
This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: http://openaccess.city.ac.uk/20383/

Link to published version: http://dx.doi.org/10.1177/1747493018806200

Copyright and reuse: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.
A core outcome set for aphasia treatment research: the ROMA consensus statement

Sarah J. Wallace¹, Linda Worrall¹, Tanya Rose¹, Guylaine Le Dorze², Caterina Breitenstein³, Katerina Hilari⁴, Edna Babbitt⁵, Arpita Bose⁶, Marian Brady⁷, Leora R. Cherney⁵, David Copland¹, Madeline Cruice⁴, Pam Enderby⁸, Deborah Hersh⁹, Tami Howe¹⁰, Helen Kelly¹¹, Swathi Kiran¹², Ann-Charlotte Laska¹³, Jane Marshall⁴, Marjorie Nicholas¹⁴, Janet Patterson¹⁵, Gill Pearl¹⁶, Elizabeth Rochon¹⁷, Miranda Rose¹⁸, Karen Sage¹, Steven Small²⁰ & Janet Webster²¹.

¹School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Australia.
²School of Speech-Language Pathology and Audiology, University of Montreal, Montreal, Canada.
³Department of General Neurology with Institute of Translational Neurology, University of Münster, Münster, Germany.
⁴School of Health Sciences, City University of London, London, UK.
⁵Department of Physical Medicine and Rehabilitation, Feinberg School of Medicine, Northwestern University, Chicago, USA & Shirley Ryan AbilityLab, Chicago, USA.
⁶School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK.
⁷Nursing, Midwifery and Allied Health Professionals Research Unit, Glasgow Caledonian University, Glasgow, UK.
⁸School of Health and Related Research, University of Sheffield, Sheffield, UK.
9School of Psychology and Social Science, Edith Cowan University, Perth, Australia.

10School of Audiology and Speech Sciences, University of British Columbia, Vancouver, Canada.

11School of Clinical Therapies, University College Cork, Cork, Republic of Ireland.

12Department of Speech, Language, and Hearing Sciences, Boston University, Boston, USA

13Karolinska Institutet, Department of Clinical Sciences, Danderyd Hospital, Division of Internal Medicine, Stockholm, Sweden.

14School of Health and Rehabilitation Sciences, MGH Institute of Health Professions, Boston, USA.

15California State University East Bay, Hayward, USA.

16Speakeasy, UK.

17Department of Speech-Language Pathology, University of Toronto, Toronto, Canada.

18School of Allied Health, La Trobe University, Bundoora, Australia.

19Centre for Health and Social Care Research, Sheffield Hallam University, Sheffield, UK.

20Department of Neurology, University of California, Irvine, USA.

21School of Education, Communication and Language Sciences, Newcastle University, Newcastle Upon Tyne, UK.
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, 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 OMIIs 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 OMIIs (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)

<table>
<thead>
<tr>
<th>Panel Characteristics</th>
<th>n (%)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Country</th>
<th>Count (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>United Kingdom</td>
<td>9 (39)</td>
</tr>
<tr>
<td>United States of America</td>
<td>6 (26)</td>
</tr>
<tr>
<td>Australia</td>
<td>3 (13)</td>
</tr>
<tr>
<td>Canada</td>
<td>2 (9)</td>
</tr>
<tr>
<td>Germany</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Sweden</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Ireland</td>
<td>1 (4)</td>
</tr>
</tbody>
</table>

ICF component to which their own research relates (panel members could nominate more than one component)

<table>
<thead>
<tr>
<th>Component</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body functions</td>
<td>16</td>
</tr>
<tr>
<td>Activity/Participation</td>
<td>21</td>
</tr>
<tr>
<td>Environmental factors</td>
<td>10</td>
</tr>
<tr>
<td>Personal factors</td>
<td>15</td>
</tr>
<tr>
<td>Quality of life*</td>
<td>12</td>
</tr>
</tbody>
</table>

Number of treatment studies published by participants

<table>
<thead>
<tr>
<th>Range</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2-5</td>
<td>8</td>
</tr>
<tr>
<td>6-10</td>
<td>4</td>
</tr>
<tr>
<td>more than 10</td>
<td>7</td>
</tr>
<tr>
<td>not specified</td>
<td>2</td>
</tr>
</tbody>
</table>

*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.
Panel members divided into 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.

