1.0 Title

Assessments for Identifying Tactile Deficits in Individuals with Stroke: A Scoping Review

2.0 Registration

The protocol will be registered to Northwestern University's DigitalHub. All authors reviewed the protocol prior to completion.

3.0 Authors

Karan Nayak
Undergraduate Student
Northwestern University
Weinberg College of Arts and Sciences
karannayak2023@u.northwestern.edu

Lindsey Sydnor
Undergraduate Student
Virginia Polytechnic Institute and State University
School of Neuroscience
lindseycs@vt.edu

Q. Eileen Wafford, MSt, MLIS
Research Librarian
Northwestern University
Feinberg School of Medicine
g-wafford@northwestern.edu

Jane E. Sullivan, PT, DHS, MS
Professor
Northwestern University
Feinberg School of Medicine
j-sullivan@northwestern.edu

Kevin Parcetich, PT, DPT, NCS
Assistant Professor
Radford University
Waldron College of Health and Human Services
kparcetichii@radford.edu

Daniel Miner, PT, DPT, NCS
Assistant Professor
Radford University
Waldron College of Health and Human Services
dminer1@radford.edu

Arco Paul, PT, PhD, NCS
Assistant Professor
Radford University
Waldron College of Health and Human Services
apaul8@radford.edu

Netta Gurari, PhD
Assistant Professor
Northwestern University
Feinberg School of Medicine
netta.gurari@northwestern.edu

4.0 Rationale

Accurate perception of objects that one interacts with, through touch, is instrumental to successfully learn and perform physical activities.[1] Studies indicate that the perception of touch, or tactile stimuli, in the upper extremities is commonly negatively impacted after a stroke.[2] A consequence of an individual experiencing a tactile deficit is having poorer functional control of their limb movements.[3] Tactile deficits may also impact areas of the brain responsible for a wide range of other skills including body awareness, memory, and vision.[4] Despite the importance of accurate tactile perception, our understanding of why tactile deficits occur in individuals with stroke remains poor. The reasons for our poor knowledge regarding tactile deficits post-stroke include the minimal focus on somatosensory deficits and limitations of existing assessments.[2]

Approaches for quantifying tactile perception in the human stroke population include behavioral assessments, which require an individual to respond whether a tactile stimulus is felt,[5] and brain recordings, which indicate whether a tactile stimulus delivered at the finger reaches an individual's brain.[6] Behavioral assessments make it possible to determine whether a tactile stimulus is cognitively perceived. Even so, behavioral assessments have limitations, including that information regarding the processes occurring within the nervous system are not directly elucidated.[2] Brain recordings make it possible to determine whether a sensory stimulus applied at the periphery reaches the brain. Yet, this approach has limitations, including potentially not identifying the location in the nervous system where the signal transmission may be altered.

In this review, we aim to summarize existing protocols for assessing tactile deficits in the human stroke population. We will identify the extent to which currently available tactile assessments can address the processes occurring within the nervous system of individuals with stroke. For example, we will indicate whether studies have linked outcomes on behavioral tactile assessments to activity in the brain. In addition, we will comment on how the outcomes of the tactile assessments correspond to an individual's ability to perform activities of daily living. This includes describing how the tactile assessments are executed and what the outcomes may mean for one's execution of volitional sensorimotor tasks. Moreover, we will discuss the relevance of these tactile assessments for the clinical setting, including summarizing their benefits and limitations. In turn, we will highlight the strengths of existing tactile assessments as well as identify gaps in our understanding of when and why tactile perceptual deficits occur post-stroke and the implications for their use in the clinical setting. Given this knowledge, the field will be well-positioned to address areas in which our understanding of tactile perceptual deficits occurring post-stroke can be improved. In turn, this review will advance the field towards the greater goal of identifying novel biomarkers and, subsequently, treatments for tactile deficits in individuals with stroke.

4.1 Background

Reviewing existing tactile assessments and synthesizing existing knowledge may help identify gaps in our understanding. The targeted outcomes are two-fold. First, this information may lead to the development of more comprehensive assessments and, subsequently, an improved understanding of the neural processes governing tactile perception post-stroke. Second, the summarized findings may inform clinicians on the clinical value of existing tactile assessments, including the implications of deficits on an individual's function, the prognoses for improvement, and the relevance to various treatment approaches. This may lead to more informed clinical decision-making.

