## Measurement properties of performance-based instruments to assess mental function during activity and participation in persons who have survived a stroke

Kristensen, Lola Qvist; Tulder, M.W. van; Kristensen, Hanne Kaae; Muren, Marie Almkvist; Mokkink, Lidwine B; Oestergaard, Lisa Gregersen

***published in***
Scandinavian Journal of Occupational Therapy
2023

***DOI (link to publisher)***
10.1080/11038128.2022.2053740

***document version***
Publisher's PDF, also known as Version of record

Taylor & Francis
Taylor & Francis Group

Scandinavian Journal of Occupational Therapy

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/iocc20

# Measurement properties of performance-based instruments to assess mental function during activity and participation in persons who have survived a stroke: A systematic review protocol

**Lola Qvist Kristensen, Maurits W. van Tulder, Hanne Kaae Kristensen, Marie Almkvist Muren, Lidwine B. Mokkink & Lisa Gregersen Oestergaard**

**To cite this article:** Lola Qvist Kristensen, Maurits W. van Tulder, Hanne Kaae Kristensen, Marie Almkvist Muren, Lidwine B. Mokkink & Lisa Gregersen Oestergaard (2023) Measurement properties of performance-based instruments to assess mental function during activity and participation in persons who have survived a stroke: A systematic review protocol, Scandinavian Journal of Occupational Therapy, 30:2, 222-227, DOI: 10.1080/11038128.2022.2053740

**To link to this article:** https://doi.org/10.1080/11038128.2022.2053740

SCANDINAVIAN JOURNAL OF OCCUPATIONAL THERAPY
2023, VOL. 30, NO. 2, 222–227
https://doi.org/10.1080/11038128.2022.2053740

Taylor & Francis
Taylor & Francis Group

PROTOCOL FOR A REVIEW STUDY



# Measurement properties of performance-based instruments to assess mental function during activity and participation in persons who have survived a stroke: A systematic review protocol

Lola Qvist Kristensen[a], Maurits W. van Tulder[a,b,c], Hanne Kaae Kristensen[d,e], Marie Almkvist Muren[a], Lidwine B. Mokkink[f] and Lisa Gregersen Oestergaard[a,g,h]

[a]Department of Physiotherapy and Occupational Therapy, Aarhus University Hospital, Aarhus, Denmark; [b]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark; [c]Department of Human Movement Sciences and Amsterdam Research Institute MOVE, Faculty Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, the Netherlands; [d]Department of Clinical Research, University of Southern Denmark, Odense, Denmark; [e]Health Sciences Research Center, University College Lillebaelt, Odense, Denmark; [f]Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Epidemiology and Data Science, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands; [g]DEFACTUM, Central Denmark Region, Aarhus, Denmark; [h]Department of Public Health, Aarhus University, Aarhus, Denmark

**ABSTRACT**

**Background:** A frequent consequence of stroke is impaired mental function, which often affects the ability to perform activities and participate in life situation. In occupational therapy practice, performance-based instruments during activity and participation are often used. However, it is important to assess if the instruments used are valid, reliable and responsive.

**Objective:** The objective of this systematic review is to investigate measurement properties of performance-based instruments to assess mental function during activity and participation in persons who have survived stroke.

**Material and methods:** Systematic database searches of PubMed, EMBASE, CINAHL, PsycINFO, and OTseeker will be conducted. COSMIN Risk of Bias checklist will be used to evaluate methodological quality of the included articles. Measurement properties of the included studies will be rated against criteria for good measurement properties. The overall evidence of each measurement property per instrument is graded using a modified GRADE approach.

**Results:** Results will be presented in text and tables.

**Conclusions:** Conclusion will be drawn up-on the overall evidence to give recommendations on the most suitable instrument.

**Significance:** It is expected that findings of the review will provide evidence to guide professionals in the selection of a performance-based instruments to measure mental function in practice and research.

**Trial registration number:** PROSPERO CRD42018086744

**ARTICLE HISTORY**
Received 12 August 2021
Revised 20 January 2022
Accepted 11 March 2022

**KEYWORDS**
Stroke; cognitive impairment; measurement properties; validity; reliability; ICF; rehabilitation

## Background

A frequent consequence of stroke is impaired mental function [1,2]. According to The International Classification of Functioning, Disability and Health (ICF), the domain mental function includes (I) global functions, which encompass unconsciousness, regulation of arousal and mental state, and (II) specific functions, which encompass consciousness related to thought and behavioural cognition [3]. In a review of subjective cognitive complaints after stroke, the prevalence of patient-reported impaired mental function was 92% [4,5]. Examples of impairments reported by persons who have survived stroke were decreased memory, mental speed, and concentration difficulties in both the acute and chronic phase [5,6]. Furthermore, impaired mental function often affects the ability to perform activities and participate in everyday life [7–9]. Consistent with ICF, activity and participation encompass the execution of tasks and involvement in life situation [3]. In a cohort study of cognitive function in stroke including 197 persons who have survived stroke, it was reported that the presence of impaired mental functions related to language or executive functions (e.g. initiating, planning and problem solving) hampered the participation in the person's life situation even six months after discharge from hospital, rehabilitation unit or geriatric day hospital [10].

