-stand. Other opportunities to observe uninstructed movements occur when patients move in the waiting room or on their way to the clinician's office.[71] It is also essential to gain as much information as possible about the living environment of the person investigated. If the person has cluttered furniture at home, healthcare professionals might focus more on assessing mobility in small, crowded places instead of large open hospital hallways. In addition, the type of furniture, lighting, patterns and other environmental factors might be important.[72]

Considering mobility differences between the supervised and unsupervised setting can also be relevant for the measurement of other symptoms and deficits. For example, deficits in upper extremity movement occur in many neurological patients[73] and several methods have been proposed to continuously assess upper limb bradykinesia in daily-life[74]. However, a direct comparison of these various symptoms in supervised and unsupervised settings remains largely unexplored. One exception is a study that assessed habitual keyboard typing behaviour in PD patients.[75] In contrast to what we have observed for mobility, this study showed that various keystroke metrics as measured in the clinic were strongly correlated with those obtained in at-home settings. This suggests that some upper extremity performances (in this case, a measure of bradykinesia) are comparable between supervised and unsupervised conditions. These considerations again underscore the need to assess different aspects of motor functioning on a case-by-case basis.

Conclusions

There is rapidly increasing evidence that, depending on whether mobility is assessed under supervised or unsupervised conditions, the results can differ substantially.[26,37,64] Recognition of these striking differences and appreciating the importance of measurements obtained in both settings call for expanding our knowledge about unsupervised mobility (recall Table 2). With this in mind, unsupervised mobility parameters can be implemented to improve clinical care and, as primary or secondary endpoints, in intervention trials.

Contributors

EW, JH and WM developed the idea. EW performed the literature search. EW, JH, CH and WM framed the outline. BRB, JH, AA, YZ, AM, KA, AE, CH, AK, CL, AP, LR, GS and LE worked on different aspects of the manuscript. All authors commented on the manuscript and approved the final version.

Declaration of interests

EW, AA, YZ and CH have nothing to disclose.

JH has received grant support from the NIH, the Michael J Fox Foundation for Parkinson's Research, the EU (H2020), the BSF, the Israeli Science Foundation and the National Multiple Sclerosis Society. He has submitted a patent application on the use of body-fixed sensors for assessing PD symptoms, the intellectual property rights for which are held by the Tel Aviv Medical Center. He has or currently

serves on the Movement Disorders Society Technology Task Force and on Michael J Fox Foundation task force on gait, on the board of the International Society for the Measurement of Physical Behavior, and on advisory boards for Sanofi and Biogen.

AM has received grant support from the Michael J Fox Foundation for Parkinson's Research, the EU (H2020), the Israeli Science Foundation and the Ministry of Science and Technology. She has submitted a patent application on the use of body-fixed sensors for assessing PD symptoms, the intellectual property rights for which are held by the Tel Aviv Medical Center. She currently chairs the Michael J Fox Foundation task force on gait and is a consultant for Biogen Biotechnologies.

KA has received grant from Swiss National foundation, CTI, Innossuise, the EU (H2020, FP7, FP6), Lausanne Orthopedics, Research Foundations, Austria Research Promotion Agency (FFG), Swiss Multiple Sclerosis Society, Swiss heart foundations, Pierre Mercier Foundation, Swiss Federal office of sport, EHC-Hospital of Morges, Suvaliv, was co-founder of Gait Up and served as its board member until 2018.

AJE has received grant support from the NIH, Great Lakes Neurotechnologies and the Michael J Fox Foundation; personal compensation as a consultant/scientific advisory board member for Abbvie, Adamas, Acadia, Acorda, Neuroderm, Impax, Sunovion, Lundbeck, Osmotica Pharmaceutical, and USWorldMeds; publishing royalties from Lippincott Williams & Wilkins, Cambridge University Press, and Springer; and honoraria from USWorldMeds, Lundbeck, Acadia, Sunovion, the American Academy of Neurology, and the Movement Disorders Society. He serves as the chair of the MDS Technology Task Force.

AK receives grants from the Michael J Fox Foundation, the Schaller-Nikolich Foundation and is co-founder and has shares of Hummingbird Diagnostics GmbH (Heidelberg, Germany).