4.2 Objectives

We seek to examine existing tactile assessments for the upper extremity of individuals with stroke. We aim to determine whether there is a gap in existing approaches that limits our understanding of the source of tactile perceptual deficits post-stroke. This information will be beneficial to understand the clinical value of existing assessments. We will answer the following questions:

- How are tactile deficits post-stroke currently assessed?
- What can be learned from existing approaches that assess tactile deficits post-stroke?
- What gaps exist in current assessment protocols that gauge tactile perception?

4.3 Research Question

To what extent do existing upper-extremity tactile assessments i) identify the process(es) within the nervous system that elicit(s) a tactile perceptual deficit in people with stroke and ii) inform clinicians on their decision making?

5.0 Methods

The steps of the scoping review are listed and visually depicted in Figure 1 and Figure 2 of the Appendix, respectively.

This scoping review will follow the Joanna Briggs Institute (JBI) methodological framework.[7] This framework for finding literature allows for attaining both in-depth and broad outcomes. Instead of utilizing a highly focused research question that leads to searches for a particular study design, this framework is guided by a requirement to identify all relevant literature regardless of the study design. The stages of this framework include identifying a research question, finding relevant studies, selecting studies, extracting data, and summarizing/reporting results. This process prompts researchers to thoroughly engage with each stage, reflect upon changes made, and repeat steps to ensure that the literature is adequately covered.[7]

The PRISMA-ScR checklist will also be followed to complete the review.[8] This checklist has 22 sections that describe how to ensure the scoping review abides by stated standards, including reporting the data extraction process, critically appraising individual sources of evidence, and synthesizing the results.

5.1 Search

The study employed the Population-Concept-Context (PCC) framework to identify main concepts.[9] The population consists of individuals who have experienced a stroke. The concept identifies the quality of assessments that are used to elucidate tactile deficits in this population. The context is that these assessments are performed on individuals with stroke in both the clinical and research settings.

The team includes a research librarian who is developing a comprehensive search strategy. The search strategy will incorporate keywords and Medical Subject Heading terms describing individuals with stroke, upper extremities, assessments, and tactile deficits (see Section 5.4). We will exclude animal studies and limit the search to English publications (see Section 5.3). No restrictions to publication date or research design will be applied. We will adapt the search to the selected databases. Any modifications to the search strategy will be documented.

5.2 Information Sources

We will search the following databases from the date of inception to the present:

- Medline (Ovid)
- The Cochrane Library (Wiley)
- CINAHL Plus with Full Text (Ebsco)
- Scopus (Elsevier)
- PsycInfo (Ebsco)
- Proquest dissertations and theses global
- ClinicalTrials.gov

5.3 Eligibility Criteria

Inclusion

- Individuals with stroke
- Adult human population (18 years or older)
- Tactile assessment(s) at the upper extremity (e.g., fingers, wrists, forearms, upper arm, shoulder)
- All study designs with assessments, along with gray literature, systematic reviews, and scoping reviews

Exclusion

- Assessments evaluating participant responses to pain or temperature (e.g., algometer)
- Articles written in a language that is not English
- Articles including intervention trials

5.4 Ovid Medline Search

1. exp Stroke/
2. (apople* or "cerebral accident*" or "cerebrovascular accident*" or poststroke or stroke or strokes).ti,ab.
3. ((brain or cerebral or intracranial or intracerebral) adj2 (bleed* or embolism* or hemiparesis or hemorrhage* or infarct* or infract* or injur* or isch?emi* or thrombo* or "vascular accident*")).ti,ab.
4. 1 or 2 or 3
5. exp Upper Extremity/
6. (arm or arms or axilla* or elbow* or finger* or forearm* or hand or hands or metacarpus or shoulder* or thumb* or "upper extremit*" or "upper limb*" or wrist*).ti,ab.
7. 5 or 6
8. exp Touch/
9. exp Skin/
10. *Perception/
11. exp Touch Perception/
12. 8 or 9 or 10 or 11
13. exp *"Diagnostic Techniques and Procedures"/

