medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Consensus based framework for digital mobility monitoring

Felix Kluge[1,*], Silvia Del Din[2], Andrea Cereatti[3], Heiko Gaßner[4], Clint Hansen[5], Jorunn L Helbostadt[6], Jochen Klucken[4], Arne Küderle[1], Arne Müller[7], Lynn Rochester[2,8], Martin Ullrich[1], Bjoern M Eskofier[1], Claudia Mazzà[9], and on behalf of the Mobilise-D consortium[10]

[1]Machine Learning and Data Analytics Lab, Department Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany
[2]Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK
[3]Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy
[4]Department of Molecular Neurology, University Hospital Erlangen, Erlangen, Germany
[5]Department of Neurology, University of Kiel, Kiel, Germany
[6]Department of Neuromedicine and Movement Science, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway
[7]Novartis Pharma AG, Basel, Switzerland
[8]The Newcastle upon Tyne NHS Foundation Trust, Newcastle upon Tyne, UK
[9]Department of Mechanical Engineering & Insigneo Institute for in silico Medicine, The University of Sheffield, Sheffield, UK
[10]Full membership of the Mobilise-D consortium is available on the website https://www.mobilise-d.eu
[*]felix.kluge@fau.de

ABSTRACT

Digital mobility assessment using wearable sensor systems has the potential to capture walking performance in a patient's natural environment. It enables the monitoring of health status and disease progression and outcome evaluation of interventions in real-world situations. In contrast to laboratory settings, real-world walking occurs in non-conventional environments and under unconstrained and uncontrolled conditions. Despite the general understanding, there is a lack of agreed definitions about what constitutes real-world walking, impeding the comparison and interpretation of the acquired data across systems and studies. Hence, there is a need for a terminological framework for the guidance of further algorithmic implementation of digital measures for gait assessment. We used an objective methodology based on an adapted Delphi process to obtain consensus on specific terminology related to real-world walking by asking a diverse panel of clinical, scientific, and industrial stakeholders. Six constituents ('real-world', 'walking', 'purposeful', 'walking bout', 'walking speed', 'turning') have successfully been defined in two feedback rounds. The identification of a consented set of real-world walking definitions has important implications for the development of assessment and analysis protocols, as well as for the reporting and comparison of digital mobility outcomes across studies and systems. The definitions will serve as a common framework for implementing digital and mobile technologies for gait assessment.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Mobility in general, and gait in particular, can be affected by a variety of chronic health conditions, spanning from neurological, such as in multiple sclerosis (MS) and Parkinson's disease (PD), to respiratory, such as in chronic obstructive pulmonary disease (COPD), cardiac like in congestive heart failure (CHF), or musculoskeletal disorders such as in proximal femoral fracture (PFF) recovery[1–6]. Such functional mobility impairments present a great burden to patients, as they are associated with severe limitations in quality of life[7–9], increased fall risk[10–12], and mortality[13,14].

Changes in various gait measures such as cadence, gait speed, stride length amongst others may characterize those mobility impairments. The use of digital mobility outcomes (DMOs), which we refer to as digital measures acquired using digital health technology[15] has already been studied in clinical settings using brief standardized tests in a range of diseases[2,4,16,17]. However, a single observation may not be reliable for clinical characterization especially when mobility related disease symptoms fluctuate over time, sometimes even on an hourly basis. Therefore, the objective assessment of gait calls for valid and reliable methods to sensitively capture changes in gait function more frequently[18]. As it is not feasible to increase patient visits to the clinic, a more continuous monitoring outside laboratory or clinical environments is desired[19]. Thus, the continuous assessment of real-world digital measures is essential and opens the opportunity for frequent and long-term remote monitoring[18,20–22]. In the past years, real-world gait analysis has been technologically enabled by the development of lightweight and easy to use sensor-based systems that can be worn unobtrusively. Although DMOs quantified from real-world data are able to discriminate and detect gait impairments in various diseases[20,23–26], accepted and routinely used tools are not applied in practice yet[27].

Whilst real-world measurement of mobility holds promise, one fundamental reason for the missing adoption is the difficulty of comparing DMOs across studies due to the inconsistent use of terminology. As an example, a broad variety of terms describing the "real-world" context exist, including "real-life", "daily-life", "everyday-life", and "free-living"[19,20,28,29], that are used interchangeably with ambiguous definitions. This can lead to different test paradigms being considered impeding comparability across measurements, systems, and studies. Furthermore, observed DMO variations may not only be caused by disease symptoms but also environmental factors and measurement protocols, which affect the reliability of DMO assessment. Therefore, unified definitions of relevant DMOs and the context of measurement are necessary to guarantee clinical meaningfulness. As a further example, the term "walking bout" has been used in the context of real-world gait analysis and refers to the quantification of continuous periods of free-living walking[30]. However, walking bout definitions are inconsistent and may include different walking bout durations and number of strides[10,12,28,30–32]. The duration of resting periods between walking bouts[33], and whether turning is considered as part of walking[34,35] are treated differently as well. However, a clear definition of a "walking bout" is critical, since it directly affects digital measures[28,36]. Additionally, "turning" needs to be considered as a main constituent of walking, as an average of more

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

than 60 turns per hour has been reported for real-world walking[34]. Due to their high occurence, turnings are likely to break sequences of straight walking into smaller walking bouts. Therefore, the specific definition of a turn directly influences the distribution of walking bouts with regard to their duration. Furthermore, spatio-temporal parameters during straight walking and turning differ[37], such that real-world DMOs based on averages of those parameters strongly depend on whether turning is included in their estimation. Currently, different operational approaches exist for the detection of turning. As an example, turning characteristics may be based on stride to stride angular parameters using foot rotation[38] or on angular changes related to the trunk rotation[34]. This diversity highlights the need for defining "turning" in a framework of DMO assessment, which will anable a more specific operationalization of real-world DMOs.

