Received: 15 August 2022 | Accepted: 8 December 2022

DOI: 10.1111/1460-6984.12840

International Journal of Language & Communication Disorders

RESEARCH REPORT

Measuring communication as a core outcome in aphasia trials: Results of the ROMA-2 international core outcome set development meeting

Sarah J. Wallace[1,2,3] | **Linda Worrall**[1,2,3] | **Tanya A. Rose**[1,2] | **Reem S. W. Alyahya**[5,18] | **Edna Babbitt**[6] | **Suzanne Beeke**[7] | **Carola de Beer**[8] | **Arpita Bose**[9] | **Audrey Bowen**[10] | **Marian C. Brady**[11] | **Caterina Breitenstein**[12] | **Stefanie Bruehl**[13,14,15] | **Lucy Bryant**[16] | **Bonnie B. Y. Cheng**[1,2,3] | **Leora R. Cherney**[6] | **Paul Conroy**[10] | **David A. Copland**[1,2,3] | **Clare Croteau**[17] | **Madeline Cruice**[18] | **Lucy Dipper**[18] | **Katerina Hilari**[18] | **Tami Howe**[19] | **Helen Kelly**[20] | **Swathi Kiran**[21] | **Ann-Charlotte Laska**[22] | **Jane Marshall**[18] | **Laura L. Murray**[23] | **Janet Patterson**[24] | **Gill Pearl**[25] | **Jana Quinting**[26] | **Elizabeth Rochon**[27] | **Miranda L. Rose**[3,28] | **Ilona Rubi-Fessen**[26,29] | **Karen Sage**[30] | **Nina Simmons-Mackie**[31] | **Evy Visch-Brink**[32] | **Anna Volkmer**[7] | **Janet Webster**[33] | **Anne Whitworth**[4] | **Guylaine Le Dorze**[17]

[1]School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, QLD, Australia

[2]Queensland Aphasia Research Centre, Brisbane, QLD, Australia

[3]Centre of Research Excellence in Aphasia Recovery and Rehabilitation, Australia

[4]Communication and Swallowing Disorders Department, King Fahad Medical City, Riyadh, Saudi Arabia

[5]Department of Physical Medicine and Rehabilitation, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

[6]Deparment of Psychology and Language Sciences, University College London, London, UK

[7]Faculty of Linguistics and Literary Studies & Medical School OWL, University of Bielefeld, Bielefeld, Germany

[8]School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK

[9]Geoffrey Jefferson Brain Research Centre, The Manchester Academic Health Science Centre, Northern Care Alliance & University of Manchester, Manchester, UK

[10]Nursing, Midwifery and Allied Health Professionals Research Unit, Glasgow Caledonian University, Glasgow, UK

[11]Department of Neurology with Institute of Translational Neurology, University of Münster, Münster, Germany

[12]St. Mauritius Rehabilitation Centre, Meerbusch, Germany

[13]Medical Faculty, Heinrich-Heine University Duesseldorf, Duesseldorf, Germany

[14]Clinical and Cognitive Neurosciences, Department of Neurology, RWTH Aachen University, Germany

[15]University of Technology Sydney Graduate School of Health, Faculty of Health, Sydney, NSW, Australia

[16]School of Speech–Language Pathology and Audiology, Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal, Faculty of Medicine, Université de Montréal, Montreal, QC, Canada

[17]School of Health and Psychological Sciences, City, University of London, London, UK

[18]School of Audiology and Speech Sciences, University of British Columbia, Vancouver, BC, Canada

[19]School of Clinical Therapies, University College Cork, Cork, Ireland

[20]Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA, USA

[21]Department of Clinical Sciences, Karolinska Institutet Danderyd Hospital, Stockholm, Sweden

[22]School of Communication Sciences and Disorders, Western University, London, ON, Canada

[23]Research Service, VA Northern California Health Care System, Martinez, CA, USA

[24]Speakeasy Specialist Aphasia Centre, Bury, UK

[25]Speech Language Pathology, Department of Rehabilitation and Special Education, Faculty of Human Sciences, University of Cologne, Cologne, Germany

[26]Department of Speech–Language Pathology and Rehabilitation Sciences Institute, University of Toronto, Toronto, ON, Canada

[27]School of Allied Health, Human Services and Sport, La Trobe University, Bundoora, VIC, Australia

[28]RehaNova Rehabilitation Hospital, Cologne, Germany

[29]Faculty of Health and Education, Manchester Metropolitan University, Manchester, UK

[30]Communication Sciences & Disorders, Southeastern Louisiana University, Hammond, LA, USA

[31]Department of Neurology, Erasmus University, Rotterdam, the Netherlands

[32]School of Education, Communication and Language Sciences, Newcastle University, Newcastle upon Tyne, UK

[33]School of Health Sciences, College of Health and Medicine, University of Tasmania, Hobart, Australia

Correspondence
Sarah J. Wallace, School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, QLD, Australia.
Email: s.wallace3@uq.edu.au

Funding information
National Health and Medical Research Council, Grant/Award Number: 1175821; University of Manchester; A. Bowen is supported by National Institute for Health and Care Research (NIHR),Stroke Association and University of Manchester grants.; A. Volkmer is funded byan NIHR Advanced Fellowship.

Abstract

Background: Evidence-based recommendations for a core outcome set (COS; minimum set of outcomes) for aphasia treatment research have been developed (the Research Outcome Measurement in Aphasia—ROMA, COS). Five recommended core outcome constructs: communication, language, quality of life, emotional well-being and patient-reported satisfaction/impact of treatment, were identified through three international consensus studies. Constructs were paired with outcome measurement instruments (OMIs) during an international consensus meeting (ROMA-1). Before the current study (ROMA-2), agreement had not been reached on OMIs for the constructs of communication or patient-reported satisfaction/impact of treatment.