14. exp Patient Outcome Assessment/
15. exp Disability Evaluation/
16. exp Physical Examination/
17. exp Symptom Assessment/
18. exp "Surveys and Questionnaires"/
19. investigative techniques/
20. exp Somatosensory Disorders/
21. exp Sensory Thresholds/
22. exp Hemiplegia/
23. 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22
24. 12 and 23
25. ((touch* or tactile or cutaneous* or skin or haptic* or sensation* or sense or senses or sensory or somatosensory or perception*) adj6 (deficit* or hemiplegia* or impairment* or monoplegia* or dysfunction or dysfunctions or assessment* or detect* or evaluation* or examination* or index or indexes or instrument or instruments or measure* or outcome* or questionnaire* or scale or scales or score or scores or survey or surveys or test or tests)).ti,ab.
26. clinical assessment*.ti. and (touch* or tactile or cutaneous* or skin or haptic* or sensation* or sense or senses or sensory or somatosensory or perception* or pressure or vibration* or force).ti,ab.
27. (pressure assessment* not "blood pressure").ti,ab.
28. 25 or 26 or 27
29. 24 or 28
30. 4 and 7 and 29
31. exp animals/ not humans/
32. 30 not 31
33. limit 32 to english language

5.5 Citation Management

After performing the searches, references will be imported to EndNote and duplicate records will be removed.[10] The authors will use the citation list in EndNote to access selected abstracts and full-text manuscripts. EndNote will also be used for title/abstract screening. Once this step is completed, the citations will be managed within Rayyan, a software program that will help organize the full-text screening process.[11]

5.6 Reader Training

To begin, a pair of independent readers will perform a pilot test for the title/abstract screening round. Each reader will independently screen the titles and abstracts of 100 of the same articles based on the exclusion criteria. A search involving tactile deficits in the upper extremities of individuals with stroke will be conducted to find the pilot articles. The selected articles will be similar to those that will be used in the scoping review. After screening all articles, the readers will convene to discuss the results with three expert reviewers, who will resolve any conflicts.

The two independent readers will also be trained on the full-text review by reviewing three articles similar to the articles that will be reviewed in the scoping review. The three expert reviewers will select three articles and will create a key of correct responses for the eligibility criteria. Subsequently, the pair of independent readers will independently screen the same articles and determine whether the inclusion and exclusion criteria are met for each article. The three expert reviewers will discuss the decision-making process with the independent readers. If an individual reader does not reach the same conclusion as the three expert reviewers about the eligibility of two out of the three articles chosen for training, additional article(s) will be reviewed following the same process until at least two out of three articles are successfully screened. Once training is successfully completed for each independent reader, the pair of independent readers will be approved to begin screening articles returned from the database search.

5.7 Article Selection Process

In the first round of review, gathered articles will be screened within EndNote. To begin, the articles will be uploaded to EndNote, and then the two trained independent readers will screen the titles and abstracts of the articles to refine the relevant studies. Once this first review process has been completed, the second round will include a complete review of the full text of each article by each of the two independent readers. Articles that have been chosen after the title/abstract review will be exported into Rayyan for the full-text review. Each round will be completed independently and blinded. If a conflict arises it will be resolved by an expert reviewer who will serve as a tie-breaker. Screened papers must fulfill the inclusion criteria and not include the exclusion criteria to be considered. Once all papers are screened by the independent readers, an expert reviewer will review the full text of all of the screened articles to confirm that the article selection process was carried out appropriately.

5.8 Data Items

Below we identify the data items of interest for extraction from each reviewed article:

- Article Number
- Article Title
- Author(s)
- Year of Publication
- Abstract
- Number of Participants
- Age
- Gender
- Stage of Stroke (acute, sub-acute, chronic)
- Type of Stroke (ischemic, hemorrhagic)
- Brain Hemisphere Affected by Stroke (left, right, both)

- Brain Hemisphere Activated by Tactile Stimuli (left, right, both)
- Laterality of Assessment (left, right, both)
- Lesion Location(s) (e.g., subcortical, cortical)
- Assessment Approach/Technology Used
- Assessor and Training (e.g., was reliability established?)
- Authors' Claims of Assessment Type (e.g., sensory or perceptual; behavioral or brain recording)
- Type of Tactile Stimulus Delivered (e.g., pressure, vibration)
- Nature of Task (e.g., active or passive)
- Body Location of Assessment (e.g., elbow, fingertip)
- Severity of Clinically-Assessed Tactile/Somatosensory Deficit (reported by test and score)
- Assessment Setting (e.g., clinical setting, research laboratory)

