City Research Online

City, University of London Institutional Repository

A systematic review of Intensive Comprehensive Aphasia Programmes – who takes part, what is measured, what are the outcomes?

Katie Monnelly, Jane Marshall, Lucy Dipper & Madeline Cruice

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

🔓 OPEN ACCESS   Check for updates

A systematic review of Intensive Comprehensive Aphasia Programmes – who takes part, what is measured, what are the outcomes?

Katie Monnelly 🔟, Jane Marshall 🔟, Lucy Dipper 🔟 and Madeline Cruice 🔟

Department of Language and Communication Science, City, University of London, London, UK

ABSTRACT

Purpose: This study synthesizes participant and outcome data from peer-reviewed Intensive Comprehensive Aphasia Programme (ICAP) studies.

Methods: A systematic review was conducted following PRISMA guidelines. Study eligibility criteria were specified in relation to population, intervention, comparison, outcome, and design considerations. Data were extracted according to six research questions. Narrative synthesis was used.

Results: Twenty-one studies were included covering 13 ICAPs ($N = 485$, aged 18–86 years, between 11 and 335 months post-stroke). Twenty-seven participant selection criteria were identified. Fifty-six outcome measures spanning the WHO-ICF were used, with the majority assessing the body function domain. Only eight studies employed an experimental design with data appropriate for analysis and synthesis. Risk of bias was noted across this sub-group. Participants improved in word-finding, communication, activity/participation, and communication-related quality of life, and maintained their gains; however, except for word finding, evidence of effect came from isolated studies. Factors influencing outcomes were rarely considered. Some drop-outs, missed sessions, and fatigue were noted. Some studies reported IPD alongside group analyses.

Conclusions: ICAP selection criteria need justification and should contribute to the understanding of candidacy for this treatment model. Rationalisation of ICAP treatment content and outcome measurement is required, spanning all WHO-ICF domains. Employment of the core outcome set for aphasia would enable data synthesis and facilitate comparisons between the ICAP and other therapy models.

> **IMPLICATIONS FOR REHABILITATION**
> - Healthcare professionals can use this review to appreciate that the evidence base for intensive and comprehensive aphasia programmes is emerging and based on studies of varying methodological quality and thus findings are not conclusive.
> - Patients across the lifespan and across a range of aphasia severities, and patients who are independent or have support for activities of daily living, can participate in intensive and comprehensive aphasia programmes.
> - Patients can expect improved word finding ability from participation in an intensive and comprehensive aphasia programme, and some patients can experience benefits in functional communication, communication confidence, and aphasia-related quality of life.
> - Outcome measurement from intensive and comprehensive aphasia programmes should encompass language functioning, communication activities/participation, quality of life, and outcomes for family members, and ideally environmental and personal factors should be considered.

Introduction

Aphasia is an acquired language disorder that usually occurs because of a stroke. Its most obvious consequence is impaired expressive language, but in most cases, aphasia also affects the ability to understand, read, and write. As our world is mediated by language, there are significant secondary consequences of aphasia including exclusion from social interactions [1] and work [2]. These consequences and others have a negative impact on the mental health of people with aphasia (PWA) [3,4]. Families of PWA experience a high level of caregiver burden [5] exacerbated by poor public awareness of aphasia, which belies its prevalence [6]. Aphasia is usually a chronic disorder.

The diverse symptoms of aphasia, and complex secondary impacts pose a challenge for intervention. There is level 1 evidence that speech and language therapy (SLT) is effective for aphasia [7]. In many studies, treatment aims to remediate the language impairment, by, for example improving word retrieval [8]. However, attempts to improve compensatory communication skills have also been positively evaluated [9] as have psychological interventions addressing the social and emotional needs of PWA [10]. Family interventions, tackling the needs of caregivers have also been explored [11].

Intensive Comprehensive Aphasia Programmes (ICAPs) aim to address the complex needs of PWA in one treatment model. ICAPs

combine intensive impairment-based language therapy with a range of other therapy approaches, which fall under the "comprehensive" umbrella. A scoping review [12] found that about 50% of ICAPs are pure SLT in focus while the other 50% provide input from other professionals, e.g., counselling/occupational therapy.

ICAPs are premised on two assumptions. The first relates to dose. There is increasing evidence that positive outcomes, particularly with respect to the language impairment, are contingent on the amount of treatment provided [13]. Efforts are being made to separate out the importance of overall dose of intervention versus the importance of intensity with which that dose is delivered [14–16]. While we do not yet have the answers to these important questions, what is clear is that PWA worldwide receive low overall dose of SLT and that this is linked to inadequate outcomes [17–19]. For example, a recent survey of 227 UK SLTs found that the average weekly dose was 2 h a week [20]. Aphasia intervention in the UK does not typically extend beyond three months post-stroke and most intervention is weighted towards the acute stage [21]. An ICAP provides a minimum of 30 h total dose of therapy, delivered at an intensity of 3 h a day × 5 days a week [22]. This is substantially more than the typical clinical offering internationally. Provision of high intensity stroke rehabilitation models of care is a core part of the UK National Health Service long-term plan for stroke care [23].

The second underpinning assumption of the ICAP model is that aphasia intervention should be comprehensive. The International Classification of Functioning, Disability and Health (ICF) is a useful framework when considering the comprehensiveness of an ICAP [24]. ICAP interventions address both body functions (the aphasia impairment, i.e., language therapy), and the domains of activities/participation, and environmental factors. The goals of PWA and their families are known to focus more heavily on activities and participation rather than impairment-based interventions [25,26], so ICAPs are a welcome shift of focus towards more comprehensive aphasia therapy delivery. It is recommended to include families/carers of PWA in ICAPs [22], which is also welcome given their needs and the level 1 evidence for interventions that address communication between conversation partners [27].

Despite the potential positives of ICAP intervention, concerns remain. These were raised in a scoping review of 17 ICAPs [12]. The most obvious concern was the variability evident between ICAPs. ICAPs were found to interpret the comprehensive component differently and treatments for individual participants were highly tailored. Although this may have resulted in a positive individual experience for ICAP participants, it made it challenging to identify the active ingredients of an ICAP. The variation of treatment content has implications for the synthesis of outcome data, particularly if varying measures have been used to reflect that content.

Interest in ICAP research is growing with at least nine new publications since the release of the 2021 scoping review [12]. While individual studies have published positive outcomes from ICAPs, there has been no formal synthesis of ICAP outcomes. PWA and their families want to know what aphasia interventions are best [28], and outcome studies are necessary to answer those questions. This study is the first systematic review of outcomes from ICAP interventions.

ICAPs are a demanding therapy model. They require intensive attendance and, in most cases, the engagement of family members. Comprehensive treatments are, by definition, diverse, so address multiple skills. It may well be that not all PWA can profit from this approach. This review, therefore, also explored the selection of participants for ICAP studies and the nature of ICAP study samples. While very preliminary, these data may begin to illuminate candidacy for the ICAP model.

Methods

Research questions

Six research questions (RQs) were devised for the review:

1. What inclusion/exclusion criteria were listed for ICAPs?
2. Who participated in an ICAP?
3. What outcome measures were used on ICAPs?
4. Have ICAPs brought about significant improvements on outcome measures, and if so in what domains of the ICF?
5. What factors were reported to influence ICAP outcomes?
6. Are there any reported negative outcomes from ICAPs?

Review methodology

After the questions were devised, it was decided that a two-stage study selection procedure was required. Research questions 1–3 enquired about breath of knowledge and were not felt to require an assessment of methodological quality. However, RQs 4 and 5 and aspects of RQ6 enquired about outcomes. Therefore, an assessment of methodological quality was required to ensure outcomes reported were based on high quality data that could be synthesized in a meta-analysis.

The Cochrane group Cochrane Effective Practice and Organisation of Care (EPOC) [29] state is it "difficult if not impossible to attribute causation" from uncontrolled before-after studies, i.e., those without an experimental design. Given this state of the evidence, the authors decided to exclude uncontrolled before-after studies from inclusion in answering RQs 4 and 5 (those studies termed pre-post with only one data point pre-intervention).

A systematic review was conducted, but a risk of bias tool was only applied to studies with an experimental control, which were used to address RQs 4 and 5 (Table 1).

Search strategy

An initial literature search was conducted between 11 December 2019 and 18 December 2019 and covered database inception to the review dates. This search strategy and the protocol were published in 2021 [12]. The 20 results from the 2021 scoping review were considered for inclusion in the current review. Of the 20, 12 were included in the current review and eight were excluded. Three results were excluded as they reported on qualitative outcomes only [30–32]. Three were excluded as they were conference abstracts [33–35]. One was excluded as outcomes for ICAP and non-ICAP participants could not be separated [36]. One was excluded as the content did not relate to ICAP outcomes [37].

An updated search using the same protocol was conducted on 22 August 2022 and 23 August 2022 and covered 1 December 2019 to the review dates. Nine new studies were included in the current review. Figure 1 shows a PRISMA flow diagram [38] with details of the updated search and how 2021 scoping review findings were combined.

The first author conducted both searches, importing all references to RefWorks (ProQuest, Ann Arbor, MI). Inclusion and exclusion criteria developed in the protocol were used to make inclusion/exclusion decisions on abstracts. The first author made these decisions on the first 20 articles and this process was supervised by MC and JM. Thereafter, the first author made these decisions on all abstracts. MC and JM made independent inclusion/exclusion judgements on 10% of abstracts each (i.e., a total of 20%) as a quality control measure and as there was inadequate resource for two authors to

Table 1. PICOD question for the review.

PICOD criteria	Exclusion criteria
Population: People with aphasia 18+ years old or their family members/carers/equivalent.	People with primary-progressive aphasia or aphasia linked to a progressive neurological condition. There were no other limitations on aphasia presentation and no limitations on presentations of family members.
Intervention: ICAPs. ICAPs were defined by the following criteria. (1) Intensity/dose: at least 3 h/day × 5 days/week × 2 weeks or a dose of 30 h in 2 weeks. (2) Individual and group therapy provided. (3) Targeting more than the language impairment (e.g., targeting activities/participation). "Education" was not required as a core intervention component based on findings of a previous scoping review [12].	Programmes which fell below the intensity criteria. Therapy programmes which were intensive but only focused on one main approach (e.g., constraint induced language therapy) or programmes which were delivered to people with acquired brain injury where the primary focus was on cognitive rehabilitation rather than language/communication/aphasia.
Comparison: A comparison group was not required for inclusion in the study to answer RQs 1–3. However, an experimental control was required to answer RQs 4–5 (e.g., a repeated measures study/multiple baselines/interrupted time series or a control group where the control was a comparator intervention/usual care/control group, etc.).	n/a
Outcome: Any quantitative outcomes reported from ICAPs were considered eligible for inclusion in the review.	Studies that only reported qualitative outcomes, e.g., themes from focus groups or interviews or qualitative quotes reported in studies.
Design: Quantitative methodologies, peer reviewed, full text available, English language, search from database inception to present (i.e., no date limit)	No qualitative methodologies, no grey literature, no conference abstracts.

independently screen abstracts. Full texts for inclusion were selected by the first author and all full texts were independently verified for exclusion/inclusion by a combination of MC, JM, and LD with 100% consensus on the final full texts selected for inclusion.