Aim: To establish consensus on a communication OMI for inclusion in the ROMA COS.

Methods & Procedures: Research methods were based on recommendations from the Core Outcome Measures in Effectiveness Trials (COMET) Initiative. Participants with expertise in design and conduct of aphasia trials, measurement instrument development/testing and/or communication outcome measurement were recruited through an open call. Before the consensus meeting, participants agreed on a definition of communication, identified appropriate OMIs, extracted their measurement properties and established criteria for their quality assessment. During the consensus meeting they short-listed OMIs and participants without conflicts of interest voted on the two most highly ranked instruments. Consensus was defined a priori as agreement by ≥ 70% of participants.

Outcomes & Results: In total, 40 researchers from nine countries participated in ROMA-2 (including four facilitators and three-panel members who participated in pre-meeting activities only). A total of 20 OMIs were identified and evaluated. Eight short-listed communication measures were further evaluated

for their measurement properties and ranked. Participants in the consensus meeting ($n = 33$) who did not have conflicts of interest ($n = 29$) voted on the top two ranked OMIs: The Scenario Test (TST) and the Communication Activities of Daily Living—3 (CADL-3). TST received 72% ($n = 21$) of 'yes' votes and the CADL-3 received 28% ($n = 8$) of 'yes' votes.

Conclusions & Implications: Consensus was achieved that TST was the preferred communication OMI for inclusion in the ROMA COS. It is currently available in the original Dutch version and has been adapted into English, German and Greek. Further consideration must be given to the best way to measure communication in people with mild aphasia. Development of a patient-reported measure for satisfaction with/impact of treatment and multilingual versions of all OMIs of the COS is still required. Implementation of the ROMA COS would improve research outcome measurement and the quality, relevance, transparency, replicability and efficiency of aphasia treatment research.

KEYWORDS

aphasia, consensus, core outcome set, recommendations, stroke

WHAT THIS PAPER ADDS

What is already known on this subject

International consensus has been reached on five core constructs to be routinely measured in aphasia treatment studies. International consensus has also been established for OMIs for the three constructs of language, quality of life and emotional well-being. Before this study, OMIs for the constructs of communication and patient-reported satisfaction/impact of treatment were not established.

What this paper adds to existing knowledge

We gained international consensus on an OMI for the construct of communication. TST is recommended for inclusion in the ROMA COS for routine use in aphasia treatment research.

What are the potential or actual clinical implications of this work?

The ROMA COS recommends OMIs for a minimum set of outcomes for adults with post-stroke aphasia within phases I–IV aphasia treatment research. Although not intended for clinical use, clinicians may employ the instruments of the ROMA COS, considering the quality of their measurement properties. The systematic inclusion of a measure of communication, such as TST, in clinical practice could ultimately support the implementation of research evidence and best practices.

INTRODUCTION

A core outcome set (COS) is a minimum set of outcomes that should be measured and reported in research trials of a specific health condition or population (Prinsen et al., 2016). COS use increases compatibility of data across trials, enabling data-pooling, cross-study comparisons and conclusive meta-analyses based on homogenous study designs. The benefits of COS use include increased consistency and relevance of outcomes measured in research (de Wit et al., 2013; Kirkham et al., 2017). The requirement to report agreed outcomes may also act as a deterrent to the selective reporting of outcomes. COS use is increasingly recommended in guidelines, such as the Standard

Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines (Chan et al., 2013) and Cochrane Collaboration Handbook for systematic reviews (McKenzie et al., 2019), and by funding bodies, such as the National Institute for Health Research Health Technology Assessment (NIHR HTA) programme (National Institute for Health Research, 2021) and European Co-operation in Science and Technology (European Commission Horizon, 2015, 2020).

Evidence-based recommendations for a COS for aphasia treatment research have been developed (Wallace et al., 2019). The Research Outcome Measurement in Aphasia (ROMA) COS provides recommendations for the measurement of outcomes for adults with post-stroke aphasia (impaired language/communication) within phases I–IV aphasia treatment studies. Core outcome constructs were identified through a trilogy of international stakeholder consensus studies, which sought the perspectives of people with aphasia and their families (Wallace et al., 2017b), aphasia clinicians and service managers (Wallace et al., 2017a) and aphasia researchers (Wallace et al., 2016). The core outcome constructs in the ROMA COS are: (1) communication, (2) language, (3) quality of life, (4) emotional well-being and (5) patient-reported satisfaction with/impact of treatment.

Potential outcome measurement instruments (OMIs) for each construct area were identified through a scoping review (Wallace et al., 2022). During the first ROMA meeting, consensus was reached for measures of language (The Western Aphasia Battery—Revised—WAB-R), emotional well-being (General Health Questionnaire—GHQ-12); and quality of life (Stroke and Aphasia Quality of Life Scale—SAQOL-39 g, with agreement of 74%, 83% and 96%, respectively; Wallace et al., 2019). Consensus was defined a priori as an agreement on each OMI for each construct by $\geq$ 70% of meeting participants. No consensus was reached at that point for a measure of communication or patient-reported satisfaction with/impact of treatment.

The reason for the lack of consensus for a communication OMI in the original process was unclear. There are a number of possibilities, including the multifactorial nature of this construct (see Doedens & Meteyard, 2020, for a review), the lack of explicit description of what is being tested in some of the communication OMIs, and the lack of a priori agreement about how communication is best defined and operationalized in treatment research. In the present study, we aimed to identify a communication OMI which could be recommended for inclusion in the ROMA COS and routine use in aphasia treatment studies.

