Article type : Editorial

Editorial

In search of the perfect outcome in neuroanaesthesia and neurocritical care

A.M. Flexman[1] and A. Tung[2]

1 Clinical Associate Professor, Department of Anaesthesiology, Pharmacology and Therapeutics, University of British Columbia and Department of Anaesthesiology, Providence Health Care/St. Paul's Hospital, Vancouver, BC, Canada

2 Fellow, Department of Anaesthesiology, University of Michigan, Ann Arbor, MI, USA

Correspondence to: A. M. Flexman

Email to alana.flexman@ubc.ca

Twitter: @alanaflex

This issue of *Anaesthesia* includes several thorough and informative reviews on neuroanaesthesia topics, including mechanical thrombectomy [1], traumatic brain injury (TBI) [2] and peri-operative neurological monitoring [3]. These reviews provide a valuable synthesis of the current evidence base and have a common thread: much of the current literature focuses on surrogate or short term outcomes, and lacks adequate inclusion of long-term, patient-centred outcomes for neurological

conditions and procedures. Although few would disagree that there is a need to move beyond short-term, surrogate outcomes, what those outcomes should be remains debatable and should to be tailored to different neurological contexts. What is clear is that the relative importance of these outcomes should be defined in consultation with our patients, and shift towards meaningful function, experience and quality of life. In this editorial, we will provide an examples of outcomes used commonly in peri-operative neuroscience and critical care, the application of patient-reported outcome measures (PROMs) to neurological outcomes, and future directions for endpoints in clinical research.

Research outcome measures in stroke, traumatic brain injury and peri-operative cognition

Much of the research on interventions in stroke management has naturally focused on functional outcomes. The modified Rankin Scale (mRS) is a functional assessment scale that categorises a patient's degree of dependence post-stroke and ranges from 0 to 6, encompassing no symptoms, slight to severe disability and death. [4] The mRS has several strengths that have led to its frequent incorporation into most stroke trials. The scale is well-established, intuitive, grasped easily by clinicians and patients, and spans the spectrum of disability with strong validation [5]. Furthermore, the widespread use of the mRS also allows us to more easily compare outcomes between different trials. Subsequent research has shown that even a single point change in the mRS is clinically meaningful to clinicians and patients, affecting duration of hospital stay and disability-adjusted life years. [6,7]. However, the mRS has several limitations that should be considered including: subjectivity; high inter-observer variability and reproducibility; and uncertainty about how to dichotomise good vs. poor outcomes in analysis. The mRS does not include the full range of outcomes [5], nor does it detect subtle changes in neurological sequelae that may have considerable impact on function and disability, such as the ability to maintain employment or relationships. Finally, the mRS does not explicitly incorporate quality of life measures.

Outcome after acute stroke is often measured using surrogate endpoints that may not fully capture patient perspective, quality of life and disability. For example, the 15-item National Institutes of Health Stroke Scale (NIHSS) is used commonly for stroke prognostication and can be used as a surrogate measure of functional outcome [8]. A change of at least four points is considered to be clinically meaningful in intravenous thrombolysis studies [6]. Although the NIHSS has advantages such as familiarity, it does not capture all stroke-related impairments, particularly those of vertebrobasilar circulation, and does not incorporate the patient's ability to complete activities daily living and level of disability.

Another surrogate outcome measure used in trials involving mechanical thrombectomy is the Thrombolysis in Cerebral Ischemia (TICI) score. Interventional neuroradiologists assign a subjective score ranging from 0 (no reperfusion) to 3 (complete filling of the expected vascular territory) to grade cerebral perfusion using cerebral angiography. In trials investigating mechanical thrombectomy, a typical goal is to achieve a minimum TICI 2b score in patients with an ischaemic thrombotic stroke. Similar to the NIHSS, the TICI score does not incorporate functional outcomes, nor does it account for eloquent brain areas or occlusion site, which are important for determining neurological sequelae. Surrogate outcome measures have advantages in terms of availability, ease of collection and immediacy; however, they have clear limitations when used in isolation.

The review by Chung et al. on interventions to prevent postoperative delirium (POD) and postoperative neurocognitive disorder (POCD) provides an useful summary of the role of peri-operative neurological monitoring with electroencephalography and cerebral oximetry [3]. Both conditions deserve careful examination when applied to peri-operative neuroscientific research. In an effort to reduce inconsistencies in outcome measures between studies, a working party has made recommendations to standardise terminology, time frames and criteria for postoperative cognitive disorders and to align these with the Diagnostic and Statistical Manual for Mental Disorders, 5th edition) (DSM-5) [9]. This group recommended the use of the term 'delayed neurocognitive recovery' for a diagnosis made < 30 days from surgery to incorporate the potential for recovery, with the diagnosis based on DSM-5 criteria for mild and major neurocognitive disorders. [9]. Between 30 days and 12 months after surgery, POCD can be diagnosed using similar DSM-5 criteria to indicate the temporal relationship with surgery and anaesthesia, but not necessarily the aetiology [9]. The standardisation of criteria and terminology in these recommendations addresses the heterogeneity in measurement of peri-operative neurocognitive disorders, although several limitations remain. For example, POCD is a research outcome that is diagnosed by extensive psychometric assessments that are not easily applied to the clinical setting. As this working group noted, identifying changes after surgery requires defining baseline cognitive status before POCD can be diagnosed. In addition, the expert working group represented primarily highly-resourced English-speaking countries, which may not reflect the full spectrum of cultural context and application of these terms.