Overall, a guiding framework for the implementation of DMOs in real-world settings is still missing. Accordingly, the aim of this study was building a terminological framework in order to drive the development and assessment of DMOs for real-world monitoring. We aimed to reach agreement upon a set of narrative definitions within the scope of the Mobilise-D project[15], which is a five-year EU-funded IMI consortium that will build a technically and clinically-valid system for real-world digital mobility assessment across multiple populations with the goal to improve healthcare. We used an objective and systematic consensus process based on an adapted Delphi method[39]. Our results will enable operational definitions to implement mobility assessment algorithms, foster comparability across studies, and serve as a common communication framework for the scientific community. Perspectively, consensus on such a terminological framework is a prerequisite for the adoption of validated digital biomarkers characterizing mobility impairments in various diseases[15,27].

Results

Our approach of defining a terminological framework consisted of the following steps. First, relevant domains and key terms related to real-world walking for the consensus process were identified. Six terms related to four domains of real-world walking needing group consensus were selected (Table 1). For some terms, different aspects were regarded. For example, we proposed a physiological definition of "walking" and highlighted its relationship to walking bouts. For the definition of "real-world", we defined fundamental characteristics, how it is discriminated from standardized measurements and which test paradigms in the clinical context may be regarded as real-world assessment. "Walking speed" has been physically defined. Additionally, we defined the need to consider different granularities when calculating aggregated speed measures and proposed that real-world walking speed needs to be inherently connected to walking bouts. We proposed initial definitions for those eleven aspects based on the study team's expert knowledge and the current state of literature. Iterative feedback was included to improve structure and content of the definitions (for the initial definitions refer to Supplementary Table 1). We used these definitions as starting point for the subsequent consensus process.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Consensus process

We chose an adapted Delphi process to reach agreement on the initial definitions. A detailed description of the experimental proctol and the data analysis can be found in the *Methods* section. In short, a Delphi process is an objective iterative and anonymous process of asking participants for feedback on proposed statements and adjusting those statements until overall consensus is reached[39]. We defined that more than 75 % agreement to a statement was required to reach consensus, which is in line with previous literature[49,52].

In total, the consensus process required two rounds until agreement on all definitions was reached. 162 members of the Mobilise-D consortium were asked to participate in the first round of the consensus process. Of those, 79 individuals started the questionnaire. Eight individuals did not sign the participation or data usage agreement and five participants did not complete the questionnaire. Hence, their data was discarded. Data from the remaining 66 participants (40.7 % response rate in the first round) were analysed. Of the participants who completed the first round, 55 individuals started the second round. One individual did not sign the participation agreement and three individuals did not complete the questionnaire. We analysed the data of the remaining 51 participants (continuation response rate of 77.3 %). The overall response rate regarding both rounds was 31.5 %. The professional background of the panel was diverse but homogenously distributed across clinical and technical disciplines (Table 2). Only 12.1 % stated to have no experience in gait analysis. Two thirds of all individuals had expertise in free-living mobility. Most participants (88.7 %) stated to have expertise with patients in the past. As answering the background questions was not obligatory, the total number differs from the total number of individuals who participated in round one of the process.

Agreement on definitions

In round one, the definitions of "Purposeful" and "Walking speed (Relation to walking bouts)" did not reach agreement (Table 3) and were subject to modification based on participants' feedback. Although the definition of "Walking bout" reached agreement, there was inclarity regarding its inherent connection to the "Walking speed" definition. "Walking speed" initially assumed a different number of strides required to assess average walking speed (for the initial definitions refer to the Supplementary Tabe 1). Therefore, we decided to harmonize the "Walking bout" and "Walking speed (Relation to walking bouts)" definitions, which were both put to vote again in the second round. Full consensus for all definitions was reached in round two (Table 3) resulting in a final set of definitions for real-world gait analysis (Table 4).

Discussion

To the best of the authors' knowledge, this has been the first study to engage clinicians, academic researchers, and industry stakeholders working in the field of digital gait and mobility measures development to identify and agree upon a framework of narrative definitions for the assessment of DMOs acquired in real-world conditions. An adapted Delphi consensus process allowed to achieve consensus on eleven

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

statements related to six key terms of real-world walking.