Data sources

The review only considered articles in English due to the abilities of the reviewers and the lack of resource for translation. It was unclear how many papers had to be excluded on this basis as non-English language papers were screened out at the point of initial searching. All peer-reviewed study designs were accepted for inclusion and articles from database inception to the present were searched. The context included articles from all settings and geographical locations.

Patient public involvement (PPI)

Two family members of PWA who had attended ICAPs were offered payment for their expertise in reviewing the draft data extraction tool and the review questions (JS and PS). Additional questions added were 1. Who does an ICAP work for? This question was incorporated into RQ 5, and 2. Are there any reported negative side-effects to ICAPs? This question was added as RQ6. A PPI group of four PWA was conducted before the review

commenced. These individuals did not alter the review questions and were most interested in creating aphasia-friendly accessible versions of the results.

Synthesis

The results were not suitable for meta-analysis due to the heterogeneity of outcome measures used and methods of data analysis. Therefore, synthesis without meta-analysis (SWiM) guidance was used [39]. Possible methods of data synthesis, e.g., combining *p* values were not deemed feasible following consultation with a statistician. Therefore, data were synthesised by ICAP. For outcome measurement questions, outcome measures were categorised using the ICF domains using a recently published scoping review [40]. As previously noted, synthesis was done on study type with only experimentally designed studies answering RQs 4–6.

Results

The search identified nine studies to contribute to the current study, see PRISMA flow diagram in Figure 1. Combined with 12 studies from the 2021 scoping review, this produced a total of 21 papers that reported quantitative outcomes from ICAPs. These papers represented 13 ICAPs that were listed by name and location (see Table 2). In some instances, more than one paper from an ICAP was used to address a question. This is because secondary reports added relevant information. Table 2 indicates which papers contributed to each RQ.

Characteristics of papers

Of the 21 papers, two papers were non-randomised pre-post trials with comparator groups [53,54], which is NHMRC (National Health and Medical Research Council) [62] evidence level III-2. Seven were repeated measures/interrupted time series studies without comparator groups [41,44,45,49,55,57,61], which is NHMRC evidence level III-3, and the remaining 12 were pre-post studies with no comparator groups or NHMRC evidence level IV [42,43,46–48,50–52,56,58–60].

Non-randomised studies of intervention

The Cochrane Collaboration provides direction on using non-randomised studies of intervention (NRSIs) in systematic reviews [63]. They advise caution when ranking NRSIs according to evidence hierarchies, which were not designed for effectiveness questions. Additionally, they recommend not making inclusion/exclusion decisions based on study design labels, but instead advise use of a checklist of study design features [64] to assess study quality.

Study design features checklist

As advised by Reeves et al. [63], a study design features checklist was applied to the nine (9) studies at III-2 or III-3 NHMRC evidence levels. The studies showed many features in common, e.g., the intervention effect was estimated by change over time in all studies, so they were felt to be adequately alike to compare (see study design features checklist in Appendix 1). These studies all underwent quality appraisal.

Figure 1. PRISMA flow diagram.
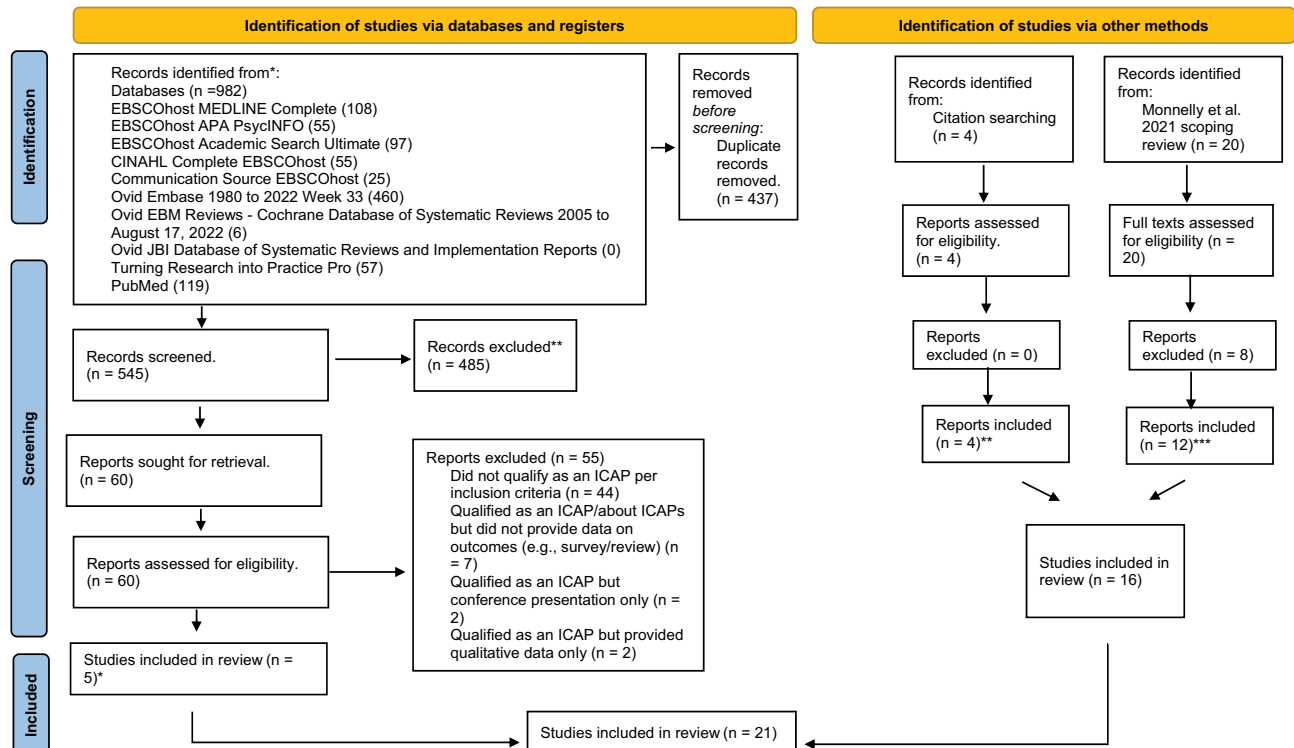
* Griffin-Musick et al. 2022, Griffin-Musick et al. 2021, Auclair-Ouellet et al. 2022, Leff et al. 2021, Nicholas et al. 2022
** Babbitt et al. 2016, Baliki et al. 2018, Wilson et al. 2012, Doogan et al. 2022
*** Code et al. 2010, Hoover et al. 2017, Escher et al. 2018, Babbitt et al. 2015, Mackenzie, 1991, Persad et al. 2013, Rodriguez et al. 2013, Dignam et al. 2015, Dignam et al. 2016, Brindley et al. 1989, Winans-Mitrik et al. 2014, Hinckley & Craig, 1998

Table 2. Mapping papers to research questions (RQs).

ICAP name	Paper #	Authors	NHMRC evidence hierarchy[a]	RQ1	RQ2	RQ3	RQ4	RQ5
1. Belfast	1	Code et al. [41]	III-3	X	X	X	X	X
2. Big Sky Montana	2	Griffin-Musick et al. [42]	IV		X	X		
	3	Griffin-Musick et al. [43]	IV	X		X		
3. Boston	4	Hoover et al. [44]	III-3	X	X	X	X	X
	5	Escher et al. [45]	III-3	X		X	X	X
4. Chicago	6	Babbitt et al. [46]	IV	X		X		
	7	Babbitt et al. [47]	IV		X			
	8	Baliki et al. [48]	IV			X		
5. Glasgow	9	Mackenzie [49]	III-3	X	X	X		
6. InteRACT	10a	Persad et al. [50][b]	IV	X	a	a		
	11	Wilson et al. [51]	IV			X		
7. LIFT	12	Rodriguez et al. [52]	IV	X		X		
	13	Dignam et al. [53]	III-2		X	X	X	X
	14	Dignam et al. [54]	III-2	X		X	X	X
8. Milton Keynes	15	Brindley et al. [55]	III-3	X	X	X	X	X
9. Montreal	16	Auclair-Ouellet et al. [56]	IV	X	X	X		
10. PIRATE	17	Winans-Mitrik et al. [57]	III-3	X	X	X	X	X
11. Queen Square	18	Leff et al. [58]	IV	X		X		
	19	Doogan et al. [59]	IV	X	X	X		
12. S-IHP's CAP Boston	20	Nicholas et al. [60]	IV	X	X	X		
13. UMAP	21	Hinckley and Craig [61]	III-3	X		X	X	X
	10b	Persad et al. [50][b]	IV		a	a		
		Papers answering this research question		16	13 Persad counted twice[a]	21 Persad counted twice[a]	8	8

[a]NHMRC levels: III-2 "a comparative study with concurrent controls"; III-3 "interrupted time series without a parallel control group"; IV "case series with pre-post-test outcomes".
[b]Persad listed twice as it reports both on InteRACT and UMAP.

Quality appraisal results

Choosing quality appraisal tools for NRSIs is a challenge as there is a lack of guidance in this area. A narrative review on methodological quality (risk of bias) assessment tools [65] advised use of the EPOC RoB (risk of bias) Tool for interrupted time-series studies [66], but the interrupted time-series/repeated measures studies included in the present review did not contain three data points pre and post and therefore the EPOC RoB Tool was not designed

for these studies. Ma et al. [65] suggest the National Heart, Lung, and Blood Institute (NIH) quality assessment tool for before-after (pre-post) studies without control group [67]. However, the studies contained two data points pre and post and sat somewhere in between a "repeated measures/interrupted time series" study design and above a "pre-post" design with no double baseline. Cochrane guidance [68] advises use of the Risk of Bias in Non-randomized Studies of Interventions (ROBINS-I) [69] and therefore the ROBINS-I tool was applied to the nine experimental design studies (see Table 3). Two raters independently rated the studies – KM and FS. KM tended to err on the side of caution rating "no information" whereas FS assigned "low risk" to these decisions for a variety of reasons, e.g., making a clinical judgement on lack of confounding variables or likely lack of deviation from intended intervention. A decision was made to rate these domains as "no information", though "low risk" was the likelihood if the information had been provided by the study authors. Other discrepancies in ratings were discussed between raters after the rating process with the final decision made with the input of a third rater (MC). A scoping review found about 50% of ICAPs provide input from SLTs only, but 50% supplement SLT with input from other professionals [12]. This was the case for two experimentally designed studies, which provided counselling [41] or physiotherapy/occupational therapy/nutrition [44], and one study that reported on the occupational therapy input into an ICAP [45]. It was decided not to judge non-SLT input as a confounding variable since there is no ICAP theory that states the intervention must be SLT only. Additionally, the outcomes reported in papers were SLT specific, rather than global measures that could have been influenced by input from other professionals.