Postoperative delirium is screened for using the Confusion Assessment Method (CAM) and its intensive care unit version (CAM-ICU). This tool can be used for screening up to 30 days after surgery and was developed for use by non-psychiatrists [10]. The CAM combines clinical features to screen for delirium including: acute onset or fluctuating mental status changes; inattention; and either altered level of consciousness or disorganised thinking. This was modified in the CAM-ICU to allow for non-verbal assessments in patients receiving mechanical ventilation [11]. Both CAM and CAM-ICU have several advantages, as these tools are easy to use, utilised widely in research and validated for use in the hospital and ICU [11]. While these tools are sensitive (93–100%) and specific (98–100%) when used by trained research assistants [11], the CAM-ICU tool did not perform as well when used in routine practice (sensitivity 47%, specificity 98%) [12]. This decreased sensitivity raises questions about the reliability of practical delirium assessment in clinical research studies, and may underestimate the incidence of delirium. The severity of delirium can also be difficult to assess, although additional tools such as the CAM-Severity (CAM-S) score can be used to provide additional weight to symptom severity [13]. Whilst validated for older adults aged ≥ 70 y, the CAM-S still requires validation in younger populations [13]. In addition, while delirium is correlated with long-term postoperative cognitive outcomes [14], this outcome measure does not measure disability or patient-centred outcomes directly.

Outcomes in traumatic brain injury (TBI) have focused on mortality, seizure incidence and neurological outcomes. The Glasgow Outcome Scale (GOS) was first proposed by Jennett and Bond as a standardised assessment of global outcome after severe brain injury [15]. Given concerns about inter-rater reliability and poor sensitivity to changes in functional status in higher functional categories, the GOS outcome categories were later expanded to become the Glasgow Outcomes Scale-Extended (GOS-E), and structured interviews were published to standardise scoring [16]. The GOS-E is the most common outcome measure used in trials of TBI management and assesses changes in a patient's level of functionality in discrete categories (domains) arranged in a hierarchy [16]. These domains represent a holistic assessment in function and include: consciousness; independence at and outside of home; work capacity; social and leisure activities; psychological disruption to family relationships or friendships; and return to normal life. Advantages of the GOS-E include a structured format with guidelines to standardise assessments and maintain consistency between trials, and inclusion of clinically meaningful patient-oriented outcomes relevant for functional capacity. Despite these advantages, there remain limitations with the GOS and GOS-E, such as concerns about subjectivity and inter-rater reliability [17]. In addition, the GOS-E and structured interviews were developed by researchers, and may not account

for the patient's view on quality of life and disability, particularly within domains that can be highly influenced by emotional factors [18]. The GOS-E may also not fully capture the entire spectrum of functional limitation [19]. Finally, most trials report GOS-E outcomes within 6 months of injury which may not encompass the patient's final neurological outcome, as disability can improve or deteriorate over subsequent years [18].

Shift towards patient-reported outcome measures

Research has seen a much-needed shift from investigator-driven outcomes towards incorporating patient-reported outcome measures (PROMs) and centring the patient in research. Although traditional outcomes such as death, incidence of complications and duration of stay are important to both patients and clinicians, they do not encompass the full spectrum of meaningful events [20]. Patient-reported outcome measures are collected directly from patients, reflecting the patients' perceptions of their health, and broadly measure general health, quality of life, disability and functional impairment [21]. They can be general or disease-specific, and typically involve questionnaires administered before and after surgery. For example, Health Measures (www.healthmeasures.net) provides a set of health-related quality of life measures, which includes the National Institutes of Health (NIH) toolbox for Assessment of Neurological and Behavioural Function [22] and Quality of Life in Neurological Disorders (Neuro-QoL) [23]. These validated and standardised PROMs were designed to be applied to a diverse range of study designs and populations across the lifespan of patients and are designed specifically to capture multiple dimensions of neurological function. Although the advantages of incorporating PROMs are clear, we must acknowledge their limitations, which include: greater subjectivity; duration of time to complete the questionnaires; difficulty with interpretation; and lack of validation in specific populations including those with different neurological conditions [24], patients from lower-income countries and patients who speak languages other than English.