CL has received funding from EU Interreg IV progamme, SNN innovation grant, Data Science support program University of Groningen, for cotutelle PhD's with the Universitė Grenobles Aples, France. Professional board member: Hanze University of Applied Science, Master Sensor System Engineering – Institute of Engineering. Speaker's fee conference geriatric care, Maastricht. Expert review-panel fee for EU ACTIVAGE project.

LE is funded by Topconsortium voor Kennis en Innovatie (TKI) Life Sciences & Health (TKI-LSH), Michael J. Fox Foundation, Philips Research and ZonMW (91215076 "Big Data for Personalised Medicine").

AP has received speaker honoraria from BioMarin Pharmaceutical, Chiesi Pharmaceuticals, Nutricia Pharmaceuticals, UCB Pharma and Zambon Pharmaceuticals; personal compensation as consultant/advisory board for Z-cube (Technology division of Zambon Pharmaceuticals); travel grants from AbbVie Pharmaceuticals, BioMarin Pharmaceutical, Nutricia Pharmaceuticals, Zambon Pharmaceuticals and the Italian movement disorder society.

LR receives grant support from the European Union, Parkinson's UK, the Stroke Association, The Medical research Council, National Institute of Health Research, New Zealand Health Research Council, and has served on Advisory Board for Biogen.

GS receives funding from the German Research Foundation (DFG).

BRB currently serves as Associate Editor for the Journal of Parkinson's disease, serves on the editorial of Practical Neurology and Digital Biomarkers, has received honoraria from serving on the scientific advisory board for Abbvie, Biogen, UCB and Walk with Path, has received fees for speaking at conferences from AbbVie, Zambon, Roche, GE Healthcare and Bial, and has received research support from the Netherlands Organization for Scientific Research, the Michael J Fox Foundation, UCB, Abbvie, the Stichting Parkinson Fonds, the Hersenstichting Nederland, the Parkinson's

Foundation, Verily Life Sciences, Horizon 2020, the Topsector Life Sciences and Health, and the Parkinson Vereniging.

WM receives or received funding from the European Union, the German Federal Ministry of Education of Research, Michael J. Fox Foundation, Robert Bosch Foundation, Neuroalliance, Lundbeck and Janssen. He received speaker honoraria from Abbvie, Bayer, GlaxoSmithKline, Licher MT, Rölke Pharma and UCB, and was invited to Advisory Boards of Abbvie, Biogen, Lundbeck and Market Access & Pricing Strategy GmbH. He serves as the co-chair of the MDS Technology Task Force. He is an advisory board member of Critical Path for Parkinson's (CPP).

Acknowledgments

EW, AA, and YZ are supported by Keep Control from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 721577. LR, AM, JH, KA, and WM are supported in part by Mobilise-D from Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 820820, Horizon2020 initiative of the European Union.

Search strategy and selection criteria

We searched multiple databases (PubMed, Web of Science and Google Scholar) with the terms: (environment* OR setting* OR compare*) AND (supervised OR lab OR laboratory OR standard* OR clinic*) AND (unsupervised OR home OR real life OR real world OR daily life OR daily living OR free living) AND (wearable sensor OR inertial sensor OR inertial measurement unit OR acceleromet* OR gyroscope OR pendant sensor) NOT (intervention [Title/Abstract] OR rehabilitation[Title/Abstract] OR heart rate[Title/Abstract] OR energy expenditure[Title/Abstract] OR classification[Title/Abstract]) from Aug 1, 20014, to Aug 1, 2019. Only articles in English, Dutch or German were taken into consideration. Articles were relevant if (i) they measured neurological patients or older adults (mean/median age 60+), (ii) measured similar mobility parameters with a wearable device in a supervised and in an unsupervised setting, and (iii) compared these results with each other. Reference lists of relevant articles were screened for additional references to generate the final reference list, and also the authors were asked to provide input.