A broad definition of walking may include various displacements of the body in space (e.g. for-, back-, or side-ward walking). We, however, define walking to be only associated with forward displacement using both legs in order to assure reliability of DMO assessment in various contexts. Stepping on the spot, side stepping, and backward walking have thus been excluded from this definition. Walking is also not defined in terms of a specific speed. The use of walking aids has been included into the definition as they may be an essential requirement for safe locomotion of people with gait impairments. Otherwise, certain patients and elderly individuals might be excluded from the DMO assessment. We acknowledged steps and strides as basic elements of walking as previously suggested[45]. Furthermore, the definitions include that walking is always made up of walking bouts. It was agreed that walking bouts represent sequences containing at least two full consecutive strides of both feet without a break (e.g. *R-L-R-L-R-L* or *L-R-L-R-L-R*, with *R/L* being the contacts of the right/left foot with the ground, respectively). Start and end of a walking bout are determined by a break that can either consist of a resting period, turning, or any other non-walking real-world activity. More specifically, the start is always defined by an initial step of a walking bout following a non-walking period, while the final step precedes the next non-walking period. Walking bouts are thus an important building block in the terminology framework for the assessment of DMOs acquired in real-world conditions.

Walking speed has been referred to as "the sixth vital sign", as slow walking has been associated with morbidity, cognitive decline, and fall risk amongst others[13,46]. Despite this, there is still no accepted common measure of mobility that serves across multiple conditions, which is underlined by a wide range of inconsistent testing procedures. With an operative definition of walking speed and a proposition of respective aggregation levels at which it is measured, we aim to provide a common framework to be used across clinical conditions. The physical definition of the walking speed definition reached high consensus. The panel agreed that walking speed will be assessed based on a minimal number of consecutive strides. According to clinical questions, walking speed needs to be assessed with regard to different aggregation levels (hourly, daily, weekly, etc.). This definition is in line with the walking bout definition. On the one hand, this specification yields a unified approach of assessing speed in our framework, but requires a stride-wise analysis of walking, which might not be feasible in all analysis cases, for example when the extraction of single strides is not feasible. The definition also suggests that walking speed derived from strides that are not part of walking bouts (i.e. short strides, shuffling, turning, etc.) should not be considered for the estimation of real-world walking speed.

Daily mobility does not only contain straight walking but also curved walking and turns. Therefore, we included a definition of turning in the framework to guide the implementation of walking bouts and the related DMO assessment. Turning can be regarded as being a deceleration of the forward motion, rotating the body as a whole, and stepping out toward the new direction[44]. It results in a change of walking direction and change in angular orientation including a rotational movement of the body around the longitudinal axis. As an example, a threshold on the rotation angle (e.g. $> 45^\circ$) at a certain turn

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

duration (e.g. between 0.5 s and 10 s) can be used to detect a turn[34]. This definition does not include all the required aspects for quantitative ambulatory mobility measurement. For example, further discussion will be necessary to define specific angular thresholds between straight and non-straight (i.e. curvilinear) walking. However, those operational definitions are not part of the narrative framework considered in this study and are, therefore, part of future evaluations based on real-world data from different clinical populations.

The environmental context of walking greatly influences DMOs. Thus, it was deemed necessary to specify inclusion and exclusion criteria of what is considered real-world. The participants agreed on different aspects of the terminology: real-world is conceptualized as free-living, unsupervised, uncontrolled and non-standardized. In the real-world context, the measurement of DMOs should not interfere with daily activities of the participant. The measurement process of real-world DMOs should thus be as non-obtrusive as possible. Accordingly, this definition is distinct from laboratory-based[43], supervised (fully controlled and observed), and semi-controlled (walking freely but with supervision) environments, in which observer and instruction effects might occur and influence DMOs. For example, walking happens in non-simulated real-world situations in unconstrained environments equivalent to actions at home or in the community over continuous periods of time[28]. Daily-living, including home and clinical environment are equivalent to real-world as long as the walking happens unsupervised. Scripted walking capacity tests such as 4x10 m walking conducted at home are excluded from the definition, as significant differences between DMOs derived from those tests and DMOs acquired during unscripted real-world walking are expected[47]. However, differences and relationships between standardized tests and real-world assessments still need to be evaluated in future research studies.

The participants agreed to the definition of purposeful as a consistent term for the assessment of DMOs acquired in real-world conditions. Purposeful walking includes an intentional component of the movement. We assume that unsupervised walking is per se purposeful and that the intentional aspect occurs especially for long walking bouts and needs to be evaluated taking for example contextual aspects of walking into account. Differences between purposeful, self-initiated movements, and movements performed in a supervised (and thus not real-world setting) are discussed in detail in[48], and should be further investigated with the consented real-world walking definitions.

The definitions agreed upon in this study are suggestions in order to capture gait analysis characteristics and properties in the real-world environment. The goal to have working definitions for various clinical populations has, on the one hand, resulted in a rather broad definition of walking (i.e. inclusion of walking aids). On the other hand, clarity on specific parts of the definition (i.e. that walking only includes forward locomotion) will allow to implement very specific digital mobility measures without restricting the application cases.

While the Delphi approach is commonly used to obtain broad consensus among experts by determining the level of agreement on a given topic[39] there is always a certain bias. We used purposive sampling under the assumption that members of the Mobilise-D consortium represented experts in the field of real-world

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

gait research. In general, we acknowledge that the choice of participants restricted to the consortium limits generalizability of the results. However, the consortium includes a large group of experts on gait analysis from Europe and the USA. The participants were homogeneously distributed regarding technical and clinical background. Their views were gathered from a wide range of clinical and academic disciplines to equally represent a breadth of expertise. Some participants stated no experience with gait analysis before. However, most of the participants explicitly mentioned having worked in the field of free-living gait analysis. Moreover, the larger proportion of the participants already had experience in clinical gait analysis. Nevertheless, further work is needed to validate the results of our study in light of an even broader international group of gait experts. For example, the survey could be opened to a wider panel to validate and refine the findings. Furthermore, the framework needs to be evaluated with regard to clinical interpretability of the acquired DMOs based on actual real-world data. Overall, given geographical spread, the Delphi consensus method conducted online was an appropriate tool for gathering the different viewpoints as compared to physical discussion rounds.