5.9 Evaluation of Articles

5.9.1 Outcomes of Interest

Article Appraisal

Each selected article will be critically appraised using a scoring system known as the COnsensus-based Standards for the Selection of Health Measurement INstruments (COSMIN).[12] This scoring system will aid in assessing the caliber of the methods and statistical analyses present in the papers. We modified the COSMIN checklist so that it is tailored to our specific outcomes (see Figure 3 in the Appendix). The modified COSMIN checklist is composed of sections that assess the following psychometric properties of interest: internal consistency, reliability, measurement error, and responsiveness. For each article reviewed, the expert reviewers will fill out a checklist with a 'yes' or 'no' response indicating whether the criteria are met.

Initially, the level of evidence provided in each article will be determined. The number of "yes" ratings in each section will be summed and reported as a percentage of the maximum points possible for that section. The section with the lowest percentage will represent the level of evidence. If the article's lowest percentage is $\geq$ 50%, then it is considered to have a level of evidence of 1 (strong), and if it is below this threshold, it is considered to have a level of evidence of 2 (moderate).[13] See Figure 4 of the Appendix for example ratings.

In addition to the level of evidence, a psychometric property rating will be given evaluating the reporting of the psychometric properties of interest, specifically the internal consistency, reliability, measurement error, and responsiveness (see Figure 5 of the Appendix). If all psychometric properties are reported, the article will be given a

psychometric property rating of (+). If any psychometric property is not reported, then the article will receive a rating of (-).

The COSMIN rating will consist of the level of evidence score and psychometric property rating . The COSMIN ratings in order of methodological quality are 1+, 1-, 2+, and 2-.

Data Extraction
The data items of interest from each article, as listed in Section 5.8, will be input into a data charting table.

5.9.2 Training
Before evaluating the selected articles, two expert reviewers will be trained for inter-rater reliability when assessing an article's methodological quality and extracting the data items of interest.

To assess inter-rater reliability of the COSMIN scoring, the two expert reviewers will review three articles independently and blindly, and complete the COSMIN scoring checklist for each. The reviewers will then discuss their scoring together with a third expert reviewer, and any discrepancies will be resolved through a triangulation of agreement process. There must be at least 80% agreement between the two reviewers on each article in order for the training to be complete. If 80% is not reached, the third reviewer will select another article and the process will be repeated until 80% agreement is met. This training mimics the triangulation process that will occur during the review, where a third expert reviewer will act as a tie-breaker if conflicts arise during the COSMIN scoring process between the other two reviewers.

To assess inter-rater reliability of data extraction, the data items of interest will be extracted from the same three articles by the two expert reviewers using a data form containing the items of interest (see Section 5.8). The reviewers will then discuss their scoring together with a third expert reviewer, and any discrepancies will be resolved through a triangulation of agreement process. The reviewers must achieve at least 80% agreement for all of the data items to complete the training. Similar to the previous stage of training, if 80% is not reached, the third expert reviewer will produce another article, and the process will be repeated until 80% agreement is met.

5.10 Data Synthesis
The data items of interest will be summarized and analyzed by a team member to gain insights on common trends present throughout the literature, such as the type of assessment (e.g., behavioral or brain recording), nature of the task (e.g., active or

passive), and location of the assessment (e.g., fingertip or elbow). The data will first be summarized using statistics such as the range, mean/standard deviation, median/lower and upper quartiles, and sum. Analyses will be run on a subset of these summary statistics using primarily non-parametric tests, such as the Mann-Whitney test, Kruskal-Wallis test, and Spearman rank correlation. Analyzed data will be used to develop conclusions that will be presented in the discussion section of the manuscript. Data relevant to existing assessments addressing tactile perception post-stroke will be synthesized to gain knowledge that will potentially spur the development of new targeted assessments and, eventually, therapeutic approaches to evaluate and treat millions of survivors of stroke.

5.11 Study Appraisal

The Joanna Briggs Institute (JBI) critical appraisal tool for systematic reviews will be used to evaluate the methodological quality of the study and address the possibility of bias in its design, conduct, and analysis.[15] This tool, although originally intended for systematic reviews, addresses all the important components of a scoping review. Questions raised in the checklist specifically relate to topics such as the appropriateness of the search strategy, potential errors arising during data extraction, and likelihood of publication bias. The review will be appraised by two content experts of the research team to ensure that the study abides by official standards for scoping reviews.