As reported in Table 3, there was evidence of risk of bias across all studies particularly with likely lack of blinding to measurement of outcomes and the possibility of bias in reporting due to lack of pre-published protocols. Six/nine studies were rated moderate risk of bias for these reasons, but all analyses were clearly defined and consistent within papers. Three/nine studies had serious

overall risk of bias due to either allocation of participants to intervention [53,54] or classification of the intervention and issues in reporting outcomes [49]. The Mackenzie paper [49] only presented therapy outcomes for individuals as a visual summary of raw data with no group or statistical analysis. This meant the data could not reliably be interpreted and were not comparable with the other eight papers. Therefore, this paper was excluded from answering RQs 4 and 5.

Research questions

1. What inclusion/exclusion criteria are listed for ICAPs?
Data from 16 papers contributed to this question (see Table 2). The findings are listed in Table 4, from most to least frequently cited criteria and detail on specific ICAPs listing these inclusion/exclusion criteria is noted.

2. Who participates in an ICAP?
Data from 13 papers answered this question (i.e., one paper per ICAP). When there was more than one publication for an ICAP, the paper with the largest number of reported participants for that ICAP was selected. See details in Table 2 and Appendix 2.

A recent e-Delphi to establish reporting standards for aphasia participants called the DESCRIBE project listed 11 criteria, which should be reported for PWA in aphasia research [70]. Eight of the 11 criteria were reported in studies and can be found in Appendix 2. Criteria that were not reported were "language of treatment", which was English in all studies and "primary language [spoken by the PWA]", which was likely to have been English for most participants given this was the language of treatment. "Conditions arising from the neurological event" are partially reported in Appendix 2, e.g., aphasia severity/type/motor speech disorders, but the ICAP studies in this review did not provide extensive information on additional sensory/motor/psychological side-effects for each participant.

Table 3. ROBINS-I risk of bias ratings on nine experimental design studies (low, moderate, and serious).

	Bias due to confounding	Bias in selection of participants into the study	Bias in classification of interventions	Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported result	Overall risk of bias
1. Brindley et al. [55]	No information	Low	Low	No information	Low	Low or moderate[e]	Moderate	Moderate
2. Code et al. [41]	No information	No information	No information	No information	Low	Low or moderate[e]	Moderate	Moderate
3. Dignam et al. [53]	No information	Serious[a]	Low	No information	Low	Low or moderate[e]	Moderate	Serious
4. Dignam et al. [54]	No information	Serious[a]	Low	No information	Low	Low or moderate[e]	Moderate	Serious
5. Escher et al. [45]	No information	Low	Low	No information	Low	Moderate risk[f]	Moderate	Moderate
6. Hinckley and Craig [61]	No information	Low	Low	No information	Low	Low or moderate[e]	Moderate	Moderate
7. Hoover et al. [44]	No information	Low	Low	No information	Low	Low or moderate[e]	Moderate	Moderate
8. Mackenzie [49]	No information	Moderate[b]	Serious[c]	No information	Low	Moderate[g]	Moderate or serious[h]	Serious
9. Winans-Mitrik et al. [57]	No information	Low	Low	No information	Low[d]	Low or moderate[e]	Moderate	Moderate

[a]"Group allocation was based on individuals' geographic location, the availability of a position within the research programme and personal factors" (page 188).
[b]"Additional referral criteria were that subjects were considered to have high motivation and sufficient stamina for a daily programme" (page 277).
[c]Intervention status not well defined (who got what why).
[d]The analysis addressed missing data.
[e]These studies report on outcome methods but do not report if the assessors were blinded. If blinded, the risk would be low. If not blinded, the risk would be moderate.
[f]The SIS was completed "with family assistance" (page 3), which would have minimally influenced the outcome measure.
[g]Study reports the assessor was not involved in therapy but was not blinded to the participants having had therapy.
[h]Did not have an *a priori* plan (moderate) but does not fit the descriptor for seriously biased.

Table 4. RQ1 what inclusion/exclusion criteria are listed for ICAPs?

	Criteria (ICAPs listing this criterion)	Number of ICAPs listing this as inclusion/exclusion criteria
1	Time post-onset	7
	>6 MPO (months post-onset) [44,52,55,60,61]	
	>3 MPO [59]	
	"Recency of onset" unspecified detail [46]	
2	Independent for ADLs (activities of daily living)	6
	PWA must be independent for ADLs [45,57,59,61]	
	or caregiver must be present if not (e.g., for walking/eating/toileting) [46,60]	
3	Endurance/stamina/tolerance [43,45,46,49,57,60]	6
4	Medically stable	4
	Generally medically stable [43,57,61]	
	Permission from physician [46]	
	Stable mental health [57]	
5	Motivation/commitment to attend	3
	Motivation [49,57]	
	Commitment [58]	
6	Over 18 [43,59,60]	3
7	Caregiver involvement and support, e.g., for carryover of skills [50,57] *(Persad for InteRACT)*	2
8	No cognitive deficits predicted to limit a candidate's potential to participate in or benefit from treatment [43,57]	2
9	Sufficient auditory comprehension skills to provide informed consent/participate [44,49]	2
10	Aphasia due to a single stroke [44,52]	2
11	No significant behavioural issues [46,57]	2
12	No history of other neurological disorders [44,52]	2
13	No vision and hearing deficits that would impact engagement [52,61]	2
14	No severe motor speech disorders [54,59]	2
15	English speakers [60,61]	2
16	Own transportation to/from ICAP [59,60]	2
17	Current rehab needs/presence of significant aphasia that affects participation in life activities [43,60]	2
18	Compatibility with cohort/group	2
	Fit of PWA within cohort [43]	
	Comfortable doing group activities [60]	
19	Access to therapy and resources currently [43]	1
20	Potential for improvement [46]	1
21	Previous intensive experience [43]	1
22	Upper age limit of 65 [55]	1
23	Diagnosed with aphasia by SLP/language battery [44]	1
24	Severity of aphasia (at both ends of the spectrum) [unspecified] [46]	1
25	No history of learning difficulties [44]	1
26	Not participating in other rehabilitation therapies during the programme [60]	1
27	Access restricted to veterans and active-duty military personnel [57]	1

Summary of findings: Four hundred and eighty-five individuals were reported on in the 13 ICAPs including 333 males, 125 females, and 27 with no data on sex. Age and time post-onset were reported by all studies and thereafter reported characteristics varied (see Appendix 2). Age range across studies was 18–86. Eleven studies reported mean age. It was not possible to devise a mean age of the average ICAP participant by combining and averaging means. The mean was mid-50s (54/55/56) in 6/11 studies [44,47,50,57,60] (with Persad representing both InteRACT and UMAP). Two studies had mean ages that were higher [42,53] and three had mean ages that were lower [41,55,59]. Months post-onset ranged from 1 to 335. It was not possible to select a mean time post-onset across all studies, but 4/5 of the largest studies (*n* = 52-83 participants) had a mean time post-onset of 15–24 months. Aphasia severity across studies ranged from mild to severe aphasia. Two participants exceeded the non-aphasic cut-off but were included in one study [60]. Where average pre-treatment scores were given, these were either moderate or severe. All aphasia types were reported apart from jargon aphasia. Broca's or nonfluent aphasia were the most common classifications. Stroke was the most common aetiology, but a variety were reported (e.g., head injury). A variety of educational levels were also reported. Mean years of education where reported was at least 14 years. Whether or not participants had attended a previous ICAP was reported in only four studies. These indicated that

rates of ICAP repetition varied from 4 to 40% [42,57]. Where reported, most participants were white/Caucasian with highest prevalence of 96% [47] to lowest prevalence of 77% [60]. Black/African American participants made the second largest demographic with as many as 17% [60] to as little as 3% of sample [47]. Motor speech disorders including apraxia of speech (AOS) and dysarthria were not often reported, but AOS was present in 55% of participants attending the Chicago ICAP [47]. Languages spoken by participants were infrequently noted. Where reported, the most common language was English. Four studies reported on occupation. Two studies [47,53] reported that their participants were predominantly (94%) right-handed.

3. What outcome measures are used on ICAPs?

Data from 21 papers answered this question. Measures were categorised according to ICF domain following a recently published scoping review [40]. Categorisation was performed by the first author and ratified by the full authorship team. There was one deviation from the scoping review that classified the CAT under "multiple ICF categories", likely due to the disability questionnaire section of the CAT. However, in ICAP research, only the language battery of the CAT has been used, not the disability questionnaire. Therefore, we categorised the CAT as a "comprehensive language test" under "body functions". Additionally, the scoping review [40]

did not categorise discourse measures. Discourse measures were allocated to ICF domains in the current study as three authors (MC, JM, LD) have authored a systematic review on discourse treatments in aphasia [71]. See categorisation of measures in Appendix 3.

There were 56 different outcome measures used on the 13 ICAPs. Each ICAP used between 2 and 12 outcome measures but did not always report the results of all measures administered. Thirty-eight measures were in the domain of "body functions" (33 standardised and five bespoke). Ten measures captured "activities/participation" (nine standardised and one bespoke). Of the remaining eight measures, three covered multiple ICF categories, three measured quality of life, and two were neurological measures. The three most popular measures were the CETI (a measure of activities/participation completed by a significant other) used by seven ICAPs, the WAB-R (a comprehensive test of language) used by six ICAPS, and the BNT (a naming test) used by five ICAPs.

All 13 ICAPs measured "body functions". Twelve ICAPs measured "activities/participation". Two ICAPs measured "environmental factors" and "personal factors" *via* the same single assessment – the Assessment of Living with Aphasia [72]. Four ICAPs measured quality of life. Two ICAPs incorporated neurological measures, e.g., imaging.

4. Have ICAPs brought about significant improvements on outcome measures, and if so in what domains of the ICF?

To answer this question, it was decided to separate reported outcomes by evidence level. Only evidence at level III-2 or III-3 was considered (see Table 2) and ultimately eight papers of sufficient methodological quality answered this RQ covering six ICAPs (see Table 3).

There were 33 different outcome measures used on the six ICAPs. Despite 33 measures being listed, only data from 19 measures were available for analysis at evidence level III. There were three reasons for this: (1) outcome measures were listed with no data provided (these measures are highlighted in red in Appendix 3); (2) some measures only had pre-post data provided, so they did not reach level III evidence; and (3) on one measure, results for ICAP and non-ICAP participants were collapsed. See Appendix 5 for detail. Therefore, it is important to note that in analysing outcomes from ICAPs, only two measures were explored in more than one ICAP. Specifically, there are BNT data from studies [53,61] and CETI (significant other) data from studies [41,53]. This makes synthesis of ICAP outcome data challenging as different measures are used in each study. It was decided to combine outcome measures listed under the same ICF category to report on outcomes.