The low response rates observed in this study, especially in the first round, are typical for consensus processes and have previously been reported[49]. Especially with large sample sizes, low response rates are considered to be a drawback[50]. Specifically to this study, we invited all members of the Mobilise-D consortium to take part in the study, if they could comment on the topic. Some of the participants invited might not have had a real-world gait analysis background or interest and did therefore not participate in the process. As discussed, the analysis of the participants' professional background showed that the included participants had relevant experience in the field of interest. Furthermore, the final sample size of 51 participants was higher than the lower threshold of 12, which has been regarded as minimal number to ensure reliability of results in a consensus process[51].

Only group feedback in the questionnaires were provided, as individual feedback was not possible due to anonymity. However, the participants were sent an email with their individual responses and comments after completion of a round. This allowed them to reflect on their own ratings in the subsequent round.

One limitation of this study is that the definitions are only of narrative nature. While the obtained definitions have been objectively derived, some may need refinement according to the practical needs to directly guide algorithm implementation (e.g. thresholds on differentiating turning from curvilinear or straight walking need to be derived from further consensus or based on real-world data). Whilst extracting and analysing DMOs, more detailed definitions need to be derived from the initial framework to enlarge the scope and ensure applicability across different technologies and solutions for real-world gait assessment.

This work was conducted as part of the Mobilise-D project[15] with the aim to guide the data analysis process regarding real-world walking analysis with a focus on the assessment of real-world walking speed. It has to be noted, that different ways of assessing real-world mobility exist, such as analyzing daily activity patterns (e.g. daily step count, physical activity, energy expenditure amongst others). Related digital measures are of high interest for some diseases and might benefit from similar terminological frameworks.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

In conclusion, consensus on narrative definitions for the assessment of gait related DMOs acquired in real-world conditions was obtained through an objective consensus process, which has important consequences for clinical gait and mobility research, and the implementation of digital measures. Future work within the community may evaluate DMOs based on those definitions, which will improve comparability between studies and settings. The results of this study have thus important implications for the development of standardized analyzig protocols, as well as for the reporting, and comparison of DMOs. Overall, the definitions will allow a standardized and more precise use of those terms in future studies, enabling a stronger congruence of clinical, technical, and regulatory activities in this field.

Methods

We adopted an objective consensus building process for the study based on the Delphi methodology[39]. In contrast to quantitative methods such as systematic reviews or meta-analyses, which are based on available literature and studies, the Delphi process allows to obtain consensus among experts by determining the level of agreement on a given topic[39]. Specifically, the Delphi method is characterized by anonymity to avoid dominance of single experts, multiple iterations, and feedback to the group. As such, a basic Delphi technique can contain any type of self-administered questionnaire with no meetings[52], which is the approach that was used in this explorative study to quantitatively assess the agreement on the initial definitions.

A consensus process may consist of multiple rounds until agreement on the definitions is reached. Based on a 5-point Likert scale (Figure 1), agreement was quantitatively assessed amongst the participants[49]. A neutral statement (“No opinion”) was included. In our study, agreement to a given definition was defined a-priori[52,53] as more than 75 % of the answers belonging to the categories "Somewhat agree" or "Strongly agree". In the first round, the participants were asked to independently rate all eleven statements across the six key terms “Walking”, “Purposeful”, Real-world”, Walking bout”, “Walking speed”. Additionally, participants were asked to provide free-text comments for each item in order to capture input for the improvement of definitions[49]. In subsequent rounds, we presented modified definitions that previously did not reach agreement and assessed the agreement.

The consensus process was performed among member of the Mobilise-D consortium. The project includes technical, clinical, and regulatory expertise of 34 partners from Europe and the USA. All 162 members of the consortium were asked for participation via email. Participants who did not respond in the first round were not invited to participate in the second round. We did not define any exclusion criteria. Data on participants’ “technical or clinical background”, “gait expertise”, “free-living expertise”, and “expertise with patients” were collected in the first round to analyse the panel’s background.

The consensus process was implemented as a series of questionnaires based on the “Survey” feature of the ILIAS E-learning platform (version 5.4.5, ILIAS open source e-Learning e. V.). It allowed the anonymous acquisition of responses. The participants’ email addresses were linked to access codes, which were provided to start the questionnaire. The use of access codes allowed sending reminders

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

to the participants and preventing double participation. The acquisition and analysis of the data was anonymous. Descriptive statistics was used to investigate participants' background information and agreement responses in each round. Analyses were conducted using R version 4.0.3[54].

Ethical approval for this study was granted by the ethics committee of the University Hospital Erlangen (Re.-No. 241_19 Bc). All participants provided informed consent for inclusion in the study. Participation in this study was voluntary. All data were handled in accordance with European data protection regulations.