6.0 Implications and Dissemination

This scoping review is intended to summarize existing assessments measuring tactile perception post-stroke. By examining gaps in current approaches, we will be better positioned to effectively employ existing methods and/or develop novel approaches for assessing tactile perception in the stroke population. In turn, we can identify when and why deficits occur and use that information to inform clinical practice. This increased knowledge may lead to the identification of biomarkers that will help the field address when tactile perception is negatively impacted after a stroke. It may also lead to future research involving the development of treatments for those with deficiencies in tactile perception. This review can also potentially highlight the need for implementing a standard core set of outcome measures. Establishing a core set of outcomes will allow researchers to create and analyze larger data sets and assess changes in the progression of people with stroke across a larger population. It also has important implications for standardized practice in the clinical setting and allows clinicians to make more informed decisions.

7.0 References

1. Cole M. Pride and a Daily Marathon. *Neurology*. 1996;47(3):856-857. doi:10.1212/WNL.47.3.856-c
2. Sullivan JE, Hedman LD. Sensory dysfunction following stroke: incidence, significance, examination, and intervention. *Top Stroke Rehabil*. 2008;15(3):200-217. doi:10.1310/tsr1503-200
3. Johansson RS, Flanagan JR. Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nat Rev Neurosci*. 2009;10(5):345-359. doi:10.1038/nrn2621
4. Bremner AJ, Spence C. Chapter Seven - The development of tactile perception. In: Benson JB, ed. *Advances in Child Development and Behavior*. Vol 52. JAI; 2017:227-268. doi:10.1016/bs.acdb.2016.12.002
5. Wu C Yi, Chuang I Ching, Ma H Ing, Lin K Chung, Chen C ling. Validity and responsiveness of the revised Nottingham sensation assessment for outcome evaluation in stroke rehabilitation. *Am J Occup Ther*. 2016;70(2):7002290040p1-7002290040p8. doi:10.5014/ajot.2016.018390
6. Connell L, Lincoln N, Radford K. Somatosensory impairment after stroke: frequency of different deficits and their recovery. *Clin Rehabil*. 2008;22(8):758-767. doi:10.1177/0269215508090674
7. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005;8(1):19-32. doi:10.1080/1364557032000119616
8. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169(7):467-473. doi:10.7326/M18-0850
9. JBI Manual for Evidence Synthesis - JBI Global Wiki. https://jbi-global-wiki.refined.site/space/MANUAL
10. EndNote | The best reference management tool. EndNote. https://endnote.com/
11. Rayyan – Intelligent Systematic Review - Rayyan. https://www.rayyan.ai/
12. Smit Liesbeth. COSMIN - Improving the selection of outcome measurement instruments. COSMIN. https://www.cosmin.nl/
13. Moore JL, Potter K, Blankshain K, Kaplan SL, O'Dwyer LC, Sullivan JE. A core set of outcome measures for adults with neurologic conditions undergoing rehabilitation: A clinical practice guideline. *J Neurol Phys Ther JNPT*. 2018;42(3):174-220. doi:10.1097/NPT.0000000000000229
14. Adams J, Bateman B, Becker F, et al. *Triangulation and Integration of Results*. NIHR Journals Library; 2015. https://www.ncbi.nlm.nih.gov/books/NBK326956/
15. Martin J. © Joanna Briggs Institute 2017 Critical appraisal checklist for systematic reviews and research syntheses. Published online 2017:7.

8.0 Appendix

Figure 1. This figure shows the full sequence of steps that will be followed for the review.