4a. Body functions

Body functions covers "individual language modalities, comprehensive tests of language, psychological functioning, and cognitive functioning". There were 24 measures of body functions listed in level III-2 and III-3 studies but only 15 measures with data eligible for analysis coming from eight studies [41,44,45,53–55,57,61] (see Appendix 5).

Six studies reported outcomes relating to "individual language modalities", such as spoken naming. Of these, five studies reported significant gains following treatment [44,53–55,61]. One reported that the gains made during treatment were not greater than gains made pre-therapy [57]. Maintenance of gains was achieved in four studies [44,53–55] and not in two [57,61].

In relation to "comprehensive tests of language", participants in two studies demonstrated significant gains in comprehensive language functioning following treatment as assessed by the CAT, and gains were maintained at follow-up [41] or further improved at follow-up [57]. However, neither of these studies demonstrated a stable baseline pre-therapy that means the findings from post-therapy and follow-up cannot reliably be attributed to treatment.

With respect to psychological functioning, one study assessed and reported significant gains following treatment using the CCRSA that were maintained at follow-up [53].

With respect to cognitive functioning, one study reported significant gains following treatment using the FAS word fluency task, which were maintained at follow-up [44].

In summary, eight papers have explored this domain. The most common positive outcome was in word finding, with maintenance effects. ICAP outcomes relating to gains on comprehensive language tests come from studies where pre-therapy baseline functioning was variable and outcomes relating to psychological function/cognition are limited to a single study each.

4b. Activities and participation

There were six measures of activity/participation listed in level III-2 and III-3 studies but only three measures with data eligible for analysis coming from four studies [41,45,53,55] (see Appendix 5).

First to "communication". Of three papers, two reported significant gains following treatment that were maintained after a follow-up period of no treatment [53,55]. The third paper reported that the improvement from pre- to post-therapy was not significant [41].

Second to "other" areas of activity/participation measured by one study [45] that reported significant changes following treatment, which were maintained at follow-up.

In summary, three papers found significant and maintained gains in activities/participation following ICAP treatment. Additionally, one paper found improvement was non-significant.

4c. Multiple ICF domains

There were three measures of activity/participation listed in level III-2 and III-3 studies but only one measure with data eligible for analysis from a single study [53] (see Appendix 5).

A single paper [53] reported data on an outcome measure relating to this domain (the ALA) and found significant gains following treatment, which were maintained at follow-up.

4d. How ICAPs compare to different intensity models

Two studies [53,54] from the same author group compared outcomes from an ICAP (LIFT) to a distributed model of therapy (D-LIFT). Both LIFT and D-LIFT provided 48 h of total therapy, but LIFT delivered this intensively at ICAP standard (16 h a week over 3 weeks) and D-LIFT provided therapy in a distributed fashion (6 h a week over 8 weeks). In the first study [53], using linear mixed models, the authors found those receiving distributed therapy did significantly better than LIFT (ICAP) participants at post-therapy ($p = 0.04$) and follow-up ($p = 0.002$) on a measure of naming (BNT). In the second study [54] also using linear mixed models, no significant difference was found between LIFT and D-LIFT participants at post-therapy ($p = 0.44$) or follow-up ($p = 0.31$) on their performance on words treated in therapy. The first study [53] also found no difference between LIFT and D-LIFT in psychological outcomes (on the CCRSA), on a measure of communication activity/participation (CETI – proxy rated), or across multiple ICF domains (on the ALA). In short, the only studies to compare ICAP with a distributed model of therapy found the

distributed model was superior for naming outcomes and both models achieved the same improvements in psychological functioning, communication activity, and wider changes, e.g., environmental/personal.

5. What factors are reported to influence ICAP outcomes?

There were only three papers [54,55,57] at level III-2 or III-3 evidence that ran statistical analysis of variables, which could have been factors upon ICAP outcomes.

Authors reported that participants with milder aphasia were more successful on "body function" outcomes whether that be the naming therapy component of ICAPs [54] or outcomes on a comprehensive language test or story retell procedure [57]. No relationship was found between time post-onset and the degree of improvement on either a discourse measure or a measure of communication activity [55].

6. Are there any reported negative outcomes from ICAPs?

In this section, negative outcomes from ICAPs will be considered. These include adverse events, attrition, lack of change, decline, and comparison with other therapy models.

6a. Adverse events – from all 21 studies in the review

Five studies noted adverse events which occurred during the ICAP but may not have been related to the ICAP. Escher et al. reported one participant (of 19) missed a week due to significant health problems [45]. All others were reported to attend at least 85% of sessions. Babbitt et al. [46] noted one (of 74 participants) missed two days during the last week due to illness. On LIFT2, two sessions were missed in the last week due to fatigue [52]. Fatigue and exhaustion were noted in both studies by Griffin-Musick et al., resulting in discontinuation of assessment at times, but not on completion of the ICAP [42,43]. Brindley et al. noted that participants expressed "disenchantment" with returning to conventional therapy and a desire for more intensive therapy when the ICAP ended [55].

6b. Attrition – from all 21 studies in the review

Six studies reported on attrition. Two studies from the same author group reported that all recruited LIFT participants completed therapy [53,54]. Code et al. [41] reported that one participant (of eight) withdrew at an unspecified timepoint due to a seizure. No data were provided on this participant. Rodriguez et al. reported one participant (of eight) was unable to complete LIFT2 due to a prolonged cold/flu [52]. Winans-Mitrik et al. [57] reported three participants (of 73) did not complete the ICAP, one due to multiple falls, one due to unfounded concern for acute stroke, and one for personal reasons. Leff et al. reported one participant (of 47) dropped out after a week for an unspecified reason and two were lost to follow-up due to COVID-19 [58]. Babbitt et al. [46] had 84 participants but reported on only 74 as 10 did not complete either pre- or post-therapy testing for a variety of reasons including "the severity of aphasia, hospitalization prior to the end of the program, fatigue, or concomitant cognitive deficits" [p.S857]. Wilson et al. [51] reported 15 participants were recruited but only provided results for nine participants with no reason provided.

6c. Lack of change – from eight experimental design studies only

Body functions. There were four studies that reported lack of change in "body function" outcomes. One study found that the change that occurred during therapy on a story retell measure (SRP) was not greater than change experienced pre-therapy [57]. The same study noted "for a small minority of participants, no change or a decline in performance was observed for all outcomes" [57,p.S337]. At level III evidence, this included a comprehensive test of language (the CAT), and the story retell measure, both "body function" measures. Individual participant data (IPD) from one study [41] revealed that although group improvement was noted on a comprehensive language test (AAT, English version), 1/7 participants did not improve "for DO, the most severe case, no progress can be observed" [41,p.25]. IPD from another study [54] revealed 3/13 ICAP participants did not experience a statistically significant improvement on treated naming items post-therapy (see Table 2 in this study). In one study, IPD displayed on tables demonstrated that during the first 6-week intensive intervention period anywhere from 7 to 13% of participants did not improve on a measure of naming (BNT) and 20–45% did not improve on a picture description task (CIU analysis) [61].

Activities/participation. Two studies reported lack of change in this domain. One found non-significant improvement post-therapy on a measure of communication activity (the CETI) rated by a significant other [41]. Another study found no significant improvement post-therapy on some sections of a communication assessment [55].

6d. Decline – from eight experimental design studies only

Only one instance of individual participant decline was noted. In Hinckley and Craig [61], one participant demonstrated deterioration during the first 6-week intensive intervention period on CIU analysis (see Table 3, subject 5).

Discussion

1. What inclusion/exclusion criteria are listed for ICAPs?

There were 27 criteria listed for ICAP inclusion/exclusion. There may have been additional criteria not listed by studies, e.g., due to publication constraints. We focus our discussion on the top four commonly cited criteria, which were (1) time post-onset (being mostly 6 months post-onset, seven ICAPs), (2) being independent or having caregiver support for ADLs (activities of daily living) (six ICAPs), (3) having endurance/stamina/tolerance (six ICAPs), and (4) being physically and mentally medically stable (four ICAPs). How do these criteria fit with the existing evidence base? An international survey [73] specifically asked about ICAP admission criteria on 14 ICAPs and found the top criteria were: (1) being independent or having caregiver support for ADLs (11 ICAPs), (2) age with unspecified detail (nine ICAPs), (3) time post-onset with unspecified detail (five ICAPs), and (4) no history of dementia/cognitive impairment *or* primary aphasia diagnosis with no concurrent diagnoses (four ICAPs each). In terms of speech and language therapist (SLT) perspective, a survey of UK SLTs found 33% felt *any time* post-onset was acceptable to engage in an ICAP, and 60% felt support for toileting was necessary [20].

Therefore, being independent/having support for ADLs and time post-onset (TPO) seem important to discuss as they are raised both in studies and in surveys. Being independent/having support for ADLs may need to be implemented as an inclusion criterion for a variety of reasons. For example, ICAP staff may not be trained in ADL assistance or there may be lack of funding or a need to increase ICAP costs to have qualified staff provide toileting – i.e., practical constraints. It is important to acknowledge this criterion

limits access to ICAPs and should be seen as a provider and accessibility issue rather than patient suitability issue. The rationale for being a certain time post-onset to engage in an ICAP is not provided by studies and six ICAPs included participants who were less than six months post-onset [42,47,50,53,57,60] (*Persad for UMAP*). Therefore, there is no consensus in the ICAP community on this criterion. There could be two reasons for imposing a TPO constraint. One relates to the research design, particularly if there is no control group and a need to demonstrate a stable baseline. International consensus is that six months post-onset (MPO) signifies the chronic stage in stroke research [74], which may be why this timeframe is selected. The second reason may be clinical. ICAPs may be more suitable for those in the chronic phase, e.g., because they are medically more stable and more robust. But it could be that intensive and comprehensive input is potentially most effective in the acute phase. The VERSE trial found that non-ICAP SLT delivered in the acute stage (starting within 2 weeks post-stroke) at an intensity of 5 h/week × 4 weeks (mean total dose of 23 h) did not produce additional benefit for PWA when compared with usual care (mean total dose of 9.5 h) [75]. Whether ICAP levels of intensity (15 h/week × 2 weeks) and the provision of more comprehensive aphasia therapy would produce additional benefit in the acute stage is unknown. This requires testing through a design that compares ICAP with an alternative model at the same dose/intensity/amount.

Future ICAP research should consider TPO as a criterion and rationalise any constraints employed.