The modification of the mRS is an example of how patient-centred outcomes can be integrated into existing, established neurological outcome measures. Integration of external PROMs into the mRS can provide additional weighting of outcomes based on patient-reported quality of life and function. Utility values provide a quantitative valuation of the quality of life expected given a specific health rating. These have been calculated by mapping responses from the European Quality of Life scale (EQ-50) [25] and the World Health Organization Global Burden of Disease [26] onto the mRS levels in patients who have suffered a stroke, which then provides an estimate of disability weights for various mRS levels. As a

result of these analyses, the utility-weighted modified Rankin Scale (UW-mRS) was developed to incorporate patient-centred outcomes and quality of life [25]. The UW-mRS also provides efficiency gains in trials as it is able to detect more granular treatment effects [27]. Despite the many advantages of the UW-mRS, this tool may be limited by high variability in utility values between and within mRS categories over time post-stroke [28]. In addition, these measures were again predominantly developed in high-income countries and in patients who spoke English.

The future of outcomes in neuroanaesthesia and neurocritical care

Future trials should continue to incorporate patient-reported outcomes, in addition to other clinically important measures currently used such as conventional neurological outcome scales and mortality. Patients who have had a stroke are concerned about social functioning, mood, self-care and the ability to return to usual activities, not just mortality [29]. When studying interventions for delirium, a composite of outcomes relevant to clinicians (e.g. delirium incidence, severity and duration of hospital stay) and to patients (e.g. emotional distress, ability to return to previous cognitive abilities) should be considered [30]. After TBI, patient-centred outcomes should include the quality of life post-trauma of both the patient and their family [2]. Resources for researchers on recommended trial outcomes include the Core Outcome Measures in Effectiveness Trials (COMET) initiative (www.comet-initiative.org) and the Patient-Reported outcomes Measurement Information System (PROMIS) (www.healthmeasures.net).

The standardisation of outcomes is equally important to facilitate comparisons of interventions between different trials and to increase the validity of systematic reviews and meta-analyses. As noted above, POD and neurocognitive dysfunction have been defined recently and standardised for this purpose.[9] The Core Outcome Measures in Peri-operative and Anaesthetic Care (COMPAC) and parallel Standardised Endpoints for Peri-operative Medicine (StEP) groups were created to develop recommendations for standardised end points in anaesthesia and peri-operative research, and of note include a working group on cognition and stroke [31]. Finally, long-term data are needed, particularly in patients who experience neurological conditions such as delirium, cognitive impairment, stroke, TBI and seizures. These outcomes will guide clinicians in not only on determining the effectiveness of interventions like mechanical thrombectomy and anaesthetic choices, but also on fulfilling the goals that matter to patients and their families.

Conclusion

The outcomes used in neuroanaesthesia and neurocritical care research have traditionally be represented using ordinal scales of disability such as the mRS, surrogate outcomes and clinician-centred outcomes such as duration of hospital stay and cost measured over the short term. Patient-reported outcomes such as quality of life have been integrated increasingly in clinical trials, and traditional outcomes modified to include long-term patient perspectives. Several standardised tools and metrics to measure neurological outcomes and PROMs are now available to researchers, and are measured over a longer period of time. Finally, many currently available outcome measures reflect the perspectives of patients from high-income countries and who are English language speakers, and this gap needs to be addressed in the future.

Acknowledgments

AF is an associate editor of the *Canadian Journal of Anesthesia* and *Journal of Neurosurgical Anesthesiology* and receives compensation for contributions to UpToDate. No other competing interests declared.

REFERENCES

1. Dinsmore JE, Tan A. Anaesthesia for mechanical thrombectomy: a narrative review. *Anaesthesia* 2022. https://doi.org/10.1111/anae.15586
2. Wiles MD. Management of traumatic brain injury: a narrative review of current evidence. *Anaesthesia* 2022. https://doi.org/10.1111/anae.15608
3. Chung CKE, Poon CCM, Irwin MG. Peri-operative neurological monitoring with electroencephalography and cerebral oximetry: a narrative review. *Anaesthesia* 2022. https://doi.org/10.1111/anae.15616
4. Rankin J. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. *Scottish Medical Journal* 1957; **2:** 200-15.
5. Broderick JP, Adeoye O, Elm J. Evolution of the Modified Rankin Scale and its use in future stroke trials. *Stroke* 2017; **48:** 2007-12.

