Mobile health technologies (wearable, portable, body-fixed sensors, or domestic-integrated devices) that quantify mobility in unsupervised, daily living environments are emerging as complementary clinical assessments. Data collected in these ecologically valid, patient-relevant settings can overcome limitations of conventional clinical assessments, as they capture fluctuating and rare events. These data could support clinical decision making and could also serve as outcomes in clinical trials. However, studies that directly compared assessments made in unsupervised and supervised (eg, in the laboratory or hospital) settings point to large disparities, even in the same parameters of mobility. These differences appear to be affected by psychological, physiological, cognitive, environmental, and technical factors, and by the types of mobility and diagnoses assessed. To facilitate the successful adaptation of the unsupervised assessment of mobility into clinical practice and clinical trials, clinicians and researchers should consider these disparities and the multiple factors that contribute to them.

Introduction

Deficits in mobility are common in patients with neurological disorders and often affect activities of daily living, work, and socialisation.1 These deficits predict morbidity, cognitive decline, and mortality3,4 and negatively affect quality of life, especially in patients with movement disorders.6,7 For example, in patients with Parkinson’s disease, health-related quality of life is strongly associated with the activities and participation components of the International Classification Of Functioning, Disability, and Health model.4 Therefore, it is crucial for health-care professionals to obtain a full and objective evaluation of a patient’s mobility as a basis for individually tailored clinical decision making and prognostication. Mobility assessments are mainly done under supervised conditions in a laboratory or hospital using standardised, mostly qualitative or semi-structured evaluations (panel).2,11 However, many patients do paradoxically well when they know that they are being observed. Various clinically relevant events are also difficult to capture during these snapshot observations, because they take place over long periods of time (eg, the total amount of physical activity), are rare (eg, falls or freezing episodes),7 occur at night (eg, sleep disturbances), or have complex fluctuating patterns (eg, the response to dopaminergic treatment in Parkinson’s disease). To reliably evaluate such events, it is important to measure patients unobtrusively and for longer periods of time, while they move about freely and unsupervised in their daily-living environment.

Several reviews describe the promise of unsupervised assessments of mobility using novel technologies.9,10 Unsupervised assessments of mobility using novel technology, although very different from other daily living acquired parameters that are already used in clinical routine (such as the Holter electrocardiogram12 and blood glucose monitoring),13 could soon be essential for the long-term evaluation of mobility and personalised clinical decision making in neurology.10,14 Unsupervised assessments might save time and cost by capturing health-related data since these assessments would be largely independent of the availability of health-care services. These assessments are particularly important for patients living in rural areas or developing countries, where the number of health-care professionals is small relative to the population size.14 Finally, unsupervised assessments offer patients an opportunity to become more actively involved by, for example, using their own devices such as smartphones and receiving feedback about their own daily living performance.15

Unsupervised assessments of mobility can provide additional and, at least partly, complementary information compared with supervised assessments. However, differences with respect to the conventional evaluation need to be considered. In this Personal View, we summarise the evidence of the weak association between mobility assessed in the two settings and discuss potential reasons for the observed differences. We also present suggestions to facilitate the implementation of unsupervised mobility assessment in clinical care and future research.

Unsupervised mobility assessment

Unsupervised assessments are usually done with mobile health technologies11 that can measure physical activity,16–22 evaluate mobility or specific movements such as gait,23–26 or detect specific symptoms in unsupervised environments.27–29 The potential added value of unsupervised assessments in patients with mobility deficits has been shown in several studies. For example, both predicting the risk of future falls and discriminating fallers from non-fallers in older adults (>60 years of age)23–25 and stroke survivors26 appears to be more accurate when using data collected in the unsupervised environment. Indeed, the relevance of unsupervised mobility parameters was acknowledged by the US Food and Drug Administration30 and the European Medicines Agency,31 both of which encourage the inclusion of parameters from unsupervised mobility assessments as exploratory endpoints in clinical trials.

Long-term unsupervised mobility assessment in movement disorders
Mobile health technologies

Umbrella term for wearable, portable, or 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 (eg, smartphone and keyboard).

We focus on technologies that can measure the frequency and quality of movement, and mobility characteristics.

Supervised assessment

Refers to the traditional, conventional mode of assessing mobility in a laboratory or clinical setting.

Typically, a qualitative or semi-quantitative one-time snapshot evaluation of mobility by a trained health-care professional.

Unsupervised assessment

Refers to the quantitative assessment of mobility in the home and daily living environment that is done continuously with new, mainly mobile, health technologies over relatively long periods of time.

Wearables

Mobile devices worn on the body, such as inertial measurement units, smartwatches, or Holter electrocardiogram monitors.