References

1 Zijlstra W, Bisseling RW, Schlumbohm S, Baldus H. A body-fixed-sensor-based analysis of power during sit-to-stand movements. *Gait Posture* 2010; **31**: 272–8.
2 Paradis E, Sutkin G. Beyond a good story: from Hawthorne Effect to reactivity in health professions education research. *Med Educ* 2017; **51**: 31–9.
3 Espay AJ, Hausdorff JM, Sanchez-Ferro A, *et al.* A Roadmap for Implementation of Patient-Centered Digital Outcome Measures in Parkinson's disease Obtained Using Mobile Health Technologies. *Mov Disord* 2019; **34**: 657–63.
4 World Health Organization. International Classification of Functioning, Disability and Health. 2001 https://psychiatr.ru/download/1313?view=name=CF_18.pdf (accessed Jan 3, 2019).
5 Tolea MI, Galvin JE. Relationship between Mobility Dysfunction Staging and Global Cognitive Performance. *Alzheimer Dis Assoc Disord* 2016; **30**: 230–6.
6 Bergland A, Emaus N, Strand BH. Mobility as a predictor of all-cause mortality in older men

and women: 11.8 year follow-up in the Tromsø study. *BMC Heal Serv Res* 2017; **17**. DOI:10.1186/s12913-016-1950-0.
7 Mahmoudi R, Novella JL, Manckoundia P, *et al.* Is functional mobility an independent mortality risk factor in subjects with dementia? *Maturitas* 2017; **103**: 65–70.
8 Wallace LMK, Theou O, Godin J, Andrew MK, Bennett DA, Rockwood K. Investigation of frailty as a moderator of the relationship between neuropathology and dementia in Alzheimer's disease: a cross-sectional analysis of data from the Rush Memory and Aging Project. *Lancet Neurol* 2019; **18**: 177–84.
9 Shafrin J, Sullivan J, Goldman DP, Gill TM. The association between observed mobility and quality of life in the near elderly. *PLoS One* 2017; **12**. DOI:10.1371/journal.pone.0182920.
10 Davis JC, Bryan S, Li LC, *et al.* Mobility and cognition are associated with wellbeing and health related quality of life among older adults: a cross-sectional analysis of the Vancouver Falls Prevention Cohort. *BMC Geriatr* 2015; **15**. DOI:10.1186/s12877-015-0076-2.
11 van Uem JMT, Marinus J, Canning C, *et al.* Health-Related Quality of Life in patients with Parkinson's disease—A systematic review based on the ICF model. *Neurosci Biobehav Rev* 2016; **61**: 26–34.
12 Goetz CG, Tilley BC, Shaftman SR, *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**: 2129–70.
13 Kurtzke JF. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology* 1983; **33**: 1444–52.
14 Van Lummel RC, Walgaard S, Hobert MA, *et al.* Intra-Rater, Inter-Rater and Test-Retest Reliability of an Instrumented Timed Up and Go (iTUG) Test in Patients with Parkinson's Disease. *PLoS One* 2016; **11**. DOI:10.1371/journal.pone.0151881.
15 Mancini M, Curtze C, Stuart S, *et al.* The Impact of Freezing of Gait on Balance Perception and Mobility in Community-Living with Parkinson'S Disease. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS* 2018; : 3040–3.
16 Dorsey ER, Glidden AM, Holloway MR, Birbeck GL, Schwamm LH. Teleneurology and mobile technologies: the future of neurological care. *Nat Rev Neurol* 2018; **14**: 285–97.
17 Holter NJ. Radioelectrocardiography: A New Technique for Cardiocasular Studies. *Ann New York Acad Sci* 1957; **65**: 913–23.
18 Wachter R, Gröschel K, Gelbrich G, *et al.* Holter-electrocardiogram-monitoring in patients with acute ischaemic stroke (Find-AFrandomised): an open-label randomised controlled trial. *Lancet Neurol* 2017; **16**: 282–90.
19 Bally L, Thabit H, Hartnell S, *et al.* Closed-Loop Insulin Delivery for Glycemic Control in Noncritical Care. *N Engl J Med* 2018; **379**: 547–56.
20 Van Uem JMT, Isaacs T, Lewin A, *et al.* A Viewpoint on Wearable Technology-Enabled Measurement of Wellbeing and Health-Related Quality of Life in Parkinson's Disease. *J Parkinsons Dis* 2016; **6**: 279–87.
21 Block VAJ, Pitsch E, Tahir P, Cree BAC, Allen DD, Gelfand JM. Remote Physical Activity Monitoring in Neurological Disease: A Systematic Review. *PLoS One* 2016; **11**: e0154335.
22 Ryan JM, Fortune J, Stennett A, *et al.* Changing physical activity behaviour for people with multiple sclerosis: Protocol of a randomised controlled feasibility trial (iStep-MS). *BMJ Open* 2017; **7**: 1–13.
23 Barrett TM, Liebert MA, Schrock JM, *et al.* Physical function and activity among older adults in Jodhpur, India. *Ann Hum Biol* 2016; **43**: 488–91.
24 Pham MH, Elshehabi M, Haertner L, *et al.* Algorithm for Turning Detection and Analysis Validated under Home-Like Conditions in Patients with Parkinson's Disease and Older Adults using a 6 Degree-of-Freedom Inertial Measurement Unit at the Lower Back. *Front Neurol* 2017; **8**. DOI:10.3389/fneur.2017.00135.
25 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* 2016; **13**. DOI:10.1186/s12984-016-0154-5.