References

1. Magaziner, J. *et al.* Recovery From Hip Fracture in Eight Areas of Function. *The Journals Gerontol. Ser. A: Biol. Sci. Med. Sci.* **55**, M498–M507 (2000), https://doi.org/10.1093/gerona/55.9.M498.

2. Schlachetzki, J. C. M. *et al.* Wearable sensors objectively measure gait parameters in Parkinson's disease. *PLOS ONE* **12**, e0183989 (2017), https://doi.org/10.1371/journal.pone.0183989.

3. Karpman, C. & Benzo, R. Gait speed as a measure of functional status in COPD patients. *Int. J. Chronic Obstr. Pulm. Dis.* **9**, 1315 (2014), https://doi.org/10.2147/COPD.S54481.

4. Spain, R. I. *et al.* Body-worn motion sensors detect balance and gait deficits in people with multiple sclerosis who have normal walking speed. *Gait & Posture* **35**, 573–578 (2012), https://doi.org/10.1016/j.gaitpost.2011.11.026.

5. Martin, C. L. *et al.* Gait and balance impairment in early multiple sclerosis in the absence of clinical disability. *Multiple Scler.* **12**, 620–628 (2006), https://doi.org/10.1177/1352458506070658.

6. van den Berg-Emons, H. R., Bussmann, J. H., Balk, A., Keijzer-Oster, D. & Stam, H. Level of Activities Associated With Mobility During Everyday Life in Patients With Chronic Congestive Heart Failure as Measured With an Activity Monitor. *Phys. Ther.* **81**, 1502–1511 (2001), https://doi.org/10.1093/ptj/81.9.1502.

7. Kohn, C. G., Baker, W. L., Sidovar, M. F. & Coleman, C. I. Walking Speed and Health-Related Quality of Life in Multiple Sclerosis. *The Patient - Patient-Centered Outcomes Res.* **7**, 55–61 (2014), https://doi.org/10.1007/s40271-013-0028-x.

8. Gimeno-Santos, E. *et al.* Determinants and outcomes of physical activity in patients with COPD: a systematic review. *Thorax* **69**, 731–739 (2014), https://doi.org/10.1136/thoraxjnl-2013-204763.

9. Schmid, A. *et al.* Improvements in Speed-Based Gait Classifications Are Meaningful. *Stroke* **38**, 2096–2100 (2007), https://doi.org/10.1161/STROKEAHA.106.475921.

10. Brodie, M. A., Lord, S. R., Coppens, M. J., Annegarn, J. & Delbaere, K. Eight-Week Remote Monitoring Using a Freely Worn Device Reveals Unstable Gait Patterns in Older Fallers. *IEEE Transactions on Biomed. Eng.* **62**, 2588–2594 (2015), https://doi.org/10.1109/TBME.2015.2433935.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

11. Brodie, M. A. *et al.* Comparison between clinical gait and daily-life gait assessments of fall risk in older people. *Geriatr. & Gerontol. Int.* **17**, 2274–2282 (2017), https://doi.org/10.1111/ggi.12979.

12. Weiss, A. *et al.* Does the Evaluation of Gait Quality During Daily Life Provide Insight Into Fall Risk? A Novel Approach Using 3-Day Accelerometer Recordings. *Neurorehabilitation Neural Repair* **27**, 742–752 (2013), https://doi.org/10.1177/1545968313491004.

13. Studenski, S. Gait Speed and Survival in Older Adults. *JAMA* **305**, 50 (2011), https://doi.org/10.1001/jama.2010.1923.

14. Andrianopoulos, V. *et al.* Prognostic value of variables derived from the six-minute walk test in patients with COPD: Results from the ECLIPSE study. *Respir. Medicine* **109**, 1138–1146 (2015), https://doi.org/10.1016/j.rmed.2015.06.013.

15. 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), https://doi.org/10.1159/000512513.

16. Annegarn, J. *et al.* Differences in Walking Pattern during 6-Min Walk Test between Patients with COPD and Healthy Subjects. *PLoS ONE* **7**, e37329 (2012), https://doi.org/10.1371/journal.pone.0037329.

17. 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. *Park. & Relat. Disord.* **62**, 85–90 (2019), https://doi.org/10.1016/j.parkreldis.2019.01.022.

18. Espay, A. J. *et al.* Technology in Parkinson's disease: Challenges and opportunities. *Mov. Disord.* **31**, 1272–1282 (2016), https://doi.org/10.1002/mds.26642.

19. Dobkin, B. H. & Dorsch, A. The Promise of mHealth. *Neurorehabilitation Neural Repair* **25**, 788–798 (2011), https://doi.org/10.1177/1545968311425908.

20. Weiss, A. *et al.* Toward Automated, At-Home Assessment of Mobility Among Patients With Parkinson Disease, Using a Body-Worn Accelerometer. *Neurorehabilitation Neural Repair* **25**, 810–818 (2011), https://doi.org/10.1177/1545968311424869.

21. 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), https://doi.org/10.1002/mds.26718.

22. Hillel, I. *et al.* Is every-day walking in older adults more analogous to dual-task walking or to usual walking? Elucidating the gaps between gait performance in the lab and during 24/7 monitoring. *Eur. Rev. Aging Phys. Activity* **16**, 6 (2019), https://doi.org/10.1186/s11556-019-0214-5.