We did a systematic search to compare the same features of mobility (ie, gait, turns, and postural transitions) in supervised and unsupervised assessments. 12 studies done in three different populations—adults older than 60 years, patients with Parkinson’s disease, and patients with multiple sclerosis—were identified (appendix pp 1–3). Strikingly, the same mobility parameters obtained in different settings with identical participants differed from −40% (eg, gait speed and cadence in patients with Parkinson’s disease) to 180% (end turn angular velocity in healthy older adults, figure 1). These differences are much larger than the effects usually measured after interventions. Thus, small and even moderate treatment effects might be buried under the variations introduced by the measurement techniques themselves if the differences between supervised and unsupervised assessments are not appropriately considered.

Differences between supervised and unsupervised assessments

Several reasons could explain the substantial differences in mobility parameters when comparing supervised with unsupervised assessments (table 1). Unsupervised movements are typically self-initiated, embedded in a rich behavioural context, and goal directed. By contrast, movements in a supervised setting are usually triggered by a command and done in an isolated, standardised setting with limited ecological validity.6 For example, self-initiated finger movements activate different brain structures compared with externally triggered movements, suggesting that the brain generates supervised movements using networks that differ from those that generate unsupervised movements.6 Moreover, with an external focus, attention is directed to the outcome of the action (eg, leaving the room), while with an internal focus, attention is directed to controlling the body parts while executing the movement.49 An external focus of attention results, at least sometimes, in more fluent movements.50

Performance can be affected by several psychological and physiological processes that might differ across settings. These factors include alertness, motivation, the white-coat effect (a change [typically worsening] in a parameter because it is measured in a clinical setting), the reverse white-coat effect (a change [typically an improvement] in a clinical parameter because it is measured in a clinical setting), the Hawthorne effect (the change in behaviour of participants because of the awareness of being studied),51 fatigue, pain, and stress. These effects might explain why patients rise from a chair with lower peak power in unsupervised assessments than during supervised assessments, even when these movements are done in an identical environment and with the same equipment.52 Similar disparities have been identified for other gait parameters.53 Supervised assessments seemingly provide a measure of someone’s best, rather than their usual performance; that is, they capture capacity rather than performance.53,54

The environment is usually standardised in supervised conditions (eg, walking in a clean and sterile environment without distractions), but much more variable in unsupervised conditions (eg, furniture, lighting, patterns, colour of the environment, and obstacles). Unsupervised environments can induce large variability and asymmetry in mobility patterns, as shown by studies that assessed walking through busy corridors and through a city centre.55-57 Different types of seats and couches (eg, firm chair or armchair) in unsupervised conditions can also partly explain the greater variability observed in postural transitions (ie, sit-to-stand and stand-to-sit movements or turning over in bed) in daily living.56,58-60 Moreover, asymmetry can be introduced through a constrained environment that requires gait adaptation or turning in the same direction.

Furthermore, multitasking situations are common in unsupervised environments (eg, walking and texting), but uncommon in supervised assessments, which could further contribute to the observed differences. Even during supervised dual-task walking, the gait quality was usually better than that during unsupervised walking.61 The presence of a partner or caregiver can also affect mobility in unsupervised conditions. Social interactions...
are common during everyday walking: for example, spouses who act as an external cue to improve walking in patients with Parkinson’s disease or to relieve anxiety in people with a cautious gait disorder.  

Technical limitations might also add to the differences observed. Most algorithms have been developed and validated in supervised environments. Because the amount and variability of activities and mobility are much larger in unsupervised than in supervised environments, these algorithms might have difficulties differentiating similar movements (such as picking something up from the floor and sit-to-stand movements) that were not evaluated in the supervised assessment. Notably, only one study found in our systematic search used algorithms that were explicitly validated in both standardised and non-standardised settings. A further bias might be introduced by the use of different device locations on the body (eg, waist or ankle). The use of distinct mobile health technologies (eg, hardware or algorithms) could also play a part, but this aspect is limited as a change in hardware will not have a large influence on the results of a validated algorithm, because the data collected are the same (appendix p 4). The validation of algorithms for unsupervised daily living assessments brings new challenges as gold-standard references are currently absent, and urgently needed.

Finally, the statistical approaches for the analysis of supervised assessments (eg, means and SDs), might not be optimal for characterisation of complex data obtained from unsupervised settings. The supervised assessment typically involves one test, whereas the unsupervised evaluation might include thousands of walking bouts, turns, and transitions. It is yet to be determined how to best compare a single value with values obtained from a distribution (or histogram; figure 2; appendix pp 1–3). Several studies showed that the tails of an individual’s distribution correspond better to supervised assessments and therefore to clinical endpoints, such as risk of falls, limitation in activities, frailty, and supervised gait speed, compared with mean and median values.

