NINDS CDE Project Traumatic Brain Injury Version 3.0
Instrument Selection Criteria

Instruments were evaluated by the TBI v3.0 Working Group using Instrument Selection Criteria, which were informed by the PhenX Toolkit (Hamilton et al., 2011).

The six criteria below are required for CDE designation:

- **Clearly defined:** An unambiguous statement of the purpose of the measure.

- **Well-established, broadly applicable to the intended population (e.g., adult and/or pediatric), and generally accepted by the scientific community:** The instrument has been used in well-designed studies published in at least two peer-reviewed journals.

- **Broadly validated with demonstrated utility:** Psychometric data are available showing adequate diagnostic or prognostic accuracy (the degree to which the scores of an instrument are an adequate reflection of a "reference standard"); it has been published in at least two peer-reviewed journals.

- **Specific:** Clear description of the concept of interest for this measure (i.e., GOSE measures global disability and recovery after TBI). Concept of interest is defined as the trait, skill, knowledge, sign, symptom, perception, limitation or any “thing” that represents a meaningful aspect to the patient experience and is intended to be measured by the instrument (Christoforou et al., 2020).

- **Reliable:** Psychometric data are available showing good to excellent reliability (internal consistency, test/retest reliability, interrater reliability (IRR), etc.) and have been published in at least two peer-reviewed journals.

- **Standard measurement protocols exist:** A written protocol or manual of operating procedures (MOP) is publicly available.

The four criteria below are preferred CDE characteristics but are not required for CDE designation:

- **Low burden to participants and investigators:** The instrument does not require specialized training, and cost (open source vs. copyrighted/licensed), completion time, and equipment needed (e.g., technology requirements like smart phones, activity monitors, or Wi-Fi) would be considered reasonable by most users.

  Determining if an instrument is low burden requires making a judgement based on available information. When evaluating burden, all criteria in the definition must be met. That is, if any one of the following 4 items are true, then the response would be “No”:
  1) Specialized training for raters is required
  2) Completion time per instrument would not be considered reasonable by most users
  3) Specialized equipment is required for administration or scoring (e.g., laptop computer, software program) that would not be considered reasonable by most users.
  4) There is an associated cost that would not be considered reasonable by most users. For a validated instrument, cost should not be the primary factor in deciding whether to include it on the list, especially if the instrument is considered the gold standard. Reviewers should base their ratings on scientific evidence and their professional expertise rather than on cost alone.

- **Crosscutting relevance for population groups as well as diseases and conditions:** The instrument is validated in different age groups, in more than one language and must be culturally acceptable or culturally adapted (from original version) and across different diseases and conditions. It is designed to avoid inadvertent biases.

Evidence of validation in at least two age groups, languages (with cultural adaptation), and diseases or conditions is needed to meet the definition for this criterion.

- **Rural vs. Urban (Feasibility of Acquisition):** Details are provided on whether the instrument is available to be administered in rural and urban settings.

  Reviewers should evaluate the feasibility of acquisition in any setting based on considerations like technology and training requirements. Evidence of use is not required.

- **International harmonization (International applicability):** Whenever possible, CDE/instrument should be applicable to international populations. International protocols are preferred but the emphasis is on US populations. If applicable to international populations, cross-cultural validity should be established.

  Distinguishing between 'Crosscutting relevance for population groups as well as diseases and conditions' and 'International harmonization (International applicability)':
  While crosscutting relevance addresses broad applicability within a specific population and disease area, international harmonization specifically evaluates the suitability of the instrument for use in international contexts (across different countries). International applicability is not a required component of crosscutting relevance; an instrument may be crosscutting within a country or region but not necessarily harmonized for global use. Instruments should be validated in the translated language.

REFERENCES

Christoforou AN, Armstrong MJ, Bergin MJG, Robbins A, Merillat SA, Erwin P, Getchius TSD, McCrea M, Markowitz AJ, Manley GT, Giacino JT. An evidence-based methodology for systematic evaluation of clinical outcome assessment measures for traumatic brain injury. PLoS One. 2020 Dec 14;15(12):e0242811.

Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, Hammond JA, Huggins W, Jackman D, Pan H, Nettles DS, Beaty TH, Farrer LA, Kraft P, Marazita ML, Ordovas JM, Pato CN, Spitz MR, Wagener D, Williams M, Junkins HA, Harlan WR, Ramos EM, Haines J. The PhenX Toolkit: get the most from your measures. Am J Epidemiol. 2011 Aug 1;174(3):253-60.