Quality of Life Research (2019) 28:1119–1128
https://doi.org/10.1007/s11136-018-2053-7

REVIEW

Patient-reported outcomes in stroke clinical trials 2002–2016: a systematic review

Eboni G. Price-Haywood[1,2] · Jewel Harden-Barrios[1] · Christopher Carr[3] · Laya Reddy[2] · Lydia A. Bazzano[3] · Mieke L. van Driel[4]

Accepted: 12 November 2018 / Published online: 21 November 2018

Abstract

Objective Given the global and economic burden of stroke as a major cause of long-term disability, patient-reported outcomes (PRO) data from clinical trials can elucidate differential benefits/harms of interventions from patients' perspectives and influence clinical decision making in stroke care management.

Methods This systematic review examines stroke-related randomized controlled trials (RCT) published in 12 high-impact journals between 2002 and 2016 for (1) associations between trial characteristics and the reporting of PRO measures; and (2) psychometric properties of PRO instruments used in these studies. The study combines clinical trials identified in a prior review with trials identified with an updated literature search.

Results Only 34 of 159 stroke-related RCTs reported PRO measures. Among the 34 trials, most were published in rehabilitation and general medical journals, were conducted in the United States or Europe, were funded by government/non-industry sponsors, and focused on post-stroke care. Thirty-one PRO instruments were employed in these studies. Only 5 instruments were stroke-specific measures, whereas the remaining 26 instruments were generic measures. Eight instruments assessed functional status, 9 measured quality of life, and 14 assessed symptoms. The most common health domains measured were emotional status and physical function.

Conclusions This study highlights the paucity of information from patients' perspective in stroke-related RCTs. This trend may change over time as researchers increase adherence to reporting guidelines for clinical trial protocols.

Keywords Stroke · Randomized controlled trial · Patient-reported outcome measures

Introduction

The global burden of stroke is well documented with disease prevalence reaching 25.7 million in 2013 [1]. Stroke is a major cause of long-term disability. Catastrophic healthcare expenditures and productivity losses related to disabilities, time away from work for medical care, and/or unemployment create tremendous economic burdens for individuals and society. In 2016, the International Consortium of Health Outcomes Measurement (ICHOM) developed a standard set of patient-centered stroke outcome measures to enable assessment of value-based healthcare in the management of patients who present with ischemic stroke or intracerebral hemorrhage [2]. Outcomes recommended for assessment include survival, disease control, acute complications, and patient-reported outcomes.

A patient-reported outcome (PRO) is "any report of the status of a patient's health condition that comes directly from

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11136-018-2053-7) contains supplementary material, which is available to authorized users.

✉ Eboni G. Price-Haywood
eboni.pricehaywood@ochsner.org

[1] Center for Applied Health Services Research, Ochsner Health System, Academic Building – 2nd Floor, 1514 Jefferson Highway, New Orleans, LA 70121, USA

[2] Ochsner Clinical School, University of Queensland, New Orleans, LA, USA

[3] Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA

[4] Primary Care Clinical Unit, Faculty of Medicine, University of Queensland, Brisbane, Australia

the patient, without interpretation of the patient's response by a clinician or anyone else" [3]. PRO data from clinical trials are increasingly influencing patient care, clinical decision making, health policy, and/or reimbursement decisions. If captured in a scientifically rigorous manner, PRO measures can help elucidate differential benefits and/or harms of interventions from patients' perspective beyond traditional clinical and/or biomedical stroke care outcomes (e.g., time to thrombolytics; modified Rankin Scale score at discharge; proportion of stroke patients prescribed antiplatelet or statin drugs). Post-stroke problems with mobility, activities of daily living, cognitive function, and psychological distress are well documented [1]. Measures that capture these challenges from the patients' perspective are especially important during rehabilitation, community reintegration, and recovery. PRO measures can further inform clinicians how to incorporate patients' needs and preferences into individualized stroke care management plans to generate health outcomes that matter most to patients.

Given the ICHOM recommendations for stroke management, this study examines reporting of PROs among stroke randomized controlled trials published in high-impact journals between 2002 and 2016. The objectives of this systematic review are to assess the associations between study characteristics and the reporting of PRO measures, and to examine the properties of various PRO measures that were identified from stroke clinical trials. We hypothesized that PRO measures were more likely to be collected among clinical trials funded by non-industry sponsors, focused on post-stroke care (e.g., rehabilitation), and published within recent years.

Methods

Journal and study selection

The study combines clinical trials identified in a prior review with trials identified with an updated literature search. The prior review included a cohort of 99 randomized controlled trials (RCTs) in a systematic review of under-reporting of socioeconomic status of patients in stroke trials [4]. The authors of the previous review identified RCTs via Medline searches that were conducted between 2002 and 2012 with stroke or transient ischemic attack (TIA) defining recruitment or study outcome. Inclusion criteria were RCTs recruiting patients after a stroke or TIA; prevention studies with stroke or TIA as the primary or secondary outcome; RCTs with clinical outcomes (no biochemical or surrogate outcomes); and studies that included at least 30 days of post-intervention follow-up.