Returning briefly to the 27 listed criteria, rationales for the criteria were rarely given, and in some cases might be challenged. For example, one study [55] imposed an upper age limit of 65 without citing evidence that age affects prognosis in aphasia. It was also difficult to determine how some criteria were, or indeed could be assessed, "motivation" and "endurance" being examples. Two qualitative studies found that motivation was important to stroke rehabilitation professionals, but was affected by cultural factors (e.g., a belief that a stroke was God's will) [76] and by age (e.g., older patients being less likely to demonstrate motivation through facial expression and self-directed rehabilitation activity) [77]. A recent scoping review found motivation has only been measured in two aphasia studies and its links to treatment outcomes have not been explored [78]. It is a concern that judgements on subjective factors may be negative when a person is from a different cultural background or is elderly, underscoring the need for selection criteria that can be objectively assessed. The likelihood of increased medical co-morbidities with age may also be a limiting factor, i.e., elderly people may be less likely to meet a criterion of medical stability, and older participants may not have a living caregiver to support their attendance.

Recommendations for future ICAPs
(1) Reporting should be clear when criteria are driven by provider/resource needs. (2) Research based rationales for exclusionary criteria should be given. (3) Criteria must be measurable, to avoid subjective judgements. (4) Further work is needed on candidacy for ICAPs so that future researchers can employ evidence-based selection criteria, e.g., using designs that compare outcomes across participant factors.

2. Who participates in an ICAP?

The current study found that most ICAP participants were male (at least 68%), in their mid-50s (range 18–86), mostly white/Caucasian (77–96%), and mean years of education was 14. TPO was 1–335 months (no mean could be established though it was noted several of the largest studies had a mean time post-onset of 15–24 months). Winans-Mitrik et al. [57] reported "the current program admission rate is 50% of referrals" [57,p.S332]. How do these demographics relate to participants in aphasia research generally? The RELEASE collaboration [13] analysed IPD of 928 PWA who took part in randomised controlled trials. Fifty-eight percent were male, the median age was 63 (IQR 54.1–74), 94.7% were white, and median years of education was 11 (IQR 8–14). TPO was median 61 days (range 7–487). A scoping review of 139 aphasia treatment studies from the US [79] noted under-representation of women (61% of participants were male), older stroke survivors (mean age 58), and ethnic minorities (86% white when 75% of the US population is white) [80]. Looking with a wider lens, women are more likely to experience strokes than men [81] though this varies based on age range (with older women more likely to experience a stroke than older men), the median age of stroke in the UK was 70 for men and 76 for women [82] and an average age of 69 for both sexes has been reported from US cohort data [83], and black people also experience worse stroke rehabilitation outcomes than white people [84]. This indicates that PWA taking part in any type of aphasia research are a subset of stroke survivors, but ICAP participants are more likely to be male, younger, more highly educated, and (though time post-onset is tricker to compare) significantly longer post-onset of aphasia than PWA taking part in aphasia research. Why might this be the case?

When it comes to sex and the under-representation of people from ethnic minority background, financial constraints could be a factor. Six of the ICAPs included in this study (five in the US and one in Canada) were fee paying (BSAP, Chicago, InteRACT, PIRATE for some participants, S-IHP's cap, and UMAP), and two of these required out of pocket payments (i.e., not covered by health insurance) of over $20 000. US Census data from 2020 indicated the median earnings of Black and Hispanic households was lower than that of White households, and the female-to-male earnings ratio was 0.83 (i.e., the average women earned 83 cents for each dollar earned by a man) [85]. Analysis of data from the UK Office of National Statistics (as the equivalent US data was not available) showed that the gender pay gap was greatest for women in their 50s and 60s [86]. Therefore, women are less likely to have disposable income to pay for ICAP attendance. All six fee-paying ICAPs had more male participants. Of the remaining seven non-fee paying ICAPs, 4/7 were substantially more male dominated [41,49,53,59], one was even [55], one had more females [56], and one had no data [44]. Therefore, cost may not be the primary factor limiting female attendance. Perhaps the role of women as caregivers limits their participation? Four ICAPs required that participants be accompanied by caregivers either if not independent for ADLs (two ICAPs) or for carryover of skills (two ICAPs) (see Table 4). As women are more likely to be caregivers [87] they may facilitate and support male attendees, but the reverse may not be true, particularly if the man is the main breadwinner and must continue to work. Perhaps if ICAP staff or students could be upskilled to support ADLs and facilitate people to attend without a caregiver, this would help boost female attendees?

Recommendations for future ICAPs
(1) Several methods to broaden representation of the wider aphasia population on ICAPs could include targeted and funded culturally sensitive recruitment campaigns for minority groups, e.g., Nguy et al. [79]. (2) Conducting more studies within a non-fee-paying structure to ensure inclusion of those who cannot pay for attendance. (3) Exploring what barriers might prevent women from engaging on ICAPs.

3. What outcome measures are used on ICAPs?

A large number of outcome measures (56) were used across ICAPs and even when limited to experimental designs only, the range was still wide (33) with very little overlap. ICAPs are comprehensive therapy programmes, so the diverse range of measures partially reflects efforts to capture comprehensive changes. However, of the 56 measures, 38 (of 68%) were in the domain of "body functions", demonstrating that the focus of ICAP outcome measurement is mostly on impairment-based outcomes. As outlined in the "Introduction", goals of PWA and their families do not focus primarily on "body functions" [25,26], so their goal priorities are not well reflected in the outcome measures employed by ICAPs. The variation in measures also means it is challenging to pool data from ICAPs. The issue of too many measures has previously been raised in the aphasia literature (44 formal outcome measures were identified in randomised controlled trials in the Cochrane review of SLT for aphasia [7] and 143 standardised outcome measures for aphasia were identified in a scoping review [40]). Both reviews identified that most outcome measures focused on the domain of "body functions", e.g., 66% of measures [40]. The need for homogeneity in aphasia outcome measurement was the driver behind establishing a core outcome set (COS) for aphasia in 2019 [88], which recommended use of the WAB-R, GHQ-12, and SAQOL-39 (see Appendix 5 for full names and references). The WAB was included in six ICAPs, but the GHQ and SAQOL only in one [56], likely due to the recency of the COS recommendations. Now that COS recommendations are available, including a recent addition of the Scenario Test as a measure of communication [89], these should be included in future ICAP research. Should ICAPs have their own COS? Part of the issue in establishing a COS for ICAPs is that ICAP content varies widely, both on a programme and individual participant level [12]. There are no existing COS recommendations to cover the ICF domains of "environmental factors" or "personal factors" due to lack of consensus. Therefore, establishing an ICAP-specific set of measures may prove equally difficult.

The diversity in ICAP content and outcome measurement points to a need for rationalisation. This might be achieved by developing a logic model of the intervention. Such models specify the underlying theory of an intervention by outlining how the intervention components link to outcomes [90]. A logic model would assist in rationalising ICAP content/components. This would be a useful springboard for conversations on the merit of standardising ICAP therapy content. Then linked outcome measures for each ICAP component and the feasibility of a COS for ICAPs could be considered. For example, ICAPs as originally conceptualised recommended inclusion of family members [22], and many ICAPs require family/caregiver attendance. However, there was only one ICAP that included a measure of outcomes for families [59] though the data were not reported.

Recommendations for future ICAPs
(1) If family/carers are required to attend, it should be clarified whether that is for their benefit with expected improved outcomes for families, or whether family attendance is primarily for the benefit of PWA. If the former is intended, a relevant family outcome measure should be included. (2) Exploration of a logic model and possible COS for ICAPs.

4. Have ICAPs brought about significant improvements on outcome measures, and if so in what domains of the ICF?

Synthesising outcome data across ICAP studies was difficult, owing to the diversity of and lack of overlap between measures. For this reason, outcomes were collapsed into ICF domains. Despite the range of outcome measures used on ICAPs, there were only four ICAPs at level III-3 or III-2 evidence that produced experimental evidence on outcomes in more than one ICF domain – "body functions" and "activities/participation" (Belfast, Boston, LIFT, and Milton Keynes). Therefore, though ICAPs contain a variety of therapy content [12], the evidence-base to argue that ICAPs deliver comprehensive changes for PWA is lacking. This is disappointing as ICAPs are comprehensive therapy programmes. It is desirable to have outcomes on a range of ICF domains to differentiate the additional benefit of comprehensive aphasia therapy provided on an ICAP from purely language-based intensive programmes (e.g., as delivered by intensive language-action therapy/constraint-induced aphasia therapy).

Outcome measurement on ICAPs is skewed towards "body functions". 5/6 studies found pre- to post-therapy gains on naming and positive maintenance of naming gains was shown in 4/6 studies. Evidence for outcomes in relation to cognition or psychological function came only from single studies. Evidence for change in activities/participation was equivocal with 2/3 studies finding pre- to post-therapy gains on a communication measure. The fact that this domain was so rarely measured was disappointing, given that this is a key priority for PWA and their family members [26]. The focus on outcome measures in the domain of "body function" may reflect ICAP content. For instance, participants on Dignam et al. [54] had 30 targeted words which were treated in 14 h of impairment-based therapy and up to 14 h of computer therapy out of a total of 48 h of therapy time. If most of the treatment time is focused on language impairment, then the outcome measures will likely be weighted towards this domain. In another example of how ICAP content may influence outcomes, Escher et al. [45] found significant improvement post-ICAP on the Canadian Occupational Performance Measure (COPM) but was the only ICAP with occupational therapy input. There are no recommendations on how to weight ICAP treatment time between ICF domains, and therefore each ICAP interprets this differently on a programme and individual level [12].

This review did not set out to evaluate whether ICAPs are more/less effective than other models of therapy provision. However, two studies included comparative data (ICAPs versus dispersed therapy models) with no evidence of an ICAP advantage [53,54] and an advantage for a dispersed model on a measure of naming [53]. Psychological interventions for aphasia have typically been delivered in a more dispersed model, e.g., over a three-month period [91].

Recommendations for future ICAPs
(1) Publish ICAP outcomes from across ICF domains. (2) Future research should explore whether ICAPs are superior to dispersed models for a *range* of therapy outcomes.

5. What factors are reported to influence ICAP outcomes?

There is evidence from two ICAP studies [54,57] that those with milder aphasia had better outcomes on naming therapy/language assessment, and evidence from one study that time post-onset did not relate to outcomes in "body function" or "activities/participation" [55]. With only three studies providing answers to this question in the current study, conclusions must remain tentative. However, the findings from this review that (1) milder aphasia may be superior for "body function" outcomes and that (2) time post-onset did not link to outcomes, aligns with the existing evidence base. A narrative review by Kristinsson et al. found that aphasia severity "is generally considered the single strongest predictor of response to impairment-based therapy" [92]. While there

is more evidence to support better outcomes for those with milder aphasia, the authors cite one study showing those with more severe aphasia do better [50]. However, it should be noted this study which was an ICAP was not deemed of sufficient experimental quality to contribute to RQ5 in the current review. A review of 23 single subject aphasia studies, where all participants were at least one year post-onset found no relationship between TPO and impairment-based language outcomes [93]. These findings indicate that therapy induced change can occur in the chronic phase of aphasia, in line with the available ICAP data.

