University of Reading

A core outcome set for aphasia treatment research: the ROMA consensus statement

Article

Accepted Version

Wallace, Sarah J, Worrall, Linda, Rose, Tanya, Le Dorze, Guylaine, Breitenstein, Caterina, Hilari, Katerina, Babbitt, Edna, Bose, Arpita, Brady, Marian, Cherney, Leora R, Copland, David, Cruice, Madeline, Enderby, Pam, Hersh, Deborah, Howe, Tami, Kelly, Helen, Kiran, Swathi, Laska, Ann-Charlotte, Marshall, Jane, Nicholas, Marjorie, Patterson, Janet, Pearl, Gill, Rochon, Elizabeth, Rose, Miranda, Sage, Karen, Small, Steve and Webster, Janet (2019) A core outcome set for aphasia treatment research: the ROMA consensus statement. International Journal of Stroke, 14 (2). pp. 180-185. ISSN 1747-4930 doi: https://doi.org/10.1177/1747493018806200 Available at https://centaur.reading.ac.uk/80729/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1177/1747493018806200

Publisher: SAGE Publications

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other

University of Reading

copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Accepted in Journal of Stroke

Guidelines

International Journal of Stroke WSO

International Journal of Stroke
0(0) 1–6
© 2018 World Stroke Organization
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1747493018806200
journals.sagepub.com/home/wso
SAGE

A core outcome set for aphasia treatment research: The ROMA consensus statement

Sarah J Wallace[1], Linda Worrall[1], Tanya Rose[1], Guylaine Le Dorze[2], Caterina Breitenstein[3], Katerina Hilari[4], Edna Babbitt[5], Arpita Bose[6], Marian Brady[7], Leora R. Cherney[5], David Copland[1], Madeline Cruice[4], Pam Enderby[8], Deborah Hersh[9], Tami Howe[10], Helen Kelly[11], Swathi Kiran[12], Ann-Charlotte Laska[13], Jane Marshall[4], Marjorie Nicholas[14], Janet Patterson[15], Gill Pearl[16], Elizabeth Rochon[17], Miranda Rose[18], Karen Sage[1], Steven Small[20] and Janet Webster[21]

Abstract

Background: A core outcome set (COS; an agreed, minimum set of outcomes) was needed to address the heterogeneous measurement of outcomes in aphasia treatment research and to facilitate the production of transparent, meaningful and efficient outcome data.

Objective: The Research Outcome Measurement in Aphasia (ROMA) consensus statement provides evidence-based recommendations for the measurement of outcomes for adults with post-stroke aphasia within phase I-IV aphasia treatment studies**.**

Methods: This statement was informed by a four-year program of research which comprised investigation of stakeholder-important outcomes using consensus processes, a scoping review of aphasia outcome measurement instruments, and an international consensus meeting. This paper provides an overview of this process and presents the results and recommendations arising from the international consensus meeting.

Results: Five essential outcome constructs were identified: Language, communication, patient-reported satisfaction with treatment and impact of treatment, emotional wellbeing, and quality of life. Consensus was reached for the following measurement instruments: Language: The Western Aphasia Battery Revised (WAB-R) (74% consensus); emotional well-being: General Health Questionnaire (GHQ)-12 (83% consensus); quality of life: Stroke and Aphasia Quality of Life Scale (SAQOL-39) (96% consensus). Consensus was unable to be reached for measures of communication (where multiple measures exist) or patient-reported satisfaction with treatment or impact of treatment (where no measures exist).

Discussion: Harmonisation of the ROMA COS with other core outcome initiatives in stroke rehabilitation is discussed. Ongoing research and consensus processes are outlined.

Conclusion: The WAB-R, GHQ-12, and SAQOL-39 are recommended to be routinely included within phase I-IV aphasia treatment studies. This consensus statement has been endorsed by the Collaboration of Aphasia Trialists, the British Aphasiology Society, the

German Society for Aphasia Research and Therapy, and the Royal College of Speech Language Therapists.

A core outcome set for aphasia treatment research: the ROMA consensus statement

The Research Outcome Measurement in Aphasia (ROMA) consensus statement provides recommendations for a core outcome set (COS) for use in aphasia treatment studies. A COS is a minimum set of outcomes that should be measured and reported in research trials of a specific health condition or population (1). The use of a COS does not preclude the measurement of additional outcomes, but rather represents the minimum outcomes that should be collected and reported (2). A COS for aphasia was developed in response to a trend of heterogeneous outcome measurement in research and the merits of this initiative were debated in a published forum in 2014 (3-7). The ROMA consensus statement was informed by a four-year program of research in three phases: (1) investigation of stakeholder-important outcomes using consensus processes (8-11); (2) a scoping review to identify aphasia outcome measurement instruments (OMIs) and their psychometric properties (12); and (3) an international consensus meeting (results reported herein). The ROMA COS is intended to complement other existing and ongoing initiatives to standardise the measurement of stroke recovery (13-15).

Objective

The ROMA consensus statement provides evidence-based recommendations for the measurement of outcomes for adults with post-stroke aphasia within phase I-IV aphasia treatment studies.

Target users

The primary users of this consensus statement will be researchers involved in the design and conduct of aphasia treatment studies.

Methods

The research methods are based on the recommendations of the Core Outcome Measures in Effectiveness Trials (COMET) Initiative (2, 16) and are reported in alignment with the COS-STAR (Core Outcome Set-STAndards for Reporting) statement (17). The World Health Organization International Classification of Functioning, Disability and Health (ICF) (18) has been used as a conceptual framework and classification tool. This project is registered with the COMET Initiative (http://www.comet-initiative.org/studies/details/287).

Stage 1: Identification of Core Outcome Constructs

Outcome constructs were derived from three separate stakeholder consensus studies conducted with: people with aphasia and their families (9); aphasia clinicians and managers (8); and aphasia researchers (10). Outcomes prioritised by stakeholder groups were integrated using the framework of the ICF (19). Essential constructs were identified as: Language, communication, patient-reported satisfaction with treatment and impact of treatment, emotional wellbeing, and quality of life (11).

Stage 2: Identification of Outcome Measurement Instruments

A scoping review was conducted to identify OMIs which have been validated with people with aphasia. Primary searches were run using PUBMED, EMBASE, and CINAHL databases on 10 November 2015. The search strategy incorporated filters developed for the identification of studies reporting the measurement properties of health OMIs (see 20 and supplementary file). Inclusion criteria required that studies focused on the psychometric properties of measurement instrument and included participants with aphasia or stroke patients where participants with aphasia were not specifically excluded. Studies reporting measurement instruments which primarily measure neurological function associated with, but not central to aphasia: e.g., consciousness; health; motor speech; cognition; memory; were excluded. Secondary searches were conducted for each OMI identified in the first search. In

total, 184 references for 79 measurement instruments were identified (12). No measures of patient-reported treatment impact or patient-reported satisfaction were identified through this search.

Stage 3. Formation of Consensus Panel

Researchers who participated in the first phase of this project (n=80) (10) were invited to participate in the final consensus meeting. These researchers were purposively sampled from researchers whose trials were included with the Cochrane Collaboration review of "Speech and language therapy for aphasia following stroke"(21) and the 100 most highly published aphasia treatment researchers in the Web of Science database. In total, 23 researchers participated in a consensus meeting in London, UK (December, 2016). Panel members were experienced researchers with expertise in: the design and conduct of aphasia trials; measurement instrument development and testing; and clinical guidelines development (see table 1 and supplementary table 1). Authors Wallace, Worrall, Le Dorze and T. Rose facilitated the COS development process and did not participate in COS voting.

Table 1

Characteristics of researchers who participated in the international consensus panel (n=23)

Panel Characteristics	n (%)
Country	
United Kingdom	9 (39)
United States of America	6 (26)
Australia	3 (13)
Canada	2 (9)
Germany	1 (4)
Sweden	1 (4)
Ireland	1 (4)

ICF component to which their own research relates (panel members could nominate more than one component)	
Body functions	16
Activity/Participation	21
Environmental factors	10
Personal factors	15
Quality of life*	12
Number of treatment studies published by participants	
1	2
2-5	8
6-10	4
more than 10	7
not specified	2

*****nb. Quality of life is not defined as a component of the ICF

Stage 4. International Consensus Meeting

Ethical approval for the consensus meeting was gained from the Behavioural and Social Sciences Ethical Review Committee at The University of Queensland, Australia. The following process was used:

Prior to meeting

(1) Panel members generated consensus-based criteria to enable an initial reduction of OMIs (see table 2).

(2) The consensus-based criteria were applied to the list of OMIs identified in the stage 2 scoping review (n=79) to produce a short-list (n=50) (see supplementary table 2).

(3) Panel members generated consensus-based feasibility criteria (see table 3).

(4) The short-listed OMIs (see supplementary table 2) were assigned to panel members, who reviewed OMI feasibility and measurement properties prior to the consensus meeting.

During the meeting

(1) Panel members engaged in a whole-group discussion using an iterative process to apply feasibility criteria and eliminate OMIs.

(2) Panel members divided to smaller groups to review the measurement properties for each OMI in the target population (people with aphasia). Properties considered included: acceptability/feasibility of use with people with aphasia, reliability (test-retest, inter- and intra- as applicable), construct validity, and sensitivity to change.

(3) Each small group recommended two OMIs for voting. Panel members voted YES/NO for each OMI in a closed voting process with consensus defined a priori as agreement on each OMI for each outcome construct by ≥ 70% of meeting participants, as suggested by the COMET initiative and GRADE working group (2). Potential conflicts of interest were managed through agreement that authors of OMIs under consideration could not participate in voting for that construct area.

Table 2

Criteria for initial reduction of outcome measurement instruments

Measures were excluded if:
1. The purpose of the measurement instrument was to screen for the presence of aphasia, rather than to measure outcomes.
2. The measurement instrument was published more than thirty years ago (i.e., prior to 1986) without subsequent revision and/or was not in current use.
3. The measurement instrument targeted only one severity level of aphasia.
4. For measures of language: the measurement instrument did not assess all modalities of language (e.g. reading only, writing only, comprehension only, verbal output only).

Table 3

Feasibility criteria

1. Availability in different languages or ease of translation/adaptation.
2. Cost.
3. Burden to respondents or researchers (ease of administration, length of outcome measurement instrument, completion time).
4. Ease of score calculation and provision of an aggregate score.

Results

After compilation of votes, panel members reached consensus for measures of language, emotional wellbeing, and quality of life (refer to table 4). A consensus of ≥ 70% was not reached for a measure of communication. Inability to gain consensus on a measure of communication may relate to the multi-factorial nature of this construct, as well a lack of understanding and consensus around how 'effective communication' is best operationalised in treatment research.