Journals were selected based on aiming to include a range of general medical, stroke, general neurology, and rehabilitation journals. Journals were then selected from each of these categories with the highest impact factor listed on Web of Science in 2012 at the time the original review was conducted. The current impact factor for included journals are as follows: Annals of Neurology, 5.7; Archives of Physical Medicine Rehabilitation, 3.07; British Medical Journal, 23.3; Brain, 10.84; Cerebrovascular Disease, 2.97; International Journal of Stroke, 3.86; Lancet, 53.26; Lancet Neurology, 11.96; New England Journal of Medicine, 79.26; Neurorehabilitation and Neural Repair, 2.27; Physical Therapy, 2.53; and Stroke, 3.52. They had to have at least ten RCTs on stroke or TIA published since 2001. Journals that did not focus on clinical medicine were excluded. Using PUBMED, the current review updated the search and was restricted to RCTs published in English between 2012 and 2016 within the same 12 journals. The search MESH terms included "adult," "stroke," and "human."

Data extraction

Information on general study characteristics was extracted (e.g., journal; year of publication). Each study's purpose or objective was classified as stroke prevention (e.g., treatment of carotid stenosis), acute stroke management, post-stroke care (e.g., physical rehabilitation), or multi-purpose. Study location was classified as occurring in the United States, Europe, Asia, Australia, multi-national, other, or not reported. Source of funding was identified as government, industry, non-industry (e.g., private, non-commercial), multi-source, unfunded, or not reported. The presence/absence of PRO measures (as defined by the International Society for Quality of Life Research [ISOQOL]) [5] was noted for each study. Data were independently extracted by two investigators. Disagreements were adjudicated by a third reviewer (EPH).

For studies that reported PRO measures, each instrument was classified as condition-specific (stroke or neurologic disorder) versus generic. PRO measures were then classified by the type of outcome assessed in accordance with ISOQOL classifications: [5] symptoms, functional status, health-related quality of life, and/or health behaviors. The Medical Outcomes Study (MOS) Framework of Health Indicators was used to classify the domains of health measured by each instrument [6]. The ISOQOL and MOS were selected because their working definitions of health concepts were useful as guides for organizing PRO measures. Information on the psychometric properties (validity, reliability) of each PRO measure and domains of health assessed was acquired through a literature review of validation studies for each instrument used. Each PRO measure's copyright status as well as requirements for user fees and registration was collected via Google, e-PROVIDE, Canadian Stroke

Engine, and requests for information sent to authors of the instrument original validation studies.

Data analysis

The frequency count of the study characteristics was examined. The year of publication was categorized as 2002–2008, 2009–2012, and 2013–2016 to account for publication of the 2009 US Food and Drug Administration guidance for industry [3], the 2013 ISOQOL standards [5], and the 2016 ICHOM standards [2]. A bivariate analysis of study characteristics by PRO measure collection status using the Fisher's exact test was conducted. A logistic regression model was constructed with binary response reflecting the presence of PROs and included fixed effects for medical journal, study purpose, study location, and study funding. Firth's penalized likelihood approach was used to account for complete separation (i.e., no PRO collected) in some levels of the study characteristics [7]. Records for manuscripts with unknown purpose, location, or funding were excluded from the analysis. All analyses were conducted using SAS/STAT ® software, Version 9.4 of the SAS System for Windows (Cary, NC, USA).

Results

The additional search yielded 99 articles of which 39 were excluded for the following reasons: 19 were not an original report of a trial; 14 studies targeted populations or measured outcomes that were not stroke related; 4 were baseline methods papers; 1 study had less than 30 days follow-up; and 1 study only reported cost-effectiveness analysis. The remaining 60 articles were included in this review yielding a total of 159 articles for data extraction (see Supplementary material e-References).

Table 1 summarizes the characteristics of the 159 stroke RCTs included in this review. Only 34 (21%) studies collected or reported PRO measures. Most of these studies were published in rehabilitation, and general medical journals between 2009 and 2012 were conducted in the United States or Europe, were funded by government or non-industry sponsors, and focused on post-stroke care. Few multinational studies and none of the industry-sponsored studies included in this review collected or reported PRO measures. In the multivariate logistic regression analysis, the study purpose was the only significant factor associated with collection of PRO measures.