More research should explore participant factors that may affect responses to ICAP therapy. There are complexities involved with this, e.g., it may be that predictors are outcome specific. A greater understanding of participant factors might allow for improved tailoring of therapy, with greater emphasis given to the domains that are most likely to change.

There were 27 inclusion/exclusion criteria listed for ICAPs, but the current state of the evidence does not provide sufficient evidence to link pre-morbid factors to ICAP outcomes. It is likely clinical experience comes into play in determining suitability for ICAPs. The weight placed on clinical experience should not be undermined. Sackett et al. clearly highlighted that clinical expertise was an essential component of evidence-based medicine, particularly as only clinical expertise can guide whether external clinical evidence should be applied to certain patients [94].

Stronger evidence that severity affects impairment-based outcomes could be used in future to refine and individualise ICAP treatment content. For example, therapy for those with severe aphasia might focus more on compensatory communication strategies than specific language targets.

Recommendations for future ICAPs
(1) Describe ICAP participant characteristics, e.g., using DESCRIBE criteria [70]. (2) Explore the feasibility and validity of measuring and linking ICAP participant characteristics to outcomes across ICF domains.

6. Are there any reported negative outcomes from ICAPs?

It is important to track any negative side effects of ICAPs, given the demanding nature of the therapy model. There were few signs of such effects. While some adverse events were reported, none were related to therapy content. Rates/nature of events were like those reported in other therapy trials [95,96]. Only one participant was reported to decline following therapy, and then only on one measure of "body function" [61].

A further concern is that attrition in ICAPs might be high due to the demanding nature of therapy. For example, the Cochrane review [7] found higher rates of drop-out from low versus high intensity aphasia trials, but only four trials were included in the analysis. An analysis of the significance of drop-out rates was not conducted in the current study, so the significance of the attrition rate cannot be calculated. The only study comparing ICAP with distributed therapy found in favour of the intensive regime with 16/16 completing the ICAP programme (LIFT) and 16/18 completing the distributed therapy programme (D-LIFT) [53].

Although decline was barely reported, lack of change was noted at group level on an outcome measure in three studies [41,55,57] and from IPD in three studies [41,54,61]. Lack of change or decline is not often reported in aphasia treatment. This is perhaps not unusual given the known bias in healthcare towards publishing significant or positive results [97]. Menahemi-Falkov et al. [98] found that aphasia trials often reported positive

significant group level change. However, an exploration of IPD revealed this group level change was driven by a minority of participants (about 33% who experienced post-therapy positive change and only 22% maintaining the change at follow-up).

The current review did not include non-experimental designs when enquiring about lack of change in response to ICAPs. However, there are indications from the non-experimental ICAP literature (NHMRC Level IV pre-post studies) of lack of change experienced by participants. For example, a third of participants on InteRACT, 31% of ICAP participants on the Chicago ICAP, and 10/54 on UMAP did not respond to treatment, i.e., did not show a five-point change in their WAB-R AQ score [47,50]. Additionally, half of participants on InteRACT did not experience clinically significant change on measures of activity/participation (the CETI or CADL-2) [50]. Lack of change data is particularly important when testing a new treatment model as there may be important indications for candidacy.

Recommendations for future ICAPs
(1) Publish participant flow through the ICAP to clearly illustrate recruitment and attrition with associated reasons. (2) It would be useful for experimentally designed ICAPs to analyse both group and individual level change, and to share whether any participants experienced decline during the intervention.

Limitations

This review did not employ two authors to independently screen abstracts but instead used random checking by a second and third reviewer of 20% of the results at abstract screening level and 100% of results by all authors at full text level. Although two independent reviewers are not required for a systematic review [99], use of a completely independent second reviewer can increase the number of results found at abstract screening level by up to 9% [100]. Therefore, use of a completely independent second reviewer would have been desirable. There is a potential risk of bias in this dataset because non-English papers were excluded at the search stage. The number of studies excluded using this characteristic could not be determined as this criterion was applied at the level of the search within each database. Exclusion of non-English papers was found to have a minimal effect on the conclusions of systematic reviews of clinical interventions [101]. A 2011 review found publications about aphasia were published in English-language journals though only 61% of participants were English speakers (the other popular languages being German 7%, Italian 6.5%, Dutch 4.6%, and French 3.8%) [102]. This suggests aphasia research is likely to have been published in the English language. Additionally, the concept and definition of an ICAP came from the English-speaking aphasia world (US, Canada, and Australia), presented at the American Speech-Language-Hearing Association conference in 2011 [103]. These considerations hopefully reduce the risk of excluding non-English ICAP studies in the review.

Only eight studies were of adequate experimental design to answer research questions 4 and 5. A meta-analysis of results was not possible due to an inability to meaningfully pool and analysis quantitative data from these studies. A qualitative analysis was provided instead.

Conclusions

ICAPs are intensive and comprehensive service delivery models. Emphasis on increasing therapy intensity is a central focus of

aphasia research and a focus of governmental healthcare strategy in the UK. Delivery of comprehensive therapy is a key concern of patients, families, and clinicians, and for these reasons, the UK national audit that has previously focused on therapy intensity is exploring whether therapy content can also be detailed. ICAPs offer a potential model for increasing the intensity and comprehensiveness of therapy. However, they are not yet mainstream service delivery models. The results of this systematic review suggest that ICAP research is still under-developed. The model has only been tested with samples that are poorly representative of the aphasia population. A wide range of measures are used on ICAPs, leading to problems with data synthesis. Measures do not fully reflect the comprehensive ambitions of the ICAP model. Most are focused on body functions. While there is some assessment of activities/participation, the environmental and personal factor domains of the ICF are neglected. This prioritisation of outcomes contrasts with outcomes desired from aphasia therapy by PWA and their families. There were eight ICAP studies of adequate methodological quality, which answered questions about ICAP outcomes. The most robust finding was that attending an ICAP results in group-level improvements in naming, which are usually maintained post-therapy. However, ICAP therapy is not superior to dispersed therapy models in achieving these naming gains. Although there are indications that ICAPs may produce positive outcomes in other domains, e.g., communication/cognition/psychological functioning, there are insufficient data in these areas to support strong conclusions. There is also evidence from IPD that some ICAP participants will not experience statistically significant improvement in outcomes. Aphasia severity and time post-onset are the only two predictive factors that have been explored in relation to ICAP outcomes, but only in a total of three studies making any firm conclusions inappropriate. Future ICAP research must have higher quality intervention design to be able to adequately answer outcome questions, e.g., employing time-interrupted/multiple baseline/randomised control design rather than pre-post designs. Larger *n* studies would also allow further exploration of candidacy. More comparative studies between ICAPs and non-ICAPs (e.g., more dispersed models) would be useful. This review did not include outcomes from pre-post studies as they were deemed to be inadequately experimental in design to reliably answer RQs on ICAP outcomes. Published research on ICAPs likely consisted of many pre-post reviews as the programmes were designed for clinical rather than research purposes. For example, most of the pre-post papers contained assessment data from fee paying attendees. This scenario does not allow for double baseline/follow-up assessment or control group data due to time and financial constraints. Additionally, an international survey found 25% of ICAPs had ceased operation over a 7-year period for reasons including financial stability [73]. When the running of non-experimentally designed ICAPs is vulnerable, it is no wonder the data published are in the form of pragmatic pre-post programme evaluations, rather than higher quality research designs. It is positive the pre-post data on ICAPs has been published, and future work will explore the outcomes from pre-post ICAP studies. This avenue of research may provide useful directions for future experimentally designed ICAP studies, particularly as factors influencing outcomes are explored in pre-post studies. ICAPs should employ the aphasia core outcome measurement set to ensure all ICF domains are addressed and to allow for comparison across ICAPs. Rationalisation of measures may well demand further consensus work on treatment content, hence a need for an ICAP logic model to justify treatment content linked to outcomes. Input from PWA and families in a co-designed ICAP is a logical next step.

ORCID

Katie Monnelly http://orcid.org/0000-0002-3112-9830
Jane Marshall http://orcid.org/0000-0002-6589-221X
Lucy Dipper http://orcid.org/0000-0002-5918-3898
Madeline Cruice http://orcid.org/0000-0001-7344-2262

References

[1] Hilari K, Northcott S. Social support in people with chronic aphasia. Aphasiology. 2006;20(1):17–36. doi: 10.1080/02687030500279982.

[2] Graham JR, Pereira S, Teasell R. Aphasia and return to work in younger stroke survivors. Aphasiology. 2011;25(8):952–960. doi: 10.1080/02687038.2011.563861.

[3] Morris R, Eccles A, Ryan B, et al. Prevalence of anxiety in people with aphasia after stroke. Aphasiology. 2017;31(12):1410–1415. doi: 10.1080/02687038.2017.1304633.

[4] Wang S, Wang C, Zhang N, et al. The association between post-stroke depression, aphasia, and physical independence in stroke patients at 3-month follow-up. Front Psychiatry. 2018;9:374. doi: 10.3389/fpsyt.2018.00374.

[5] Grawburg M, Howe T, Worrall L, et al. Describing the impact of aphasia on close family members using the ICF framework. Disabil Rehabil. 2014;36(14):1184–1195. doi: 10.3109/09638288.2013.834984.

[6] Code C. The implications of public awareness and knowledge of aphasia around the world. Ann Indian Acad Neurol. 2020;23(Suppl. 2):S95–S101. doi: 10.4103/aian.AIAN_460_20.

[7] Brady MC, Kelly H, Godwin J, et al. Speech and language therapy for aphasia following stroke. Cochrane Database Syst Rev. 2016;6:CD000425.

[8] Wisenburn B, Mahoney K. A meta-analysis of word-finding treatments for aphasia. Aphasiology. 2009;23(11):1338–1352. doi: 10.1080/02687030902732745.

[9] Rose ML, Raymer AM, Lanyon LE, et al. A systematic review of gesture treatments for post-stroke aphasia. Aphasiology. 2013;27(9):1090–1127. doi: 10.1080/02687038.2013.805726.

[10] Baker C, Worrall L, Rose M, et al. A systematic review of rehabilitation interventions to prevent and treat depression in post-stroke aphasia. Disabil Rehabil. 2018;40(16):1870–1892. doi: 10.1080/09638288.2017.1315181.

[11] Fox L, Poulsen S, Clark Bawden K, et al. Critical elements and outcomes of a residential family-based intervention for aphasia caregivers. Aphasiology. 2004;18(12):1177–1199. doi: 10.1080/02687030444000525.

[12] Monnelly K, Marshall J, Cruice M. Intensive Comprehensive Aphasia Programmes: a systematic scoping review and analysis using the TIDieR checklist for reporting interventions. Disabil Rehabil. 2022;44(21):6471–6496. doi: 10.1080/09638288.2021.1964626.