26 Zhang W, Regterschot GRH, Geraedts H, Baldus H, Zijlstra W. Chair Rise Peak Power in Daily Life Measured With a Pendant Sensor Associates With Mobility, Limitation in Activities, and Frailty in Old People. *IEEE J Biomed Heal Informatics* 2017; **21**: 211–7.
27 Ramsperger R, Meckler S, Heger T, *et al.* Continuous leg dyskinesia assessment in Parkinson's disease -clinical validity and ecological effect. *Park Relat Disord* 2016; **26**: 41–6.
28 Teshuva I, Hillel I, Gazit E, Giladi N, Mirelman A, Hausdorff JM. Using wearables to assess bradykinesia and rigidity in patients with Parkinson's disease: a focused, narrative review of the literature. *J Neural Transm* 2017; **126**: 699–710.
29 Silva de Lima AL, Evers LJW, Hahn T, *et al.* Freezing of gait and fall detection in Parkinson's disease using wearable sensors: a systematic review. *J Neurol* 2017; **264**: 1642–54.
30 Van Schooten KS, Pijnappels M, Rispens SM, *et al.* Daily-life gait quality as predictor of falls in older people: A 1-year prospective cohort study. *PLoS One* 2016; **11**. DOI:10.1371/journal.pone.0158623.
31 Weiss A, Brozgol M, Dorfman M, *et al.* Does the Evaluation of Gait Quality During Daily Life Provide Insight Into Fall Risk? A Novel Approach Using 3-Day Accelerometer Recordings. *Neurorehabil Neural Repair* 2013; **27**: 742–52.
32 Iluz T, Weiss A, Gazit E, *et al.* Can a Body-Fixed Sensor Reduce Heisenberg's Uncertainty When It Comes to the Evaluation of Mobility? Effects of Aging and Fall Risk on Transitions in Daily Living. *J Gerontol A Biol Sci Med Sci* 2016; **71**: 1459–65.
33 Del Din S, Galna B, Godfrey A, *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 A Biol Sci Med Sci* 2017; **74**: 500–6.
34 Punt M, Bruijn SM, Wittink H, Van De Port IG, Van Dieën JH. Do clinical assessments, steady-state or daily-life gait characteristics predict falls in ambulatory chronic stroke survivors? *J Rehabil Med* 2017; **49**: 402–9.
35 CTTI Recommendations: Developing Novel Endpoints Generated by Mobile Technology for Use in Clinical Trials. 2017 https://www.ctti-clinicaltrials.org/files/novelendpoints-recs.pdf (accessed March 25, 2019).
36 Cerreta F. EMA experience on mHealth technology. 2018 https://www.ema.europa.eu/en/documents/presentation/presentation-ema-experience-mhealth-technology-emas-pcwp-hcpwp-joint-meeting-francesca-cerreta_en.pdf (accessed March 25, 2019).
37 Van Ancum JM, van Schooten KS, Jonkman NH, *et al.* Gait speed assessed by a 4-meter walk test is not representative of daily-life gait speed in community-dwelling adults. *Maturitas* 2019; **121**: 28–34.
38 Jenkins IH, Jahanshahi M, Jueptner M, Passingham RE, Brooks DJ. Self-initiated versus externally triggered movements. II. The effect of movement predictability on regional cerebral blood flow. *Brain* 2000; **123**: 1216–28.
39 Wessel K, Zeffiro T, Toro C, Hallett M. Self-Paced Versus Metronome-Paced Finger Movements. *J Neuroimaging* 1997; **7**: 145–51.
40 Hunt C, Paez A, Folmar E. Literature Review the Impact of Attentional Focus on the Treatment of Musculoskeletal and. *Int J Sports Phys Ther* 2017; **12**: 901–7.
41 Kal EC, Van Der Kamp J, Houdijk H. External attentional focus enhances movement automatization: A comprehensive test of the constrained action hypothesis. *Hum Mov Sci* 2013; **32**: 527–39.
42 Robles-Garcia V, Corral-Bergantinos Y, Espinosa N, *et al.* Spatiotemporal Gait Patterns During Overt and Covert Evaluation in Patients With Parkinson s Disease and Healthy Subjects: Is There a Hawthorne Effect? *J Appl Biomech* 2015; **31**: 189–94.
43 Brodie MAD, Coppens MJM, Lord SR, *et al.* Wearable pendant device monitoring using new wavelet-based methods shows daily life and laboratory gaits are different. *Med Biol Eng Comput* 2016; **54**: 663–74.
44 van Lummel RC, Walgaard S, Pijnappels M, *et al.* Physical Performance and Physical Activity in Older Adults: Associated but Separate Domains of Physical Function in Old Age. *PLoS One*