Tables 2 and 3 describe the 31 PRO measures collected. Only 5 instruments were stroke-specific PRO measures [8–13], whereas the remaining 26 instruments were generic measures [14–55]. Eight instruments assessed functional status, 9 measured quality of life, and 14 assessed symptoms. The most common health domains measured were psychological/emotional status ($n=19$) and physical function ($n=16$). Most PRO measures utilized had been previously evaluated for instrument reliability (internal consistency, $n=23$; test–retest reliability, $n=21$) as well as construct ($n=15$), convergent ($n=12$), predictive ($n=4$), and/or criterion ($n=14$) validity. Some generic measures have been psychometrically tested among patients with a history of stroke in addition to the general population: Activity-Specific Balanced Confidence Scale [14, 15], Frenchay Activity Index [16–18], Motor Activity Log [19, 20], Nottingham Extended Activities Daily Living [21, 22], EuroQol [27, 28], SF-12 [31–33], Fatigue Severity Scale [43, 44], and Mental Slowness Questionnaire [51]. Finally, 20 instruments were copyrighted ($n=20$), 13 required registration for use, and 9 required fees for use.

Discussion

Cardiovascular diseases are the major causes of global public health crises. Stroke is a major cause of long-term disability [1], yet less than 25% of stroke clinical trials included in this review collected and/or reported PROs. To our knowledge, our study is the first to associate the use of PRO instruments with funding source for stroke clinical trials. Industry-funded studies collected the fewest PRO measures. Study location and journal of publication are also associated with the collection and/or reporting of PROs. Generic PRO measures as opposed to stroke-specific measures were more often used. Studies that included PRO measures targeted post-stroke care and frequently assessed emotional status (e.g., depression) and physical function using valid and reliable instruments. Study purpose had the strongest association with the collection/reporting of PRO measures. These findings, in the case of a disease entity with potentially devastating effects on functional outcomes of patients as well as major societal burden, highlight the need for increased use of PRO measures in stroke clinical research.

The utility and applicability of PRO data from stroke clinical trials depends on accurate, valid reporting. Ten of the 12 medical journals targeted for this review endorse the 2010 Consolidating Standard of Reporting Trials (CONSORT) statement. However, endorsement of the 2013 CONSORT-PRO extension was not clear. The CONSORT-PRO includes five recommendations relevant to identifying PROs as primary or secondary outcomes; describing the PRO hypothesis; providing or citing the evidence of the PRO instruments psychometric properties; describing statistical approaches for dealing with missing data; and discussing PRO-specific limitation of study findings and generalizability of results to other populations and clinical practice [56]. Lack of clear journal endorsement

Table 1 Characteristics of stroke randomized clinical trials published between 2002 and 2016

Study characteristic	PRO measures collected or reported		*P* value*
	Yes (*N* = 34)	No (*N* = 125)	
Medical Journal, *n* (%)			0.0030
Annals of Neurology	0 (0.0)	6 (4.8)	
Archives Physical Medicine Rehabilitation	5 (14.7)	5 (4.0)	
BMJ	4 (11.8)	6 (4.8)	
Brain	2 (5.9)	2 (1.6)	
Cerebrovascular Disease	4 (11.8)	7 (5.6)	
International Journal of Stroke	0 (0.0)	7 (5.6)	
Lancet	1 (2.9)	15 (12.0)	
Lancet Neurology	1 (2.9)	11 (8.8)	
New England Journal of Medicine	2 (5.9)	25 (20.0)	
Neurorehabilitation and Neural Repair	5 (14.7)	4 (3.2)	
Physical Therapy	3 (8.8)	5 (4.0)	
Stroke	7 (20.6)	32 (25.6)	
Publication year, *n* (%)			0.0009
2002–2008	9 (26.5)	16 (12.8)	
2009–2012	23 (67.7)	61 (48.8)	
2013–2016	2 (5.9)	48 (38.4)	
Study purpose, *n* (%)			< 0.0001
Stroke prevention	1 (2.9)	46 (36.8)	
Acute stroke management	3 (8.8)	48 (38.4)	
Post-stroke care	29 (85.3)	28 (22.4)	
Multiple	0 (0.0)	3 (2.4)	
Unknown	1 (2.9)	0 (0.0)	
Study location, *n* (%)			0.0005
United States	8 (23.5)	12 (9.6)	
Europe	16 (47.1)	41 (32.8)	
Asia	1 (2.9)	17 (13.6)	
Australia	3 (8.8)	7 (5.6)	
Other	2 (5.9)	2 (1.6)	
Multi-national	3 (8.8)	46 (36.8)	
Not reported	1 (2.9)	0 (0.0)	
Study funding, *n* (%)			0.0001
Government	14 (41.2)	30 (24.0)	
Industry	0 (0.0)	28 (22.4)	
Non-industry	10 (29.4)	10 (8.0)	
Multiple	9 (26.5)	49 (39.2)	
None	1 (2.9)	4 (3.2)	
Not reported	0 (0.0)	4 (3.2)	
PRO construct measured, *n* (%)			
Functional status	6 (17.6)	–	
Health behaviors	1 (2.9)	–	
Quality of life	4 (11.8)	–	
Symptoms	3 (8.8)	–	
Multiple	10 (29.4)	–	
None/unknown	2 (5.9)	125 (100.0)	

of CONSORT-PRO may explain why this study did not observe a consistent upward trend in the reporting of PRO measures in stroke RCTs beyond 2012.