CONTACT Lola Qvist Kristensen lola.qvist@rm.dk Department of Physiotherapy and Occupational Therapy, Aarhus University Hospital, Aarhus, Denmark

Neuropsychological tests of mental function are often observation-based using pen and paper tests like the Mini-Mental State Examination (MMSE), Montreal Cognitive Screen (MoCA) or drawing tasks, which may be useful for a quick screen [11,12]. In the literature, a lack of generalisation to real-life settings in pen and paper tests has been reported, because they fail to evaluate the persons' ability to perform real-life activities [9,13,14]. In occupational therapy practice, performance-based instruments during activity and participation, like ADL-focused Occupation-based Neurobehavioral Evaluation (A-ONE) and Executive Function Performance Test (EFPT), are often used [15,16]. This to assess mental function and evaluate its impact on real-life activities, e.g. ADL performance [14,17]. In assessing this, reliable, valid and responsive instruments are needed to limit insufficient mental function assessment and consequences such as adaptation of rehabilitation and reduced autonomy in ADL [18–20]. A similar review on individuals with traumatic brain injury has been published showing a lack of evidence of reliable, valid and responsive performance-based measurements instruments to assess mental function during activity and participation [21]. The lacking use of valid measurement instruments can both be due to lack of skills of the occupational therapists, and due to lack of validation of the instruments available for use in occupational therapy. The objective of this systematic review is to investigate measurement properties of performance-based instruments to assess mental function during activity and participation persons who have survived stroke.

## Material and methods

### Study design and registration

This systematic review is registered in the International Prospective Register of systematic reviews (PROSPERO) (registration number CRD42018086744). The Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 checklist has guided this protocol [22]. This review will be conducted in accordance with the COnsensus-based Standards for the selection of health Measurements INstruments (COSMIN) methodology [23,24].

### Search strategy

The search strategy described is used for a series of systematic reviews on different types of acquired brain injury, however this protocol only includes studies targeting persons who have survived a stroke [21]. Systematic database searches of PubMed, EMBASE, CINAHL, PsycINFO, and OTseeker will be conducted from their date of inception. The search strategy includes a mixture of three literature search blocks: 'acquired brain injury', 'mental function' and 'method of assessment'. These search blocks will be combined with a search filter to identify studies on measurement properties in the databases PubMed, CINAHL and EMBASE (www.cosmin.nl) [25]. A translation of the published search filter will be applied in PsycINFO and OTseeker. Database searches will be conducted with controlled vocabulary such as Medical Subjects Headings (MeSH), EMTREE, CINAHL headings and Thesaurus, supplemented with applicable free-text terms to identify relevant articles that are not categorised. In addition to the database search, a hand search will be conducted by screening references from included studies for additional records meeting the inclusion criteria. To gather all identified studies, the online bibliographic program RefWorks (www.refworks.com) will be used for study upload.

### Selection of studies

For inclusion, studies should report on performance-based outcome measurement instruments during activity and participation and within at least one mental function subdomain [3,26] (see Table 1). Studies need to report on the development of the outcome measurement instruments or on at least one of the measurement properties defined in the COSMIN taxonomy [27] (Table 1). The development process is not a measurement property, but it is still recommended to evaluate the development of an instrument as this information is required when assessing the content validity of an instrument. Studies will be included regardless of severity of stroke (mild, moderate, severe). No language restrictions will be applied to the search strategy. If grey literature is located during free-text or reference search, it will be included. Exclusion criteria are studies not available in full-text or studies with mixed populations and no separate analysis of persons who have survived stroke.