METHODS

Research methods were based on the recommendations of the Core Outcome Measures in Effectiveness Trials (COMET) Initiative (Prinsen et al., 2016; Williamson et al., 2012) and are reported in alignment with the Core Outcome Set-STAndards for Reporting (COS-STAR) statement (Kirkham et al., 2016). Guidelines produced by the COMET initiative and COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) (Mokkink et al., 2016) outline four main steps when selecting OMIs for outcomes (i.e., constructs) included in a COS: (1) conceptual considerations; (2) finding existing OMIs; (3) quality assessment of OMIs; and (4) generic recommendations on the selection of OMIs for a COS (i.e., select only one OMI; apply minimum requirements for acceptable content validity, internal consistency (if applicable), and feasibility; use a consensus procedure to obtain final agreement).

Formation of the consensus panel

Researchers (n = 23) who participated in the inaugural ROMA meeting (London, UK, 2016; Wallace et al., 2019) were invited to participate in the current study. In addition to these researchers, an open call for expressions of interest was circulated through The Collaboration of Aphasia Trialists (https://www.aphasiatrials.org/), international professional networks and social media (Twitter). Panel members were sought with expertise in: (1) design and conduct of aphasia trials; (2) measurement instrument development and testing; (3) outcome measurement with particular focus on communication, patient-reported outcomes and satisfaction; and (4) implementation science.

Procedures

Step 1. Conceptual considerations

The ROMA COS is intended for use in phases I–IV aphasia treatment studies for adults with post-stroke aphasia. Although developed for treatment research, the ROMA COS OMIs may be useful inclusions within other types of aphasia research. To mitigate potential issues preventing consensus on a communication OMI in the initial ROMA meeting, we sought agreement on a definition of communication at the outset. Through panel e-mail discussion, agreement was reached to define communication

as per Clarke's (1996) definition of 'situated language use' which is described in the context of aphasia rehabilitation by Doedens and Meteyard (2018). Situated language use is defined as: (1) interactive, that is, involving two or more people; (2) multimodal, for example, using a combination of facial expression, gesture, prosody, speech and body movement to communicate; and (3) reliant on 'common ground', that is, the context which allows a speaker 'to assume a degree of "givenness" of information or directly use physical referents during communication' (Doedens & Meteyard, 2018: 37). This definition was chosen as it recognizes that everyday communication is co-constructed and that language production and comprehension do not occur in isolation. We contextualize the components of this definition with the following examples. In conversation, *interaction* can be exemplified by the use of minimal cues (i.e., nodding, smiling, use of acknowledgements such as 'a-huh', 'mmm')—feedback which allows the listener to signal understanding in the co-construction of meaning. *Multi-modality* recognizes that gesture, facial expression and body movements may supplement and complement and even entirely replace speech to create meaning. For example, when verbalizing the word 'no', a raised eyebrow and rising intonation can shift a speaker's intended meaning from a statement to a question. *Common ground* acknowledges the role of context in constructing meaning. Referring to a photo during conversation, for example, can provide situational context, orienting the listener to a topic and increasing understanding and the ease with which communication occurs (Doedens & Meteyard, 2018).

Step 2. Identifying existing OMIs

In the absence of a published systematic review of communication OMIs for aphasia and in alignment with COMET guidelines, we identified OMIs from an existing scoping review and searched the reference lists of relevant publications to compile relevant information. OMIs were identified through: (1) a scoping review of aphasia measurement instruments (Wallace et al., 2022); (2) systematic reviews of discourse measures (Bryant et al., 2016; Pritchard et al., 2017, 2018); and (3) discussion amongst expert panel members. Panel members were allocated an OMI and extracted evidence regarding the instrument (i.e., type of OMI, aim, items, duration, scoring system, training requirements, cost/availability, language translations; see Table S2 in the additional supporting information) and its measurement properties (internal consistency, reliability, validity, responsiveness; Table 4).

Step 3. Establishing criteria for quality assessment of OMIs

Feasibility. Essential feasibility criteria were established through an online survey with panel members. Participants rated the importance of twelve feasibility criteria (Table 3) using a modified version of the nine-point rating scale developed by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group (http://www.gradeworkinggroup.org). The scale was modified to use the descriptor 'essential' rather than 'critical', as researchers from the United States suggested that the word 'critical' may be interpreted as adverse or disapproving, rather than crucial, by that audience. The criteria related to: (1) availability of translations/adaptations; (2) ability to be easily translated/adapted to other languages; (3) ease of administration; (4) length; (5) completion time; (6) ease of scoring; (7) ease of interpretation; (8) provision of an aggregate score; (9) cost of OMI; (10) equipment or resources requirements for administration or interpretation; (11) applicability to different phases of post-stroke recovery; and (12) requirement for training to be completed in order to administer instrument. The presentation of these criteria was randomized to prevent any order effect. The number and percentage of participants rating each criterion as of limited importance (1–3), important but not essential (4–6) and essential (7–9) were calculated. Consistent with methods employed in ROMA-1 (Wallace et al., 2019) and as suggested by the COMET initiative (Williamson & Clarke, 2012), consensus was predefined as a rating of seven to nine by $\geq$ 70% participants.

Measurement properties. In the absence of existing systematic synthesized information about the properties of communication OMIs in people with post-stroke aphasia, we undertook a best-available evidence synthesis. In the evaluation of the measurement properties of OMIs being considered for inclusion in a COS, the COSMIN group has recommended a predefined order of importance:

1. Content validity (including face validity).
2. Internal consistency.
3. Remaining measurement properties (i.e., test–retest reliability, measurement error, hypotheses testing, cross-cultural validity, criterion validity and responsiveness).