[13] REhabilitation and recovery of peopLE with Aphasia after StrokE (RELEASE) Collaborators. Dosage, intensity, and frequency of language therapy for aphasia: a systematic review-based, individual participant data network meta-analysis. Stroke. 2022;53(3):956–967. doi: 10.1161/STROKEAHA.121.035216.

[14] Harvey SR, Carragher M, Dickey MW, et al. Treatment dose in post-stroke aphasia: a systematic scoping review. Neuropsychol Rehabil. 2021;31(10):1629–1660. doi: 10.1080/09602011.2020.1786412.

[15] Cherney LR, Patterson JP, Raymer AM. Intensity of aphasia therapy: evidence and efficacy. Curr Neurol Neurosci Rep. 2011;11(6):560–569. doi: 10.1007/s11910-011-0227-6.

[16] Pierce JE, O'Halloran R, Menahemi-Falkov M, et al. Comparing higher and lower weekly treatment intensity for chronic aphasia: a systematic review and meta-analysis. Neuropsychol Rehabil. 2021;31(8):1289–1313. doi: 10.1080/09602011.2020.1768127.

[17] Katz RC, Hallowell B, Code C, et al. A multinational comparison of aphasia management practices. Int J Lang Commun Disord. 2000;35(2):303–314. doi: 10.1080/136828200247205.

[18] Palmer R, Witts H, Chater T. What speech and language therapy do community dwelling stroke survivors with aphasia receive in the UK? PLOS One. 2018;13(7):e0200096. doi: 10.1371/journal.pone.0200096.

[19] Code C, Petheram B. Delivering for aphasia. Int J Speech Lang Pathol. 2011;13(1):3–10. doi: 10.3109/17549507.2010.520090.

[20] Monnelly K, Marshall J, Dipper J, et al. Intensive and comprehensive aphasia therapy—a survey of the definitions, practices and views of speech and language therapists in the United Kingdom. Int J Lang Commun Disord. 2023. doi: 10.1111/1460-6984.12918.

[21] Sentinel Stroke 18 National Audit Programme. Sentinel Stroke National Audit Programme: post-acute audit – services [dataset]; 2015. Available from: https://www.strokeaudit.org/results/PostAcute/services.aspx

[22] Rose ML, Cherney LR, Worrall LE. Intensive comprehensive aphasia programs: an international survey of practice. Top Stroke Rehabil. 2013;20(5):379–387. doi: 10.1310/tsr2005-379.

[23] NHS Long Term Plan. Chapter 3. Stroke care; 2019 [Internet] [cited 2023 Jan 31]. Available from: https://www.longtermplan.nhs.uk/online-version/chapter-3-further-progress-on-care-quality-and-outcomes/better-care-for-major-health-conditions/stroke-care/

[24] World Health Organization. Framework on integrated, people-centred health services. Geneva (Switzerland): World Health Organization; 2016.

[25] Worrall L, Sherratt S, Rogers P, et al. What people with aphasia want: their goals according to the ICF. Aphasiology. 2011;25(3):309–322. doi: 10.1080/02687038.2010.508530.

[26] Wallace SJ, Worrall L, Rose T, et al. Which outcomes are most important to people with aphasia and their families? An international nominal group technique study framed within the ICF. Disabil Rehabil. 2017;39(14):1364–1379. doi: 10.1080/09638288.2016.1194899.

[27] Simmons-Mackie N, Raymer A, Cherney LR. Communication partner training in aphasia: an updated systematic review. Arch Phys Med Rehabil. 2016;97(12):2202–2221.e8. doi: 10.1016/j.apmr.2016.03.023.

[28] Priorities in stroke rehabilitation and long-term care [Internet]; 2021. Stroke Association [cited 2023 Jan 30]. Available from: https://www.stroke.org.uk/sites/default/files/research/priorities_in_stroke_rehabilitation_and_lon-term_care.pdf

[29] What study designs can be considered for inclusion in an EPOC review and what should they be called? [Internet]. Cochrane Effective Practice and Organisation of Care (EPOC); 2017 [cited 2023 Jan 23]. Available from: https://zenodo.org/record/5106085#.Y72Fh3bP1D8

[30] Helm-Estabrooks N, Whiteside J. Use of life interests and values (LIV) cards for self-determination of aphasia rehabilitation goals. Perspect Neurophysiol Neurogenic Speech Lang Disord. 2012;22(1):6–11. doi: 10.1044/nnsld22.1.6.

[31] Off CA, Griffin JR, Murray KW, et al. Interprofessional caregiver education, training, and wellness in the context of a cohort model for aphasia rehabilitation. Top Lang Disord. 2019;39(1):5–28. doi: 10.1097/TLD.0000000000000171.

[32] Hoover EL, Carney A. Integrating the iPad into an intensive, comprehensive aphasia program. Semin Speech Lang. 2014;35(1):25–37. doi: 10.1055/s-0033-1362990.

[33] Whiteside J, Pak Hin Kong A. Therapeutic effect of an intensive comprehensive aphasia program. Paper presented at: Clinical Aphasiology Conference; 2013 May; Tucson, AZ.

[34] Jensen LR, Lønnberg C. Return to work with aphasia: review of the literature and preliminary vocational outcome of an intensive rehabilitation programme. Paper presented at: 2nd Nordic Aphasia Conference; 2009 May 14–16; Copenhagen, DE.

[35] Ferdinandi A, Duke W. Intensive comprehensive aphasia treatment: a pilot project for British Columbia and preliminary study. Paper presented at: 3rd Canadian Stroke Congress; 2012 Sep 29 to Oct 2; Calgary, AB.

[36] Prigatano GP, Fordyce DJ, Zeiner HK, et al. Neuropsychological rehabilitation after closed head injury in young adults. J Neurol Neurosurg Psychiatry. 1984;47(5):505–513. doi: 10.1136/jnnp.47.5.505.

[37] Fitzgerald-DeJean DM, Rubin SS, Carson RL. An application of the experience sampling method to the study of aphasia: a case report. Aphasiology. 2012;26(2):234–251. doi: 10.1080/02687038.2011.621208.

[38] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi: 10.1136/bmj.n71.

[39] Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ. 2020;368:l6890. doi: 10.1136/bmj.l6890.

[40] Wallace SJ, Worrall L, Le Dorze G, et al. Many ways of measuring: a scoping review of measurement instruments for use with people with aphasia. Aphasiology. 2022;36(4):401–466. doi: 10.1080/02687038.2020.1836318.

[41] Code C, Torney A, Gildea-Howardine E, et al. Outcome of a one-month therapy intensive for chronic aphasia: variable individual responses. Semin Speech Lang. 2010;31(1):21–33. doi: 10.1055/s-0029-1244950.

[42] Griffin-Musick J, Jakober D, Sallay A, et al. Cognitive-linguistic outcomes from an intensive comprehensive aphasia program implemented by graduate student clinicians. Aphasiology. 2022;36(9):1015–1029. doi: 10.1080/02687038.2021.1937920.

[43] Griffin-Musick J, Off CA, Milman L, et al. The impact of a university-based intensive comprehensive aphasia program (ICAP) on psychosocial well-being in stroke survivors with

aphasia. Aphasiology. 2021;35(10):1363–1389. doi: 10.1080/02687038.2020.1814949.

[44] Hoover EL, Caplan DN, Waters GS, et al. Communication and quality of life outcomes from an interprofessional intensive, comprehensive, aphasia program (ICAP). Top Stroke Rehabil. 2017;24(2):82–90. doi: 10.1080/10749357.2016.1207147.

[45] Escher AA, Amlani AM, Viani AM, et al. Occupational therapy in an intensive comprehensive aphasia program: performance and satisfaction outcomes. Am J Occup Ther. 2018;72(3):7203205110p1. doi: 10.5014/ajot.2018.026187.

[46] Babbitt EM, Worrall L, Cherney LR. Structure, processes, and retrospective outcomes from an intensive comprehensive aphasia program. Am J Speech Lang Pathol. 2015;24(4):S854–S863. doi: 10.1044/2015_AJSLP-14-0164.

[47] Babbitt E, Worrall L, Cherney L. Who benefits from an intensive comprehensive aphasia program? Top Lang Disord. 2016;36(2):168–184. doi: 10.1097/TLD.0000000000000089.

[48] Baliki MN, Babbitt E, Cherney L. Brain network topology influences response to intensive comprehensive aphasia treatment. NeuroRehabilitation. 2018;43(1):63–76. doi: 10.3233/NRE-182428.

[49] Mackenzie C. An aphasia group intensive efficacy study. Br J Disord Commun. 1991;26(3):275–291. doi: 10.3109/13682829109012015.

[50] Persad C, Wozniak L, Kostopoulos E. Retrospective analysis of outcomes from two intensive comprehensive aphasia programs. Top Stroke Rehabil. 2013;20(5):388–397. doi: 10.1310/tsr2005-388.

[51] Wilson KR, O'Rourke H, Wozniak LA, et al. Changes in N400 topography following intensive speech language therapy for individuals with aphasia. Brain Lang. 2012;123(2):94–103. doi: 10.1016/j.bandl.2012.06.005.

[52] Rodriguez AD, Worrall L, Brown K, et al. Aphasia LIFT: exploratory investigation of an Intensive Comprehensive Aphasia Programme. Aphasiology. 2013;27(11):1339–1361. doi: 10.1080/02687038.2013.825759.

[53] Dignam J, Copland D, McKinnon E, et al. Intensive versus distributed aphasia therapy: a nonrandomized, parallel-group, dosage-controlled study. Stroke. 2015;46(8):2206–2211. doi: 10.1161/STROKEAHA.115.009522.

[54] Dignam J, Copland D, Rawlings A, et al. The relationship between novel word learning and anomia treatment success in adults with chronic aphasia. Neuropsychologia. 2016;81:186–197. doi: 10.1016/j.neuropsychologia.2015.12.026.

[55] Brindley P, Copeland M, Demain C, et al. A comparison of the speech of ten chronic Broca's aphasics following intensive and non-intensive periods of therapy. Aphasiology. 1989;3(8):695–707. doi: 10.1080/02687038908249037.

[56] Auclair-Ouellet N, Tittley L, Root K. Effect of an intensive comprehensive aphasia program on language and communication in chronic aphasia. Aphasiology. 2022;36(11):1312–1332. doi: 10.1080/02687038.2021.1959016.

[57] Winans-Mitrik RL, Hula WD, Dickey MW, et al. Description of an intensive residential aphasia treatment program: rationale, clinical processes, and outcomes. Am J Speech Lang Pathol. 2014;23(2):S330–S342. doi: 10.1044/2014_AJSLP-13-0102.

[58] Leff AP, Nightingale S, Gooding B, et al. Clinical effectiveness of the Queen Square intensive comprehensive aphasia service for patients with poststroke aphasia. Stroke. 2021;52:e594–e598.