In a recent review, Vodicka et al. report that the number of clinical trials that collect patient-reported outcome measures is increasing [57]. This trend is especially notable

Table 2 Domains of health assessed by patient-reported outcomes measures used in stroke clinical trials

PRO measure	Item #	Physical function	Psych. / emotion	Cognitive function	Impairment	Social support	Social activity	Role function	Sexual function	Pain	Energy/fatigue
Stroke specific											
Stroke-adapted sickness impact profile [8]	30	x	x	x			x	x			
Stroke and Aphasia Quality of Life Scale [9]	39	x	x								x
Stroke Expectations Questionnaire [10, 11]	15		x								
Stroke Impact Scale 2.0 [12]	64	x	x	x			x				
Stroke Impact Scale 3.0 [13]	59	x	x	x							
Generic											
Activities-Specific Balance Confidence Scale [14, 15][a,b]	16	x									
Frenchay Activities Index [16–18]	15	x									
Motor Activity Log [19, 20]	28/14	x									
Nottingham Extended ADL Scale [21, 22][a]	26	x									
Rivermead Mobility Index [23][a]	15	x									
Nottingham leisure questionnaire [24][a]	30	x					x				
Physical activity and disability survey [25][a,b]	28	x									
Physical Activity Scale for the elderly [26][a,b,c]	12	x									
EuroQol EQ5D [27, 28] [a,b]	15	x	x		x		x	x		x	
General health questionnaire [29, 30][a,b,c]	12		x								
Health Utility Index [28][a,b,c]	8	x	x	x	x		x	x	x	x	x
Medical outcome study SF12 [31–33]	12	x	x				x	x		x	x
Medical outcome study SF36 [34–36]	36	x	x		x		x	x		x	x
Beck Depression Scale [37–39][a,b,c]	21		x								
Center for Epidemiologic Studies Depression Scale Revised [40]	20		x								x
Checklist individual strength—fatigue [41, 42][a,b]	20										x
Fatigue Severity Scale [43, 44] [a,d]	9										x
Geriatric Depression Scale—short form [45, 46]	15		x								
Hospital anxiety and depression scale [47–49][a,b,c]	14		x								
Irritability, depression and anxiety scale [50][d]	18		x								
Mental slowness questionnaire [51][a]	21			x							
Montgomery–Åsberg Depression Scale [52][a]	10		x								x
Numerical rating for pain	1									x	
Profile of mood states [53][a,b,c]	65		x	x							x
State Trait Anger Expression Scale [54][a,b,c]	8		x								
Yale depression screener [55]	1		x								

Superscript numbers correspond to reference numbers

[a]Copyright

[b]Register for use

[c]Use fee

[d]Unknown

Table 3 Psychometric properties of patient-reported outcomes measures included in stroke clinical trials

Patient-reported outcomes		Reliability		Validity				
Type of PRO	PRO instrument	Internal consistency[a]	Test–retest[b]	Construct[c]	Convergent[d]	Divergent/ discriminant[e]	Predictive[f]	Criterion[g]
Functional status	Activities-Specific Balance Confidence Scale	0.94	0.80–0.878	x				x
	Frenchay Activities Index	0.78–0.87	0.96	x	x	x		x
	Motor Activity Log	Amount of use (AOU) 0.95; Quality of movement (QOM) 0.95	AOU 0.70, QOM 0.82		x	x		x
	Nottingham Extended Activity of Daily Living scale		0.62–1.00	x			x	x
	Rivermead Mobility Index	0.93	0.732					
Health behaviors	Nottingham Leisure Questionnaire		0.44–0.94					
	Physical Activity and Disability Survey	0.66–0.77	0.78–0.95				x	x
	Physical Activity Scale for the Elderly	0.69	0.75	x				
	Stroke-Adapted Sickness Impact Profile	0.85		x	x			
	Stroke and Aphasia Quality of Life Scale	0.93 scale; 0.74–0.94 subdomains	0.98 scale; 0.89–0.98 subdomains	x	x	x		
Quality of life	Stroke Impact Scale Versions 2.0 & 3.0	0.83–0.90	0.70–0.92, except emotion (0.57)		x	x		x
	EuroQol EQ5D		0.61–0. 72	x	x	x		x
	General Health Questionnaire	0.82–0.86						x
	Health Utility Index		0.75–0.76	x				
	Medical Outcome Study Short Form 12	Total 0.83–0.89, PCS 0.88; MCS 0.82	0.78 PCS; 0.60 MCS	x	x	x	x	x
	Medical Outcome Study Short Form 36	0.70–0.93 (Physical, PCS and mental MCS scales)		x	x	x		x

Table 3 (continued)