Content validity is considered the most important property of an outcome measure, because if it is unclear what the outcome measure is actually evaluating, the assessment of the other measurement properties may not be of

value. In light of a lack of published evidence regarding the content validity of aphasia OMIs, content validity in the current study was primarily assessed only in terms of face validity. Panel members assessed face validity using the definition of 'situated language use' (Clarke, 1996; Doedens & Meteyard, 2018).

Step 4. International consensus meeting: shortlisting, quality assessment and voting

An international consensus meeting was held during the International Aphasia Rehabilitation Conference in Aveiro, Portugal (September 2018). There was provision for in-person and online participation. Panel members were asked to declare conflicts of interest (e.g., authorship of OMIs under consideration) at the beginning of the meeting. Panel members with conflicts of interest were excluded from voting on construct areas in which they had authored OMIs. A three-stage process was undertaken during the meeting. (1) The panel discussed each OMI, assessing face validity against the definition of 'situated language use' and assessing feasibility against the criteria which had reached group consensus (Table 4). OMIs not meeting these initial criteria were not considered further. (2) Short-listed measures were evaluated against available evidence for: internal consistency, test–retest reliability, responsiveness and feasibility. (3) In light of this evaluation, each participant ranked the short-listed measures and then voted on the top two. Consistent with methods employed in ROMA-1 (Wallace et al., 2019) and as suggested by the COMET initiative (Williamson & Clarke, 2012), consensus was predefined as a vote of yes by ≥ 70% participants. Ranking and voting were conducted online using the survey platform SurveyMonkey (www.surveymonkey.com).

RESULTS

In total, 40 international researchers in aphasia rehabilitation took part in ROMA-2 (Table 1). Included within this number were four facilitators (the first, second, third and last authors) and three-panel members who participated in the pre-meeting activities only. A total of 33 participants took part in the consensus meeting, and 29 participated in voting (four abstained due to conflicts of interest). A comprehensive description of consensus meeting facilitators and panel members is provided in Table S1 in the additional supporting information.

Identification of OMIs

A total of 20 communication OMIs were identified (Table 2). Of these OMIs, five were performance-based outcome measures, four were clinician-reported outcome measures, eight were patient-reported outcomes measures and three were observer-reported outcome measures. While all were identified as measures of communication, some assessed additional constructs.

Consensus on feasibility criteria

Only one feasibility property met the criterion for consensus: 'Able to be easily translated/adapted to other languages' (rated essential by 74% of survey respondents). Completion time (burden on person with aphasia) and applicability to different phases of post-stroke recovery (e.g., hyperacute, acute, subacute and chronic periods post-stroke) were considered essential by 68% of survey respondents, falling short of the 70% requirement for inclusion (Table 3).

Shortlisting of OMIs

A total of 12 OMIs did not meet the criteria for face validity against the definition of situated language, that is, (1) interactive; (2) multimodal; and (3) reliant on common ground), as rated by the expert panel and were excluded: ANELT, AIQ-21, ALA, ACESA, COMACT, SOCACT, CAL, COAST, FOQ-A, TSQ, TOM and discourse analysis.

Quality assessment, ranking and voting

Following quality assessment, The Scenario Test (TST) and Communication Activities of Daily Living—3 (CADL-3) were the most highly ranked OMIs by panel members (Table 4). A total of 29 panel members voted on the two-top ranked OMIs to establish a consensus (defined as ≥ 70%; see the methods section). TST received 72.41% (n = 21) of 'yes' votes and CADL-3 received 27.59% (n = 8) of 'yes' votes.

DISCUSSION

We established consensus on TST (Van der Meulen et al., 2010) as the preferred measure of communication within the ROMA COS. TST is a Dutch OMI based on the

TABLE 1 Nationality and participation of researchers in the consensus process ($n = 40$)

Panel characteristics	*n*
UK	15
Australia	7
United States	5
Canada	5
Germany	4
Sweden	1
Ireland	1
Netherlands	1
Saudi Arabia	1
Role within the panel	
Participant in consensus meeting (in person)	30
Participant in consensus meeting (via videoconferencing)	3
Participant in pre-consensus meeting activities only	3
Facilitators of consensus process	4

TABLE 2 Communication OMIs considered for inclusion in the ROMA COS ($n = 20$)

Outcome measurement instrument	Classification
1 American Speech–Language and Hearing Association Functional Assessment of Communication Skills for Adults (ASHA-FACS)	ClinROM
2 Amsterdam–Nijmegen Everyday Language Test (ANELT)	PerBOM
3 Aphasia Communication Outcome Measure (ACOM)	PROM
4 Aphasia Impact Questionnaire (AIQ-21)	PROM
5 Assessment for Living with Aphasia (ALA)	PROM
6 The Assessment of Communicative Effectiveness in Severe Aphasia (ACESA)	PerBOM
7 The Communicative Activities Checklist (COMACT)	PROM
8 Communication Activities of Daily Living (CADL-3)	PerBOM
9 The Communication Activity Log (CAL)	ObsROM
10 The Communication Confidence Rating Scale for Aphasia (CCRSA)	PROM
11 The Communication Effectiveness Index (CETI)	ObsROM
12 The Communication Outcome After Stroke (COAST)	PROM
13 The Communication Participation Item Bank (CPIB)	PROM
14 The Functional Outcome Questionnaire for Aphasia (FOQ-A)	ObsROM
15 Measure of Participation in Conversation (MPC)	ClinROM
16 The Scenario Test (TST)	PerBOM
17 The Social Activities Checklist (SOCACT)	PROM
18 The Speech Questionnaire (TSQ)	ClinROM
19 Therapy Outcome Measures (TOM)	ClinROM
20 Discourse analysis (e.g., Story Grammar measures; Utterance/propositional level information measures; Correct Information Units)	PerBOM

Notes: See Table S2 in the additional supporting information for references.