Table 4

Results of final voting to decide core outcome measurement instruments

Construct	Measure*	Votes for inclusion
Language	**The Western Aphasia Battery Revised (WAB-R)**	**74% (n=17)**
	The Comprehensive Aphasia Test (CAT)	22% (n=5)
	Neither	4% (n=1)
Communication	The Scenario Test	57% (n=13)
	The Communication Effectiveness Index (CETI)	39% (n=9)
	Abstained	4% (n=1)
Emotional well-being	**General Health Questionnaire (GHQ)-12**	**83% (n=19)**

	Stroke Aphasic Depression Questionnaire (SADQ)	17% (n=4)
Quality of life	**Stroke and Aphasia Quality of Life Scale (SAQOL-39)**	**96% (n=22)**
	Burden of Stroke Scale (BOSS)	0% (n=0)
	Abstained	4% (n=1)

Bolded figures indicate consensus criteria (≥70%) reached and OMI included in COS
*Refer to supplementary tables 3 & 4 for OMI characteristics, properties and references.

Recommendations

It is recommended that the WAB-R, GHQ-12 and SAQOL-39 be included as core outcome measurement instruments in phase I-IV aphasia treatment studies for adults with post-stroke aphasia. These outcome measurement instruments and their psychometric properties are described in supplementary tables 3 & 4.

Discussion

The importance of implementing standardised approaches to outcome measurement in research trials is increasing acknowledged. In the field of stroke rehabilitation, the Stroke Recovery and Rehabilitation Roundtable (SRRR) (13) have provided consensus-based core recommendations for the measurement of sensorimotor recovery after stroke. Other initiatives have addressed the measurement of stroke outcomes in clinical practice (15) and there are ongoing works to standardise measures in arm rehabilitation trials after stroke (14). The ROMA COS has sought to provide recommendations specifically for the measurement of aphasia recovery post-stroke. Accordingly, some frequently used measures of global disability and health-related quality of life (e.g., EQ-5D) which do not contain communication-specific items or which have not been validated with stroke survivors with aphasia were not considered within this process. The ROMA COS seeks to harmonise with other existing stroke rehabilitation initiatives in addressing the need for standardised

approaches to research trial outcomes measurement and its supplementary use may therefore be considered in any stroke study where people with aphasia are included.

Future Directions

The ROMA COS will be reviewed biennially. The next consensus meeting will focus on measures of communication and consider the development of measures of patient-reported satisfaction with treatment / impact of treatment. Factors relating to international COS implementation will be considered. New publications, initiatives and user feedback will also be considered in each review to: align this COS with other COSs; consider new OMIs; and to review the choice of OMIs based on user feedback.

Limitations

Participants in the international consensus meeting were predominately from English speaking countries. This may have impacted the consensus process and findings. Future meetings will seek to increase the diversity of participants with respect to cultural and linguistic background.

Funding

This work was supported by a British Aphasiology Society Initiatives in Aphasia Seed Fund grant. The Collaboration of Aphasia Trialists provided operational support for the international consensus meeting.

Conflicts of Interest

Authors Babbit, Breitenstein, Cherney, Cruice, Enderby, and Hilari authored or adapted OMIs considered in this consensus process. These authors declared their conflict of interest during the meeting and did not participate in voting which related to their authored OMIs. Authors Wallace, Worrall, Le Dorze and T. Rose did not participate in voting on OMIs.

References

1. Prinsen CAC, Vohra S, Rose MR, King-Jones S, Ishaque S, Bhaloo Z. Core Outcome Measures in Effectiveness Trials (COMET) Initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a 'core outcome set'. Trials. 2014;15.
2. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E. Developing core outcome sets for clinical trials: Issues to consider. Trials. 2012;13.
3. Wallace SJ, Worrall L, Rose T, Le Dorze G. Measuring outcomes in aphasia research: A review of current practice and an agenda for standardisation. Aphasiology. 2014;28(11):1364-84.
4. Wallace SJ, Worrall L, Rose T, Le Dorze G. A good outcome for aphasia. Aphasiology. 2014;28(11):1400-4.
5. Brady MC, Ali M, Fyndanis C, Kambanaros M, Grohmann KK, Laska A-C, et al. Time for a step change? Improving the efficiency, relevance, reliability, validity and transparency of aphasia rehabilitation research through core outcome measures, a common data set and improved reporting criteria. Aphasiology. 2014:1-8.
6. Hula WD, Fergadiotis G, Doyle PJ. A core outcome set for aphasia treatment research: Obstacles, risks, and benefits. Aphasiology. 2014;28(11):1396-9.
7. MacWhinney B. Challenges facing COS development for aphasia. Aphasiology. 2014:1-3.
8. Wallace SJ, Worrall L, Rose T, Le Dorze G. Which treatment outcomes are most important to aphasia clinicians and managers? An international e-Delphi consensus study. Aphasiology. 2017;31(6):643-73.

9. Wallace SJ, Worrall L, Rose T, Le Dorze G, Cruice M, Isaksen J, et al. Which outcomes are most important to people with aphasia and their families? An international nominal group technique study framed within the ICF. Disability and rehabilitation. 2017;39(14):1364-79.
10. Wallace SJ, Worrall L, Rose T, Le Dorze G. Core Outcomes in Aphasia Treatment Research: An e-Delphi Consensus Study of International Aphasia Researchers. American Journal of Speech-Language Pathology. 2016;25(4S):S729-S42.
11. Wallace SJ, Worrall L, Rose T, Le Dorze G. Using the International Classification of Functioning, Disability, and Health to identify outcome domains for a core outcome set for aphasia: a comparison of stakeholder perspectives. Disability and rehabilitation. 2017:1-10.
12. Wallace SJ, Worrall L, Rose T, Le Dorze G, Brandenburg C. A scoping review of standardised aphasia assessments. In preparation.
13. Kwakkel G, Lannin NA, Borschmann K, English C, Ali M, Churilov L, et al. Standardized measurement of sensorimotor recovery in stroke trials: Consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. International Journal of Stroke. 2017;12(5):451-61.
14. Duncan Millar J, Ali M, van Wijck F, Pollock A. Outcome measurement in stroke upper limb rehabilitation trials: What tools are used and how often?2016.
15. Salinas J, Sprinkhuizen SM, Ackerson T, Bernhardt J, Davie C, George MG, et al. An International Standard Set of Patient-Centered Outcome Measures After Stroke. Stroke; a journal of cerebral circulation. 2016;47(1):180-6.
16. Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a “Core Outcome Set” – a practical guideline. Trials. 2016;17(1):1-10.

17. Kirkham JJ, Gorst S, Altman DG, Blazeby J, Clarke M, Devane D. COS-STAR a reporting guideline for studies developing core outcome sets (protocol). Trials. 2015;16.
18. World Health Organization. World Health Organization: International Classification of Functioning, Disability and Health (ICF). Geneva, Switzerland: World Health Organization; 2001.
19. Wallace SJ, Worrall L, Rose T, Le Dorze G. Using the International Classification of Functioning, Disability and Health to identify outcome domains for a core outcome set for aphasia: A comparison of stakeholder perspectives. Disability & Rehabilitation. In press.
20. Terwee CB, Jansma EP, Riphagen I, Vet HW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Quality of Life Research. 2009;18(8):1115-23.
21. Brady MC, Kelly H, Godwin J, Enderby P. Speech and language therapy for aphasia following stroke. Cochrane Database of Systematic Reviews [Internet]. 2012; (5). Available from: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD000425.pub3/abstract.

Supplementary Table 1

ROMA consensus meeting facilitators

Sarah J. Wallace PhD BSpPath(Hons) GradCert Gerontology CPSP Certified Practising Speech Pathologist and Teaching and Research Academic, School of Health and Rehabilitation Sciences, The University of Queensland. *Expertise: post-stroke aphasia rehabilitation, core outcome set development, stakeholder perspectives, consensus processes, ICF.*	**Linda Worrall PhD BSpThy FSPA** Speech Pathologist, Teaching and Research Academic, School of Health and Rehabilitation Sciences, The University of Queensland, Australia. *Expertise: post-stroke aphasia rehabilitation, ICF, aphasia trial design and conduct, consumer perspective, aphasia rehabilitation guideline development.*	**Guylaine Le Dorze Ph.D MSc (A)** Teaching and Research Academic, Speech-Language Pathologist, School of Speech-Language Pathology and Audiology, Faculty of Medicine, Université de Montréal. Expertise: *post-stroke aphasia rehabilitation, participation, single-subject designs, qualitative methods.*	**Tanya Rose PhD BSpPath(Hons) GradCert Higher Ed CPSP** Certified Practising Speech Pathologist and Teaching and Research Academic, School of Health and Rehabilitation Sciences, The University of Queensland. *Expertise: Post-stroke aphasia rehabilitation, paediatric and adult language, accessible health information, mixed-methods research.*

ROMA consensus panel

Edna Babbitt PhD CCC-SLP BC-ANCDS Research Speech-Language Pathologist Assistant Research Professor, Department of Physical Medicine and Rehabilitation, Feinberg School of Medicine, Northwestern University, Chicago, USA & Shirley Ryan AbilityLab, Chicago, USA. *Expertise: Post-stroke aphasia assessment and rehabilitation.*	**Arpita Bose PhD MSc (Speech and Hearing) BSc (Audiology and Speech Rehabilitation).** Speech and Language Therapist, Teaching and Research Academic, School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK. *Expertise: Post-stroke aphasia assessment and rehabilitation, bilingualism, single subject experimental designs, quality of life issues in aphasia, SLT training in decision-making in aphasia.*	**Marian Brady PhD BSc** Speech and language therapist, Director Stroke Rehabilitation Research, NMAHP Research Unit, Glasgow Caledonian University, Glasgow, Scotland. *Expertise: Stroke rehabilitation, design, development and evaluation of complex multidisciplinary interventions, survey, mixed methods, systematic review, meta-analyses and the use of randomised controlled trial archives.*	**Caterina Breitenstein PhD academic degrees in Clinical Psychology and Cognitive Neuroscience.** Teaching and Research Academic, Dept. of Neurology, University of Muenster, Germany. *Expertise: Development and national adaptations of communication outcome measures, clinical trials methodology, intervention studies in post-stroke aphasia rehabilitation.*
Leora R. Cherney PhD CCC-SLP BC-ANCDS. Research Scientist and Speech and Language Pathologist. Shirley Ryan AbilityLab (formerly the Rehabilitation Institute of Chicago) and Northwestern University, Chicago, IL USA. *Expertise: Post-stroke aphasia*	**David Copland PhD BSpPath (Hons)** Speech Pathologist, Principal Research Fellow, School of Health & Rehabilitation Sciences and Centre for Clinical Research, The University of Queensland, Brisbane, Australia. *Expertise: Post-stroke aphasia assessment and rehabilitation, aphasia*	**Madeline Cruice PhD BSpPath (Hons)** Registered Speech and Language Therapist, Reader, Teaching and Research Academic, School of Health Sciences, City University of London, London, UK. *Expertise: Post-stroke aphasia rehabilitation, therapeutic process and evaluation, quality of*	**Pam Enderby PhD MBE DSc (Hons) MSc FRCSLT** Speech and Language Therapist, Professor Emeritus of Community Rehabilitation, University of Sheffield, Sheffield. UK. *Expertise: Aphasia management, Clinical Evaluation of Interventions, RCTs, Psychometric Properties of Outcome Measures.*