23. Del Din, S., Godfrey, A. & Rochester, L. Validation of an Accelerometer to Quantify a Comprehensive Battery of Gait Characteristics in Healthy Older Adults and Parkinson's Disease: Toward Clinical and

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

at Home Use. *IEEE J. Biomed. Heal. Informatics* **20**, 838–847 (2016), https://doi.org/10.1109/JBHI.2015.2419317.

24. Sosnoff, J. J., Goldman, M. D. & Motl, R. W. Real-life walking impairment in multiple sclerosis: preliminary comparison of four methods for processing accelerometry data. *Multiple Scler. J.* **16**, 868–877 (2010), https://doi.org/10.1177/1352458510373111.

25. Pericleous, P. & Van Staa, T. P. The use of wearable technology to monitor physical activity in patients with COPD: a literature review. *Int. J. Chronic Obstr. Pulm. Dis.* **Volume 14**, 1317–1322 (2019), https://doi.org/10.2147/COPD.S193037.

26. Taraldsen, K. *et al.* The long-term effect of being treated in a geriatric ward compared to an orthopaedic ward on six measures of free-living physical behavior 4 and 12 months after a hip fracture - a randomised controlled trial. *BMC Geriatr.* **15**, 160 (2015), https://doi.org/10.1186/s12877-015-0153-6.

27. Boehme, P. *et al.* How soon will digital endpoints become a cornerstone for future drug development? *Drug Discov. Today* **24**, 16–19 (2019), https://doi.org/10.1016/j.drudis.2018.07.001.

28. 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. NeuroEngineering Rehabil.* **13**, 46 (2016), https://doi.org/10.1186/s12984-016-0154-5.

29. Khandelwal, S. & Wickstrom, N. Gait Event Detection in Real-World Environment for Long-Term Applications: Incorporating Domain Knowledge Into Time-Frequency Analysis. *IEEE Transactions on Neural Syst. Rehabil. Eng.* **24**, 1363–1372 (2016), https://doi.org/10.1109/TNSRE.2016.2536278.

30. Orendurff, M. S. How humans walk: Bout duration, steps per bout, and rest duration. *The J. Rehabil. Res. Dev.* **45**, 1077–1090 (2008), https://doi.org/10.1682/JRRD.2007.11.0197.

31. Danks, K. A., Roos, M. A., McCoy, D. & Reisman, D. S. A step activity monitoring program improves real world walking activity post stroke. *Disabil. Rehabil.* **36**, 2233–2236 (2014), https://doi.org/10.3109/09638288.2014.903303.

32. Hickey, A., Del Din, S., Rochester, L. & Godfrey, A. Detecting free-living steps and walking bouts: validating an algorithm for macro gait analysis. *Physiol. Meas.* **38**, N1–N15 (2017), https://doi.org/10.1088/1361-6579/38/1/N1.

33. Aminian, K., Najafi, B., Büla, C., Leyvraz, P.-F. & Robert, P. Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes. *J. Biomech.* **35**, 689–699 (2002), https://doi.org/10.1016/S0021-9290(02)00008-8.

34. El-Gohary, M. *et al.* Continuous Monitoring of Turning in Patients with Movement Disability. *Sensors* **14**, 356–369 (2013), https://doi.org/10.3390/s140100356.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

35. Haji Ghassemi, N. *et al.* Turning Analysis during Standardized Test Using On-Shoe Wearable Sensors in Parkinson's Disease. *Sensors* **19**, 3103 (2019), https://doi.org/10.3390/s19143103.

36. 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), https://doi.org/10.1371/journal.pone.0196463.

37. Gulley, E., Ayers, E. & Verghese, J. A comparison of turn and straight walking phases as predictors of incident falls. *Gait & Posture* **79**, 239–243 (2020), https://doi.org/10.1016/j.gaitpost.2020.05.002.

38. Mariani, B., Jiménez, M. C., Vingerhoets, F. J. G. & Aminian, K. On-Shoe Wearable Sensors for Gait and Turning Assessment of Patients With Parkinson's Disease. *IEEE Transactions on Biomed. Eng.* **60**, 155–158 (2013), https://doi.org/10.1109/TBME.2012.2227317.

39. Jones, J. & Hunter, D. Qualitative Research: Consensus methods for medical and health services research. *BMJ* **311**, 376–380 (1995), https://doi.org/10.1136/bmj.311.7001.376.

40. Perry, J. *Ganganalyse: Norm und Pathologie des Gehens* (Urban & Fischer, München [u.a.], 2003), 1 edn.

41. Whittle, M. W. *Gait analysis: An introduction* (Butterworth-Heinemann, Edinburgh, 2005), 3. edn.

42. Dijkstra, B., Kamsma, Y. P. & Zijlstra, W. Detection of Gait and Postures Using a Miniaturized Triaxial Accelerometer-Based System: Accuracy in Patients With Mild to Moderate Parkinson's Disease. *Arch. Phys. Medicine Rehabil.* **91**, 1272–1277 (2010), https://doi.org/10.1016/j.apmr.2010.05.004.

43. 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. NeuroEngineering Rehabil.* **11**, 48 (2014), https://doi.org/10.1186/1743-0003-11-48.