PROM = Patient-reported outcome measure: A measurement based on a report that comes from the patient about the status of a patient's health condition without amendment or interpretation of the patient's report by a clinician or anyone else. ClinROM: Clinician-reported outcome measure: A measurement based on a report that comes from a trained health professional after observation of a patient's health condition. ObsROM: Observer-reported outcome measure: A measurement based on observation by someone other than the patient or a health professional. This may be a parent, spouse or another non-clinical caregiver who can observe and report on a specific aspect of the patient's health. PerBOM: Performance-based outcome measure: A measurement based on a task(s) performed by a patient according to instructions that are administered by a health professional (FDA-NIH Biomarker Working Group, 2016).

TABLE 3 Ratings to establish essential feasibility criteria

	Participant ratings ($n = 31$[a])										
	Of limited importance				Important but not essential				Essential			
Feasibility criteria	**1**	**2**	**3**	**1–3 total *n* (%)**	**4**	**5**	**6**	**4–6 total *n* (%)**	**7**	**8**	**9**	**7–9 total *n* (%)**
Available in multiple languages or has been translated into more than one language	0	0	6	6 (19)	3	6	6	15 (48)	7	2	1	10 (32)
Able to be easily translated/adapted to other languages	**0**	**0**	**0**	**0 (0)**	**1**	**1**	**6**	**8 (26)**	**14**	**7**	**2**	**23 (74)** [b]
Ease of administration	0	1	1	2 (6)	3	4	2	9 (29)	8	7	5	20 (65)
Length of outcome measurement instrument	0	1	0	1 (3)	4	2	9	15 (48)	9	4	2	15 (48)
Completion time (burden on person with aphasia)	0	0	0	0 (0)	1	3	6	10 (32)	12	6	3	21 (68)
Ease of scoring	0	1	2	3 (10)	4	5	2	11 (35)	10	3	4	17 (55)
Ease of interpretation	0	0	0	0 (0)	2	4	5	11(35)	7	9	4	20 (65)
Provision of an aggregate score	0	1	0	1 (3)	5	6	5	16 (52)	7	3	4	14 (45)
Cost of outcome measurement instrument	0	2	4	6 (19)	5	4	6	15 (48)	6	2	2	10 (32)
Additional equipment or resources are required for administration or interpretation	1	9	6	16 (52)	1	9	3	13 (42)	1	1	0	2 (6)
Applicability to different phases of post-stroke recovery (e.g., hyperacute, acute, subacute and chronic periods post-stroke)	0	1	0	1 (3)	2	1	6	9 (29)	11	8	2	21 (68)
Requirement for training to be completed in order to administer instrument	1	2	5	8 (26)	3	7	3	13 (42)	6	2	2	10 (32)

Notes: [a] $n = 4$ facilitators did not participate in voting; $n = 5$ panel members did not complete the survey.
[b] Criteria reaching consensus.

TABLE 4 Evaluation of measures meeting minimum criteria in ranked preference order

Rank	Outcome measures	Internal consistency	Reliability	Responsiveness	Feasibility
1	The Scenario Test (TST) (van der Meulen et al., 2010)	+	+	+	+
2	Communication Activities of Daily Living—3 (CADL-3) (Holland et al., 2018)	+	+	n.r. (CADL-2 +)	+
3	The Communication Effectiveness Index (CETI) (Lomas et al., 1989)	+	+	+	+
4	Measure of Participation in Conversation (MPC) (Kagan et al., 2004)	+	?	+	+
5	The Communication Confidence Rating Scale for Aphasia (CCRSA) (Babbitt et al., 2011)	+	n.r.	+	+
6	American Speech–Language and Hearing Association Functional Assessment of Communication Skills for Adults (ASHA-FACS) (Frattali et al., 1995)	+	+	n.r.	+
7	Aphasia Communication Outcome Measure (ACOM) (Doyle et al., 2008)	+	?	?	+
8	The Communication Participation Item Bank (CPIB) (Baylor et al., 2013)	+	n.r.	n.r.	+

Note: +, Positive evidence; ?, indeterminate or conflicting evidence; −, negative evidence.
n.r., Not reported.

Amsterdam–Nijmegen Everyday Language Test (ANELT; Blomert et al. 1994). While both assessments use role-play in daily life situations, the ANELT measures the adequacy of verbal reactions to orally presented scenarios and the TST measures multimodal communication in an interactive setting. TST has 18 items, grouped into six scenarios intended to represent everyday communicative situations (shopping, doctor, taxi, friend, domestic help, restaurant). Each scenario is presented in a pictorial and spoken format and the person with aphasia is asked to respond as if they were in that situation. Responses can be multimodal, and the scenarios are interactive allowing for hierarchical communication support. Scenario responses are video recorded and scored following the assessment using a four-point scale which assesses the amount of information conveyed and the amount of assistance required. TST is currently available in the original Dutch (van der Meulen et al., 2010), and has been adapted into English UK (Hilari et al., 2018), German (Nobis-Bosch et al., 2020) and Greek (Chalarambous et al., 2022). Several further adaptations (e.g., Italian, Japanese, Norwegian) are in progress with the support of the Collaboration of Aphasia Trialists Outcomes and Assessment working group (https://www.aphasiatrials.org/aphasia-assessments/). The only feasibility criterion to reach consensus in this study related to ease of translation/adaptation of OMIs. This reflects the importance of including panel representatives from countries where English is not the dominant language, as well as the inherent challenge of translating aphasia assessments into numerous languages. Importantly, the use of OMIs that can be translated into many languages is crucial for international implementation of the COS and thus for cross-national data pooling and secondary analysis of large data sets in aphasia treatment research. Further details on the test including where it can be purchased are available in Table S2 in the additional supporting information.