Step	General Process	Specific Tasks
1.	Team Recruitment	Recruit expert reviewers and librarian
2.	Identification of Research Question and Focus	Initial review of available literature to understand the breadth of information present on the topic of tactile deficits post-stroke
	Creation of Inclusion and Exclusion Criteria	Deliberate over important ideas/concepts that must be present in papers for inclusion
3.	Formation of Search Strategy	Apply population-concept-context (PCC) framework to establish important parameters that will guide database search
4.	Development of Protocol	Document procedures for performing the review, including rationale, background, objectives, etc.
	Conduction of Database Search	Search for appropriate papers across a variety of databases by incorporating relevant keywords as defined by the eligibility criteria
5.	Management of Citations	Import references gathered from database search into EndNote software to remove duplicate records
6.	Reader Training	Readers train on the title/abstract and full-text screenings
7.	Article Selection Process	Readers complete the title/abstract screening based on the exclusion criteria, followed by the full-text screening based on inclusion and exclusion eligibility criteria
8.	Inter-rater Reliability Training	Expert reviewers train to achieve inter-rater reliability for the COSMIN appraisal and extraction of data items of interest
9.	Assessment of Methodological Quality of Selected Articles	Expert reviewers use the COSMIN tool to assess the methodological quality of each article
	Extraction of Data Items	Expert reviewers extract the data items of interest from each selected article
10.	Data Synthesis	Statistical analyses run on the extracted data items to discover trends
	Results/Discussion	Results are summarized and discussed, and future research recommendations are made
11.	Study Appraisal	Scoping review is appraised using the JBI critical appraisal tool

Figure 2. This figure gives a condensed chronological overview of the steps that will be taken in the scoping review.

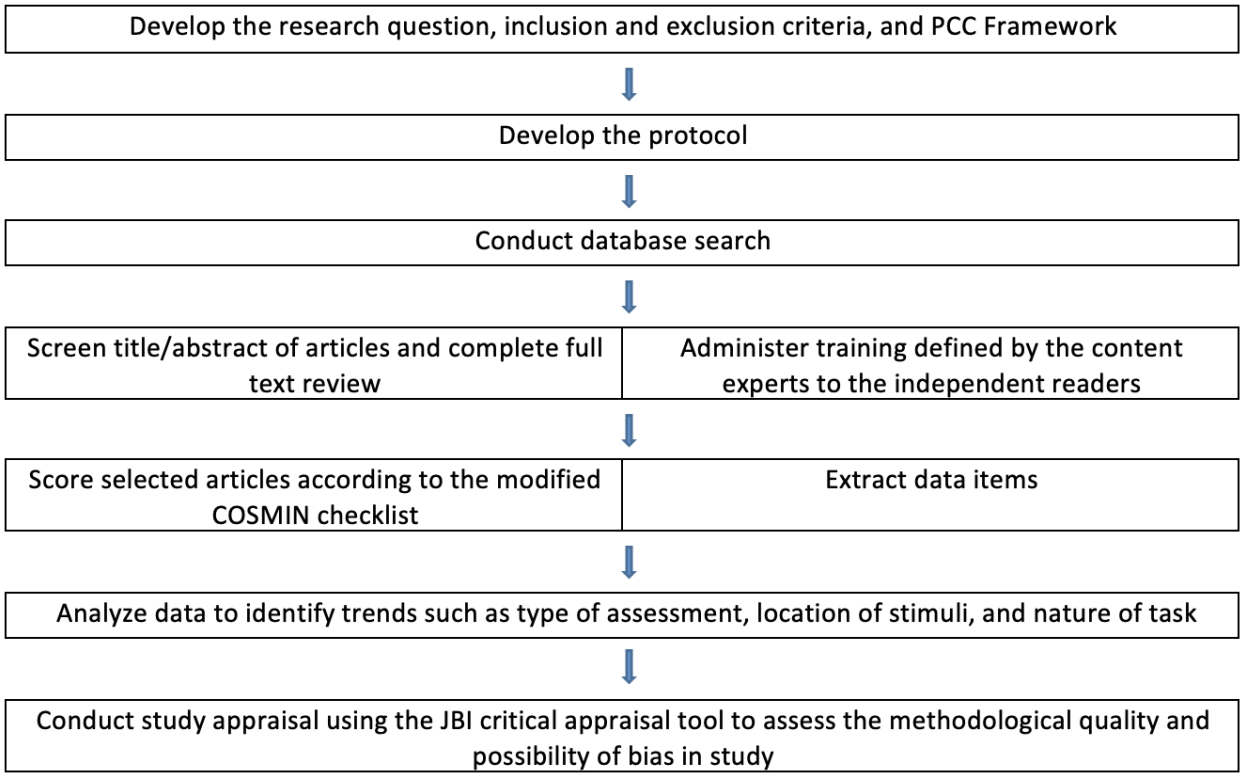

Figure 3. This figure shows the modified COSMIN checklist that will be used to evaluate the methodological quality of each article.