### Selection procedure

Selection procedure will be performed by three of the review authors, all with clinical and research experience in the field of stroke. Duplicates will be removed

**Table 1.** Terms of mental functions, activity and participation and measurement properties.

| ICF subdomains of mental functions | ICF classification of activity and participation | COSMIN risk of bias checklist of measurement properties |
|---|---|---|
| Attention functions | Learning and applying knowledge | * Development |
| Memory functions | General tasks and demands | Content validity |
| Psychomotor functions | Communication | Structural validity |
| Emotional functions | Mobility | Internal consistency |
| Perceptual functions | Self care | Cross-cultural validity |
| Thought functions | Domestic life | Reliability |
| Higher-level cognitive functions | Interpersonal interactions and relationships | Measurement error |
| Mental functions of language | Major life areas and community | Criterion validity |
| Calculation functions | Social and civic life | Hypothesis testing |
| Mental functions of sequencing complex movements | | Responsiveness |
| Experience of self and time functions | | |

*Not a measurement property but recommended to be evaluated according to the COSMIN manual.
References: [3,25,27].

using the duplicates function of the reference manager RefWorks. After removing duplicates, two review authors will independently screen all titles and abstracts of the retrieved articles and read the full text of all potentially eligible studies. If disagreement occurs, a third review author will be consulted to reach consensus. Details on the process of study selection will be illustrated in a flowchart.

### *Appraisal of the methodological quality of included studies*

The methodological quality assessment of the included studies will be performed independently by two review authors using the COSMIN Risk Of Bias (RoB) checklist [24,28,29]. The COSMIN RoB includes following measurement properties: Development, content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypotheses testing and/or responsiveness (Table 1) [23,28]. For the evaluation of hypothesis testing and responsiveness, hypotheses for each study will be developed prior to the evaluation, depending on the focus of the study. Assuming no gold standard is available, hypotheses will be formulated using a construct approach. This includes (1) the expected direction of correlations between scores of the instrument of interest and instruments with similar or unrelated constructs; we expect correlations between scores of instruments with similar constructs and no correlations between scores of instruments measuring different constructs. (2) The expected difference between subgroups; we expect the instrument to distinguish between persons, who have survived a stroke – or specified impaired mental functions related to this (e.g. aphasia) – and persons without stroke or related specified impaired mental functions (e.g. aphasia) [30].

The ten boxes for measurement properties of the COSMIN RoB checklist (Table 1) consist of multiple standards, which are each rated on a four-point scale as either very good, adequate, doubtful or inadequate, with sometimes the possibility to rate a standard not applicable. The lowest score within a box determines the overall quality of the study [24,28,29]. If disagreement occurs, a third review author will be consulted.

The COSMIN checklist was original developed for self-reported measurement instruments, but has now been expanded to also include performance-based and clinician-reported measurement instruments [24]. Yet, for evaluation of development and content validity, the COSMIN RoB checklist has been adapted by the authors of this review in order to fit when using it on performance-based instruments (see adapted version in appendix). The main adaptation in the box on Development included adding standards about the involvement of professionals in the development of the design and in the pilot test. The inclusion of professionals means that the standards can be used on a study including qualitative information from either patients or professionals (e.g. focus group interviews), as we do not expect both groups to be included in one study. The box content validity was extended by adding the view of the professionals to the standards about assessing relevance and comprehensiveness.

### *Data extraction*

Two review authors will independently complete a standardised data extraction form. Variables of the included studies will be extracted and presented in tables:

- A table of the study characteristics including author/year of publication, country for the

publication/language of the instrument, sample size, stroke severity, age of the study population, activity and/or participation, measurement properties assessed and mental function subdomains assessed [3,23,26].

- A table of the methodological quality of each study per measurement property, including ratings of good measurement properties per study result [23].
- A table of the overall evidence grade of each measurement property per instrument using the modified GRADE approach.

### Data synthesis of included outcome measures

Measurement properties of the included studies will be rated against criteria for good measurement properties as either sufficient (+), insufficient (−) or indeterminate (?) (Table 2) [31]. All individual results will be summarised for each instrument per measurement property. The summarised results are rated against the same criteria for good measurement properties whereas possible inconsistency is detected (Table 2). When inconsistency is detected between results of the same measurement property of an instrument, it can be rated as inconsistent results, or results are summarised in subgroups of studies, if applicable.

The overall evidence of each measurement property per instrument is graded using a modified GRADE approach. The evidence is graded as high, moderate, low or very low. The overall evidence is based on the confidence of the results, which will be lowered in case of increased risk of bias, inconsistency, imprecision and indirectness of all studies on each measurement property of an instrument (Table 3) [23].