TST is valid, reliable and sensitive to change (van der Meulen et al., 2010). Like all OMIs, it has both strengths and limitations. Amongst its strengths are its ecological validity. TST uses everyday scenarios and allows for verbal and non-verbal communication and interaction between the person with aphasia and the interlocuter. TST satisfies the definition of situated language use, drawing on all three components of interaction, multimodality and context. While the use of roleplay is a strength in terms of validity, it is also a potential limitation in terms of imposing cognitive demands that exceed or differ from those in real-life interactions (Doedens & Meteyard, 2020). This assessment format where both the examiner and examinee interact during role play may also be more susceptible to administration variability compared to more structured formats, thus potentially compromising assessment fidelity within aphasia treatment trials (Richardson

et al., 2016). Given the interaction of communication and cognition, the potential for TST results to be confounded by cognitive ability needs to be considered (Wall et al., 2017). However, cognitive demands may confound performance on any communication test, given the intricate interaction of communication and cognition. By the same token, the interactive nature of TST promotes ecological validity, yet may pose a risk to assessment fidelity. Introducing greater structure to such an assessment would undoubtedly have the effect of lessening its ecological validity and in the case of the TST, its inherent value in capturing situated language use. A further strength and limitation of TST is its appropriateness for use with individuals with severe aphasia (Hilari et al., 2018). TST, while supporting inclusion, tests comparatively basic communicative competencies. It was originally designed to measure changes in multimodal communication in patients with moderate to severe aphasia. A ceiling effect has been reported for people with mild aphasia who communicate primarily through spoken modality (van der Meulen et al., 2010), therefore consideration of the optimal way to measure communication with people with mild language impairments is needed. TST shows more variation in people with severe aphasia than the ANELT, and the correlation with the ANELT is higher in people using only verbal communication, compared with people with severe aphasia and a very restricted verbal output (Van der Meulen et al., 2010). An additional limitation of the TST is the lack of differentiation between verbal and nonverbal communication performance. It should however be noted that the ROMA COS represents a set of OMIs (Wallace, 2019) and use of other recommended measures (e.g., WAB-R) may assist with differentiating communicative and linguistic competence.

The recommendation of a core outcome measure should also consider how readily evidence from research can be translated into practice. As consumers of aphasia research, clinicians working with people with aphasia use evidence to inform treatment planning among other aspects of clinical management (Foster et al., 2015). The utility of evidence for these purposes partly depends on whether the OMI captures change in a clinically meaningful way. We are currently establishing clinically relevant benchmarks of change (minimal important change) for TST and all of the ROMA COS measures. This will enable treatment success on these OMIs to be defined from the consumer perspective.

Study limitations

Our recommendation of TST for inclusion in the ROMA COS is based on a rigorous methodology, however, we acknowledge several limitations to our study. The only feasibility criterion considered 'essential' based on our consensus criterion was the requirement that OMIs be easily adaptable across languages. As such, other potentially important criteria, such as applicability across post-stroke timepoints, were not assessed.

Among the shortlisted communication OMIs meeting our minimum criteria (Table 4), some instruments test the performance of the person with aphasia, some use self-rating, and others use the caregiver or clinician's rating as an observer. Not only do these OMIs differ in their method of measurement, but also they differ in the way they were designed. Some were designed by researchers and health professionals and others included people with aphasia in the process. As such, there is variability in both what they measure and how they measure. With this in mind, we recognize that all communication OMIs considered in the present study (Table 2) have distinct merits and may be suited to additional study-specific outcome measurement.

Our expert panel did not include representatives from every country and continent where aphasia treatment research is conducted. However, the inclusion of many international aphasia rehabilitation researchers in the ROMA process serves the goal of dissemination of the ROMA COS and supports the development of the best COS for aphasia rehabilitation research.

Finally, while the early phases of ROMA drew heavily on the lived experience of aphasia to establish consensus on core outcome constructs, we have opted for the inclusion of a single consumer representative in our meetings seeking consensus on core OMIs for pragmatic reasons. We acknowledge that the inclusion of people with aphasia in these meetings could have imbued this process with a different perspective.

CONCLUSIONS

Based on this international, multidisciplinary consensus, we recommend TST to be included in the ROMA COS for routine use in aphasia treatment research. The choice of an OMI for a COS happens at a discrete point in time and is based on the evidence and expert opinion available at that time. Thus, all measures selected for inclusion in the ROMA COS are subject to ongoing review. Importantly, further consideration must also be given to the best way to measure communication in people with mild aphasia. The next step in this program of research is to establish consensus on an OMI capturing patient-reported satisfaction with/impact of treatment for inclusion in the ROMA COS. Multilingual translation, adaptation and implementation of the ROMA COS measures is ongoing, as is work to

establish benchmarks of clinically meaningful change for these OMIs.

ACKNOWLEDGEMENTS

The authors acknowledge the support of the Collaboration of Aphasia Trialists, which is funded by COST and The Tavistock Trust for Aphasia, in fostering international aphasia research collaboration and in the organization and running of the ROMA-2 International Consensus meeting. The contribution of Madeleine Pritchard and Aura Kagan to the ROMA-2 consensus meeting is acknowledged.