assessment and rehabilitation, development and evaluation of novel aphasia treatments, single subject and RCT design, systematic reviews.	*trial design and conduct, neuroimaging in aphasia.*	*life evaluation in research and clinical practice, behaviour change.*	
Deborah Hersh PhD MSc BSc(Hons) GradCert Higher Ed FSPA. Speech Pathologist, Teaching and Research Academic, School of Medical and Health Sciences, Edith Cowan University, Perth, Australia. *Expertise: Post-stroke aphasia rehabilitation, consumer perspective, aphasia rehabilitation guideline development.*	**Katerina Hilari PhD MRCSLT MHPC** Psychologist, Registered Speech and Language Therapist, Teaching and Research Academic, School of Health Sciences, City, University of London, UK. *Expertise: Outcome measurement development, validation and cultural adaptation, post-stroke aphasia rehabilitation, feasibility RCTs, clinical guideline development.*	**Tami Howe PhD MHSc BEd SLP(C)** Speech Pathologist and Teaching and Research Academic, School of Audiology and Speech Sciences, University of British Columbia, Vancouver, Canada. *Expertise: Aphasia rehabilitation, ICF, accessibility, goal setting, social participation, impact of aphasia on family members.*	**Helen Kelly PhD MRCSLT** Registered Speech and Language Therapist, Teaching and Research Academic, Department of Speech and Hearing Sciences, University College Cork, Cork, Ireland. *Expertise: Post-stroke aphasia assessment and management, single subject and RCT feasibility aphasia trial design and conduct, consumer perspective.*
Swathi Kiran PhD CCC-SLP Speech Language pathologist, Teaching and Research Academic. Professor, Associate Dean for Research Sargent College of Health and Rehabilitation Sciences, Boston University, Boston, MA, USA. *Expertise: Aphasia rehabilitation, neuroimaging, bilingualism, single subject experimental design.*	**Ann-Charlotte Laska MD A/Professor** Department of Clinical Science Karolinska Institutet Danderyd Hospital, Sweden *Expertise: Post-stroke aphasia, study design and conduct, RCT.*	**Jane Marshall PhD Post Grad Diploma in Clinical Communication Studies BA FRCSLT** Registered Speech and Language Therapist, Teaching and Research Academic, School of Health Sciences, City, University of London, UK. *Expertise: Post-stroke aphasia rehabilitation, the development and evaluation of novel treatments.*	**Marjorie Nicholas PhD CCC-SLP** Professor and Interim Chair Dept. of Communication Sciences and Disorders, MGH Institute of Health Professions, Boston, MA, USA. *Expertise: Aphasia rehabilitation, nonverbal cognition in aphasia, Life Participation Approach to Aphasia and community aphasia program design, ICAP design.*
Janet Patterson PhD CCC-SLP ASHA Fellow Chief, Audiology & Speech-Language Pathology Service, VA Northern California Health Care System Practicing Speech-Language Pathologist, Teaching and Research Academic. *Expertise: Post-stroke*	**Gill Pearl MPhil Dip Hum Commun.** Certified practicing speech and language therapist in role as Chief Executive Officer of Speakeasy - specialist aphasia centre, UK. *Expertise: Development and evaluation of novel approaches to providing long term aphasia support and therapy,*	**Elizabeth Rochon PhD MSc (A) Reg CASLPO SLP(c)** Speech Pathologist, Teaching and Research Academic, Department of Speech-Language Pathology and Rehabilitation Sciences Institute, University of Toronto, Canada. *Expertise: Post-stroke aphasia assessment and*	**Miranda Rose PhD BSpPath FSPA** Speech pathologist, Teaching and Research Academic, School of Allied Health, La Trobe University, Victoria, Australia. *Expertise: Post-stroke aphasia rehabilitation, aphasia trial design and conduct, single subject designs, consumer*

aphasia rehabilitation, systematic reviews of literature, single subject designs.	*facilitator of consumer involvement in research, feasibility studies, case series studies, RCT design and conduct.*	*rehabilitation, development of aphasia treatment studies, feasibility studies, single subject and RCT design, systematic reviews.*	*perspective, aphasia rehabilitation guideline development.*
Karen Sage PhD Dip DisHumComm BA (Hons) HCPC Registered Speech and Language Therapist, MRCSLT; Teaching and Research Academic, Department of Allied Health Professions, Sheffield Hallam University, Sheffield, UK. *Expertise: Aphasia assessment and management, stroke rehabilitation, single case, case series, mixed methods.*	**Steven L. Small PhD MD** Professor of Neurology, University of California, Irvine *Expertise: Neurobiology of Language, Cognitive Neurology.*	**Janet Webster PhD MRCSLT** Registered Speech and Language Therapist, Teaching and Research Academic, Newcastle University, UK *Expertise: Post-stroke aphasia assessment and management, single subject design.*	

Supplementary Table 2

OMIs (n=50) identified in scoping review and retained following application of the consensus-based criteria

Construct	Outcome measurement instrument
Language	• The Comprehensive Aphasia Test (CAT) (1) • The Western Aphasia Battery Revised (WAB-R) (AQ+LQ) (2) • Therapy Outcome Measures (TOM) (3-5) • The Aphasia Checklist (ACL) (6) • Aachen Aphasia Test (AAT) (7) • Aphasia Language Assessment Test (ALA) (8) • The Thai Aphasia Language Performance Scales (ALPS) (9) • Bilingual Aphasia Test (BAT) (10) • The Boston Diagnostic Aphasia Examination (BDAE) (11) • Ege Aphasia Test (12) • Kentucky Aphasia Test (KAT) (13) • Montreal-Toulouse Language Assessment Battery (MTL) (14) • The Norsk Grunntest for Afasi (NGTA) (15)
Emotional well-being	• Communication Confidence Rating Scale for Aphasia (CCRSA) (16) • Hospital Anxiety and Depression Scale (HADS) (17) • Montgomery-Asberg Depression Rating Scale (MADRS) (18) • Geriatric Depression Scale (GDS) 15 item / 30 item (19, 20) • Warwick and Edinburgh mental well-being scale (21) • Geriatric anxiety scale (22) • Stroke and Aphasia (SAD) Scale (23) • Signs of Depression Scale (SODS) (24) • Stroke Aphasic Depression Questionnaire (SADQ) (25) • Visual Analogue Self-Esteem Scale (VASES) (26) • Centre for Epidemiology Depression Scale –Revised (27) • General Health Questionnaire (GHQ) 12 item (28) • Therapy Outcome Measures (TOM) (29-31) • Patient Health Questionnaire 2 item / 9 item (32, 33) • Visual Analogue Mood Scale (VAMS) (34)

Communication	• Aphasia Communication Outcome Measure (ACOM) (35) • American Speech-Language and Hearing Association Functional Assessment of Communication Skills for Adults (ASHA-FACS) (36) • Amsterdam-Nijmegen Everyday Language Test (ANELT) (37) • The Communication Activity Log (CAL) (38) • The Communication Outcome After Stroke (COAST) (39) • The Communicative Activities Checklist (COMACT) (40) • The Social Activities Checklist (SOCACT) (40) • The Communication Disability Profile (CDP) (41) • The Communication Effectiveness Index (CETI) (42) • Community Integration Questionnaire (CIQ-R) (43) • Communication Activities of Daily Living (CADL) (44) • The Functional Outcome Questionnaire for Aphasia (FOQ-A) (45) • Measure of participation in conversation (MPC) (46) • The Scenario Test (47) • The Speech Questionnaire (48) • Therapy Outcome Measures (TOM) (29-31) • The Communication Participation Item Bank (49)
Quality of Life	• Aachen Life Quality Inventory (ALQI) (50) • Burden of Stroke Scale (BOSS) (51) • The Newcastle Stroke-Specific Quality of Life Measure (NEWSQOL) (52) • Short Form 36 Health Survey (SF-36) (53) • Stroke and Aphasia Quality of Life Scale (SAQOL-39) (54, 55)

Supplementary Table 3

Description of recommended outcome measurement instruments

Outcome instrument and abbreviation	Development / alternate versions	Aims/instrument description	Number of items	Duration	Scoring system	Training	Cost*/ availability	Language translations
Western Aphasia Battery Revised (WAB-R) (2)	Developed by Kertesz in 1979 based on the original format of the Boston Diagnostic Aphasia Examination (56). Revisions published in 1982 and 2006 (WAB-R): Supplemental tasks, revision of 15 items and testing materials (e.g. spiral-bound stimulus book replacing loose stimulus cards), as well as revised directions and scoring guidelines for clarity. The WAB-R also includes a bedside screening tool (Bedside WAB-R).	**Primary:** Assessment of linguistic skills in aphasia: 1. Spontaneous speech 2. Auditory verbal comprehension 3. Repetition 4. Naming and word finding 5. Reading 6. Writing 7. Apraxia 8. Constructional, visuospatial, and calculation tasks 9. Supplemental writing and reading tasks: reading and writing of irregular and non-words (WAB-R only) **Secondary:** Assessment of non-linguistic skills in aphasia: drawing, block design, calculation, and praxis 1. Additional aims: Classification of 8 aphasia types: Global, Broca's, Transcortical motor, Wernicke's,	>300	• Bedside WAB-R: 15 min (comprises half of the items of WAB-R Part 1) • Part 1: 30-45 min • Part 2: 45-60 min	• Aphasia Quotient (AQ): a weighted average of the WAB spoken language subtest scores. • Cortical Quotient (CQ): a weighted average of both the language and non-language subtest scores. • The Language Quotient (LQ): reflects auditory comprehension, oral expression, reading, and writing performance.	Administration: "some training" required according to developers. Scoring procedures require training.	Testing materials: +++ Available from: https://www.pearsonclinical.com	Cantonese (57) Korean (58) Bangla (59) Tagalog (60) Brazilian Portuguese (61) Japanese (62) Hungarian French Turkish (63) Hebrew Spanish (64)