Effect of movement type and disease on mobility assessment

Some types of mobility (eg, postural transitions) show seemingly larger differences than others (eg, walking) when comparing supervised with unsupervised conditions (figure 1). This difference might even depend on specific parameters. In a study of patients with Parkinson’s disease, 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 under the unsupervised condition.

Notably, the type and severity of a disease might also have an effect on the differences between supervised and unsupervised assessments (figure 1). For example, the differences in stand-to-sit duration between both settings were smaller in older adults than in patients with Parkinson’s disease. Patients with multiple sclerosis showed an even more surprising pattern. Different to patients with Parkinson’s disease and older adults, their performance was comparable under supervised and unsupervised assessments (gait speed), while showing...
the opposite behaviour of what was seen in patients with Parkinson’s disease and older adults (ie, for stance, step, and swing time, which were all lower in unsupervised conditions). The reasons for these observations are not yet clear, but differences in physical, attentional, and cognitive capabilities might contribute. These differences between supervised and unsupervised performance might even be relevant at the subgroup level. The reported changes in turning parameters in patients with Parkinson’s disease 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, than did patients in other Parkinson’s disease subgroups.

Implementation of unsupervised assessments in clinical practice and future research

As we anticipate that unsupervised assessments will become a prerequisite for future clinical decision making and clinical trials, in this section we provide directions to help move this emerging field forward (table 2). Although there is still insufficient understanding of the association between supervised and unsupervised mobility when interpreting data obtained from unsupervised environments, studies suggest that any extrapolation of unsupervised mobility based on findings from supervised mobility might be substantially influenced by the type, subtype and stage of the disease, as well as type of mobility extracted from the data. The duration of the unsupervised assessments should be also addressed, for example, by using the same mobile health technologies, located in the same place, for both supervised and unsupervised measurements. The algorithms used to calculate mobility parameters should be validated, to the highest degree possible, in both settings. Moreover, algorithms for mobility assessments should be validated separately for each type of neurological movement disorder as they might be associated with distinct movement patterns. Notably, even healthy people move differently at different ages and fitness levels. Another requirement to increase the usefulness of unsupervised measures is harmonised reporting of parameters (eg, as a core dataset across studies), and should include the reporting of metadata (ie, data that accompany and describe the primary data). The duration of the unsupervised assessments should be standardised and the type of movement assessed should be reported in detail.

Special emphasis should also be placed on more sophisticated analyses of unsupervised data. A promising approach is to consider and leverage specific episodes of mobility (eg, turning, sit-to-stand, and stand-to-sit movements, and other movements used regularly during the day) and novel parameters, such as the distribution and extreme values of mobility parameters (figure 2). Special emphasis should also be placed on more sophisticated analyses of unsupervised data. A promising approach is to consider and leverage specific episodes of mobility (eg, turning, sit-to-stand, and stand-to-sit movements, and other movements used regularly during the day) and novel parameters, such as the distribution and extreme values of mobility parameters (figure 2).

So far, these analyses have been done only for healthy older adults and not for patients with neurological disorders.
An example could be the evaluation of the effects of an experimental therapy. The effects might be measured best in the optimum state (improvement in supervised assessment and the best 10% of an individual’s distribution of the unsupervised assessment), while the median and lower range of an individual’s distribution might be informative of changes throughout the day (figure 2). Future trials could use this information as outcomes.

Variability measures can serve as a useful example of how important it is for clinicians and researchers to have a deep understanding of how their treatment and compounds influence mobility in daily life. Some variability measures (eg, stride length variability) are highly affected by the environment and should be measured in a supervised setting, which better reflects the patient’s capacity. 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 Parkinson’s disease. In a trial investigating patients with suboptimal treatment, 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. Thus, the context is crucial for evaluating the effect of an intervention.

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, should be evaluated in future studies. Trait markers could still be good measures of progression, but appropriate interpretation is key for practical use. For example, the actions done during daily living are very different per individual, but show a surprisingly similar pattern within an individual.

Future statistical analyses should take advantage of the high number of repeated, specific movements occurring during long-term observation periods in unsupervised environments (figure 2). Deep learning, machine learning, and artificial intelligence approaches should be applied. Algorithms that learn from data have shown remarkable success in making accurate predictions for complex problems that previously depended on human skills (eg, referral for eye diseases, detection of Parkinson’s disease motor fluctuations).

Future work should further explore the associations between objective digital measures with conventional measures of mobility, and with patient-reported outcome measures (PROMs) and caregiver-reported outcomes. 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 are 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 (appendix pp 1–3).