CONFLICTS OF INTEREST

E. Babbitt, A. Bowen, C. Breitenstein, S. Bruehl, L.R. Cherney, M. Cruice, L. Dipper, K. Hilari and N. Simmons-Mackie authored or adapted the 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 measures. S. J. Wallace, L. Worrall, G. Le Dorze and T. Rose facilitated this process and did not participate in voting on OMIs.

DATA AVAILABILITY STATEMENT

Supporting data are provided in the supplementary files. Any additional data are available on request from the first author.

ORCID

Sarah J. Wallace https://orcid.org/0000-0002-0600-9343
Edna Babbitt https://orcid.org/0000-0002-8965-7994
Arpita Bose https://orcid.org/0000-0002-0193-5292
Marian C. Brady https://orcid.org/0000-0002-4589-7021
Lucy Bryant https://orcid.org/0000-0001-8497-7406
Bonnie B. Y. Cheng https://orcid.org/0000-0002-3433-8486
Madeline Cruice https://orcid.org/0000-0001-7344-2262
Katerina Hilari https://orcid.org/0000-0003-2091-4849
Swathi Kiran https://orcid.org/0000-0003-1586-8913
Anna Volkmer https://orcid.org/0000-0002-4149-409X

REFERENCES

Ali, M., Soroli, E., Jesus, L.M.T., Cruice, M., Isaksen, J., Visch-Brink, E., et al. (2022) An aphasia research agenda—A consensus statement from the collaboration of aphasia trialists. *Aphasiology*, 36(4), 55.

Babbitt, E.M., Heinemann, A.W., Semik, P. & Cherney, L.R. (2011) Psychometric properties of the Communication Confidence Rating Scale for Aphasia (CCRSA): phase 2. *Aphasiology*, 25(6/7), 727–735. https://doi.org/10.1080/02687038.2010.5373475-574.

Baylor, C., Yorkston, K., Eadie, T., Kim, J., Chung, H. & Amtmann, D. (2013) The Communicative Participation Item Bank (CPIB): item bank calibration and development of a disorder—Generic short form. *Journal of Speech, Language, and Hearing Research*, 56(4), 1190–1208. https://doi.org/10.1044/1092-4388(2012/12-0140

Blomert, L., Kean, M.L., Koster, C. & Schokker, J. (1994) Amsterdam–Nijmegen everyday language test: construction, reliability and validity. *Aphasiology*, 8, 381–407.

Breitenstein, C., Hilari, K., Menahemi-Falkov, M., Rose, M.L., Wallace, S.J., Brady, M.C., et al. (2022) Operationalising treatment success in aphasia rehabilitation. *Aphasiology*, 1–40. https://doi.org/10.1080/02687038.2021.2016594

Bryant, L., Ferguson, A. & Spencer, E. (2016) Linguistic analysis of discourse in aphasia: a review of the literature. *Clinical Linguistics & Phonetics*, 30, 489–518.

Chan, A.-W., Tetzlaff, J.M., Gøtzsche, P.C., Altman, D.G., Mann, H., Berlin, J.A., et al. (2013) SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ: British Medical Journal*, 346, e7586.

Charalambous, M., Phylactou, T., Elriz, T., Psychogios, L., Annoni, J.-M. & Kambanaros, M. (2022) Adaptation of The Scenario Test for Greek-speaking people with aphasia: a reliability and validity study. *International Journal of Language & Communication Disorders*, 57, 865–880. https://doi.org/10.1111/1460-6984.12727

Clark, H.H. (1996) *Using language*. Cambridge: Cambridge University Press.

De Wit, M., Abma, T., Koelewijn-Van Loon, M., Collins, S. & Kirwan, J. (2013) Involving patient research partners has a significant impact on outcomes research: a responsive evaluation of the international OMERACT conferences. *BMJ Open*, 3(5), e002241.

Doedens, W.J. & Meteyard, L. (2018) The importance of situated language use for aphasia rehabilitation. 31234/osf.io/svwp

Doedens, W.J. & Meteyard, L. (2020) Measures of functional, real-world communication for aphasia: a critical review. *Aphasiology*, 34(4), 492–514. https://doi.org/10.1080/02687038.2019.1702848

Doyle, P.J., McNeil, M.R., Le, K. & Et, A.L. (2008) Measuring communicative functioning in community dwelling stroke survivors: conceptual foundation and item development. *Aphasiology*, 22, 718–728.

European Commission Horizon 2020. (2015) Comparing the effectiveness of existing healthcare interventions in the adult population (SCI-PM-10-2017).

FDA-NIH Biomarker Working Group. *BEST (Biomarkers, EndpointS, and other Tools) Resource [Internet] 2016. Food and Drug Administration (US)*. Bethesda (MD): Co-published by National Institutes of Health (US).

Foster, A., Worrall, L., Rose, M. & O'Halloran, R. (2015) 'That doesn't translate': the role of evidence-based practice in disempowering speech pathologists in acute aphasia management. *International Journal of Language & Communication Disorders*, 50, 547–63.

Frattali, C.M., Thompson, C.M., Holland, A.L., Wohl, C.B. & Ferketic, M.M. (1995) The FACS of life ASHA facs—A functional outcome measure for adults. *Asha*, 37, 40–6.

Hilari, K., Galante, L., Huck, A., Pritchard, M., Allen, L. & Dipper, L. (2018) Cultural adaptation and psychometric testing of The Scenario Test UK for people with aphasia. *International Journal of Language & Communication Disorders*, 53(4), 748–760. https://doi.org/10.1111/1460-6984.12379

Holland, A., Fromm, D. & Wozniak, L. (2018) *Communication activities of daily living*, 3rd edition. Austin, Texas: Pro-Ed.