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 $\geq 70\%$ 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

<table>
<thead>
<tr>
<th>Construct</th>
<th>Measure</th>
<th>Votes for inclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language</strong></td>
<td><strong>The Western Aphasia Battery Revised (WAB-R)</strong></td>
<td>74% (n=17)</td>
</tr>
<tr>
<td></td>
<td>The Comprehensive Aphasia Test (CAT)</td>
<td>22% (n=5)</td>
</tr>
<tr>
<td></td>
<td>Neither</td>
<td>4% (n=1)</td>
</tr>
<tr>
<td><strong>Communication</strong></td>
<td><strong>The Scenario Test</strong></td>
<td>57% (n=13)</td>
</tr>
<tr>
<td></td>
<td>The Communication Effectiveness Index (CETI)</td>
<td>39% (n=9)</td>
</tr>
<tr>
<td></td>
<td>Abstained</td>
<td>4% (n=1)</td>
</tr>
<tr>
<td><strong>Emotional well-being</strong></td>
<td><strong>General Health Questionnaire (GHQ)-12</strong></td>
<td>83% (n=19)</td>
</tr>
<tr>
<td></td>
<td>Stroke Aphasic Depression Questionnaire (SADQ)</td>
<td>17% (n=4)</td>
</tr>
<tr>
<td><strong>Quality of life</strong></td>
<td><strong>Stroke and Aphasia Quality of Life Scale (SAQOL-39)</strong></td>
<td>96% (n=22)</td>
</tr>
<tr>
<td></td>
<td>Burden of Stroke Scale (BOSS)</td>
<td>0% (n=0)</td>
</tr>
<tr>
<td></td>
<td>Abstained</td>
<td>4% (n=1)</td>
</tr>
</tbody>
</table>

**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 increasingly 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


Supplemental File Search Strategy

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

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 'intrerrater':ab,ti OR 'intragrader':ab,ti OR 'intrarater':ab,ti OR 'intra-rater':ab,ti OR 'intertester':ab,ti OR 'interrater':ab,ti OR 'inter tester':ab,ti OR 'interobserver':ab,ti OR 'intraobserver':ab,ti OR 'interevaluator':ab,ti OR 'intratechnician':ab,ti OR 'intertechnician':ab,ti OR 'interexaminer':ab,ti OR 'intraexaminer':ab,ti OR 'interassay':ab,ti OR 'intraassay':ab,ti OR 'intraindividual':ab,ti OR 'interindividual':ab,ti OR 'intratrace':ab,ti OR 'intratrace':ab,ti OR 'intraindividual':ab,ti OR 'interindividual':ab,ti OR 'kappa':ab,ti OR 'kappas':ab,ti OR 'coefficient of variation':ab,ti OR 'repeatab*:ab,ti OR 'repeatability':ab,ti OR 'repeatability':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 'generalizab*:ab,ti OR 'generalization':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 'subscales':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 un reliab* 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 un reliab* 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 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 intra-technician OR TI inter-examiner OR TI inter-examiner OR TI intra-examiner OR TI inter-assay OR TI inter-assay OR TI intra-assay OR TI inter-individual OR TI inter-individual OR TI intra-individual OR TI inter-participant OR TI intra-participant 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 inter-tester OR AB intratester OR AB intra-tester OR AB inter-observer OR AB inter-observer OR AB intraobserver OR AB intra-observer OR AB inter-technician OR AB inter-technician OR AB intra-technician OR AB inter-examiner OR AB intra-examiner OR AB inter-assay OR AB inter-assay OR AB intra-assay OR AB inter-individual OR AB inter-individual OR AB intra-individual OR AB inter-participant OR AB intra-participant OR AB kappas 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 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
### Supplementary Table 1

#### ROMA consensus meeting facilitators

<table>
<thead>
<tr>
<th>Name</th>
<th>Degree</th>
<th>Expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sarah J. Wallace PhD BSpPath(Hons)</td>
<td>GradCert Gerontology CPSP</td>
<td>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.</td>
</tr>
<tr>
<td>Linda Worrall PhD BSpThy FSPA</td>
<td></td>
<td>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.</td>
</tr>
<tr>
<td>Guylaine Le Dorze Ph.D MSc (A)</td>
<td></td>
<td>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.</td>
</tr>
<tr>
<td>Tanya Rose PhD BSpPath(Hons)</td>
<td>GradCert Higher Ed CPSP</td>
<td>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.</td>
</tr>
</tbody>
</table>