Question	Scoring
Internal Consistency	
1. Was internal consistency assessed?	Yes = 1 No = 0
2. Was Cronbach's alpha calculated? If not, what statistic was calculated to determine internal consistency?	NS
3. Did the authors report whether any data was missing?	Yes = 1 No = 0
4. If data was missing did the authors describe how this was handled in the analysis?	Yes = 1 No = 0
5. Was the total sample size $\geq$ 30?	NS
6. If the sample size was less than 30, was a power analysis done to determine the number of subjects needed?	NS
7. If a sample size power analysis was done, was the size requirement met?	NS
8. What was the sample size?	NS
9. Input internal consistency value(s).	Yes = 1 No = 0
10. Were there any important flaws in the design or methods of the study (for the internal consistency calculation)? If yes, please describe.	Yes = 0 No = 1
Maximum points for the section (Note: counted question 5 OR 7)	**5**
Reliability	
11. Did the study evaluate the reliability of the measure?	Yes = 1 No = 0
12. Did the authors report whether any data was missing?	Yes = 1 No = 0
13. If data was missing did the authors describe how this was handled in the analysis?	Yes = 1 No = 0

14. Was the total sample size ≥ 30?	NS
15. If the sample size was less than 30, was a power analysis done to determine the number of subjects needed?	NS
16. If a sample size power analysis was done, was the size requirement met?	NS
17. Were at least two measurements taken and reported?	Yes = 1 No = 0
18. Was the time interval between measurements stated? If yes, what was the time interval?	Yes = 1 No = 0
19. Were the test conditions (e.g. type of administration, environment, instructions) similar for both measurements?	Yes = 1 No = 0

	Not stated = 0
20. Did the article state that each administration was done without knowledge of the other set of scores?	Yes = 1 No = 0
21. Are you reasonably confident that the patients were stable between test administrations? If not, please describe.	Yes = 1 No = 0
22. Were there any important flaws in the design or methods of the study that would affect the reliability analysis? If yes, please describe.	Yes = 0 No = 1
23. For ratio or interval level data, was an intraclass correlation coefficient calculated?	Yes = 1 No = 0 NA = 1
24. For dichotomous/nominal/ordinal scores, was kappa calculated?	Yes = 1 No = 0 NA = 1
25. For each type of reliability value calculated, enter the type and value.	NS
Maximum points for the section (Note: counted question 14 OR 16)	**11**
Responsiveness / Interpretability	
26. Were any of the following values assessed: standard error of measurement (SEM), smallest real difference (SRD), minimal detectable change (MDC), minimal important change (MIC), and minimal important difference (MID)? Questions 27 – 43 pertain to these psychometric variables.	Yes = 1 No = 0

27. Did the authors report whether any data was missing?	Yes = 1 No = 0
28. If data was missing did the authors describe how this was handled in the analysis?	Yes = 1 No = 0
29. Was the total sample size $\geq$ 30?	NS
30. If the sample size was less than 30, was a power analysis done to determine the number of subjects needed?	NS
31. If a sample size power analysis was done, was the size requirement met?	NS
32. Were at least two measurements taken and reported?	Yes = 1 No = 0
33. Was the time interval stated?	Yes = 1 No = 0
34. Were the test conditions (e.g. type of administration, environment, and instructions) similar for both measurements?	Yes = 1 No = 0
35. Did the article state that each administration was done without knowledge of the other set of scores?	Yes = 1 No = 0
36. Are you reasonably confident that the patients were stable between test administrations?	Yes = 1 No = 0
37. Was more than one subgroup (e.g., diagnosis, condition, acuity or impairment level) included in the sample?	NS
38. Was data reported for each subgroup?	Yes = 1 No = 0

	NA = 1
39. Which of the following values were assessed? (Listed all values) Report where the values are found in the manuscript.	NS
40. Was the distribution of the (total) scores in the study sample described?	Yes = 1 No = 0
41. Were there any important flaws related to interpretability or responsiveness in the design or methods of the study? If yes, please describe the flaws.	Yes = 0 No = 1

