PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Walking-related digital mobility outcomes as clinical trial endpoint measures: Protocol for a scoping review
AUTHORS	Polhemus, Ashley; Bergquist, Ronny; Bosch de Basea, Magda; Brittain, Gavin; Buttery, Sara; Chynkiamis, Nikolaos; dalla Costa, Gloria; Delgado Ortiz, Laura; Demeyer, Heleen; Emmert, Kirsten; Garcia Aymerich, Judith; Gassner, Heiko; Hansen, Clint; Hopkinson, Nicholas; Klucken, Jochen; Kluge, Felix; Koch, Sarah; Leocani, Letizia; Maetzler, Walter; Micó-Amigo, M. Encarna; Mikolaizak, A. Stefanie; Piraino, Paolo; Salis, Francesca; Schlenstedt, Christian; Schwickert, Lars; Scott, Kirsty; Sharrack, Basil; Taraldsen, Kristin; Troosters, Thierry; Vereijken, Beatrix; Vogiatzis, Ioannis; Yarnall, Alison; Mazza, Claudia; Becker, Clemens; Rochester, Lynn; Puhan, Milo; Frei, Anja

VERSION 1 - REVIEW

REVIEWER	Paolo Eusebi Regional Health Authority of Umbria
REVIEW RETURNED	04-Apr-2020

GENERAL COMMENTS	The protocol is clear and well-written. Research findings would add clear value.

REVIEWER	Pier Spinazze Imperial College London
REVIEW RETURNED	13-Apr-2020

GENERAL COMMENTS	Overall, well structured and defined protocol. The one aspect missing for me, is an indication of what wearable sensors include in the background i.e. does this include mobile phones and digital phenotyping, IoT, clinical grade sensors and consumer devices etc. Would also be good to highlight potential limitations of the study especially potentially found in the heterogeneity of assessment models and how these will be compared.

REVIEWER	Dr Rajesh Raj Launceston General Hospital Australia
REVIEW RETURNED	13-Apr-2020

GENERAL COMMENTS	The area selected for scoping review is relevant; no similar scoping reviews have been published. A large group of researchers are involved in the design and conduct of this review and this is evident in the extensive documentation. However, I feel that this protocol paper will benefit from being rewritten with an emphasis on simplicity. As discussed under eligibility criteria below, it appears that the authors are quite restrictive in their selection process for articles to include in the review. This appears to be in contrast to the usual paradigm for a scoping review which seeks to include all the relevant material pertaining to the topic at hand. Please see below for more detailed comments. Page 1: Line 37: proximal femoral fractures are not included in the keywords. Please correct if this is omission ; otherwise justify. Page 4, line 4 The phrase “exploring the potential of“ is unnecessary in the title. By definition, the scoping review explores the potential of the subject of the review. Similarly, for clarity, I suggest you consider including the term “walking-related digital mobility outcomes” in the title. Your scoping review does not include all digital mobility outcomes; only those related to walking. Page 5, Line 20-21: “multi-diagnostic approach” is not a standard phrase. Please re-word this. Page 6 Lines 32-38: this paragraph, seeking to explain the limitations of existing measures can be much improved by providing a few examples, if needed within parentheses. As it stands, the concept is rather abstract. Page 7: Line 40: in this section, please begin by describing how the research questions were identified. Line 43: please provide“box2”. In the absence of clearly demarcated research questions, this protocol is very heavy on text, making it difficult for the reader to clearly identify what the research questions are. If the authors intended to refer to table 2, that table is insufficient for the purpose of listing the research questions. Table 2 only suggests the aims of the research questions, not the research questions themselves. Please correct this. For clarity, you should include a table that poses the 4 research questions as simple questions rather than as deducible from the aims. Line 58: “Prognosis of clinically relevant outcomes” needs to be succinctly defined. Page 8: Line 53-55: while this definition describes the mobility outcomes, it does not define what you mean by “digital”. Please include this. Page 10: Line 50: this section on eligibility criteria needs to be rewritten for