#### ROMA consensus panel

<table>
<thead>
<tr>
<th>Name</th>
<th>Degree</th>
<th>Expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edna Babbitt PhD CCC-SLP BC-ANCDS</td>
<td></td>
<td>Research Speech-Language Pathologist Assistant Research Professor, Department of Physical Medicine and Rehabilitation, Feinberg School of Medicine, Northwestern University, Chicago, USA &amp; Shirley Ryan AbilityLab, Chicago, USA. Expertise: Post-stroke aphasia assessment and rehabilitation.</td>
</tr>
<tr>
<td>Arpita Bose PhD MSc (Speech and Hearing) BSc (Audiology and Speech Rehabilitation)</td>
<td></td>
<td>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.</td>
</tr>
<tr>
<td>Marian Brady PhD BSc</td>
<td></td>
<td>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.</td>
</tr>
<tr>
<td>Caterina Breitenstein PhD academic degrees in Clinical Psychology and Cognitive Neuroscience.</td>
<td></td>
<td>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.</td>
</tr>
<tr>
<td>Leora R. Cherney PhD CCC-SLP BC-ANCDS</td>
<td></td>
<td>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</td>
</tr>
<tr>
<td>David Copland PhD BSpPath (Hons)</td>
<td></td>
<td>Speech Pathologist, Principal Research Fellow, School of Health &amp; Rehabilitation Sciences and Centre for Clinical Research, The University of Queensland, Brisbane, Australia.</td>
</tr>
<tr>
<td>Madeline Cruice PhD BSpPath (Hons)</td>
<td></td>
<td>Registered Speech and Language Therapist, Reader, Teaching and Research Academic, School of Health Sciences, City University of London, London, UK.</td>
</tr>
<tr>
<td>Pam Enderby PhD MBE DSc (Hons)</td>
<td></td>
<td>Speech and Language Therapist, Professor Emeritus of Community Rehabilitation, University of Sheffield, Sheffield, UK.</td>
</tr>
<tr>
<td>Name</td>
<td>Title</td>
<td>Expertise</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>-----------</td>
</tr>
<tr>
<td>Deborah Hersh PhD MSc BSc(Hons) GradCert Higher Ed FSPA</td>
<td>Speech Pathologist, Teaching and Research Academic, School of Medical and Health Sciences, Edith Cowan University, Perth, Australia.</td>
<td>Post-stroke aphasia assessment and rehabilitation, aphasia trial design and conduct, neuroimaging in aphasia.</td>
</tr>
<tr>
<td>Katerina Hilari PhD MRCST MHPC</td>
<td>Psychologist, Registered Speech and Language Therapist, Teaching and Research Academic, School of Health Sciences, City, University of London, UK.</td>
<td>Outcome measurement development, validation and cultural adaptation, post-stroke aphasia rehabilitation, feasibility RCTs, clinical guideline development.</td>
</tr>
<tr>
<td>Tami Howe PhD MHSc BEd SLP(C)</td>
<td>Speech Pathologist and Teaching and Research Academic, School of Audiology and Speech Sciences, University of British Columbia, Vancouver, Canada.</td>
<td>Aphasia rehabilitation, ICF, accessibility, goal setting, social participation, impact of aphasia on family members.</td>
</tr>
<tr>
<td>Helen Kelly PhD MRCST</td>
<td>Registered Speech and Language Therapist, Teaching and Research Academic, Department of Speech and Hearing Sciences, University College Cork, Cork, Ireland.</td>
<td>Post-stroke aphasia assessment and management, single subject and RCT feasibility aphasia trial design and conduct, consumer perspective.</td>
</tr>
<tr>
<td>Swathi Kiran PhD CCC-SLP</td>