6. Lees KR, Bath PM, Schellinger PD, et al. Contemporary outcome measures in acute stroke research: choice of primary outcome measure. *Stroke* 2012; **43:** 1163-70.
7. Dawson J, Lees JS, Chang TP, et al. Association between disability measures and healthcare costs after initial treatment for acute stroke. *Stroke* 2007; **38:** 1893-8.
8. Chalos V, van der Ende NAM, Lingsma HF, et al. National Institutes of Health Stroke Scale: an alternative primary outcome measure for trials of acute treatment for ischemic stroke. *Stroke* 2020; **51:** 282-90.
9. Evered L, Silbert B, Knopman DS, et al. Recommendations for the Nomenclature of Cognitive Change Associated with Anaesthesia and Surgery-2018. *Anesthesiology* 2018; **129:** 872-9.
10. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Annals of Internal Medicine* 1990; **113:** 941-8.
11. Ely EW, Inouye SK, Bernard GR, et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). *Journal of the American Medical Association* 2001; **286:** 2703-10.
12. van Eijk MM, van den Boogaard M, van Marum RJ, et al. Routine use of the confusion assessment method for the intensive care unit: a multicenter study. *American Journal of Respiratory and Critical Care Medicine* 2011; **184:** 340-4.
13. Inouye SK, Kosar CM, Tommet D, et al. The CAM-S: development and validation of a new scoring system for delirium severity in 2 cohorts. *Annals of Internal Medicine* 2014; **160:** 526-33.
14. Huang H, Li H, Zhang X, et al. Association of postoperative delirium with cognitive outcomes: A meta-analysis. *Journal of Clinical Anesthesia* 2021; **75:** 110496.
15. Jennett B, Bond M. Assessment of outcome after severe brain damage. *Lancet* 1975; **1:** 480-4.
16. Wilson L, Boase K, Nelson LD, et al. A Manual for the Glasgow Outcome Scale-Extended Interview. *Journal of Neurotrauma* 2021; **38:** 2435-46.
17. Anderson SI, Housley AM, Jones PA, Slattery J, Miller JD. Glasgow Outcome Scale: an inter-rater reliability study. *Brain Injury* 1993; **7:** 309-17.
18. Whitnall L, McMillan TM, Murray GD, Teasdale GM. Disability in young people and adults after head injury: 5-7 year follow up of a prospective cohort study. *Journal of Neurology, Neurosurgery and Psychiatry* 2006; **77:** 640-5.

19. Nelson LD, Brett BL, Magnus BE, et al. Functional status examination yields higher measurement precision of functional limitations after traumatic injury than the Glasgow Outcome Scale-Extended: a preliminary study. *Journal of Neurotrauma* 2020; **37:** 675-9.
20. Ladha KS, Wijeysundera DN. Role of patient-centred outcomes after hospital discharge: a state-of-the-art review. *Anaesthesia* 2020; **75 (Suppl 1):** e151-e7.
21. Reeves M, Lisabeth L, Williams L, et al. Patient-Reported Outcome Measures (PROMs) for acute stroke: rationale, methods and future directions. *Stroke* 2018; **49:** 1549-56.
22. Gershon RC, Wagster MV, Hendrie HC, Fox NA, Cook KF, Nowinski CJ. NIH toolbox for assessment of neurological and behavioral function. *Neurology* 2013; **80:** S2-6.
23. Gershon RC, Lai JS, Bode R, et al. Neuro-QOL: quality of life item banks for adults with neurological disorders: item development and calibrations based upon clinical and general population testing. *Quality of Life Research* 2012; **21:** 475-86.
24. Shah A, Bailey CR. Outcomes following surgery: are we measuring what really matters? *Anaesthesia* 2019; **74:** 696-9.
25. Chaisinanunkul N, Adeoye O, Lewis RJ, et al. Adopting a patient-centered approach to primary outcome analysis of acute stroke trials using a utility-weighted Modified Rankin Scale. *Stroke* 2015; **46:** 2238-43.
26. Hong KS, Saver JL. Quantifying the value of stroke disability outcomes: WHO global burden of disease project disability weights for each level of the modified Rankin Scale. *Stroke* 2009; **40:** 3828-33.
27. Wang X, Moullaali TJ, Li Q, et al. Utility-weighted modified Rankin Scale Scores for the assessment of stroke outcome: pooled analysis of 20 000+ patients. *Stroke* 2020; **51:** 2411-7.
28. Rethnam V, Bernhardt J, Dewey H, et al. Utility-weighted modified Rankin Scale: Still too crude to be a truly patient-centric primary outcome measure? *International Journal of Stroke* 2020; **15:** 268-77.
29. Salinas J, Sprinkhuizen SM, Ackerson T, et al. An international standard set of patient-centered outcome measures after stroke. *Stroke* 2016; **47:** 180-6.
30. Rose L, Burry L, Agar M, et al. A Core outcome set for research evaluating interventions to prevent and/or treat delirium in critically ill adults: an international consensus study (Del-COrS). *Critical Care Medicine* 2021; **49:** 1535-46.

31. Myles PS, Grocott MP, Boney O, Moonesinghe SR. Standardizing end points in perioperative trials: towards a core and extended outcome set. *British Journal of Anaesthesia* 2016; **116:** 586-9.