	simplicity and ease of reading and understanding. I would like to remind the authors that the purpose of publishing a protocol paper is to make your protocol easily understood by the reader such that your scoping review can be easily replicated. In the current format, there is extensive mixing of general and question-specific eligibility criteria, inclusion and exclusion characteristics and minimum data sets which differ according to each research question. It might be worthwhile reflecting on the rather stringent inclusion criteria and the impact these stringent criteria may have on the "scoping" approach to the literature. By convention, the scoping review should attempt to collect all the available information on the topic. Please consider whether, with the stringent criteria that you have, your review is more suited to a systematic review process. The stringency in your inclusion/exclusion/eligibility criteria at the moment apply to the kinds of studies that you include, the criteria by which you consider a digital mobility outcome suitable for inclusion, the minimum number of participants in a study to be eligible and the lack of consideration of grey literature in your inclusion characteristics. All these are fairly restrictive and not typical for a scoping review format. Please justify. Similarly, you are only considering DMOs for 4 clinical conditions. This needs to be emphasised more prominently in relevant area in the text, including in the introduction section of the abstract.. You need to provide justification for picking these four diseases; you also need to ensure that the reader is aware that only four clinical conditions are being considered. Page 11: Line 6-7: please define "free –living conditions". This phrase is repeatedly used in the text without a clear definition. Line 34: please include justification for "minimum dataset" in the scoping review. Ideally, a scoping review summarises all the available knowledge. If these stringent criteria for selection of studies was employed in order to avoid unnecessary crowding of results, then this ought to be mentioned as a limitation of the study – that because restrictive criteria were used, some aspects of the literature could have been missed.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Paolo Eusebi

Institution and Country: Regional Health Authority of Umbria, Italy

Please state any competing interests or state 'None declared': None declared

The protocol is clear and well-written. Research findings would add clear value.

We thank Dr. Eusebi for these comments and for his consideration of our manuscript.

Reviewer: 2

Reviewer Name: Pier Spinazze

Institution and Country: Imperial College London, UK

Please state any competing interests or state ‘None declared’: None declared

Overall, well structured and defined protocol. The one aspect missing for me, is an indication of what wearable sensors include in the background i.e. does this include mobile phones and digital phenotyping, IoT, clinical grade sensors and consumer devices etc. Would also be good to highlight potential limitations of the study especially potentially found in the heterogeneity of assessment models and how these will be compared.

We thank Dr. Spinazze for this feedback and for his consideration of the manuscript. We plan to include any type of wearable or non-wearable digital technologies to capture DMOs. This could include mobile phones, wearables, instrumented walkways, and even stopwatches or pedometers. We have clarified this on page 12, line 16, of the proof, which now reads:

“We will include DMOs produced through any digital or electronic measurement method, including wearable sensors, instrumented walkways or treadmills, optometric systems, force plates, mobile phones, stopwatches, and pedometers, among others.”

We have also added a section to discuss the potential limitations of the study, entitled “Discussion and Limitations” on p. 16. This includes a discussion of the limits we set on our scope and how we will manage/interpret heterogeneity the studies that we do identify.

Reviewer: 3

Reviewer Name: Dr Rajesh Raj

Institution and Country: Launceston General Hospital, Australia

Please state any competing interests or state ‘None declared’: None declared

The area selected for scoping review is relevant; no similar scoping reviews have been published. A large group of researchers are involved in the design and conduct of this review and this is evident in the extensive documentation. However, I feel that this protocol paper will benefit from being rewritten with an emphasis on simplicity. As discussed under eligibility criteria below, it appears that the authors are quite restrictive in their selection process for articles to include in the review. This appears to be in contrast to the usual paradigm for a scoping review which seeks to include all the relevant material pertaining to the topic at hand. Please see below for more detailed comments.

We thank Dr. Raj for these valuable comments. Our intention was not to be restrictive; on the contrary, we aim to map as much of the literature as is feasible. Our original searches showed that the terminology associated with this field are highly inconsistent, and that analyses that are of interest for this review are often not reported at the abstract level. Therefore, a very broad search strategy resulting in tens of thousands of references is necessary. We strongly believe that a systematic review approach would be inappropriate here, since the literature is unmapped and our research questions and eligibility criteria remain broad.

The diversity represented in the literature posed several challenges during screening, which we attempted to address through very detailed, though still inclusive, criteria. Due to the scope of this review, our reviewers represent a variety of clinical and technical backgrounds. To ensure that our references are assessed correctly and consistently, we opted to provide reviewers with explicit and exhaustive instructions on how to interpret as many scenarios as possible. These explicit instructions were intended to make the review more inclusive and more rigorous (since many abstracts do not report relevant analyses, and many reviewers specialized in only one of the included disease areas). For clarity, we have simplified the criteria listed in this paper to reflect the core criteria we intended to apply, rather than the list of all scenarios that could meet each criterion. These are noted in response to specific comments below.