<td>Speech Language pathologist, Teaching and Research Academic. Professor, Associate Dean for Research Sargent College of Health and Rehabilitation Sciences, Boston University, Boston, MA, USA.</td>
<td>Aphasia rehabilitation, neuroimaging, bilingualism, single subject experimental design.</td>
</tr>
<tr>
<td>Ann-Charlotte Laska MD A/Professor</td>
<td>Department of Clinical Science Karolinska Institutet Danderyd Hospital, Sweden</td>
<td>Post-stroke aphasia, study design and conduct, RCT.</td>
</tr>
<tr>
<td>Jane Marshall PhD Post Grad Diploma in Clinical Communication Studies BA FRCSLT</td>
<td>Registered Speech and Language Therapist, Teaching and Research Academic, School of Health Sciences, City, University of London, UK.</td>
<td>Post-stroke aphasia rehabilitation, the development and evaluation of novel treatments.</td>
</tr>
<tr>
<td>Marjorie Nicholas PhD CCC-SLP</td>
<td>Professor and Interim Chair Dept. of Communication Sciences and Disorders, MGH Institute of Health Professions, Boston, MA, USA.</td>
<td>Aphasia rehabilitation, nonverbal cognition in aphasia, Life Participation Approach to Aphasia and community aphasia program design, ICAP design.</td>
</tr>
<tr>
<td>Janet Patterson PhD CCC-SLP ASHA Fellow</td>
<td>Chief, Audiology &amp; Speech-Language Pathology Service, VA Northern California Health Care System Practicing Speech-Language Pathologist, Teaching and Research Academic.</td>
<td>Post-stroke aphasia</td>
</tr>
<tr>
<td>Gill Pearl MPhil Dip Hum Commun</td>
<td>Certified practicing speech and language therapist in role as Chief Executive Officer of Speakeasy - specialist aphasia centre, UK.</td>
<td>Development and evaluation of novel approaches to providing long term aphasia support and therapy, facilitator of consumer involvement in aphasia.</td>
</tr>
<tr>
<td>Elizabeth Rochon PhD MSc (A) Reg CASLPO SLP(e)</td>
<td>Speech Pathologist, Teaching and Research Academic, Department of Speech-Language Pathology and Rehabilitation Sciences Institute, University of Toronto, Canada.</td>
<td>Post-stroke aphasia assessment and rehabilitation, development of aphasia treatment.</td>
</tr>
<tr>
<td>Miranda Rose PhD BSpPath FSPA</td>
<td>Speech pathologist, Teaching and Research Academic, School of Allied Health, La Trobe University, Victoria, Australia.</td>
<td>Post-stroke aphasia rehabilitation, aphasia trial design and conduct, single subject designs, consumer perspective, aphasia rehabilitation guideline development.</td>
</tr>
<tr>
<td>Karen Sage PhD Dip DisHumComm BA (Hons) HCPC</td>
<td>Steven L. Small PhD MD Professor of Neurology, University of California, Irvine</td>
<td>Janet Webster PhD MRCSLT Registered Speech and Language Therapist, Teaching and Research Academic, Newcastle University, UK</td>
</tr>
<tr>
<td>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, single case, case series, mixed methods.</td>
<td><em>Expertise: Neurobiology of Language, Cognitive Neurology.</em></td>
<td><em>Expertise: Post-stroke aphasia assessment and management, single subject design.</em></td>
</tr>
<tr>
<td><strong>rehabilitation, systematic reviews of literature, single subject designs.</strong></td>
<td><strong>research, feasibility studies, case series studies, RCT design and conduct.</strong></td>
<td><strong>studies, feasibility studies, single subject and RCT design, systematic reviews.</strong></td>
</tr>
</tbody>
</table>