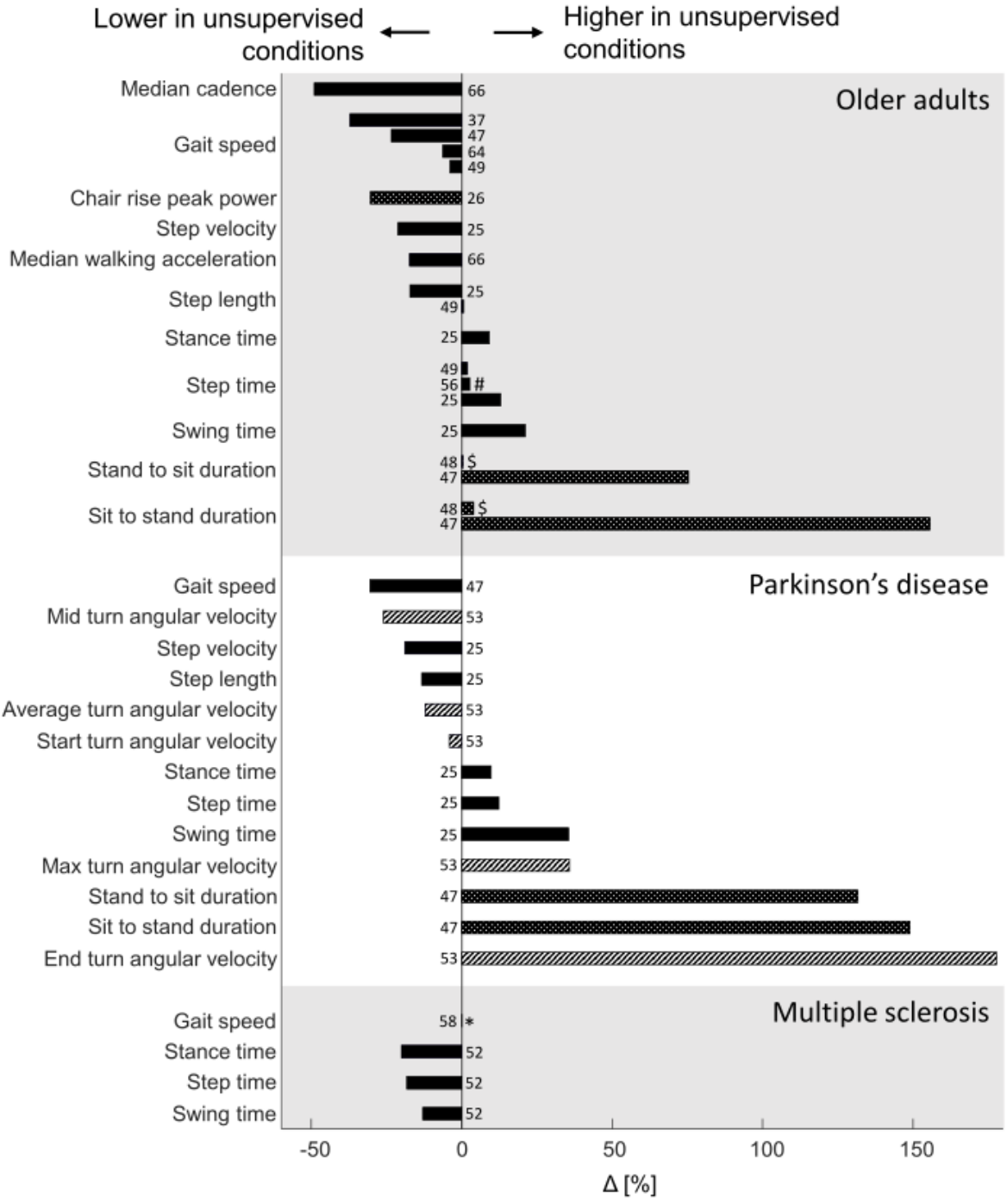
2015; **10**. DOI:10.1371/journal.pone.0144048.
45 Storm FA, Buckley CJ, Mazzà C. Gait event detection in laboratory and real life settings: Accuracy of ankle and waist sensor based methods. *Gait Posture* 2016; **50**: 42–6.
46 Tamburini P, Storm F, Buckley C, Bisi MC, Stagni R, Mazzà C. Moving from laboratory to real life conditions: Influence on the assessment of variability and stability of gait. *Gait Posture* 2018; **59**: 248–52.
47 Toosizadeh N, Mohler J, Lei H, Parvaneh S, Sherman S, Najafi B. Motor Performance Assessment in Parkinson's Disease: Association between Objective In-Clinic, Objective In-Home, and Subjective/Semi-Objective Measures. *PLoS One* 2015; **10**. DOI:10.1371/journal.pone.0124763.
48 Moufawad El Achkar C, Lenbole-Hoskovec C, Paraschiv-Ionescu A, Major K, Büla C, Aminian K. Classification and characterization of postural transitions using instrumented shoes. *Med Biol Eng Comput* 2018; **56**: 1403–12.
49 Hillel I, Gazit E, Nieuwboer A, *et al.* Is every-day walking in older adults more analogous to dual-task walking or to usual walking? Elucidating the gap between gait performance in the lab and during 24/7 monitoring. *Eur Rev Aging Phys Act* 2019; Epub ahead of print.
50 Lee JA, Williams SM, Brown DD, Laurson KR. Concurrent validation of the Actigraph gt3x+, Polar Active accelerometer, Omron HJ-720 and Yamax Digiwalker SW-701 pedometer step counts in lab-based and free-living settings. *J Sports Sci* 2015; **33**: 991–1000.
51 Pham MH, Warmerdam E, Elshehabi M, *et al.* Validation of a lower back 'wearable'-based sit-to-stand and stand-to-sit algorithm for patients with Parkinson's disease and older adults in a home-like environment. *Front Neurol* 2018; **9**. DOI:10.3389/fneur.2018.00652.
52 Storm FA, Nair KPS, Clarke AJ, Van der Meulen JM, Mazzà C. Free-living and laboratory gait characteristics in patients with multiple sclerosis. *PLoS One* 2018; **13**: 1–15.
53 Haertner L, Elshehabi M, Zaunbrecher L, *et al.* Effect of Fear of Falling on Turning Performance in Parkinson's Disease in the Lab and at Home. *Front Aging Neurosci* 2018; **10**. DOI:10.3389/fnagi.2018.00078.