We should keep in mind that mobile health technologies might alone cause behavioural changes, even when no feedback is provided (eg, Hawthorne effect), but especially if feedback is provided (eg, to induce compliance). Studies are needed to investigate if and when the performance of the user in the unsupervised setting becomes similar to that in the supervised setting, and whether the induced behavioural changes themselves might have therapeutic effects that could interfere with...
the evaluation. For example, patients who know that they are equipped with mobile health technologies might increase their level of physical activity, particularly when feedback about their own performance is provided.

Health-care professionals should also interpret their supervised assessments cautiously, as these findings could have limited ecological value. To improve their value, we suggest to provide natural, everyday life-like situations and instructions during supervised assessments. Explicit goals should be given to the patients, forcing them to focus on the goal instead of on the actual movements that must be performed to reach the goal. 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. It is also essential to gain as much information as possible about the living environment of the person being assessed. If the person has cluttered furniture at home, health-care professionals might focus more on assessing mobility in small, crowded places instead of large, open hospital hallways. Additionally, the type of furniture, lighting, patterns, and other environmental factors might be important.

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 patients with neurological disorders; and several methods have been proposed to continuously assess upper limb bradykinnesia in daily life. However, a direct comparison of these various symptoms in supervised and unsupervised settings remains absent. One exception is a study that assessed habitual keyboard typing behaviour in patients with Parkinson’s disease. This study showed that various key-stroke metrics as measured in the clinic were strongly correlated with those obtained at the patient’s home, suggesting that some upper extremity performances (in this case, a measure of bradykinesia) are similar under supervised and unsupervised conditions. This finding underscores the need to assess different aspects of motor functioning on a case-by-case basis.

Conclusions
There is increasing evidence that, depending on whether mobility is assessed under supervised or unsupervised conditions, the results can differ substantially. These striking differences and the importance of measurements obtained in both settings call for expanding our knowledge about unsupervised mobility (table 2). Unsupervised mobility parameters could be implemented to improve clinical care and could act as primary or secondary endpoints in future intervention trials.

Contributors
EW, JMH, and WM developed the idea. EW did the literature search. EW, JMH, CH, and WM framed the outline. BB, JMH, AA, YZ, AM, KA, AJE, CH, AK, CL, AP, LR, GS, and LJWE worked on different aspects of the manuscript. All authors commented on the manuscript and approved the final version.

Declaration of interests
EW, AA, and YZ are supported by Keep Control from the EU’s Horizon 2020 (H2020) research and innovation programme under the Marie Skłodowska-Curie grant agreement number 721577. JMH, AM, KA, LR, and WM are supported by Mobilise-D from the Innovative Medicines Initiative 2 Joint Undertaking (grant number R20280); this Joint Undertaking receives support from the EU’s H2020 research and innovation programme and the European Federation of Pharmaceutical Industries and Associations. JMH has received grant support from the US National Institutes of Health (NIH), the Michael J Fox Foundation for Parkinson’s Research, the EU (H2020), and the US-Israel Binational Science Foundation, and has served or currently serves on the Movement Disorders Society Technology (MDS) Task Force, Michael J Fox Foundation task force on gait, board of the International Society for the Measurement of Physical Behavior, and advisory boards for Sanofi and Biogen. AM has received grant support from the Michael J Fox Foundation for Parkinson’s Research, EU (H2020). Israeli Science Foundation, and Ministry of Science and Technology; chairs the Michael J Fox Foundation task force on gait; and is a consultant for Biogen Biotechnologies. AM and JMH have submitted a patent application on the use of body-fixed sensors for assessing Parkinson’s disease symptoms, the intellectual property rights for which are held by the Tel Aviv Medical Center (Tel Aviv, Israel). KA has received a grant from the Swiss National Science, CTT, Innsbruck, EU (H2020, FP7, and FP6), Lausanne Orthopedics Research Foundation, Austria Research Promotion Agency, Swiss Multiple Sclerosis Society, Swiss Heart Foundation, Pierre Mercier Foundation, Swiss Federal Office of Sport, EHC-Hospital of Morges (Morges, Switzerland), and Suvalv, and was co-founder of Gait Up, serving as its board member until 2018. AJE has received grant support from the NIH, Great Lakes NeuroTechnologies, and the Michael J Fox Foundation for Parkinson’s Research; personal fees as a consultant and scientific advisory board member for AbbVie, Adams, Acadia Pharmaceuticals, Acorda Therapeutics, NeuroDerm, Impax, Sunovion, Lundbeck, Otsotica Pharmaceutical, and US WorldMeds; publishing royalties from Lippincott Williams & Wilkins, Cambridge University Press, and Springer; honoraria from US WorldMeds.
References