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, 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

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

<table>
<thead>
<tr>
<th>Outcome instrument and abbreviation</th>
<th>Development / alternate versions</th>
<th>Aims/instrument description</th>
<th>Number of items</th>
<th>Duration</th>
<th>Scoring system</th>
<th>Training</th>
<th>Cost*/availability</th>
<th>Language translations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western Aphasia Battery Revised (WAB-R) (2)</td>
<td>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).</td>
<td><strong>Primary:</strong> 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) <strong>Secondary:</strong> 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,</td>
<td>&gt;300</td>
<td>• 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</td>
<td>• 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.</td>
<td>“some training” required according to developers. Scoring procedures require training.</td>
<td>Testing materials: +++ Available from: <a href="https://www.pearsonclinical.com">https://www.pearsonclinical.com</a></td>
<td>Cantonese (57) Korean (58) Bangla (59) Tagalog (60) Brazilian Portuguese (61) Japanese (62) Hungarian French Turkish (63) Hebrew Spanish (64)</td>
</tr>
<tr>
<td>Stroke and Aphasia Quality of Life Scale (SAQOL-39; SAQOL-39g) (54, 55)</td>
<td>Transcortical sensory, Mixed transcortical, Conduction, and Anomic 2. Assessment of aphasia severity 3. Used to determine the location of the lesion</td>
<td>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</td>
<td>39</td>
<td>• 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</td>
<td>3. Stroke and Aphasia Quality of Life Scale (SAQOL-39; SAQOL-39g) (54, 55) Administration: Guidance is provided in administration guidelines. Administrators need to have skills in communicating with people with aphasia. Scoring procedures: no training required</td>
<td>Free. Available from: <a href="https://blogs.city.ac.uk/cityaccess/saqol-description/">https://blogs.city.ac.uk/cityaccess/saqol-description/</a></td>
<td>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)</td>
<td></td>
</tr>
</tbody>
</table>
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.

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) | 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 | 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

<table>
<thead>
<tr>
<th></th>
<th>Western Aphasia Battery – Revised (WAB-R)</th>
<th>Stroke and Aphasia Quality of Life Scale (SAQOL-39/39g)</th>
<th>General Health Questionnaire (GHQ-12)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objectivity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• During assessment: Limited because no</td>
<td>• During assessment: Moderate (interaction between</td>
<td>• During assessment: High if assessor</td>
</tr>
<tr>
<td></td>
<td>audio recordings of verbal stimulus</td>
<td>assessors and patient frequently required because of</td>
<td>does not interact with patient</td>
</tr>
<tr>
<td></td>
<td>material available</td>
<td>physical stroke symptoms (arm paresis) and lack of</td>
<td>• During scoring: High</td>
</tr>
<tr>
<td></td>
<td>• During scoring: Limited for spontaneous</td>
<td>pictorial task instructions (written sentences only)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>speech and written output subtests</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Internal consistency</strong></td>
<td>High: Cronbach’s alpha of total score= 0.91 (93).</td>
<td>High: Cronbach’s alpha of total score= 0.93; Cronbach’s alpha of subscale scores= 0.74–0.94 (54).</td>
<td>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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SAQOL-39g: High: Cronbach’s alpha of total score= 0.95; Cronbach’s alpha for subscale scores= 0.92-0.95 (55)</td>
<td></td>
</tr>
<tr>
<td>*<em>Test-retest reliability</em></td>
<td>Excellent test-retest reliability: $r &gt;0.90$</td>
<td>Good to excellent test-retest reliability ICC=0.89-0.98</td>
<td>Acceptable to excellent test-retest reliability</td>
</tr>
<tr>
<td></td>
<td>Acute stage post stroke:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Korean version; (58); 5-day test–retest</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>interval (n=20 people with aphasia;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Aphasia Quotient: $r=0.976$; Language</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quotient: $r=0.977$; Cortical Quotient:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$r=0.920$; Spontaneous Speech: $r=0.96$;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Auditory Comprehension: $r=0.967$;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Repetition: $r=0.952$; Naming: $r=0.934$;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reading: $r=0.986$; Writing: $r=0.988$;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Praxis, $r=0.908$; Construction: $r=0.922$.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chronic stage post stroke:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 1 year test–retest interval (97), n=22</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>patients, $r=0.992$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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)
- 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=0.974). 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)

<table>
<thead>
<tr>
<th>Responsiveness</th>
<th>Sub-/acute phase (up to 1 month post-onset):</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>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&lt;0.01). (102)</td>
</tr>
<tr>
<td></td>
<td>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):</td>
</tr>
<tr>
<td></td>
<td>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)</td>
</tr>
<tr>
<td></td>
<td>Post-acute to chronic (3 months to 1 year)</td>
</tr>
<tr>
<td></td>
<td>Cohort study of stroke sample with and without aphasia, n=78. Effect size r=0.22. MID estimated 0.21. (107)</td>
</tr>
<tr>
<td></td>
<td>Chronic phase (at least 6 months post-onset):</td>
</tr>
<tr>
<td></td>
<td>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, (109)</td>
</tr>
</tbody>
</table>

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)
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)

- 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**


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