Our final searches yielded approximately 50,000 references, which was reduced to approximately 20,000 after de-duplication. Even with this broad strategy, we suspect that it is not possible to identify all literature measuring gait speed and other walking-related DMOs, simply because they are not consistently reported. Therefore, we do not necessarily intend to produce an exhaustive list of all previous studies: this is simply not feasible. Instead, we will adopt a semi-structured approach to map clinically-relevant trends across this large, fragmented body of literature. To do this, we will limit some dimensions of study scope to lengthy lists (i.e., the 32 DMOs, the 113 measures assessed in RQ2, and the 29 outcomes assessed in RQ3) and will apply basic quality thresholds (i.e., a minimum number of participants). In alignment with the reflexive approach outlined by Arksey and O'Malley, we defined a systematic method to amend these lists if additional instruments meeting these criteria are identified during study conduct. This approach allows us to remain inclusive with regard to terminology and methodology while ensuring feasibility. We have explained this more clearly in the section entitled "Definitions and Study Scope" (p. 8, line 52) and in our responses below. We also added a section addressing study limitations on p. 16 of the proof

Page 1, Line 37: proximal femoral fractures are not included in the keywords. Please correct if this is omission ; otherwise justify.

This was an omission and has been corrected. We now include "orthopaedic & trauma surgery", which was the closest available option to proximal femoral fracture.

Page 4, line 4

The phrase "exploring the potential of" is unnecessary in the title. By definition, the scoping review explores the potential of the subject of the review.

Similarly, for clarity, I suggest you consider including the term "walking-related digital mobility outcomes" in the title. Your scoping review does not include all digital mobility outcomes; only those related to walking.

As suggested, we changed the title of the manuscript to "Walking-related digital mobility outcomes as clinical trial endpoint measures: Protocol for a scoping review"

Page 5, Line 20-21: "multi-diagnostic approach" is not a standard phrase. Please re-word this.

The phrase “multi-diagnostic approach” was removed, and the abstract now reads “We will include four disease areas: Parkinson’s disease, multiple sclerosis, chronic obstructive pulmonary disease, and proximal femoral fracture.”

Page 6 Lines 32-38: this paragraph, seeking to explain the limitations of existing measures can be much improved by providing a few examples, if needed within parentheses. As it stands, the concept is rather abstract.

To address this, we added multiple examples of walking tests and specific scenarios demonstrating the listed limitations. The text now reads (p. 6, line 23):

Unfortunately, current mobility measures pose critical limitations. Clinical trials traditionally employ two types of mobility assessments: patient reported outcome instruments (PROs) and clinical gait assessments. PROs enable patients to report perceptions of their own mobility in a standardised manner,[13] though results may be subject to recall bias.[14–16] Clinical assessments, such as timed walking tests, are typically more objective. However, many still require clinical interpretation and are subject to high inter-rater variability.[17,18] For example, Zhang et al. conducted a sensitivity analysis to demonstrate the potential impact of inter-rater variability in clinical trials by assessing a trial’s primary outcome, the Expanded Disability Status Score (a common measure of function and ambulation in multiple sclerosis) in duplicate. [19] Duplicated ratings differed in over 30% of patients, affecting estimates of treatment effect. Additionally, clinical assessments are often infrequently acquired and may not be representative of real-world behaviour.[14,20,21] Compared to real-world walking, patients consistently walk faster and produce higher-quality gait patterns during “normal” walking in laboratory settings.[20,22,23] These challenges have prompted calls for more sensitive, reliable mobility measures in clinical trials.[21,24]

Page 7:

Line 40: in this section, please begin by describing how the research questions were identified.

We have added a brief explanation for the rationale for the research questions to supplement the “Study Rationale and Objectives” section (page 7, line 33). The section now reads, “To be used as clinical trial endpoints, measures must be valid, clinically meaningful, and responsive to change. Preliminary searches revealed a highly fragmented body of literature, with no overarching review describing these characteristics. Therefore, this study will map the literature across four research questions (Box 1) in a set of walking-related DMOs (Table 1).”