**Table 2.** Criteria for good measurement properties.

| Measurement property | Rating | Criteria |
|---|---|---|
| Structural validity | + | CTT:<br>CFA: CFI or TLI or comparable measure $>$ 0.95 OR RMSE$A$ $<$ 0.06 OR SRM$R$ $<$ 0.08<br>IRT/Rasch:<br>No violation of unidimensionality: CFI or TLI or comparable measure $>$ 0.95 OR RMSE$A$ $<$ 0.06 OR SRM$R$ $<$ 0.08<br>*AND*<br>no violation of local independence: residual correlations among the items after controlling for the dominant factor $<$ 0.20 OR Q3's $<$ 0.37<br>*AND*<br>no violation of monotonicity: adequate looking graphs OR item scalability $>$ 0.30<br>*AND*<br>adequate model fit:<br>IRT: $\chi^2$ $>$0.01<br>Rasch: infit and outfit mean squares $\geq$ 0.5 and $\leq$ 1.5 OR Z-standardised values $>$ −2 and $<$2 |
| | ? | CTT: Not all information for '+' reported<br>IRT/Rasch: Model fit not reported |
| | – | Criteria for '+' not met |
| Internal consistency | + | At least low evidence for sufficient structural validity AND Cronbach's alpha(s) $\geq$ 0.70 for each unidimensional scale or subscale |
| | ? | Criteria for "At least low evidence for sufficient structural validity" not met |
| | – | At least low evidence for sufficient structural validity AND Cronbach's alpha(s) $<$ 0.70 for each unidimensional scale or subscale |
| Reliability | + | ICC or weighted Kappa $\geq$ 0.70 |
| | ? | ICC or weighted Kappa not reported |
| | – | ICC or weighted Kappa $<$ 0.70 |
| Measurement error | + | SDC or LoA $<$ MIC |
| | ? | MIC not defined |
| | – | SDC or LoA $>$ MIC |
| Hypotheses testing | + | The result is in accordance with the hypothesis |
| | ? | No hypothesis defined (by the review team) |
| | – | The result is not in accordance with the hypothesis |
| Cross-cultural validity | + | No important differences found between group factors (such as age, gender, language) in multiple group factor analysis OR no important DIF for group factors (McFadden's $R^2$ $<$ 0.02) |
| | ? | No multiple group factor analysis OR DIF analysis performed |
| | – | Important differences between group factors OR DIF was found |
| Criterion validity | + | Correlation with gold standard $\geq$ 0.70 OR AUC $\geq$ 0.70 |
| | ? | Not all information for '+' reported |
| | – | Correlation with gold standard $<$ 0.70 OR AUC $<$ 0.70 |
| Responsiveness | + | The result is in accordance with the hypothesis OR AUC $\geq$ 0.70 |
| | ? | No hypothesis defined (by the review team) |
| | – | The result is not in accordance with the hypothesis OR AUC $<$ 0.7 |

References: [23].

**Table 3.** GRADE approach.

| Quality of evidence | Lower if |
|---|---|
| High | Risk of bias |
| *We are very confident that the true measurement property lies close to that of the estimate of the measurement property* | -1 Serious |
| | -2 Very serious |
| Moderate | -3 Extremely serious |
| *We are moderately confident in the measurement property estimate: the true measurement property is likely to be close to the estimate of the measurement property, but there is a possibility that it is substantially different* | Inconsistency |
| | -1 Serious |
| | -2 Very serious |
| Low | Imprecision |
| *Our confidence in the measurement property estimate is limited: the true measurement property may be substantially different from the estimate of the measurement property* | -1 Total $n = 50–100$ |
| | -2 Total $n < 50$ |
| | Indirectness |
| Very low | -1 Serious |
| *We have very little confidence in the measurement property estimate: the true measurement property is likely to be substantially different from the estimate of the measurement property* | -2 Very serious |

References: [23].

## Acknowledgements

The authors thank the health science librarian Marie Oesterbye, who assisted with the development of the search strategy, as well as Frederik Rosenbaek who assisted the protocol development.

## Ethics statement

No ethical approval is required.

## Disclosure statement

The review authors of this review declare no conflicts of interests.

## Funding

The review is initiated and funded by the Department of Physiotherapy and Occupational Therapy, Aarhus University Hospital, Denmark, and the Research Foundation of Aarhus University Hospital.

## ORCID

Hanne Kaae Kristensen http://orcid.org/0000-0002-9596-5571

## References

[1] Jaillard A, Naegele B, Trabucco-Miguel S, et al. Hidden dysfunctioning in subacute stroke. Stroke. 2009;40:2473–2479.

[2] Barker-Collo S, Feigin V. The impact of neuropsychological deficits on functional stroke outcomes. Neuropsychol Rev. 2006;16:53–64.