Patient-reported outcomes		Reliability		Validity				
Type of PRO	PRO instrument	Internal consistency[a]	Test–retest[b]	Construct[c]	Convergent[d]	Divergent/ discriminant[e]	Predictive[f]	Criterion[g]
Symptom	Stroke Expectations Questionnaire	0.89					x	
	Beck Depression Scale	0.73–0.95	0.48–0.83	x	x	x		x
	Center for Epidemiologic Studies Depression Scale Revised 20-item	0.85–0.93	0.32–0.67	x	x	x		
	Checklist Individual Strength-Fatigue	0.62–0.94	0.85		x	x		
	Fatigue Severity Scale	0.86–0.96	0.84–0.93					x
	Geriatric Depression Scale—Short Form	0.75–0.94	0.81–0.85	x				
	Hospital Anxiety and Depression scale	Anxiety 0.68–0.93, Depression 0.67–0.90 (English)	Anxiety 0.54, Depression 0.79	x		x		x
	Irritability, Depression, and Anxiety scale		Depression 0.81, Anxiety 0.87, Outward Irritability 0.88, Inwardly Irritability 0.93					x
	Mental Slowness Questionnaire	0.91	0.85–0.90					
	Montgomery–Åsberg depression rating scale	0.76						
	Numerical rating scale for pain	Single-item question						
	Profile of Mood States	0.63–0.96						
	State Trait Anger Expression Scale	0.73–0.96		x	x	x		
	Yale Depression Screener	Single-item question						

[a]Internal consistency: degree to which individual items are associated with each other in a multi-item measure (measure of precision at a single time point)

[b]Test–retest: reproducibility of scores over time

[c]Construct: measures theoretically intended constructs

[d]Convergent: correlates with other instruments that evaluate the same construct

[e]Divergent: differentiates between groups that past evidence has shown to be different

[f]Predictive: predictive of a certain state or condition

[g]Criterion: degree to which scores of the instrument reflect a "gold standard"

in cancer-related trials; studies sponsored by university/research organizations or the United States government; studies evaluating devices, procedures, or behaviors; and studies published in oncology, internal medicine, and general medical journals [57, 58]. The authors postulate that the upward trend may be partially attributable to explicit

journal endorsement of the CONSORT-PRO extension [58]. Likewise, we postulate that research published in journals whose targeted readers specialize in the care of patients with neurological disorders will also demonstrate an increase in reporting of PRO measures as these journals increasingly endorse CONSORT-PRO and the more recent SPIRIT-PRO extension statement [59].

Few clinical trials included in this review used instruments that captured outcomes specific to patients with strokes or neurologic disorders. Condition-specific health status measures are intuitively appealing to practicing clinicians. They are more responsive to small treatment effects and assess domains of greatest importance to a specific condition [60]. In addition to stroke-specific PRO measures described in this review, the Neurology Quality of Life (Neuro-QOL) [61] and National Institute of Health Toolbox [62] for assessment of neurological and behavioral function became available during the later years of our study period. None of these instruments were utilized in the RCTs included in this review. Collection of these measures in research or even clinical practice is likely influenced by whether PRO measures are incorporated into widely accepted clinical practice guidelines.

The ICHOM standards recommend using the Patient-Reported Outcomes Measurement Information System 10 (PROMIS-10) short form questionnaire (10-items) to measure multiple dimensions of health status in stroke care management [2]. It is publicly available in paper and digital format with no associated fees, has been translated into multiple languages, and the scores can be converted to scores of other validated PRO measures such as the SF-36 and EuroQOL-5D instruments. Its short length makes it relatively easy to adopt use in routine clinical practice and research.

In contrast, most of the generic PRO measures reviewed in this study were lengthy with a range of 12–65 questions in a single instrument. Given that clinical trials often include multiple assessments, investigators may opt to avoid collecting PROs that may increase patient and investigator burden in favor of other clinical measures. The International Society for Quality of Life Research developed minimum measurement standards by which PRO measures can be judged as acceptable [63]. These standards include documenting patient and investigator burden (length and literacy demand).

This systematic review has several limitations. Most of the studies included in this review were published prior to the ICHOM standards of stroke management as well as the CONSORT-PRO and SPIRIT-PRO standards of clinical trial protocol design and reporting. Restriction of the included studies to RCTs published in 12 select journals may have eliminated other study types in which collection of PRO measures may be more frequent or journals which may have a higher proclivity for publishing such studies. Selective reporting of outcomes among RCTs included in the review as well as publication bias pose additional limitations. Notwithstanding these limitations, the purpose of this review was to sample trials published in high-impact journals and which have influenced clinical guidelines for stroke care management. This study highlights the infrequent utilization of PRO measures in stroke-related randomized clinical trials, and hence lack of information from the perspective of patients dealing with the sequelae of stroke disease.

Conclusion

One would anticipate improvement in the reporting of PRO results in stroke clinical trials over time as more researchers adhere to the ICHOM, SPIRIT-PRO, and CONSORT-PRO standards. Future research must examine whether formal medical journal endorsement of these standards will help drive this trend independent of study funding sources.