Line 43: please provide “box2”. In the absence of clearly demarcated research questions, this protocol is very heavy on text, making it difficult for the reader to clearly identify what the research questions are. If the authors intended to refer to table 2, that table is insufficient for the purpose of listing the research questions. Table 2 only suggests the aims of the research questions, not the research questions themselves. Please correct this. For clarity, you should include a table that poses the 4 research questions as simple questions rather than as deducible from the aims.

Box 1, containing the research questions as simple questions, is provided on page 7 (line 43) of the proof, under the section “Identifying the Research Question.” It describes the following objective and research questions:

Objective: Map existing evidence describing the discriminant ability, construct validity, prognostic value, and responsiveness of walking-related digital mobility outcomes (DMOs)

- RQ1: What differences in DMOs have been identified between the four included populations and healthy controls?
- RQ2: What is the evidence on the associations between DMOs and clinically-relevant measures of physical function, health-related quality of life, symptoms, and disease severity in each of the included populations?
- RQ3: What is the evidence on the prognostic value of DMOs in each of the included populations?
- RQ4: In which contexts and for what purposes have DMOs been used as endpoints in interventional studies in each of the included populations?

Line 58: “Prognosis of clinically relevant outcomes” needs to be succinctly defined.

On page 8, (line 36) we now briefly define “the prognostic value of DMOs” as “their ability to predict future health outcomes.” Our definition of clinically-relevant outcomes is further discussed in the “Definitions and Study Scope” section of the manuscript, which is introduced immediately after the research questions are explained.

Page 8:

Line 53-55: while this definition describes the mobility outcomes, it does not define what you mean by “digital”. Please include this.

To clarify this, we have added the following text to Box 2 on p. 9, line 52: “In this case “digital” measures refer to those objectively derived from electronic systems, as opposed to qualitative, paper-based, or self-reported measures.”

Page 10:

Line 50: this section on eligibility criteria needs to be rewritten for simplicity and ease of reading and understanding. I would like to remind the authors that the purpose of publishing a protocol paper is to make your protocol easily understood by the reader such that your scoping review can be easily replicated. In the current format, there is extensive mixing of general and question-specific eligibility criteria, inclusion and exclusion characteristics and minimum data sets which differ according to each research question. It might be worthwhile reflecting on the rather stringent inclusion criteria and the impact these stringent criteria may have on the “scoping” approach to the literature.

By convention, the scoping review should attempt to collect all the available information on the topic. Please consider whether, with the stringent criteria that you have, your review is more suited to a systematic review process.

We strongly believe that a systematic review would be inappropriate given the current status of the literature. It is unmapped, highly fragmented, and it is unclear which relationships are important or relevant, either in clinical practice or for systematic review. Terms and norms vary by disease area and between functional circles (i.e., technologists vs. clinical researchers). A systematic review would be premature. However, scoping review methodology is the perfect tool to devise a path forward in such circumstances. Though we perhaps implement stricter criteria than other scoping reviews, this serves a purpose: The review must be feasible, interpretable, and must be conducted in a reasonable timeframe. An exhaustive review, as valuable as it may be, would take several years and would quickly become obsolete. The scoping review presented here represents a compromise: We do not necessarily intend or anticipate study scope to exhaustively identify all previous literature. Rather, it is intended to identify trends in the literature to enable targeted systematic reviews in the future. It is much broader than permitted under systematic review methodology, and will map the literature at a higher-level than is typical of a systematic review. However, due to the vast amount of literature and limitations in searching due to diverse terminology, it was necessary to apply some structure to the review a-priori (i.e., the DMOs, the measures, the outcomes). We have amended the manuscript to clarify this position on p.8, in the section entitled “Definitions and Study Scope” and again in the discussion of limitations on p. 16.

The stringency in your inclusion/exclusion/eligibility criteria at the moment apply to the kinds of studies that you include, the criteria by which you consider a digital mobility outcome suitable for inclusion, the minimum number of participants in a study to be eligible and the lack of consideration of grey literature in your inclusion characteristics.

All these are fairly restrictive and not typical for a scoping review format. Please justify.

We have attempted to simplify and more clearly justify our scope and eligibility criteria. Some of these criteria, such as the DMOs, were necessary to set search terms: searches with terms related to digital methods and technologies were unreliable because of inconsistent terminology and poor reporting. Others were set to ensure that the highly heterogenous results would be interpretable. While they may appear restrictive, they are necessary to ensure that the study is feasible. Without them, the search and subsequent review would encompass hundreds of thousands of references, which is not possible even with this large study team. This is further discussed in the section “Definitions and Study Scope” (p. 8).