		Transcortical sensory, Mixed transcortical, Conduction, and Anomic 2. Assessment of aphasia severity 3. Used to determine the location of the lesion						
Stroke and Aphasia Quality of Life Scale (SAQOL-39; SAQOL-39g) (54, 55)	The SAQOL-39 is the short form of the SAQOL (53 items), which is itself an adaptation of the SS-QOL (Stroke-specific Quality of life scale). The SAQOL-39 was originally tested in people with chronic aphasia (the measure had four domains: physical, psychosocial communication, energy.	Interview-administered self-report measure, SAQOL-39 comprises 39 questions, in four quality of life (QoL) domains: 1. Physical (17 items) 2. Communication (7 items) 3. Psychosocial (11 items) 4. Energy (4 items) SAQOL 39g comprises the same 39 questions, in three quality of life (QoL) domains: 1. Physical (16 items) 2. Communication (7 items) 3. Psychosocial (16 items) Timeframe for all questions is the past week	39	• 15-20 min (depending on severity of aphasia)	• Twenty-one of the items ask the respondents how much trouble they have had with activities (e.g., getting dressed, speaking). The response format for these questions is a 5-point scale that varies from 1='couldn't do it at all' to 5='no trouble at all'. The rest of the items (18) ask about feelings (e.g., 'did you feel irritable?') and other activities (e.g., 'did you see your friends less often than you would like?'). Their response format	Administration: Guidance is provided in administration guidelines. Administrators need to have skills in communicating with people with aphasia Scoring procedures: no training required	Free. Available from: https://blogs.city.ac.uk/cityaccess/saqol-description/	Chilean (68) Chinese (69) Chinese mandarin (70) Dutch (71) Greek (72, 73) Hindi (74) Italian (75) (76) Japanese (77) Kannada (78) Korean (79) Malayalam (80) Persian (81) Portuguese (82) Spanish (83) Turkish (84)

	Testing the SAQOL-39 in generic stroke population (n=87) resulted in the SAQOL-39g, which has the same items as the SAQOL-39 but three domains (all energy items groups with the psychosocial domain). There are alternative forms for proxy administration (65, 66) and for postal and telephone administration (67)	Multi-modal presentation, i.e., patients can both read and listen to the questions. People with expressive aphasia can point to their responses instead of verbally responding.			varies from 1='definitely yes' to 5='definitely no'. Calculation of: 1. total score: mean score of all 39 items 2. Domain scores: mean score of all items relating to the respective domain			
General Health Questionnaire (GHQ) 12	Developed in 1972. Current version published in 2011) Alternate versions: • GHQ-60: 60-item questionnaire • GHQ-30: a short form without items relating to	Primary: Screening device for identifying minor psychiatric disorders in the general population and within community or non-psychiatric clinical settings such as primary care or general medical out-patients. 12 questions relating to symptoms of various psychiatric conditions, assesses the respondent's	12	2 min administration time (in non-language impaired samples)	4-scale response options (exact wording depends on item): 1. 'better/healthier than normal' 2. 'same as usual' 3. 'worse/more than usual' 4. 'much worse/more than usual'	Administration: no training required. Scoring procedures: no training required.	Testing materials: + Available from: https://www.gl-assessment.co.uk	Italian (85) Arabic (86) Turkish (87) Persian (88) Portuguese (89) Kannada (90) Hindi (91) Spanish (92) A number of other unvalidated translations are available. The MAPI Research

	physical illness • GHQ-28: a 28 item scaled version – assesses somatic symptoms, anxiety and insomnia, social dysfunction and severe depression (7 items for each of the four scales)	current state and asks if that differs from his or her usual state, and is therefore sensitive to short-term psychiatric disorders.			4 possible methods of scoring. GHQ scoring (0-0-1-1) is advocated by the test author. GHQ-12 yields only an overall total score (range: 0 to 12 points with standard scoring procedure).			Trust distributes translated versions on behalf of GL Assessment. Contact: PROinformation@mapi-trust.org

* Free, + Up to US$100, ++ Up to US$200, +++ > US$200

Supplementary Table 4

Properties of recommended outcome measurement instruments

	Western Aphasia Battery – Revised (WAB-R)	Stroke and Aphasia Quality of Life Scale (SAQOL-39/39g)	General Health Questionnaire (GHQ-12)
Objectivity	• During assessment: Limited because no audio recordings of verbal stimulus material available • During scoring: Limited for spontaneous speech and written output subtests	• During assessment: Moderate (interaction between assessor and patient frequently required because of physical stroke symptoms (arm paresis) and lack of pictorial task instructions (written sentences only) • During scoring: High	• During assessment: High if assessor does not interact with patient • During scoring: High
Internal consistency	High: Cronbach's alpha of total score= 0.91 (93).	High: Cronbach's alpha of total score= 0.93; Cronbach's alpha of subscale scores= 0.74–0.94 (54). SAQOL-39g: High: Cronbach's alpha of total score= 0.95; Cronbach's alpha for subscale scores= 0.92-0.95 (55)	High (in general population): Cronbach's alpha of total score= 0.79-0.91 (94-96). Cronbach's alpha of subscale scores= 0.80-0.92.
Test-retest reliability*	Excellent test-retest reliability: r >0.90 Acute stage post stroke: • Korean version; (58); 5-day test–retest interval (n=20 people with aphasia; Aphasia Quotient: r=0.976; Language Quotient: r=0.977; Cortical Quotient: r=0.920; Spontaneous Speech: r=0.96; Auditory Comprehension: r=0.967; Repetition: r=0.952; Naming: r=0.934; Reading: r=0.986; Writing: r=0.988; Praxis, r=0.908; Construction: r=0.922). Chronic stage post stroke: • 1 year test–retest interval (97), n=22 patients, r=0.992	Good to excellent test-retest reliability ICC=0.89-0.98 • English version; 2 to 14 days; n=17 people with aphasia; ICC=0.98 overall, 0.94–0.98 subscales (54). • English generic stroke version (SAQOL-39g); 7 ± 4 day test–retest interval; n=18 people with stroke/ stroke and aphasia; ICC= 0.96 overall; ICC= 0.92–0.98 subscales (55) Other translated versions: • Chilean version; ICC=0.95 (67) • Chinese ICC=0.97(69) • Chinese mandarin version; ICC=0.98 (70) • Dutch ICC=0.9 (71) • Greek ICC=0.96 (73)	Acceptable to excellent test-retest reliability • General population: ICC=0.79-0.82 (100) • Stroke (inc. aphasia) population using GHQ-28: 2 month test-retest reliability with a sample of 20 individuals (r=0.90) (101)

	• 6 months to 6.5 test–retest interval (av. 12-23 months test–retest interval; (93)), n=38 patients with chronic aphasia; WAB-AQ (r=0.968), WAB-CQ (n=9, r=0.895), WAB-LQ subtests: Spontaneous Speech – Information Content (r=0.947) and Fluency (r=0.941), Comprehension (r=0.881), Repetition (r=0.970), Naming (r=0.923), Reading (n=32; r=0.927) and Writing (n=25; r=0.956) and the Construction subtest (n=14, r=974). Test-retest reliability was adequate for the Praxis subtest (n=18, r=0.581). • Danish version (98); 3.5 months test–retest interval; n=19, r=0.96. • Cantonese version (99); 12 to 16 months test–retest interval; n=16 patients, Spontaneous Speech subtest – Information, Fluency and total scores (r=0.83, 0.94, 0.96 respectively), Naming subtest (r=0.91), AQ (r=0.93).	• Hindi ICC=0.9 (74) • Italian ICC=0.916 (75) (76) • Japanese ICC=0.97 (77) • Kannada ICC=0.8 (78) • Korean ICC=0.909 (79) • Malayalam ICC=0.91 (80) • Persian ICC=0.93 (81) • Portuguese ICC=0.927 (82) • Spanish ICC=0.949 (83) • Turkish ICC=0.97 (84)	
Responsiveness	Sub-/acute phase (up to 1 month post-onset): • WAB-LQ: n=50 adults with aphasia secondary to acute stroke, who received treatment (n=42) or no treatment (n=8). Participants assessed at baseline (2-4 weeks post-onset of aphasia), 3 months, and at least 6 months post-baseline. Significant main effect for time ($F=43.33$, $df=2.96$, $p<0.0001$), significant differences in the mean scores for the three tests ($p<0.01$). (102) • Very Early Rehabilitation of Speech (VERSE) trial; n=20 participants with mild-severe aphasia receiving intervention (4-5 h/wk for 5 wks) achieved 18% greater recovery on the	Acute to post-acute phase (up to 6 months post-onset): • Generic stroke sample, n=87; people admitted to hospital with a first stroke were assessed two weeks, three months and six months post stroke. Moderate changes ($d = 0.35$—0.49; standardized response mean (SRM) = 0.29—0.53) from two weeks to six months support responsiveness. (55) Post-acute to chronic (3 months to 1 year) • Cohort study of stroke sample with and without aphasia, n=78. Effect size r=0.22. MID estimated 0.21. (107) Chronic phase (at least 6 months post-onset):	Acute to post-acute phase (up to 6 months post-onset): • Impact of stroke with and without aphasia across the first six months, n=87 people with stroke or stroke and aphasia; psychological distress significantly reduced with time on GHQ-12 [$F (2,140) = 7.1$, $p=0.001$] (109) Chronic phase (at least 6 months post-onset): • Effects of singing in a community choir on mood; n=13 people with aphasia; 2.8 point reduction in mean GHQ-12 score was seen by week 12,