44. Hase, K. & Stein, R. B. Turning Strategies During Human Walking. *J. Neurophysiol.* **81**, 2914–2922 (1999), https://doi.org/10.1152/jn.1999.81.6.2914.

45. Perry, J. *Gait Analysis: Normal and Pathological Function* (SLACK, Thorofare, New Jersey, 1992), 1. edn.

46. Fritz, S. & Lusardi, M. White paper: "walking speed: the sixth vital sign". *J. Geriatr. Phys. Ther.* **32**, 46–9 (2009).

47. Gaßner, H. *et al.* Clinical Relevance of Standardized Mobile Gait Tests. Reliability Analysis Between Gait Recordings at Hospital and Home in Parkinson's Disease: A Pilot Study. *J. Park. Dis.* **10**, 1763–1773 (2020), https://doi.org/10.3233/JPD-202129.

48. Warmerdam, E. *et al.* Long-term unsupervised mobility assessment in movement disorders. *The Lancet Neurol.* **19**, 462–470 (2020), https://doi.org/10.1016/S1474-4422(19)30397-7.

49. Vogel, C. *et al.* A Delphi study to build consensus on the definition and use of big data in obesity research. *Int. J. Obes.* **43**, 2573–2586 (2019), https://doi.org/10.1038/s41366-018-0313-9.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

50. Hsu, C. & Sandford, B. The Delphi Technique: Making Sense of Consensus. *Pract. Assessment, Res. & Eval.* **12**, 1–8 (2007), https://doi.org/10.7275/pdz9-th90.

51. Murphy, M. *et al.* Consensus Development Methods, and their Use in Creating Clinical Guidelines. *Heal. Technol. Assess.* **2**, 426–448 (1998), https://doi.org/10.4135/9781848608344.n24.

52. Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O. & Alberti, C. Using and Reporting the Delphi Method for Selecting Healthcare Quality Indicators: A Systematic Review. *PLoS ONE* **6**, e20476 (2011), https://doi.org/10.1371/journal.pone.0020476.

53. Diamond, I. R. *et al.* Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *J. Clin. Epidemiol.* **67**, 401–409 (2014), https://doi.org/10.1016/j.jclinepi.2013.12.002.

54. R Core Team. R: A Language and Environment for Statistical Computing (2016).

Acknowledgements

The MOBILISE-D project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No. 820820. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and the European Federation of Pharmaceutical Industries and Associations (EFPIA). BME gratefully acknowledges the support of the German Research Foundation (DFG) within the framework of the Heisenberg professorship program (grant number ES 434/8-1). SDD and LR are supported by the National Institute for Health Research (NIHR) Newcastle Biomedical Research Centre (BRC) based at Newcastle Upon Tyne Hospital NHS Foundation Trust and Newcastle University. The work was also supported by the NIHR/Wellcome Trust Clinical Research Facility (CRF) infrastructure at Newcastle upon Tyne Hospitals NHS Foundation Trust. CM is supported by the National Institute for Health Research (NIHR) Sheffield Biomedical Research Centre (BRC) based at Sheffield Teaching Hospital NHS Foundation Trust and University of Sheffield. All opinions are those of the authors and not the funders.

Author contributions

All authors contributed to the conceptualization of the study, as well as to definitions and terminology. FK conducted the study and analyzed the data. All authors wrote and/or edited the manuscript. All authors approved of the final manuscript.

Additional information

Competing interests

The authors declare no competing interests.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Data availability

The datasets generated and analysed in the study are available in the Zenodo repository, https://doi.org/10.5281/zenodo.4316564.

Code availability

We used only free and open-source software. The code is available on https://doi.org/10.5281/zenodo.4316739.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Tables

Table 1. Identified domains and terms needing consensus.

Domain	Term	Aspect
What are you doing?	Walking	Physiological
		Relation to walking bouts
Why are you doing it?	Purposeful	Characteristics
Where are you doing it?	Real-world	Characteristics
		Standardized measurement
		Clinical environment
How are you doing it?	Walking bout	Characteristics
	Walking speed	Physical aspect
		Granularity
		Relation to walking bouts
	Turning	Characteristics

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2. Participant background assessed in the first round.

Professional background (n=64)	
Both technical and clinical background	25.0 %
Clinical background	40.6 %
Technical background	34.4 %
Expertise in gait analysis (n=66)	
None	12.1 %
0-5 years	50.0 %
5-10 years	12.1 %
10-15 years	7.6 %
15-20 years	4.6 %
> 20 years	13.6 %
Free-living expertise (n=65)	
No	35.4 %
Yes	64.6 %
Clinical expertise (n=62)	
Patients and healthy participants	71.0 %
Healthy participants	11.3 %
Patients	17.7 %

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Table 3. Proportion of agreement and disagreement [%] of definitions. In round two, only those definitions were evaluated, which did not reach agreement in the first round. The lower limit of agreement was a priori defined as 75 %.