Acknowledgements The staff of the Ochsner Medical Library conducted the literature search. Jeffrey Burton, PhD (Ochsner Center for Applied Health Services Research) conducted the descriptive statistics. Richard Zweifler, MD (Ochsner Neurology Stroke Center) conducted a review of the manuscript. Abstracts of this work were presented at the 2017 Southern Society for Clinical Investigations regional conference and 2017 International Stroke Conference.

Compliance with ethical standards

Conflict of interest None of the authors have financial disclosures or conflicts of interest to report with this study.

References

1. Benjamin, E. J., Blaha, M. J., Chiuve, S. E., Cushman, M., Das, S. R., Deo, R., et al. American Heart Association Statistics Committee and Stroke Statistics Subcommittee. (2017). Heart disease and stroke statistics-2017 update: a report from the American Heart Association. *Circulation, 135*(10), e146–e603.
2. Salinas, J., Sprinkhuizen, S. M., Ackerson, T., Bernhardt, J., Davie, C., George, M. G., et al. (2016). An international standard set of patient-centered outcome measures after stroke. *Stroke, 47*(1), 180–186.
3. Food and Drug Administration. (2009). Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. U.S. Department of Health and Human Services. Retrieved 19 March, 2018, from https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm193282.pdf.
4. Magin, P., Victoire, A., Zhen, X. M., Furler, J., Pirotta, M., Lasserson, D. S., et al. (2013). Under-reporting of socioeconomic status of patients in stroke trials: Adherence to CONSORT principles. *Stroke, 44*(10), 2920–2922.
5. Ahmed, S., Berzon, R. A., Revicki, D. A., Lenderking, W. R., Moinpour, C. M., Basch, E., et al. International Society for Quality of Life Research (2012). The use of patient-reported outcomes (PRO) within comparative effectiveness research: Implications