	WAB-AQ compared to the usual care group (11 min/week for 3 wks) (103). Post-acute phase (2-6 months post-onset): • See (102) above • Prospective longitudinal study with n=75 participants with aphasia post stroke, assessments at 4, 8, 12 and 24 weeks post-stroke, significant improvement in WAB-AQ across first year post-stroke (104) Chronic phase (at least 6 months post-onset): • n=10 participants with chronic aphasia. Combination of d-amphetamine, TMS, and SLT superior to control intervention of placebo with TMS and SLT; Change in AQ (from 36.13[18.23] to 38.60[19.33], P = 0.04) and LQ (from 32.41[14.93] to 35.03[15.10], P = 0.02) showed a statistically significant increase in the active experiment. Comparison of proportional changes of AQ and LQ in the active experiment with AQ and LQ in the placebo experiment showed a significant difference (AQ, P = 0.02; LQ, P = 0.008) (105) Mixed stages • n= 50 participants with aphasia (49 secondary to subacute or chronic stroke). Participants' mean scores improved significantly from pre- to post-treatment on all WAB subtests, with absolute percentages ranging from 6.5% to 13% improvement (p<0.01 to p<0.0001) (106).	• Intensive speech and language therapy compared to a waiting list control condition; n=156; Verbal communication was significantly improved from baseline to post-treatment (mean difference 2·61 points [SD 4·94]; 95% CI 1·49 to 3·72), but not from baseline to after treatment deferral (–0·03 points [4·04]; –0·94 to 0·88; between-group difference Cohen's d=0·58; p=0·0004). F-value for the main comparison is 12.97 (df1=1, df2=153), p= 0.0004 (108)	suggesting a possible reduction in adverse mood symptoms that was sustained to week 20. (110) • Effects of solution-focused brief therapy, n=5 people with aphasia, On GHQ-12 the mean (SD) score before therapy was 4.80 (4.60) [median (IQR) = 6.00 (0–9.00)]. This was reduced after therapy to a mean (SD) score of 2.00 (2.55) [median (IQR) = 1.00 (0–4.50)]. The effect size was large: Cohen's *d* = 0.79. (111) Caregivers of people with aphasia: • Impact of a psychoeducation program on caregivers' burden and stress, n =31 caregivers of people with post stroke aphasia. Caregivers in the immediate treatment group had significant reductions in GHQ-12 measured stress (GHQ mean (SD) at baseline =6.26 (5.67), GHQ post treatment 3.21 (SD 4.20), =/0.006). (112)
Convergent validity	• Convergent validity in sample of n=15 people with aphasia (93). Comparison	• SAQOL-39: Good convergent validity (r=0.55 to 0.67)(54). Adequate correlation between	Convergent validity in post-stroke aphasia sample:

	with corresponding subtests of the Neurosensory Center Comprehensive Examination for Aphasia (NCCEA), using Pearson correlation coefficients ○ Excellent correlation between: WAB Spontaneous Speech and NCCEA Description of Use and Sentence Construction (r= 0.817); WAB Comprehension and NCCEA Identification by Name and Identification by Sentence (r= 0.915); WAB Repetition and NCCEA Sentence Repetition (r= 0.880); WAB Naming and NCCEA Visual Naming and Word Fluency (r= 0.904); WAB Reading and NCCEA Reading subtests (r=0.919); WAB Writing and NCCEA Writing subtests (r=0.905); and WAB and NCCEA total scores (r=0.973). ○ Excellent correlation between the WAB-CQ (minus the Praxis and Construction subtests) and a comparable NCCEA score (minus the Tactile Naming-Right/Left, Articulation, Digit Repetition-Forward/Backward subtests) (r=0.964). • Sample of n=45 people with aphasia. Excellent correlation between the WAB and the Czech version of the Mississippi Aphasia Screening Test (MASTcz) (r= 0.933) (113)	GHQ-12 and the SAQOL-39 mean (0.53, p<0.01). The physical, communication, and energy subscales show good convergent validity (r=0.39 to 0.67, r=0.55, r=0.32, respectively). The psychosocial subdomain shows adequate convergent (r=0.28 to 0.62) validity with only 1 correlation lower than predicted (r=0.28 with the SSS). Good correlations with Frenchay Activities Index (FAI) and ASHA Functional Assessment of Communication Skills (ASHA-FACS). • SAQOL-39g: Good/excellent convergent validity for overall scale (r=0.36–0.70); and subdomains (r=0.47–0.78) (55), evidenced by moderate to high correlations with measures of stroke severity (NIHSS), activities of daily living (Barthel Index), extended activities of daily living (Frenchay Activities Index), emotional distress (GHQ-12) and language (Frenchay Aphasia Screening Test).	• Good correlations with SAQOL 39/SAQOL-39 (English, Greek, and Turkish versions). • The GHQ-12 demonstrated good convergent validity in a sample of 83 individuals with chronic stroke and aphasia, by comparison with the SAQOL-39. The study yielded an adequate correlation between the GHQ-12 and the SAQOL-39 mean (0.53, p<0.01). Correlations between the GHQ-12 and SAQOL-39 subtests were adequate (physical r=0.39, energy r=0.32, p<0.01) to excellent (psychosocial r=0.62, p<0.01). (54)
Discriminant validity	• Sample of n=140 people with aphasia. Comparison of WAB with Raven's	SAQOL-39: Discriminant validity (r = 0.02-0.27) (54)	Excellent discriminant validity in Swedish population (n=556 patient cases surveyed in specialized psychiatric care outpatient age and n=556 sex-matched controls).

	Coloured Progressive Matrices scores Adequate correlation (r=0.547). • Sample of n=66 people with chronic aphasia. Discriminant validity of the WAB Aphasia Quotient (WAB-AQ) by comparison with the Scandinavian Stroke Scale (SSS), Barthel Index (BI) and Frenchay Activities Index (FAI). Excellent correlation between the WAB-AQ and the SSS (r=0.64), adequate correlations between the WAB-AQ and the BI (r=0.44) and the FAI (r=0.50).	SAQOL-39g: Good/excellent discriminant validity for overall scale and subdomains, evidenced by low to moderate correlations with external measures (r = 0.03-0.40). (55)	Individuals using specialized psychiatric services and healthy controls (Likert index AUC=0.86, GHQ index AUC=0.83), and between individuals with current disorder from healthy controls (Likert index AUC=0.90, GHQ index AUC=0.88). (114).

*** Test-retest reliability**: 1=perfect reliability; ≥ 0.9=excellent reliability; ≥ 0.8 < 0.9=good reliability; ≥ 0.7 < 0.8=acceptable reliability; ≥ 0.6 < 0.7=questionable reliability; ≥ 0.5 < 0.6=poor reliability; < 0.5=unacceptable reliability; 0=no reliability.

1. **References**

2. 1. Swinburn K, Porter G, Howard D. Comprehensive Aphasia Test. Hove: Psychology Press; 2004.
3. 2. Kertesz A, Raven JC. The Western Aphasia Battery Revised (WAB-R). San Antonio, TX: PsychCorp; 2007.
4. 3. Enderby P, John A. Therapy Outcome Measures (TOM): Speech and Language Therapy. San Diego, CA: Singular Pub. Group; 1997.
5. 4. Enderby P, John A, Petheram B. Therapy outcome measures for the rehabilitation professionals: Speech and language therapy, physiotherapy, occupational therapy, rehabilitation nursing, hearing therapists 2nd ed. Chichester, West Sussex, England: John Wiley; 2006.
6. 5. Enderby P, John P. Therapy Outcome Measures for Rehabilitation Professionals. 3rd edition. UK: J&R Press; 2015.
7. 6. Kalbe E, Reinhold N, Ender U, Kessler J. Die Aphasia-Check-List (ACL). Köln: Prolog; 2002.
8. 7. Huber W, Poeck K, Weniger D, Willmes K. Der Aachener Aphasie Test (AAT) [Aachener Aphasie Test (AAT)]. Göttingen: Hogrefe 1983.
9. 8. Toğram B, Maviş İ. Validity, Reliability and Standardization Study of the Language Assessment Test for Aphasia. Türk Nöroloji Dergisi. 2012;18(3):96-103.
10. 9. Manochiopinig S, Reed VA, Sheard C, Choo P. An aphasia screening assessment instrument for Thailand: Development and validation. Asia Pacific Journal of Speech, Language, and Hearing. 1996;1(1):21-41.
11. 10. Paradis M, Libben G. The Assessment of Bilingual Aphasia. Hillsdale, N.J.: Lawrence Erlbaum Associates; 1987.

12. 11. Goodglass H, Kaplan E, Barresi B. Boston Diagnostic Aphasia Examination. BDAE. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2001.

13. 12. Calis FA, On AY, Durmaz B. Validity and reliability of a new test for Turkish-speaking aphasic patients: Ege Aphasia Test. NeuroRehabilitation. 2013;32(1):157-63 7p.

14. 13. Marshall RC, Wright HH. Developing a clinician-friendly aphasia test. American Journal of Speech-Language Pathology. 2007;16(4):295-315.

15. 14. Nespoulous JL, Joanette Y, Lecours AR. Protocole Montréal–Toulouse d'examen linguistique de l'aphasie (mt-86). Isbergues: Ortho Edition; 1986.

16. 15. Reinvang I, Engvik H. Norsk Grunntest for Afasi. Oslo: Universitetsforl.; 1980.

17. 16. Cherney L, Babbitt E, Semik P, Heinemann A. Psychometric properties of the communication confidence rating scale for aphasia (CCRSA): Phase 1. Topics in Stroke Rehabilitation. 2011;18(4):352-60.

18. 17. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatrica Scandinavica. 1983;67(6):361-70.

19. 18. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. . The British Journal of Psychiatry. 1979(134):382-9.

20. 19. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. Journal of psychiatric research. 1982;17(1):37-49.

21. 20. Yesavage JA, Sheikh JI. 9/Geriatric Depression Scale (GDS). Clinical Gerontologist. 1986;5(1-2):165-73.

22. 21. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health and Quality of Life Outcomes. 2007;5(1):63.

23. 22. Segal DL, June A, Payne M, Coolidge FL, Yochim B. Development and initial validation of a self-report assessment tool for anxiety among older adults: the Geriatric Anxiety Scale. Journal of anxiety disorders. 2010;24(7):709-14.

24. 23. Smollan T, Penn C. The measurement of emotional reaction and depression in a South African stroke population. Disability and rehabilitation. 1997;19(2):56-63.

25. 24. Hammond MF, O'Keeffe ST, Barer DH. Development and validation of a brief observer-rated screening scale for depression in elderly medical patients. Age Ageing. 2000;29(6):511-5.

26. 25. Sutcliffe LM, Lincoln NB. The assessment of depression in aphasic stroke patients: the development of the Stroke Aphasic Depression Questionnaire. Clinical Rehabilitation. 1998;12(6):506-13.

27. 26. Brumfitt SM, Sheeran P. The development and validation of the Visual Analogue Self-Esteem Scale (VASES). Br J Clin Psychol. 1999;38 (Pt 4):387-400.

28. 27. Eaton WW, Muntaner C, Smith C, Tien A, Ybarra M. Center for Epidemiologic Studies Depression Scale: Review and revision (CESD and CESD-R). The Use of Psychological Testing for Treatment Planning and Outcomes Assessment Mahwah, NJ: Lawrence Erlbaum; 2004. p. 363-77.

29. 28. Goldberg D. General Health Questionnaire (GHQ-12). Windsor, UK: NFER-Nelson.; 1992.

30. 29. Enderby P, John A. Therapy Outcome Measures for Rehabilitation Professionals. J&R Press; 2015.

31. 30. Enderby P, John A. Therapy outcome measures (TOM) : speech and language therapy. John A, editor. San Diego, California.: Singular Pub. Group; 1997.