[59] Doogan C, Al Balushi R, Gooding B, et al. What do people with aphasia want from the Queen Square Intensive Comprehensive Aphasia Programme and do they achieve it? A quantitative and qualitative analysis of their short, medium, long-term and economic goals. Aphasiology. 2023;37(10):1661–1678. doi: 10.1080/02687038.2022.2118518.

[60] Nicholas M, Pittmann R, Pennington S, et al. Outcomes of an interprofessional intensive comprehensive aphasia program's first five years. Top Stroke Rehabil. 2022;29(8):588–604. doi: 10.1080/10749357.2021.1970452.

[61] Hinckley JJ, Craig HK. Influence of rate of treatment on the naming abilities of adults with chronic aphasia. Aphasiology. 1998;12(11):989–1006. doi: 10.1080/02687039808249465.

[62] NHMRC levels of evidence and grades for recommendations for developers of guidelines [Internet]. Australian Government National Health and Medical Research Council; 2009 [cited 2023 Jan 23]. Available from: https://www.nhmrc.gov.au/sites/default/files/images/NHMRC%20Levels%20and%20Grades%20(2009).pdf

[63] Reeves BC, Wells GA, Higgins JP, et al. Chapter 24: including non-randomized studies on intervention effects. In: Higgins JPT, Thomas J, Chandler J, et al. editors. Cochrane handbook for systematic reviews of interventions version 6.3. Chichester (UK): John Wiley and Sons Ltd; 2022. p. 595–620.

[64] Reeves BC, Wells GA, Waddington H. Quasi-experimental study designs series—paper 5: a checklist for classifying studies evaluating the effects on health interventions—a taxonomy without labels. J Clin Epidemiol. 2017;89:30–42. doi: 10.1016/j.jclinepi.2017.02.016.

[65] Ma L, Wang Y, Yang Z, et al. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? Mil Med Res. 2020;7(1):7.

[66] Resources for review authors. Suggested risk of bias criteria for EPOC reviews [Internet]. Cochrane Effective Practice and Organisation of Care (EPOC); 2017 [cited 30 Jan 2023]. Available from: https://epoc.cochrane.org/sites/epoc.cochrane.org/files/public/uploads/Resources-for-authors2017/suggested_risk_of_bias_criteria_for_epoc_reviews.pdf

[67] Study Quality Assessment Tools [Internet]. National Heart, Lung, and Blood Institute (NIH); 2023 [cited 2023 Jan 11]. Available from: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools

[68] Sterne JA, Hernán MA, McAleenan A, et al. Chapter 25: assessing risk of bias in a non-randomized study. In: Higgins JPT, Thomas J, Chandler J, et al. editors. Cochrane handbook for systematic reviews of interventions version 6.3. Chichester (UK): John Wiley and Sons Ltd; 2022.

[69] Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919. doi: 10.1136/bmj.i4919.

[70] Wallace SJ, Isaacs M, Ali M, et al. Establishing reporting standards for participant characteristics in post-stroke aphasia research: an international e-Delphi exercise and consensus meeting. Clin Rehabil. 2023;37(2):199–214. doi: 10.1177/02692155221131241.

[71] Dipper L, Marshall J, Boyle M, et al. Treatment for improving discourse in aphasia: a systematic review and synthesis of the evidence base. Aphasiology. 2021;35(9):1125–1167. doi: 10.1080/02687038.2020.1765305.

[72] Kagan A, Simmons-Mackie N, Victor C, et al. Assessment for living with aphasia. 2nd ed. Toronto (ON): Aphasia Institute; 2011.

[73] Rose ML, Pierce JE, Scharp VL, et al. Developments in the application of intensive comprehensive aphasia programs: an international survey of practice. Disabil Rehabil. 2022;44(20):5863–5877. doi: 10.1080/09638288.2021.1948621.

[74] Bernhardt J, Hayward KS, Kwakkel G, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: the Stroke Recovery and Rehabilitation Roundtable Taskforce. Int J Stroke. 2017;12(5):444–450. doi: 10.1177/1747493017711816.

[75] Godecke E, Armstrong E, Rai T, et al. A randomized control trial of intensive aphasia therapy after acute stroke: the Very Early Rehabilitation for SpEech (VERSE) study. Int J Stroke. 2021;16(5):556–572. doi: 10.1177/1747493020961926.

[76] Maclean N, Pound P, Wolfe C, et al. Qualitative analysis of stroke patients' motivation for rehabilitation. BMJ. 2000;321(7268):1051–1054. doi: 10.1136/bmj.321.7268.1051.

[77] Yoshida T, Otaka Y, Osu R, et al. Motivation for rehabilitation in patients with subacute stroke: a qualitative study. Front Rehabil Sci. 2021;2:664758.

[78] Weatherill M, Tibus EO, Rodriguez AD. Motivation as a predictor of aphasia treatment outcomes. Top Lang Disord. 2022;43(3):252–265. doi: 10.1097/TLD.0000000000000286.

[79] Nguy B, Quique YM, Cavanaugh R, et al. Representation in aphasia research: an examination of U.S. treatment studies published between 2009 and 2019. Am J Speech Lang Pathol. 2022;31(3):1424–1430. doi: 10.1044/2022_AJSLP-21-00269.

[80] United States Census Bureau Quick Facts Population Estimates [Internet]. United States Census Bureau; 2020 [cited 2023 Jan 23]. Available from: https://www.census.gov/quickfacts/fact/table/US/PST045221

[81] Rexrode KM, Madsen TE, Yu AYX, et al. The impact of sex and gender on stroke. Circ Res. 2022;130(4):512–528. doi: 10.1161/CIRCRESAHA.121.319915.

[82] What is the prevalence of stroke and TIA in the UK? [Internet]. NICE National Institute for Health and Care Excellence; [cited 2023 Jan 23]. Available from: https://cks.nice.org.uk/topics/stroke-tia/background-information/prevalence/

[83] Kissela BM, Khoury JC, Broderick JP, et al. Age at stroke: temporal trends in stroke incidence in a large, biracial population. Neurology. 2012;79(17):1781–1787. doi: 10.1212/WNL.0b013e318270401d.

[84] Ellis C, Hyacinth HI, Beckett J, et al. Racial/ethnic differences in poststroke rehabilitation outcomes. Stroke Res Treat. 2014;2014:950746–12. doi: 10.1155/2014/950746.

[85] Shrider EA, Kollar M, Chen F, et al. Income and poverty in the United States: 2020 [Internet]. U.S. Department of Commerce; 2023 cited [2023 Jan 23]. Available from: https://www.census.gov/content/dam/Census/library/publications/2021/demo/p60-273.pdf

[86] Gender pay gap at its widest for those in their 50s: average salary of a woman in her 50s is 28% lower than a man's [Internet]. Rest Less; 2023 [cited 2023 Jan 23]. Available from: https://restless.co.uk/press/gender-pay-gap-at-its-widest-for-those-in-their-50s/

[87] Berg JA, Woods NF. Global women's health: a spotlight on caregiving. Nurs Clin North Am. 2009;44(3):375–384. doi: 10.1016/j.cnur.2009.06.003.

[88] Wallace SJ, Worrall L, Rose T, et al. A core outcome set for aphasia treatment research: the ROMA consensus statement. Int J Stroke. 2019;14(2):180–185. doi: 10.1177/1747493018806200.

[89] Wallace SJ, Worrall L, Rose TA, et al. Measuring communication as a core outcome in aphasia trials: results of the ROMA-2 International Core Outcome Set Development Meeting. Int J Lang Commun Disord. 2023;58(4):1017–1028. doi: 10.1111/1460-6984.12840.

[90] Mills T, Lawton R, Sheard L. Advancing complexity science in healthcare research: the logic of logic models. BMC Med Res Methodol. 2019;19(1):55. doi: 10.1186/s12874-019-0701-4.

[91] Northcott S, Thomas S, James K, et al. Solution focused brief therapy in post-stroke aphasia (SOFIA): feasibility and acceptability results of a feasibility randomised wait-list controlled trial. BMJ Open. 2021;11(8):e050308. doi: 10.1136/bmjopen-2021-050308.

[92] Kristinsson S, Den Ouden DB, Rorden C, et al. Predictors of therapy response in chronic aphasia: building a foundation for personalized aphasia therapy. J Stroke. 2022;24(2):189–206. doi: 10.5853/jos.2022.01102.

[93] Moss A, Nicholas M. Language rehabilitation in chronic aphasia and time postonset: a review of single-subject data. Stroke. 2006;37(12):3043–3051. doi: 10.1161/01.STR.0000249427.74970.15.

[94] Sackett DL, Rosenberg WMC, Gray JAM, et al. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71–72. doi: 10.1136/bmj.312.7023.71.

[95] Palmer R, Dimairo M, Cooper C, et al. Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (big CACTUS): a multicentre, single-blinded, randomised controlled trial. Lancet Neurol. 2019;18(9):821–833. doi: 10.1016/S1474-4422(19)30192-9.

[96] Rose ML, Nickels L, Copland D, et al. Results of the COMPARE trial of constraint-induced or multimodality aphasia therapy compared with usual care in chronic post-stroke aphasia. J Neurol Neurosurg Psychiatry. 2022;93(6):573–581. doi: 10.1136/jnnp-2021-328422.

[97] Song F, Parekh S, Hooper L, et al. Dissemination and publication of research findings: an updated review of related biases. Health Technol Assess. 2010;14(8):iii, ix. doi: 10.3310/hta14080.

[98] Menahemi-Falkov M, Breitenstein C, Pierce JE, et al. A systematic review of maintenance following intensive therapy programs in chronic post-stroke aphasia: importance of individual response analysis. Disabil Rehabil. 2022;44(20):5811–5826. doi: 10.1080/09638288.2021.1955303.

[99] Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021;372:n160. doi: 10.1136/bmj.n160.

[100] Stoll CRT, Izadi S, Fowler S, et al. The value of a second reviewer for study selection in systematic reviews. Res Synth Methods. 2019;10(4):539–545. doi: 10.1002/jrsm.1369.

[101] Nussbaumer-Streit B, Klerings I, Dobrescu AI, et al. Excluding non-English publications from evidence-syntheses did not change conclusions: a meta-epidemiological study. J Clin Epidemiol. 2020;118:42–54. doi: 10.1016/j.jclinepi.2019.10.011.

[102] Beveridge MEL, Bak TH. The languages of aphasia research: bias and diversity. Aphasiology. 2011;25(12):1451–1468. doi: 10.1080/02687038.2011.624165.

[103] Cherney L, Doyle P, Hula W, et al. Intensive Comprehensive Aphasia Programs (ICAPs): philosophy, procedures, and outcomes. Paper presented at: American Speech-Language-Hearing Association Conference; 2011; San Diego, CA.