54 Bourke AK, Ihlen AF, Laegdheim Helbostad J. Development of a gold-standard method for the identification of sedentary, light and moderate physical activities in older adults: Definitions for video annotation. *J Sci Med Sport* 2019; **22**: 557–61.
55 Bourke AK, Ihlen EAF, Bergquist R, Wik PB, Vereijken B, Helbostad JL. A Physical Activity Reference Data-Set Recorded from Older Adults Using Body-Worn Inertial Sensors and Video Technology — The ADAPT Study Data-Set. *Sensors* 2017; **17**. DOI:10.3390/s17030559.
56 Rispens SM, van Schooten KS, Pijnappels M, Daffertshofer A, Beek PJ, van Dieën JH. Do Extreme Values of Daily-Life Gait Characteristics Provide More Information About Fall Risk Than Median Values? *JMIR Res Protoc* 2015; **4**. DOI:10.2196/resprot.3931.
57 Hausdorff JM, Hillel I, Shustak S, *et al.* Everyday Stepping Quantity and Quality Among Older Adult Fallers With and Without Mild Cognitive Impairment: Initial Evidence for New Motor Markers of Cognitive Deficits? *J Gerontol A Biol Sci Med Sci* 2018; **73**: 1078–82.
58 Supratak A, Datta G, Gafson AR, Nicholas R, Guo Y, Matthews PM. Remote Monitoring in the Home Validates Clinical Gait Measures for Multiple Sclerosis. *Front Neurol* 2018; **9**. DOI:10.3389/fneur.2018.00561.
59 Dujmovic I, Radovanovic S, Martinovic V, *et al.* Gait pattern in patients with different multiple sclerosis phenotypes. *Mult Scler Relat Disord* 2017; **13**: 13–20.
60 Lindemann U, Zijlstra W, Aminian K, *et al.* Recommendations for Standardizing Validation Procedures Assessing Physical Activity of Older Persons by Monitoring Body Postures and Movements. *Sensors* 2014; **14**: 1267–77.
61 Badawy R, Hameed F, Bataille L, *et al.* Metadata Concepts for Advancing the Use of Digital Health Technologies in Clinical Research. *Digit Biomarkers* 2019; 3: 116-32. **Accepted**.
62 Moon Y, Sung J, An R, Hernandez ME, Sosnoff JJ. Gait variability in people with neurological disorders: A systematic review and meta-analysis. *Hum Mov Sci* 2016; **47**: 197–208.
63 Schloemer SA, Thompson JA, Silder A, Thelen DG, Siston RA. Age-Related Differences in Gait Kinematics, Kinetics, and Muscle Function: A Principal Component Analysis. *Ann Biomed Eng*