42. If a longitudinal design was used, was the time interval stated? If yes, please report the interval.	Yes = 1 No = 0 NA = 1
43. If a longitudinal design was used, if anything occurred in the interim period (e.g. intervention, other relevant events), was it adequately described?	Yes = 1 No = 0 NA = 1
Maximum points for the section questions 27 - 43 (Note: counted question 29 OR 31)	**13**
44. Were any of the following values assessed: MCID, floor effect, and ceiling effect? Questions 45 – 55 pertain to these psychometric variables.	
45. Was more than one subgroup (e.g., diagnosis, condition, acuity or impairment level) included in the sample?	NS
46. Was data reported for each subgroup?	Yes = 1 No = 0 NA = 1
47. Which of the following values were assessed? (Listed all values) Report where the values are found in the manuscript.	NS
48. Was the distribution of the (total) scores in the study sample described?	Yes = 1 No = 0
49. Did the authors report whether any data was missing?	Yes = 1 No = 0
50. Was the total sample size $\geq$ 30?	NS
51. If the sample size was less than 30, was a power analysis done to determine the number of subjects needed?	NS
52. If a sample size power analysis was done, was the size requirement met?	NS
53. What was the sample size?	NS
54. If a longitudinal design was used, was the time interval stated?	Yes = 1 No = 0 NA = 1
55. If a longitudinal design was used, if anything occurred in the interim period (e.g. intervention, other relevant events), was it adequately described?	Yes = 1 No = 0 NA = 1
Maximum points for the section – questions 45 - 55 (Note: counted question 50 OR 52)	**5**

Generalizability	
56. Was the sample in which the instrument was evaluated adequately described in terms of subject disease/condition or normative sample (select as many as apply): brain injury,	NS
multiple sclerosis, Parkinson's disease, spinal cord injury, stroke, vestibular dysfunction, normative sample, other (please describe).	
57. What best describes the patient population studied: Acute/Subacute; Chronic Progressive; Chronic Stable? (select as many as apply)	NS
58. Was the setting(s) in which the study took place (e.g., acute care, in-patient rehabilitation, out-patient, community) described. If yes, please describe.	Yes = 1 No = 0
59. Was the age of the subject reported (median or mean age, with standard deviation or range) described? If yes, please describe.	Yes = 1 No = 0
60. Was the distribution of sex described? If yes, please describe (e.g., 45% male; 55% female).	Yes = 1 No = 0
Maximum points for the section	**3**
General Methodology*	
61. Was the method (e.g. convenience, consecutive, or random) used to select patients adequately described?	Yes = 1 No = 0
62. Was the total sample size $\geq$ 30?	NS
63. If the sample size was less than 30, was a power analysis done to determine the number of subjects needed?	NS
64. If a sample size power analysis was done, was the size requirement met?	NS
65. What was the sample size?	NS
66. If more than one group was included in the sample, was data reported by subgroups?	Yes = 1 No = 0 NA = 1
67. Did the authors report whether any data was missing?	Yes = 1

	No = 0
68. If data was missing did the authors describe how this was handled in the analysis?	NS
69. Was there a description of how the raters were training? If yes, please describe.	Yes = 1 No = 0
70. Were the raters (select as many as apply): research scientists, research clinicians, research assistants, practicing clinicians, students, not reported, other (please describe).	NS
71. Is there anything else you would like to tell us about this study?	NS
Maximum points for the section (Note: counted question 50 OR 52)	**4**

NA – not applicable; NS – not scored

Figure 4. This table provides examples of how the level of evidence will be determined for each article. The number of "yes" ratings in each section will be summed and reported as a percentage of the maximum points possible for that section. The level of evidence for the article will then be determined based on the lowest percentage reported from all sections. Note, the percentages used in this table are purely for demonstration purposes and do not reflect actual data.

	Sample Article 1	**Sample Article 2**
Section percentage	Internal consistency = 60% Reliability = 70% Responsiveness = 80% Generalizability = 90% General methodology = 70%	Internal consistency = 80% Reliability = 70% Responsiveness = 40% Generalizability = 80% General methodology = 50%
Level of evidence	60% = level 1 evidence (strong)	40% = level 2 evidence (moderate)

Figure 5. This diagram visually depicts how the psychometric property rating of (+) or (-) will be determined, depending on whether all psychometric properties are reported.