for clinical practice and health care policy. *Med Care, 50*(12), 1060–1070.
6. Stewart, A. L., & Ware, J. E. (1992). *Measuring functioning and well-being: The medical outcomes study approach*. Durham: Duke University Press.
7. Firth, D., & "Bias (1993). Reduction of maximum likelihood estimates. *Biometrika, 80*, 27–38.
8. van Straten, A., de Haan, R. J., Limburg, M., Schuling, J., Bossuyt, P. M., & van den Bos, G. A. M. (1997). A stroke-adapted 30-item version of the sickness impact profile to assess quality of life (sa-sip30). *Stroke, 28*, 2155–2161.
9. Hilari, K., Byng, S., Lamping, D. L., & Smith, S. C. (2003). Stroke and aphasia quality of life scale-39 (saqol-39): Evaluation of acceptability, reliability, and validity. *Stroke, 34*, 1944–1950.
10. Watkins, C., Leathley, M., & Davies, A. (1998). The stroke expectations questionnaire (seq): Identification of patients' ideas about recovery. *Clinical Rehabilitation, 12*, 173–175.
11. Watkins, C., Leathley, M., & Davies, A. (1998). The stroke expectations questionnaire (seq): A predictor of outcome? *Clinical Rehabilitation, 12*, 173.
12. Duncan, P. W., Wallace, D., Lai, S. M., Johnson, D., Embretson, S., & Laster, L. J. (1999). The stroke impact scale version 2.0. Evaluation of reliability, validity, and sensitivity to change. *Stroke, 30*, 2131–2140.
13. Jenkinson, C., Fitzpatrick, R., Crocker, H., & Peters, M. (2013). The stroke impact scale: Validation in a uk setting and development of a sis short form and sis index. *Stroke, 44*, 2532–2535.
14. Botner, E. M., Miller, W. C., & Eng, J. J. (2005). Measurement properties of the activities-specific balance confidence scale among individuals with stroke. *Disability and Rehabilitation, 27*, 156–163.
15. Pal, J., Hale, L. A., & Skinner, M. A. (2005). Investigating the reliability and validity of two balance measures in adults with stroke. *International Journal of Therapy and Rehabilitation, 12*, 308–315.
16. Turnbull, J. C., Kersten, P., Habib, M., McLellan, L., Mullee, M. A., & George, S. (2000). Validation of the frenchay activities index in a general population aged 16 years and older. *Archives of Physical Medicine and Rehabilitation, 81*, 1034–1038.
17. Wade, D. T., Legh-Smith, J., & Langton Hewer, R. (1985). Social activities after stroke: Measurement and natural history using the frenchay activities index. *International Rehabilitation Medicine, 7*, 176–181.
18. Schuling, J., de Haan, R., Limburg, M., & Groenier, K. H. (1993). The frenchay activities index. Assessment of functional status in stroke patients. *Stroke, 24*, 1173–1177.
19. Uswatte, G., Taub, E., Morris, D., Light, K., & Thompson, P. A. (2006). The motor activity log-28: Assessing daily use of the hemiparetic arm after stroke. *Neurology, 67*, 1189–1194.
20. Uswatte, G., Taub, E., Morris, D., Vignolo, M., & McCulloch, K. (2005). Reliability and validity of the upper-extremity motor activity log-14 for measuring real-world arm use. *Stroke, 36*, 2493–2496.
21. Gladman, J. R., Lincoln, N. B., & Adams, S. A. (1993). Use of the extended adl scale with stroke patients. *Age and Ageing, 22*, 419–424.
22. Nouri, F., & Lincoln, N. (1987). An extended activities of daily living scale for stroke patients. *Clinical Rehabilitation, 1*, 301–305.
23. Lennon, S., & Johnson, L. (2000). The modified rivermead mobility index: Validity and reliability. *Disability and Rehabilitation, 22*, 833–839.
24. Drummond, A. E., Parker, C. J., Gladman, J. R., Logan, P. A., & Group, T. S. (2001). Development and validation of the Nottingham leisure questionnaire (nlq). *Clinical Rehabilitation, 15*, 647–656.
25. Rimmer, J. H., Riley, B. B., & Rubin, S. S. (2001). A new measure for assessing the physical activity behaviors of persons with disabilities and chronic health conditions: The physical activity and disability survey. *American Journal of Health Promotion: AJHP, 16*, 34–42.
26. Washburn, R. A., Smith, K. W., Jette, A. M., & Janney, C. A. (1993). The physical activity scale for the elderly (pase): Development and evaluation. *Journal of Clinical Epidemiology, 46*, 153–162.
27. Dorman, P. J., Waddell, F., Slattery, J., Dennis, M., & Sandercock, P. (1997). Is the euroqol a valid measure of health-related quality of life after stroke? *Stroke, 28*, 1876.
28. Wilke, C. T., Pickard, A. S., Walton, S. M., Moock, J., Kohlmann, T., & Lee, T. A. (2010). Statistical implications of utility weighted and equally weighted hrql measures: An empirical study. *Health Economics, 19*, 101–110.
29. Goldberg, D. P., & Hillier, V. F. (1979). A scaled version of the general health questionnaire. *Psychological Medicine, 9*, 139–145.
30. Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., et al. (1997). The validity of two versions of the GHQ in the who study of mental illness in general health care. *Psychological Medicine, 27*, 191–197.
31. Cheak-Zamora, N. C., Wyrwich, K. W., & McBride, T. D. (2009). Reliability and validity of the sf-12v2 in the medical expenditure panel survey. *Quality of Life Research, 18*, 727–735.
32. Gandek, B., Ware, J. E., Aaronson, N. K., Apolone, G., Bjorner, J. B., Brazier, J. E., et al. (1998). Cross-validation of item selection and scoring for the sf-12 health survey in nine countries: Results from the iqola project. International quality of life assessment. *Journal of Clinical Epidemiology, 51*, 1171–1178.
33. Bohannon, R. W., Maljanian, R., Lee, N., & Ahlquist, M. (2004). Measurement properties of the short form (sf)-12 applied to patients with stroke. *International Journal of Rehabilitation Research, 27*, 151–154.
34. McHorney, C. A., Ware, J. E. Jr., Lu, J. F., & Sherbourne, C. D. (1994). The mos 36-item short-form health survey (sf-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care, 32*, 40–66.
35. McHorney, C. A., Ware, J. E. Jr., & Raczek, A. E. (1993). The mos 36-item short-form health survey (sf-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care, 31*, 247–263.
36. McHorney, C. A., Ware, J. E. Jr., Rogers, W., Raczek, A. E., & Lu, J. F. (1992). The validity and relative precision of mos short- and long-form health status scales and dartmouth coop charts. Results from the medical outcomes study. *Medical Care, 30*, Ms253–M265.
37. Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the beck depression inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*, 77–100.
38. Segal, D. L., Coolidge, F. L., Cahill, B. S., & O'Riley, A. A. (2008). Psychometric properties of the beck depression inventory II (BDI-II) among community-dwelling older adults. *Behavior Modification, 32*, 3–20.
39. Radloff, L. S. (1977). The ces-d scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385–401.
40. Van Dam, N. T., & Earleywine, M. (2011). Validation of the center for epidemiologic studies depression scale–revised (cesd-r): Pragmatic depression assessment in the general population. *Psychiatry Research, 186*, 128–132.
41. Worm-Smeitink, M., Gielissen, M., Bloot, L., van Laarhoven, H. W. M., van Engelen, B. G. M., van Riel, P., et al. (2017). The assessment of fatigue: Psychometric qualities and norms for the checklist individual strength. *Journal of Psychosomatic Research, 98*, 40–46.