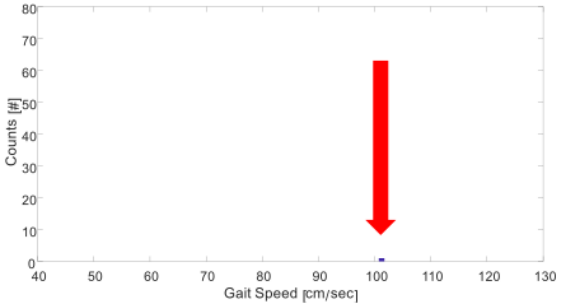
2017; **45**: 695–710.
64 Takayanagi N, Sudo M, Yamashiro Y, *et al.* Relationship between Daily and In-laboratory Gait speed among Healthy Community-dwelling older Adults. *Sci Rep* 2019; **9**. DOI:10.1038/s41598-019-39695-0.
65 Slater L V., Simpson AS, Blemker SS, *et al.* Biomechanical adaptations during running differ based on type of exercise and fitness level. *Gait Posture* 2018; **60**: 35–40.
66 Urbanek JK, Zipunnikov V, Harris T, *et al.* Validation of Gait Characteristics Extracted From Raw Accelerometry During Walking Against Measures of Physical Function, Mobility, Fatigability, and Fitness. *J Gerontol A Biol Sci Med Sci* 2018; **73**: 676–81.
67 Rispens SM, Van Dieën JH, Van Schooten KS, *et al.* Fall-related gait characteristics on the treadmill and in daily life. *J Neuroeng Rehabil* 2016; **13**. DOI:10.1186/s12984-016-0118-9.
68 Schooten KS Van, Rispens SM, Elders PJM, Lips P, Dieën JH Van, Pijnappels M. Assessing Physical Activity in Older Adults : Required Days of Trunk Accelerometer Measurements for Reliable Estimation. *J Aging Phys Act* 2015; **23**: 9–17.
69 De Fauw J, Ledsam JR, Romera-Paredes B, *et al.* Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018; **24**: 1342–50.
70 Zhan A, Mohan S, Tarolli C, *et al.* Using smartphones and machine learning to quantify Parkinson disease severity the mobile Parkinson disease score. *JAMA Neurol* 2018; **75**: 876–80.
71 Araújo R, van de Warrenburg B, Lang A, Lees A, Bloem B. The Waiting Room: neurological observations made outside the movement disorder specialist's consulting office. *Pr Neurol* 2019; **19**: 295–301.
72 Domingos J, Keus SHJ, Dean J, de Vries NM, Ferreira JJ, Bloem BR. The European Physiotherapy Guideline for Parkinson's Disease: Implications for Neurologists. *J Parkinsons Dis* 2018; **8**: 499–502.
73 Buckley AC, Galna B, Rochester L, Mazz C. Upper body accelerations as a biomarker of gait impairment in the early stages of Parkinson's disease. *Gait Posture* 2019; **71**: 289–95.
74 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**: 1293–313.
75 Arroyo-Gallego T, Ledesma-Carbayo MJ, Butterworth I, *et al.* Detecting Motor Impairment in Early Parkinson's Disease via Natural Typing Interaction With Keyboards: Validation of the neuroQWERTY Approach in an Uncontrolled At-Home Setting. *J Med Internet Res* 2018; **20**. DOI:10.2196/jmir.9462.