32. 31. Enderby P, John A, Petheram B. Therapy Outcome Measures for the rehabilitation professionals: Speech and language therapy, physiotherapy, occupational therapy, rehabilitation nursing, hearing therapists. 2nd ed.. ed. Chichester, England.: John Wiley; 2006.

33. 32. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: Validity of a Two-Item Depression Screener. Medical Care. 2003;41(11):1284-92.

34. 33. Cameron IM, Crawford JR, Lawton K, Reid IC. Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. The British journal of general practice : the journal of the Royal College of General Practitioners. 2008;58(546):32-6.

35. 34. Stern RA. Visual Analog Mood Scales professional manual. Odessa, FL: Psychological Assessment Resources; 1997.

36. 35. Doyle Patrick J, Hula William D, Austermann Hula Shannon N, Stone Clement A, Wambaugh Julie L, Ross Katherine B, et al. Self- and surrogate-reported communication functioning in aphasia. Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation. 2013;22(5):957-67.

37. 36. Frattali CM, Thompson CK, Holland AL, Wohl CB, Ferketic MM, . Functional assessment of communication skills for adults. Rockville, MD: American Speech, Language, and Hearing Association.; 1995.

38. 37. Blomert L, Kean ML, Koster C, Schokker J. Amsterdam-Nijmegen Everyday Language Test: Construction, Reliability and Validity. Aphasiology. 1994;8(4):381-407.

39. 38. Pulvermüller F, Berthier ML. Aphasia therapy on a neuroscience basis. Aphasiology. 2008;22(6):563-99.

40. 39. Long A, Hesketh A, Paszek G, Booth M, Bowen A. Development of a reliable self-report outcome measure for pragmatic trials of communication therapy following stroke: the Communication Outcome after Stroke (COAST) scale. Clin Rehabil. 2008;22(12):1083-94.

41. 40. Aujla S, Botting N, Worrall L, Hickson L, Cruice M. Preliminary psychometric analyses of two assessment measures quantifying communicative and social activities: the COMACT and SOCACT. Aphasiology. 2015:1-24.

42. 41. Chue WL, Rose ML, Swinburn K. The reliability of the Communication Disability Profile: A patient-reported outcome measure for aphasia. Aphasiology. 2010;24(6-8):940-56.

43. 42. Lomas J, Pickard L, Bester S, Elbard H, Finlayson A, Zoghaib C. The Communicative Effectiveness Index: Development and psychometric evaluation of a functional communication measure for adult aphasia. J Speech Hear Disord. 1989;54(1):113-24.

44. 43. Callaway L, Winkler D, Tippett A, Migliorini C, Herd N, Willer B. The Community Integration Questionnaire-Revised (CIQ-R). . Melbourne, Australia: Summer Foundation Ltd; 2014.

45. 44. Holland AL, Frattali C, Fromm D. CADL-2 communication activities of daily living Audrey Holland, Carol Frattali, Davida Fromm. CADL-two communication activities of daily living. 2nd ed.. ed. Austin, Tex.: Austin, Tex. : Pro-ed; 1999.

46. 45. Glueckauf RL, Blonder LX, Ecklund-Johnson E, Maher L, Crosson B, Gonzalez-Rothi L. Functional Outcome Questionnaire for Aphasia: Overview and preliminary psychometric evaluation. Neurorehabilitation. 2003;18(4):281-90.

47. 46. Kagan A, Winckel J, Black S, Duchan JF, SimmonsMackie N, Square P, et al. A Set of Observational Measures for Rating Support and Participation in Conversation Between Adults with Aphasia and Their Conversation Partners. Topics in Stroke Rehabilitation. 2004;11(1):67-83.

48. 47. Van Der Meulen I, Van De Sandt-Koenderman WME, Duivenvoorden HJ, Ribbers GM. Measuring verbal and non- verbal communication in aphasia: reliability, validity, and sensitivity to change of the Scenario Test. International Journal of Language & Communication Disorders, 2010, Vol45(4), p424-435. 2010;45(4):424-35.

49. 48. Lincoln NB. The speech questionnaire: An assessment of functional language ability. International Rehabilitation Medicine. 1982;4(3):114-7.

50. 49. Baylor C, Yorkston K, Eadie T, Kim J, Chung H, Amtmann D. The Communicative Participation Item Bank (CPIB): Item Bank Calibration and Development of a Disorder-Generic Short Form. Journal of Speech, Language, and Hearing Research. 2013;56(4):1190-208.

51. 50. Hutter BO. Sickness Impact Profile (SIP)-German version. 2001. In: Compendium of quality of life instruments [Internet]. Chichester, U.K.: John Wiley & Sons.

52. 51. Doyle PJ, McNeil MR, Mikolic JM, Prieto L, Hula WD, Lustig AP, et al. The Burden of Stroke Scale (BOSS) provides valid and reliable score estimates of functioning and well-being in stroke survivors with and without communication disorders. J Clin Epidemiol. 2004;57(10):997-1007.

53. 52. Buck D, Jacoby A, Massey A, Steen N, Sharma A, Ford GA. Development and validation of NEWSQOL, the Newcastle Stroke-Specific Quality of Life Measure. Cerebrovascular Diseases. 2004;17(2-3):143-52.

54. 53. Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. Medical Care. 1992;30(6):473-83.

55. 54. Hilari K, Byng S, Lamping DL, Smith SC. Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): evaluation of acceptability, reliability, and validity. Stroke; a journal of cerebral circulation. 2003;34(8):1944-50.

56. 55. Hilari K, Lamping DL, Smith SC, Northcott S, Lamb A, Marshall J. Psychometric properties of the Stroke and Aphasia Quality of Life Scale (SAQOL-39) in a generic stroke population. Clin Rehabil. 2009;23(6):544-57.

57. 56. Goodglass H, Kaplan E. Boston Diagnostic Aphasia Examination. Philadelphia: Lea and Febiger; 1972.

58. 57. Yiu EML. Linguistic assessment of Chinese-speaking aphasics: Development of a Cantonese aphasia battery. Journal of Neurolinguistics. 1992;7(4):379-424.

59. 58. Kim H, Na DL. Normative data on the Korean version of the Western Aphasia Battery. Journal of clinical and experimental neuropsychology. 2004;26(8):1011-20.

60. 59. Keshree NK, Kumar S, Basu S, Chakrabarty M, Kishore T. Adaptation of The Western Aphasia Battery in Bangla. 2013;17(2):189.

61. 60. Ozaeta C, Kong A, Ranoa-Javier MB. A Pilot Study of Using the Tagalog Version of the Western Aphasia Battery-Revised in the Philippines. Procedia - Social and Behavioral Sciences. 2013;94:232-3.

62. 61. Neves MB, Van Borsel J, Pereira MM, Paradela EM. Cross-cultural adaptation of the Western Aphasia Battery - Revised screening test to Brazilian Portuguese: a preliminary study. CoDAS. 2014;26(1):38-45.

63. 62. WAB Aphasia Test Construction Committee. The Japanese version of the Western Aphasia Battery. Tokyo: Igaku-Shoin Ltd.; 1986.

64. 63. Keklikoglu HD, Selcuki D, Keskin S. Should the Western Aphasia Battery be translated into Turkish?/Western Afazi Bataryasi Turkce'ye cevrilmeli mi?(Research Article/Arastirma Makalesi)(Report). 2010;47(1):40.

65. 64. Kertesz A, Pascual-Leone P, Pascual-Leone G. Western Aphasia Battery en versión y adaptación castellana. . Valencia: Nau Libres; 1990.

66. 65. Hilari K, Owen S, Farrelly SJ. Proxy and self-report agreement on the Stroke and Aphasia Quality of Life Scale-39. Journal of neurology, neurosurgery, and psychiatry. 2007;78(10):1072-5.

67. 66. Ignatiou M, Christaki V, Chelas EN, Efstratiadou EA, Hilari K. Agreement between People with Aphasia and Their Proxies on Health-Related Quality of Life after Stroke, Using the Greek SAQOL-39g. Psychology. 2012;Vol.03No.09:5.

68. 67. Caute A, Northcott S, Clarkson L, Pring T, Hilari K. Does mode of administration affect health-related quality-of-life outcomes after stroke? International journal of speech-language pathology. 2012;14(4):329-37.

69. 68. Diaz V, Gonzalez R, Salgado D, Perez D. Stroke and aphasia quality of life scale (SAQOL-39). Evaluation of acceptability, reliability and validity of Chilean version. Journal of the Neurological Sciences. 2013;333, Supplement 1:e553-e4.

70. 69. Lin R, Chen X, Feng MI, Cai LJ, Deng F. Reliability and validity of the Chinese-version Stroke and Aphasia Quality of Life Scale. . Chinese Journal of Nursing. 2013;4.

71. 70. Guo YE, Togher L, Power E, Koh GC. Validation of the Stroke and Aphasia Quality of Life Scale in a multicultural population. Disability and rehabilitation. 2016;38(26):2584-92.

72. 71. van Ewijk L, Versteegde L, Raven-Takken E, Hilari K. Measuring quality of life in Dutch people with aphasia: development and psychometric evaluation of the SAQOL-39NL. Aphasiology. 2016:1-12.

73. 72. Kartsona A, Hilari K. Quality of life in aphasia: Greek adaptation of the stroke and aphasia quality of life scale - 39 item (SAQOL-39). Europa Medicophysica. 2007;43(1):27-35.

74. 73. Efstratiadou EA, Chelas EN, Ignatiou M, Christaki V, Papathanasiou I, Hilari K. Quality of life after stroke: evaluation of the Greek SAQOL-39g. Folia phoniatrica et logopaedica : official organ of the International Association of Logopedics and Phoniatrics. 2012;64(4):179-86.

75. 74. Mitra I, Krishnan G. Adaptation and validation of stroke-aphasia quality of life (SAQOL-39) scale to Hindi. Annals of Indian Academy of Neurology. 2015;18(1):29-32.

76. 75. Posteraro L, Formis A, Bidini C, Grassi E, Curti M, Bighi M, et al. Aphasia quality of life: reliability of the Italian version of SAQOL-39. Europa Medicophysica. 2004;40(4):257-62.

77. 76. Posteraro L, Formis A, Grassi E, Bighi M, Nati P, Proietti Bocchini C, et al. Quality of life and aphasia. Multicentric standardization of a questionnaire. Europa Medicophysica. 2006;42(3):227-30 4p.

78. 77. Kamiya A, Kamiya K, Tatsumi H, Suzuki M, Horiguchi S. Japanese Adaptation of the Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Comparative Study among Different Types of Aphasia. Journal of Stroke and Cerebrovascular Diseases. 2015;24(11):2561-4.