42. Beurskens, A., Bultmann, U., Kant, I., Vercoulen, J., Bleijenberg, G., & Swaen, G. (2000). Fatigue among working people: Validity of a questionnaire measure. *Occupational and Environmental Medicine, 57*, 353–357.
43. Lerdal, A., & Kottorp, A. (2011). Psychometric properties of the fatigue severity scale-rasch analyses of individual responses in a norwegian stroke cohort. *International Journal of Nursing Studies, 48*, 1258–1265.
44. Nadarajah, M., Mazlan, M., Abdul-Latif, L., & Goh, H. T. (2016) Test-retest reliability, internal consistency and concurrent validity of fatigue severity scale in measuring post-stroke fatigue. *European Journal of Physical and Rehabilitation Medicine, 53*(5), 703–709.
45. Yesavage, J. A., & Sheikh, J. I. (2008). 9/geriatric depression scale (gds). *Clinical Gerontologist, 5*, 165–173.
46. Brown, L. M., & Schinka, J. A. (2005). Development and initial validation of a 15-item informant version of the geriatric depression scale. *International Journal of Geriatric Psychiatry, 20*, 911–918.
47. Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the hospital anxiety and depression scale. An updated literature review. *Journal of Psychosomatic Research, 52*, 69–77.
48. Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*, 361–370.
49. Johnston, M., Pollard, B., & Hennessey, P. (2000). Construct validation of the hospital anxiety and depression scale with clinical populations. *Journal of Psychosomatic Research, 48*, 579–584.
50. Snaith, R. P., & Taylor, C. M. (1985). Rating scales for depression and anxiety: A current perspective. *British Journal of Clinical Pharmacology, 19*(Suppl 1), 17 s–20 s.
51. Winkens, I., Van Heugten, C. M., Fasotti, L., & Wade, D. T. (2009). Reliability and validity of two new instruments for measuring aspects of mental slowness in the daily lives of stroke patients. *Neuropsychological Rehabilitation, 19*, 64–85.
52. Wikberg, C., Nejati, S., Larsson, M. E., Petersson, E. L., Westman, J., Ariai, N., et al. Comparison between the montgomery-asberg depression rating scale-self and the beck depression inventory ii in primary care. *The Primary Care Companion for CNS Disorders*. 2015;17.
53. McNair, D. M., Lorr, M., & Droppleman, L. F. (1971). *Manual for the profile of mood states*. San Diego: Educational and Industrial Testing Services.
54. Lievaart, M., Franken, I. H., & Hovens, J. E. (2016). Anger assessment in clinical and nonclinical populations: Further validation of the state-trait anger expression inventory-2. *Journal of Clinical Psychology, 72*, 263–278.
55. Mahoney, J., Drinka, T. J., Abler, R., Gunter-Hunt, G., Matthews, C., Gravenstein, S., et al. (1994). Screening for depression: Single question versus GDS. *Journal of the American Geriatrics Society, 42*, 1006–1008.
56. Calvert, M., Blazeby, J., Altman, D. G., Revicki, D. A., Moher, D., & Brundage, M. D., CONSORT PRO Group. (2013). Reporting of patient-reported outcomes in randomized trials: The CONSORT PRO extension. *JAMA, 309*(8), 814–822.
57. Vodicka, E., Kim, K., Devine, E. B., Gnanasakthy, A., Scoggins, J. F., & Patrick, D. L. (2015). Inclusion of patient-reported outcome measures in registered clinical trials: Evidence from ClinicalTrials.gov (2007–2013). *Contemporary Clinical Trials, 43*, 1–9.
58. Mercieca-Bebber, R., Rouette, J., Calvert, M., King, M. T., McLeod, L., Holch, P., Palmer, M. J., & Brundage, M. (2017). International Society for Quality of Life Research (ISOQOL) best practice for PROs—reporting taskforce. Preliminary evidence on the uptake, use and benefits of the CONSORT-PRO extension. *Quality of Life Research, 26*(6), 1427–1437.
59. Calvert, M., Kyte, D., Mercieca-Bebber, R., Slade, A., Chan, A. W., King, M. T., et al. (2018). Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: The SPIRIT-PRO extension. *JAMA, 319*(5), 483–494.
60. Kane, R. L., & Radosevich, D. M. (LLC 2011). *Conducting health outcomes research*. Sudbury: Johns & Bartlett Learning.
61. Cella, D., Nowinski, C., Peterman, A., Victorson, D., Miller, D., Lai, J.-S., et al. (2011). The neurology quality of life measurement initiative. *Archives of Physical Medicine and Rehabilitation, 92*(10 Suppl), S28–S36.
62. Gershon, R. C., Wagster, M. V., Hendrie, H. C., Fox, N. A., Cook, K. F., & Nowinski, C. J. (2013). NIH toolbox for assessment of neurological and behavioral function. *Neurology, 80*(11), S2–S6.
63. Reeve, B. B., Wyrwich, K. W., Wu, A. W., Velikova, G., Terwee, C. B., Snyder, C. F., et al. (2013). ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Quality of Life Research, 22*(8), 1889–1905.