FIGURE LEGENDS

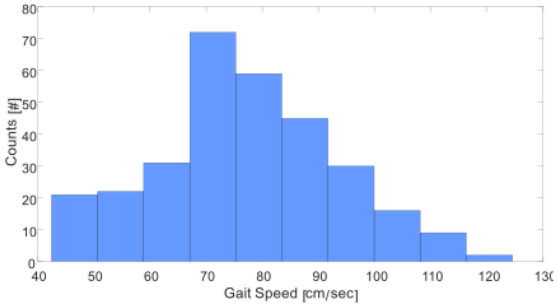
Figure 1. Percentage change from parameters measured in unsupervised conditions relative to supervised conditions*. Compared to supervised conditions, the bars to the right of zero indicate a higher value and to the left of zero indicate a lower value in unsupervised conditions. Number next to the bar is the reference number. The solid filled bars reflect parameters related to gait, the diagonal stripes reflect parameters related to turns and the dots reflect parameters related to postural transitions. See search strategy and selection criteria box for the study selection criteria and see glossary for explanation of parameters. Older adults are defined by mean/median age of the respective cohorts ≥60ys. Data was obtained from the papers or from the corresponding author. Note that we did not include variability and asymmetry parameters in the figure as they are especially sensitive to the environment and are likely higher for unsupervised assessments because of the non-instructed performance and more variable physical nature of the environment.*[67] ****** *Instructions in the supervised setting were as fast as possible.* [#] *Supervised assessment was performed on a treadmill with fixed speed, the unsupervised parameters used for the comparison were matched to the treadmill speed.* [$] *Only the best postural transitions resported were used to calculate the duration.*

Figure 2. Gait speed measures based on evaluation in the lab (A) and based on daily-living evaluation (B) in one older adult *(78-year-old woman with a history of falls). The supervised testing yields a single value (101 cm/sec), as indicated by the red arrow. In contrast, the daily-living, unsupervised testing yields hundreds of "tests" of gait speed and a distribution of values. Multiple measurements, in contrast to a single, one-time snapshot, may be highly valuable for the improvement of assessment protocols. In many of these tests, gait speed is lower than that seen during supervised testing. The daily-living values are based on 30-second walking bouts from a one-week recording, as in Hillel et al., 2019.*[49]


Lower in unsupervised conditions
Higher in unsupervised conditions
Older adults
Median cadence
Gait speed
Chair rise peak power
Step velocity
Median walking acceleration
Step length
Stance time
Step time
Swing time
Stand to sit duration
Sit to stand duration
66
37
47
64
49
26
25
66
25
49
25
49
56
25
25
48
47
48
47
Parkinson's disease
Gait speed
Mid turn angular velocity
Step velocity
Step length
Average turn angular velocity
Start turn angular velocity
Stance time
Step time
Swing time
Max turn angular velocity
Stand to sit duration
Sit to stand duration
End turn angular velocity
47
53
25
25
53
53
25
25
25
53
47
47
53
Multiple sclerosis
Gait speed
Stance time
Step time
Swing time
58
52
52
52
-50
0
50
100
150
Δ [%]



A



B