79. 78. Kiran S, Krishnan G. Stroke and aphasia quality of life scale in Kannada-evaluation of reliability, validity and internal consistency. Annals of Indian Academy of Neurology. 2013;16(3):361-4.

80. 79. Gimsujeong S, Gimdeokyong G. A measure of quality of life after stroke aphasia (SAQOL-39). Validity and Reliability of the Korean Version. . Rehabilitation and Welfare. 2012;16:245–65.

81. 80. Raju R, Krishnan G. Adaptation and validation of stroke-aphasia quality of life (SAQOL-39) scale to Malayalam. Annals of Indian Academy of Neurology. 2015;18(4):441.

82. 81. Mazdeh M, Yaghobi A. The Study of Quality of Life in Aphasic Stroke Patients in University- Medical Centers of Hamedan. Qom Univ Med Sci J. 2009;3(1):21-8.

83. 82. Rodrigues IT, Leal MG. Tradução portuguesa e análise de aspectos psicométricos da escala "Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39)". Audiology - Communication Research. 2013;18:339-44.

84. 83. Lata-Caneda MC, Pineiro-Temprano M, Garcia F, Garcia-Armesto I, Barrueco-Egido JR, Meijide-Failde R. Spanish adaptation of the stroke and aphasia quality of life scale-39 (SAQOL-39). European Journal of Physical and Rehabilitation Medicine. 2009;45(3):379-84.

85. 84. Atamaz Calis F, Celik S, Demir O, Aykanat D, Yagiz On A. The psychometric properties of the Turkish Stroke and Aphasia Quality Of Life Scale-39. International Journal of Rehabilitation Research. 2016;39(2):140-4.

86. 85. Politi PL, Piccinelli M, Wilkinson G. Reliability, validity and factor structure of the 12-item General Health Questionnaire among young males in Italy. Acta psychiatrica Scandinavica. 1994;90(6):432-7.

87. 86. Daradkeh TK, Ghubash R, el-Rufaie OE. Reliability, validity, and factor structure of the Arabic version of the 12-item General Health Questionnaire. Psychological reports. 2001;89(1):85-94.

88. 87. Kihç C, Rezaki M, Rezaki B, Kaplan I, Özgen G, Sagduyu A, et al. General Health Questionnaire (GHQ12 & GHQ28): psychometric properties and factor structure of the scales in a Turkish primary care sample. Social Psychiatry and Psychiatric Epidemiology. 1997;32(6):327-31.

89. 88. Montazeri A, Harirchi AM, Shariati M, Garmaroudi G, Ebadi M, Fateh A. The 12-item General Health Questionnaire (GHQ-12): translation and validation study of the Iranian version. Health and quality of life outcomes. 2003;1:66-.

90. 89. Laranjeira CA. General health questionnaire--12 items: adaptation study to the Portuguese population. Epidemiologia e psichiatria sociale. 2008;17(2):148-51.

91. 90. Javaregowda DC, Parthasarathy B, Suresh A, Lokanath D, Govind Babu K, Ullas B, et al. Validation of 12-item general health questionnaire into Kannada language. Journal of Clinical Oncology. 2007;25(18_suppl):17077-.

92. 91. Jacob KS, Bhugra D, Mann AH. The validation of the 12-item General Health Questionnaire among ethnic Indian women living in the United Kingdom. Psychol Med. 1997;27(5):1215-7.

93. 92. Sanchez-Lopez Mdel P, Dresch V. The 12-Item General Health Questionnaire (GHQ-12): reliability, external validity and factor structure in the Spanish population. Psicothema. 2008;20(4):839-43.

94. 93. Shewan CM, Kertesz A. Reliability and validity characteristics of the Western Aphasia Battery (WAB). J Speech Hear Disord. 1980;45(3):308-24.

95. 94. Hankins M. The reliability of the twelve-item general health questionnaire (GHQ-12) under realistic assumptions. BMC public health. 2008;8(1):355.

96. 95. Gnambs T, Staufenbiel T. The structure of the General Health Questionnaire (GHQ-12): two meta-analytic factor analyses. Health Psychology Review. 2018;12(2):179-94.

97. 96. Shevlin M, Adamson G. Alternative Factor Models and Factorial Invariance of the GHQ-12: A Large Sample Analysis Using Confirmatory Factor Analysis. . Psychological Assessment. 2005;17(2):231-6.

98. 97. Kertesz A, McCabe P. Recovery patterns and prognosis in aphasia. Brain. 1977;100 Pt 1:1-18.

99. 98. Pedersen PM, Vinter K, Olsen TS. The Communicative Effectiveness Index: Psychometric properties of a Danish adaptation. Aphasiology. 2001;15(8):787-802.

100. 99. Kong AP. The main concept analysis in cantonese aphasic oral discourse: external validation and monitoring chronic aphasia. Journal of speech, language, and hearing research : JSLHR. 2011;54(1):148-59.

101. 100. Piccinelli M, Bisoffi G, Bon MG, Cunico L, Tansella M. Validity and test-retest reliability of the Italian version of the 12-item General Health Questionnaire in general practice: a comparison between three scoring methods. Comprehensive psychiatry. 1993;34(3):198-205.

102. 101. Robinson RG, Price TR. Post-stroke depressive disorders: a follow-up study of 103 patients. Stroke; a journal of cerebral circulation. 1982;13(5):635.

103. 102. Shewan CM. The Language Quotient (LQ): A new measure for the Western Aphasia Battery. Journal of communication disorders. 1986;19(6):427-9.

104. 103. Godecke E, Ciccone NA, Granger AS, Rai T, West D, Cream A, et al. A comparison of aphasia therapy outcomes before and after a Very Early Rehabilitation programme following stroke. International journal of language &

communication disorders / Royal College of Speech & Language Therapists. 2014;49(2):149-61.

105. 104. Bakheit AMO, Shaw S, Barrett L, Wood J, Carrington S, Griffiths S, et al. A prospective, randomized, parallel group, controlled study of the effect of intensity of speech and language therapy on early recovery from poststroke aphasia. Clinical Rehabilitation. 2007;21(10):885-94.

106. 105. Keser Z, Dehgan MW, Shadravan S, Yozbatiran N, Maher LM, Francisco GE. Combined Dextroamphetamine and Transcranial Direct Current Stimulation in Poststroke Aphasia. American Journal of Physical Medicine & Rehabilitation. 2017;96(10):S141-S5.

107. 106. Aftonomos LB, Steele RD, Appelbaum JS, Harris VM. Relationships between impairment-level assessments and functional-level assessments in aphasia: Findings from LCC treatment programmes. Aphasiology. 2001;15(10-11):951-64.

108. 107. Guo YE, Togher L, Power E, Heard R, Luo N, Yap P, et al. Sensitivity to change and responsiveness of the Stroke and Aphasia Quality-of-Life Scale (SAQOL) in a Singapore stroke population. Aphasiology. 2017;31(4):427-46.

109. 108. Breitenstein C, Grewe T, Flöel A, Ziegler W, Springer L, Martus P, et al. Intensive speech and language therapy in patients with chronic aphasia after stroke: a randomised, open-label, blinded-endpoint, controlled trial in a health-care setting. The Lancet. 2017;389(10078):1528-38.

110. 109. Hilari K. The impact of stroke: are people with aphasia different to those without? Disability and rehabilitation. 2011;33(3):211-8.

111. 110. Tamplin J, Baker FA, Jones B, Way A, Lee S. 'Stroke a Chord': the effect of singing in a community choir on mood and social engagement for people living with aphasia following a stroke. NeuroRehabilitation. 2013;32(4):929-41.

112. 111. Northcott S, Burns K, Simpson A, Hilari K. 'Living with Aphasia the Best Way I Can': A Feasibility Study Exploring Solution-Focused Brief Therapy for People with Aphasia. Folia phoniatrica et logopaedica : official organ of the International Association of Logopedics and Phoniatrics. 2015;67(3):156-67.

113. 112. Draper B, Bowring G, Thompson C, Van Heyst J, Conroy P, Thompson J. Stress in caregivers of aphasic stroke patients: a randomized controlled trial. Clin Rehabil. 2007;21(2):122-30.

114. 113. Košťálová M, Bártková E, Šajgalíková K, Dolenská A, Dušek L, Bednařík J. A standardization study of the Czech version of the Mississippi Aphasia Screening Test (MASTcz) in stroke patients and control subjects. Brain Injury. 2008;22(10):793-801.

114. Lundin A, Ahs J, Asbring N, Kosidou K, Dal H, Tinghog P, et al. Discriminant validity of the 12-item version of the general health questionnaire in a Swedish case-control study. Nordic journal of psychiatry. 2017;71(3):171-9.

Search strategy (incorporates filters developed by Terwee and associates for the identification of studies reporting the measurement properties of health outcome measures; see Terwee CB, Jansma EP, Riphagen I, Vet HW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Quality of Life Research 2009;18(8):1115-23.)

PUBMED

Aphasia OR dysphasia AND stroke

AND

(instrumentation[sh] OR methods[sh] OR Validation Studies[pt] OR Comparative Study[pt] OR ‘‘psychometrics’’ [MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR clinometr*[tw] OR ‘‘outcome assessment (health care)’’[MeSH] OR outcome assessment[tiab] OR outcome measure*[tw] OR Qual Life Res (2009) 18:1115–1123 1121 123 ‘‘observer variation’’[MeSH] OR observer variation[tiab] OR ‘‘Health Status Indicators’’[Mesh] OR ‘‘reproducibility of results’’[MeSH] OR reproducib*[tiab] OR ‘‘discriminant analysis’’[MeSH] OR reliab*[tiab] OR unreliab*[tiab] OR valid*[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR ‘‘internal consistency’’[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alphas[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR reduction*[tiab])) OR agreement[tiab] OR precision[tiab] OR imprecision[tiab] OR ‘‘precise values’’[tiab] OR test– retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab* [tiab] AND (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intraobserver[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant [tiab] OR inter-participant[tiab]

OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tiab] OR ((replicab*[tiab] OR repeated[tiab]) AND (measure[tiab] OR measures[tiab] OR findings[tiab] OR result[tiab] OR results[tiab] OR test[- tiab] OR tests[tiab])) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR ''known group''[tiab] OR factor analysis[tiab] OR factor analyses[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR item discriminant[tiab] OR interscale correlation*[tiab] OR error[tiab] OR errors[tiab] OR ''individual variability''[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR ''standard error of measurement''[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR meaningful change [tiab] OR ''ceiling effect''[tiab] OR ''floor effect''[tiab] OR ''Item response model''[tiab] OR IRT[tiab] OR Rasch[tiab] OR ''Differential item functioning''[tiab] OR DIF[tiab] OR ''computer adaptive testing''[tiab] OR ''item bank''[tiab] OR ''cross-cultural equivalence''[tiab])