[3] World Health Organization. International classification of functioning, disability and health (ICF). Geneva: World Health Organization; 2001.

[4] van Rijsbergen MW, Mark RE, de Kort PL, et al. Subjective cognitive complaints after stroke: a systematic review. J Stroke Cerebrovasc Dis. 2014;23:408–420.

[5] van Rijsbergen MW, Mark RE, de Kort PL, et al. Prevalence and profile of poststroke subjective cognitive complaints. J Stroke Cerebrovasc Dis. 2015;24:1823–1831.

[6] Hoffmann T, Bennett S, Koh CL, et al. The cochrane review of occupational therapy for cognitive impairment in stroke patients. Eur J Phys Rehabil Med. 2011;47:513–519.

[7] Claesson L, Linden T, Skoog I, et al. Cognitive impairment after stroke – impact on activities of daily living and costs of care for elderly people. The Göteborg 70+ Stroke Study. Cerebrovasc Dis. 2005;19:102–109.

[8] Robinson-Smith G, Johnston MV, Allen J. Self-care self-efficacy, quality of life, and depression after stroke. Arch Phys Med Rehabil. 2000;81:460–464.

[9] Baker K, Cano SJ, Playford ED. Outcome measurement in stroke: a scale selection strategy. Stroke. 2011;42:1787–1794.

[10] Viscogliosi C, Belleville S, Desrosiers J, et al. Participation after a stroke: changes over time as a function of cognitive deficits. Arch Gerontol Geriatr. 2011;52:336–343.

[11] Nasreddine ZS, Phillips NA, Bedirian V, et al. MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc. 2005;53:695–699.

[12] Folstein MF, Folstein SE, McHugh PR. Mini-mental state". a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12:189–198.

[13] Mok VC, Wong A, Lam WW, et al. Cognitive impairment and functional outcome after stroke associated with small vessel disease. J Neurol Neurosurg Psychiatry. 2004;75:560–566.

[14] Sansonetti D, Hoffmann T. Cognitive assessment across the continuum of care: the importance of occupational performance-based assessment for individuals post-stroke and traumatic brain injury. Aust Occup Ther J. 2013;60:334–342.

[15] Arnadottir G, Fisher AG, Löfgren B. Dimensionality of nonmotor neurobehavioral impairments when observed in the natural contexts of ADL task performance. Neuro Rehabil Repair. 2009;23:579–586.
[16] Baum CM, Tabor CL, Morrison T, et al. Reliability, validity, and clinical utility of the executive function performance test: a measure of executive function in a sample of people with stroke. Am J Occup Ther. 2008;62:446–455.
[17] Arnadottir G, Lofgren B, Fisher AG. Difference in impact of neurobehavioural dysfunction on activities of daily living performance between right and left hemispheric stroke. J Rehabil Med. 2010;42:903–907.
[18] McCluskey A. Occupational therapists report on low level of knowledge, skill and involvement in evidence-based practice. Aust Occ Ther J. 2003;50:3–12.
[19] Edwards DF, Hahn MG, Baum CM, et al. Screening patients with stroke for rehabilitation needs: validation of the post-stroke rehabilitation guidelines. Neurorehabil Neural Repair. 2006;20:42–48.
[20] Bayley MT, Tate R, Douglas JM, et al. INCOG guidelines for cognitive rehabilitation following traumatic brain injury: methods and overview. J Head Trauma Rehabil. 2014;29:290–306.
[21] Kristensen LQ, Muren MA, Petersen AK, et al. Measurement properties of performance-based instruments to assess mental function during activity and participation in traumatic brain injury: a systematic review. Scand J Occup Ther. 2020;27:168–183.
[22] Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. Bmj. 2016;354:i4086.
[23] Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27:1147–1157.
[24] Mokkink LB, Boers M, van der Vleuten CPM, et al. COSMIN risk of bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. BMC Med Res Methodol. 2020;20:293.
[25] Terwee CB, Jansma EP, Riphagen II, et al. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res. 2009;18:1115–1123.
[26] Bickenbach J, Cieza A, Rauch A, Stucki G. ed. ICF core sets: manual for clinical practice. Göttingen: Hogrefe; 2012.
[27] Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63:737–745.
[28] Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27:1171–1179.
[29] Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res. 2018;27:1159–1170.
[30] Mokkink LB, Terwee CB, de Vet HCW. Key concepts in clinical epidemiology: responsiveness, the longitudinal aspect of validity. J Clin Epidemiol. 2021;140:159–162.
[31] Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res. 2012;21:651–657.