Kagan, A., Winckel, J., Black, S., Duchan, J.F., Simmons Mackie, N., Square, P., et al. (2004) A set of observational measures for rat-

ing support and participation in conversation between adults with aphasia and their conversation partners. *Topics in Stroke Rehabilitation*, 11(1), 67–83. https://doi.org/10.1310/CL3V-A94A-DE5C-CVBE

Kirkham, J.J., Clarke, M. & Williamson, P.R. (2017) A methodological approach for assessing the uptake of core outcome sets using ClinicalTrials.gov: findings from a review of randomised controlled trials of rheumatoid arthritis. *BMJ*, 357, j2262.

Kirkham, J.J., Gorst, S., Altman, D.G., Blazeby, J.M., Clarke, M., Devane, D., et al. (2016) Core outcome SET-STAndards for reporting: the COS-STAR statement. *PLoS Medicine*, 13(10), e1002148.

Nobis-Bosch, R., Bruehl, S., Krzok, F. & Jakob, H. (2020) *Szenario-Test. Testung verbaler und non-verbaler Aspekte aphasischer Kommunikation*. [Szenario-Test. Assessment of verbal and non-verbal communication in aphasia. German test publication]. Cologne: Prolog.

Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A. & Zoghaib, C. (1989) The communicative effectiveness index: development and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of Speech and Hearing Disorders*, 54(1), 113–124.

Mckenzie, J.E., Brennan, S.E., Ryan, R.E., Thomson, H.J., Johnston, R.V. & Thomas, J. (2019) Defining the criteria for including studies and how they will be grouped for the synthesis. *In Cochrane Handbook for Systematic Reviews of Interventions* (eds J.P.T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M.J. Page and V.A. Welch). https://doi.org/10.1002/9781119536604.ch3

Mokkink, L.B., Prinsen, C.A., Bouter, L.M., Vet, H.C. & Terwee, C.B. (2016) The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther*, 20, 105–13.

National Institute for Health Research. (2021) Health Technology Assessment (HTA) Programme stage 1 guidance notes (REALMS).

Prinsen, C.A.C., Vohra, S., Rose, M.R., Boers, M., Tugwell, P., Clarke, M., et al. (2016) How to select outcome measurement instruments for outcomes included in a 'Core Outcome Set'—A practical guideline. *Trials*, 17, 1–10.

Pritchard, M., Hilari, K., Cocks, N. & Dipper, L. (2017) Reviewing the quality of discourse information measures in aphasia. *International Journal of Language & Communication Disorders*, 52, 689–732.

Pritchard, M., Hilari, K., Cocks, N. & Dipper, L. (2018) Psychometric properties of discourse measures in aphasia: acceptability, reliability, and validity. *International Journal of Language & Communication Disorders*, 53, 1078–1093.

Richardson, J.D., Hudspeth Dalton, S.G., Shafer, J. & Patterson, J. (2012) Assessment fidelity in aphasia research. *American Journal of Speech–Language Pathology*, 25, S788–s797.

Van Der Meulen, I., Van De Sandt-Koenderman, W.M., Duivenvoorden, H.J. & Ribbers, G.M. (2010) Measuring verbal and non-verbal communication in aphasia: reliability, validity, and sensitivity to change of The Scenario Test. *International Journal of Language and Communication Disorders*, 45(4), 424–435. https://doi.org/10.3109/13682820903111952

Wall, K.J., Cumming, T.B. & Copland, D.A. (2017) Determining the association between language and cognitive tests in poststroke aphasia. *Front Neurol*, 8, 149.

Wallace, S.J., Worrall, L., LE Dorze, G., Brandenburg, C., Foulkes, J. & Rose, T.A. (2022) Many ways of measuring: a scoping review of measurement instruments for use with people with aphasia. *Aphasiology*, 36(4), 401–466, https://doi.org/10.1080/02687038.2020.1836318

Wallace, S.J., Worrall, L., Rose, T. & Le Dorze, G. (2016) Core outcomes in aphasia treatment research: an e-Delphi consensus study of international aphasia researchers. *American Journal of Speech–Language Pathology*, 25, S729–S742.

Wallace, S.J., Worrall, L., Rose, T. & Le Dorze, G. (2017a) Which treatment outcomes are most important to aphasia clinicians and managers? An international e-Delphi consensus study. *Aphasiology*, 31, 643–673.

Wallace, S.J., Worrall, L., Rose, T., Le Dorze, G., Breitenstein, C., Hilari, K., et al. (2019) A core outcome set for aphasia treatment research: the ROMA consensus statement. *Int J Stroke*, 14, 180–185.

Wallace, S.J., Worrall, L., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., et al. (2017b) 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*, 39, 1364–1379.

Williamson, P. & Clarke, M. (2012) The COMET (Core Outcome Measures in Effectiveness Trials) initiative: its role in improving Cochrane reviews. *Cochrane Database of Systematic Reviews (Online)*, Ed000041.

Williamson, P.R., Altman, D.G., Blazeby, J.M., Clarke, M., Devane, D., Gargon, E., et al. (2012) Developing core outcome sets for clinical trials: issues to consider. *Trials*, 13, 132.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wallace, S.J., Worrall, L., Rose, T.A., Alyahya, R.S.W., Babbitt, E., Beeke, S. et al. (2022) Measuring communication as a core outcome in aphasia trials: Results of the ROMA-2 international core outcome set development meeting. *International Journal of Language & Communication Disorders*, 1–12. https://doi.org/10.1111/1460-6984.12840