Term	Aspect	Disagreement	No opinion	Agreeement	Consensus
	Round one				
Walking	Physiological	6.1	3.0	90.9	yes
	Relation to walking bouts	0.0	3.0	97.0	yes
Purposeful	Characteristics	10.6	21.2	68.2	no
Real-world	Characteristics	4.6	1.5	93.9	yes
	Clinical environment	4.6	1.5	93.9	yes
	Standardized measurement	4.6	3.0	92.4	yes
Walking bout	Characteristics	15.1	6.1	78.8	yes
Walking speed	Physical definition	1.5	1.5	97.0	yes
	Granularity	1.5	6.1	92.4	yes
	Relation to walking bouts	15.2	24.2	60.6	no
Turning	Characteristics	3.0	3.0	94.0	yes
	Round two				
Purposeful	Characteristics	15.7	5.9	78.4	yes
Walking bout	Characteristics	11.8	5.9	82.3	yes
Walking speed	Relation to walking bouts	15.7	7.8	76.5	yes

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Table 4. Agreed definitions of terms related to real-world walking.

Term	Aspect	Definition
Walking	Physiological	Human **walking** is a method of locomotion and is defined as initiating and maintaining a forward displacement of the centre of mass in an intended direction involving the use of the two legs, which provide both support and propulsion. The feet are repetitively and reciprocally lifted and set down whereby at least one foot is in contact with the ground at all times[40,41]. **Walking** with walking aids is included in this definition. A **step** is the interval between the initial contacts of the ipsi- and contralateral foot[40] and corresponds to the forward displacement of the foot together with a forward displacement of the trunk[42]. A **stride** is the interval between two successive initial contacts of the same foot. As such, a **stride** is equivalent to the gait cycle and every stride contains two **steps**[40].
	Relation to walking bouts	**Walking** is made up of walking bouts and is equivalent to taking steps/stepping forward (thus stepping in place does not constitute walking) and is defined as starting from initial contact for the initial step until ending with full floor contact of the foot making the last step[42].
Purposeful	Characteristics	**Purposeful walking** includes an intentional component of the movement (e.g. getting to the bathroom, catching the bus, going to the grocery store, going for a walk in the park, etc.). **Purposeful walking** may constitute certain characteristics (e.g. more constant walking velocity, lower variability of gait characteristics, straighter direction of locomotion than non-purposeful walking, specific context, etc.). Those gait characteristics are quantified based on discrete walking bouts.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Real-world	Characteristics	**Real-world** relates to the context in which walking takes place – that is free-living, unsupervised, uncontrolled and non-standardised. As such, it is unscripted as there are no instructions to the subject who does not need to interact with the wearable device(s). **Real-world** actions occur in non-simulated everyday situations in unconstrained environments with minimal consciousness of being tested. It is equivalent to actions at home or in the community over continuous periods of time[28]. Synonymous terms are (environment of) **daily living**, or relating to **daily-life**. **Home environment** is used synonymously to **real-world** and **daily-life** without a separation of indoor and outdoor environment[11]. **Real-world** is distinct from laboratory-based[43], supervised (= fully controlled and observed), and semi-controlled (walking 'freely' but with supervision) tests. It also is different from scripted/instructed walking, which can take place in the home or lab (such as walking tests like the 4x10m test, 6-minute walk test (6MWT) and timed up and go (TUG)).
	Clinical environment	Free walking in hospitals is part of the **real-world** definition, but standardized supervised tests in a hospital are not. This excludes instructed actions, e.g. by medical professionals.
	Standardized measurement	Home-based tests, which are semi-standardized measurements performed in the home environment in a controlled or semi-controlled environment (such as short walk tests), are thus not regarded as being part of **real-world**. Home-based tests can nevertheless be an alternative to clinical tests and might be easier to conduct operationally and analyse than continuous monitoring (assuming standardized instructions).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Walking bout	Characteristics	A **walking bout (WB)** is a walking sequence containing at least two consecutive strides of both feet (e.g. *R-L-R-L-R-L* or *L-R-L-R-L-R*). Start and end of a **walking bout** are determined by a resting period or any other activity (non-walking period). The initial step of a **WB** follows a non-walking period and the final step precedes the next non-walking period.
Walking speed	Physical definition	**Walking speed (WS)** is the distance covered by the whole body within a certain time interval / per unit time of walking. It is measured in meters per second and is the magnitude of the velocity vector (velocity includes direction and magnitude of walking)[41].
	Granularity	**Walking speed** can be estimated at different granularities: • Instantaneous WS varies from one instant to another during the walking cycle[41] • Step-wise WS is the ratio between step distance (length) and step time[28] • Stride-wise WS[33] • Averaged over WBs • Averaged over other time intervals (hourly, daily, weekly) based on multiple WBs The granularity by which the WS is assessed should be related to clinical parameters for each population separately.
	Relation to walking bouts	**Walking speed** will be assessed with regard to walking bouts. Thus, the minimal length of one walking bout required to assess average walking speed is based on a sequence of 2 consecutive strides (e.g. *R-L-R-L-R-L* or *L-R-L-R-L-R*).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Turning	Characteristics	The process of **turning** consists of decelerating the forward motion, rotating the body as a whole, and stepping out toward the new direction[44]. Thus, **turning** includes a change of walking direction and change in angular orientation including a rotational movement of the body around the longitudinal axis. **Turning**, curvilinear walking, and straight walking involve different neuromotor strategies and need to be discriminated.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.18.20248404; this version posted December 22, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Figures

	Strongly disagree	Somewhat disagree	No opinion	Somewhat agree	Strongly agree
Do you agree to the above definition?	○	○	○	○	○

Figure 1. Five-point Likert scale for assessment of agreement.