EMBASE

aphasia OR dysphasia AND stroke

AND

'intermethod comparison'/exp OR 'data collection method'/exp OR 'validation study'/exp OR 'feasibility study'/exp OR 'pilot study'/exp OR 'psychometry'/exp OR 'reproducibility'/exp OR reproducib*:ab,ti OR 'audit':ab,ti OR psychometr*:ab,ti OR clinimetr*:ab,ti OR clinometr*:ab,ti OR 'observer variation'/exp OR 'observer variation':ab,ti OR 'discriminant analysis'/exp OR 'validity'/exp OR reliab*:ab,ti OR valid*:ab,ti OR 'coefficient':ab,ti OR 'internal consistency':ab,ti OR (cronbach*:ab,ti AND ('alpha':ab,ti OR 'alphas':ab,ti)) OR 'item correlation':ab,ti OR 'item correlations':ab,ti OR 'item selection':ab,ti OR 'item selections':ab,ti OR 'item reduction':ab,ti OR 'item reductions':ab,ti OR 'agreement':ab,ti OR 'precision':ab,ti OR 'imprecision':ab,ti OR 'precise

values':ab,ti OR 'test-retest':ab,ti OR ('test':ab,ti AND 'retest':ab,ti) OR (reliab*:ab,ti AND ('test':ab,ti OR 'retest':ab,ti)) OR 'stability':ab,ti OR 'interrater':ab,ti OR 'inter-rater':ab,ti OR 'intrarater':ab,ti OR 'intra-rater':ab,ti OR 'intertester':ab,ti OR 'inter-tester':ab,ti OR 'intratester':ab,ti OR 'intratester':ab,ti OR 'interobeserver':ab,ti OR 'inter-observer':ab,ti OR 'intraobserver':ab,ti OR 'intraobserver':ab,ti OR 'intertechnician':ab,ti OR 'inter-technician':ab,ti OR 'intratechnician':ab,ti OR 'intratechnician':ab,ti OR 'interexaminer':ab,ti OR 'inter-examiner':ab,ti OR 'intraexaminer':ab,ti OR 'intraexaminer':ab,ti OR 'interassay':ab,ti OR 'inter-assay':ab,ti OR 'intraassay':ab,ti OR 'intra-assay':ab,ti OR 'interindividual':ab,ti OR 'inter-individual':ab,ti OR 'intraindividual':ab,ti OR 'intra-individual':ab,ti OR 'interparticipant':ab,ti OR 'inter-participant':ab,ti OR 'intraparticipant':ab,ti OR 'intraparticipant':ab,ti OR 'kappa':ab,ti OR 'kappas':ab,ti OR 'coefficient of variation':ab,ti OR repeatab*:ab,ti OR (replicab*:ab,ti OR 'repeated':ab,ti AND ('measure':ab,ti OR 'measures':ab,ti OR 'findings':ab,ti OR 'result':ab,ti OR 'results':ab,ti OR 'test':ab,ti OR 'tests':ab,ti)) OR generaliza*:ab,ti OR generalisa*:ab,ti OR 'concordance':ab,ti OR ('intraclass':ab,ti AND correlation*:ab,ti) OR 'discriminative':ab,ti OR 'known group':ab,ti OR 'factor analysis':ab,ti OR 'factor analyses':ab,ti OR 'factor structure':ab,ti OR 'factor structures':ab,ti OR 'dimensionality':ab,ti OR subscale*:ab,ti OR 'multitrait scaling analysis':ab,ti OR 'multitrait scaling analyses':ab,ti OR 'item discriminant':ab,ti OR 'interscale correlation':ab,ti OR 'interscale correlations':ab,ti OR ('error':ab,ti OR 'errors':ab,ti AND (measure*:ab,ti OR correlat*:ab,ti OR evaluat*:ab,ti OR 'accuracy':ab,ti OR 'accurate':ab,ti OR 'precision':ab,ti OR 'mean':ab,ti)) OR 'individual variability':ab,ti OR 'interval variability':ab,ti OR 'rate variability':ab,ti OR 'variability analysis':ab,ti OR ('uncertainty':ab,ti AND ('measurement':ab,ti OR 'measuring':ab,ti)) OR 'standard error of measurement':ab,ti OR sensitiv*:ab,ti OR responsive*:ab,ti OR ('limit':ab,ti AND 'detection':ab,ti) OR 'minimal detectable concentration':ab,ti OR interpretab*:ab,ti OR (small*:ab,ti AND ('real':ab,ti OR 'detectable':ab,ti) AND ('change':ab,ti OR 'difference':ab,ti)) OR 'meaningful change':ab,ti OR 'minimal important change':ab,ti OR 'minimal important difference':ab,ti OR 'minimally important change':ab,ti OR 'minimally important difference':ab,ti OR 'minimal detectable change':ab,ti OR 'minimal detectable difference':ab,ti OR 'minimally detectable change':ab,ti OR 'minimally detectable difference':ab,ti OR 'minimal real change':ab,ti OR 'minimal real difference':ab,ti OR 'minimally real change':ab,ti OR 'minimally real

difference':ab,ti OR 'ceiling effect':ab,ti OR 'floor effect':ab,ti OR 'item response model':ab,ti OR 'irt':ab,ti OR 'rasch':ab,ti OR 'differential item functioning':ab,ti OR 'dif':ab,ti OR 'computer adaptive testing':ab,ti OR 'item bank':ab,ti OR 'cross-cultural equivalence':ab,ti

CINAHL

aphasia OR dysphasia AND stroke

AND

TI psychometr* OR TI observer variation OR TI reproducib* OR TI reliab* OR TI unreliab* OR TI valid* OR TI coefficient OR TI homogeneity OR TI homogeneous OR TI "internal consistency" OR AB psychometr* OR AB observer variation OR AB reproducib* OR AB reliab* OR AB unreliab* OR AB valid* OR AB coefficient OR AB homogeneity OR AB homogeneous OR AB "internal consistency" OR (TI cronbach* OR AB cronbach* AND (TI alpha OR AB alpha OR TI alphas OR AB alphas)) OR (TI item OR AB item AND (TI correlation* OR AB correlation* OR TI selection* OR AB selection* OR TI reduction* OR AB reduction*)) OR TI agreement OR TI precision OR TI imprecision OR TI "precise values" OR TI test-retest OR AB agreement OR AB precision OR AB imprecision OR AB "precise values" OR AB test-retest OR (TI test OR AB test AND TI retest OR AB retest) OR (TI reliab* OR AB reliab* AND (TI test OR AB test OR TI retest or AB retest)) OR TI stability OR TI interrater OR TI interrater OR TI intrarater OR TI intra-rater OR TI intertester OR TI inter-tester OR TI intratester OR TI intra-tester OR TI interobserver OR TI inter-observer OR TI intraobserver OR TI intra-observer OR TI intertechnician OR TI inter-technician OR TI intratechnician OR TI intra-technician OR TI interexaminer OR TI inter-examiner OR TI intraexaminer OR TI intra-examiner OR TI interassay OR TI inter-assay OR TI intraassay OR TI intra-assay OR TI interindividual OR TI inter-individual OR TI intraindividual OR TI intra-individual OR TI interparticipant OR TI inter-participant OR TI intraparticipant OR TI intra-participant OR TI kappa OR TI kappa's OR TI kappas OR TI repeatab* OR AB stability OR AB interrater OR AB inter-rater OR AB intrarater OR AB intra-rater OR AB intertester OR AB inter-tester OR AB intratester OR AB intra-tester OR AB interobserver OR AB inter-observer OR AB intraobserver OR AB intra-observer OR AB intertechnician OR AB inter-technician OR AB intratechnician OR AB intra-technician OR AB interexaminer OR AB inter-examiner OR AB intraexaminer OR AB intra-

examiner OR AB interassay OR AB inter-assay OR AB intraassay OR AB intra-assay OR AB interindividual OR AB inter-individual OR AB intraindividual OR AB intra-individual OR AB interparticipant OR AB inter-participant OR AB intraparticipant OR AB intra-participant OR AB kappa OR AB kappa's OR AB kappas OR AB repeatab* OR ((TI replicab* OR AB replicab* OR TI repeated OR AB repeated) AND (TI measure OR AB measure OR TI measures OR AB measures OR TI findings OR AB findings OR TI result OR AB result OR TI results OR AB results OR TI test OR AB test OR TI tests OR AB tests)) OR TI generaliza* OR TI generalisa* OR TI concordance OR AB generaliza* OR AB generalisa* OR AB concordance OR (TI intraclass OR AB intraclass AND TI correlation* or AB correlation*) OR TI discriminative OR TI "known group" OR TI factor analysis OR TI factor analyses OR TI dimension* OR TI subscale* OR AB discriminative OR AB "known group" OR AB factor analysis OR AB factor analyses OR AB dimension* OR AB subscale* OR (TI multitrait OR AB multitrait AND TI scaling OR AB scaling AND (TI analysis OR AB analysis OR TI analyses OR AB analyses)) OR TI item discriminant OR TI interscale correlation* OR TI error OR TI errors OR TI "individual variability" OR AB item discriminant OR AB interscale correlation* OR AB error OR AB errors OR AB "individual variability" OR (TI variability OR AB variability AND (TI analysis OR AB analysis OR TI values OR AB values)) OR (TI uncertainty OR AB uncertainty AND (TI measurement OR AB measurement OR TI measuring OR AB measuring)) OR TI "standard error of measurement" OR TI sensitiv* OR TI responsive* OR AB "standard error of measurement" OR AB sensitiv* OR AB responsive* OR ((TI minimal OR TI minimally OR TI clinical OR TI clinically OR AB minimal OR AB minimally OR AB clinical OR AB clinically) AND (TI important OR TI significant OR TI detectable OR AB important OR AB significant OR AB detectable) AND (TI change OR AB change OR TI difference OR AB difference)) OR (TI small* OR AB small* AND (TI real OR AB real OR TI detectable OR AB detectable) AND (TI change OR AB change OR TI difference OR AB difference)) OR TI meaningful change OR TI "ceiling effect" OR TI "floor effect" OR TI "Item response model" OR TI IRT OR TI Rasch OR TI "Differential item functioning" OR TI DIF OR TI "computer adaptive testing" OR TI "item bank" OR TI "cross-cultural equivalence" OR TI outcome assessment OR AB meaningful change OR AB "ceiling effect" OR AB "floor effect" OR

AB “Item response model” OR AB IRT OR AB Rasch OR AB “Differential item functioning” OR AB DIF OR AB “computer ad