The eligibility criteria section, starting on p. 11 of the proof (line 51), was also simplified and restructured. Please note that we also removed two eligibility criteria: the requirement for 20 events and adjusting for age, sex, and disease severity in RQ3, and the requirement for baseline and follow-up measures in RQ4.We have not yet reached the full-text stage of the review in which we would have applied these criteria. Instead, we now plan to compare adjusted and unadjusted models in a sensitivity analysis.

We also include grey literature in this review and include searches of several grey literature databases, as described on page 10 (line 40):

“We will include peer-reviewed and grey literature, including journal articles, reports, research letters, conference papers, doctoral theses, and other publications reporting original results. MEDLINE, CINAHL, Scopus, Web of Science, EMBASE, IEEE Digital Library, and the Cochrane Library will be

searched for eligible peer-reviewed literature. ACM Digital Library, ProQuest Dissertations, Open Grey, and the National Information Center's Health Services Research Projects in Progress Database will be searched to identify relevant grey literature."

Similarly, you are only considering DMOs for 4 clinical conditions. This needs to be emphasised more prominently in relevant area in the text, including in the introduction section of the abstract. You need to provide justification for picking these four diseases; you also need to ensure that the reader is aware that only four clinical conditions are being considered.

The limit of four clinical conditions was emphasized in several places in the manuscript, such that the reader should be clearly aware of the limited scope. We now mention 4 conditions in the "Strengths and Limitations" box, list these populations in the Study Rationale and Objectives (p. 7, line 6), methods (p. 7; line 37), and eligibility criteria (p.12, line 50) and frequently reference "the four included populations". As requested, we also moved the list of disease areas to the introduction section of the abstract.

The rationale for including these populations is justified in the "Identifying the Research Question" section (p. 7, line 39), which reads:

"The Mobilise-D consortium selected these disease areas as exemplars for DMO development due to their diverse aetiologies of mobility impairment, high public health burden, and existing evidence base. [26,45–47]"

Page 11:

Line 6-7: please define "free –living conditions". This phrase is repeatedly used in the text without a clear definition.

We removed this term, and now use the term "real-world walking" instead. We define "real-world walking" in Box 2 on page 9 according to the Mobilise-D consortium's definition:

Per the Mobilise-D consortium, "'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. 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.[23] … Real world walking is distinct from laboratory-based,[70] supervised (fully controlled and observed), and semi-controlled (walking 'freely' but with supervision) tests. It also is different from scripted or instructed walking, which can take place in the home or lab."

Line 34: please include justification for "minimum dataset" in the scoping review. Ideally, a scoping review summarises all the available knowledge.

If these stringent criteria for selection of studies was employed in order to avoid unnecessary crowding of results, then this ought to be mentioned as a limitation of the study – that because restrictive criteria were used, some aspects of the literature could have been missed.

As suggested, this was clarified in the section entitled “Definitions and Study Scope” (p. 8, line 52). It now reads:

“Preliminary searches revealed that an exhaustive review is infeasible due to inconsistent terminology and reporting practices. Thus, we do not necessarily intend to produce an exhaustive list of all previous studies. Instead, we will adopt a semi-structured approach to map clinically-relevant trends across this large, fragmented body of literature. To do this, we will limit some dimensions of study scope to lengthy lists (i.e., the DMOs, the measures assessed in RQ2, and the outcomes assessed in RQ3) and will apply basic quality thresholds (i.e., a minimum number of participants). This approach allows us to remain inclusive with regard to terminology and methodology while ensuring feasibility. The decisions used to set this scope are described below. Because understanding of seemingly common terms differs across disciplines, defining the concepts addressed by this review was not trivial. Therefore, our operational definitions of key concepts such as “mobility,” “walking,” “real-world,” and “digital mobility outcomes” are clearly defined in Box 2.”

VERSION 2 – REVIEW

REVIEWER	DR Rajesh Raj Launceston General Hospital Australia
REVIEW RETURNED	18-May-2020

GENERAL COMMENTS	I must congratulate the authors for the extensive revisions in the text. The entire article now reads much better. My previous concerns have been well addressed.