Comparing higher and lower weekly treatment intensity for chronic aphasia: A systematic review and meta-analysis.

John E. Pierce[a,b,c], Robyn O'Halloran[a,c], Maya Menahemi-Falkov[a,c], Leanne Togher[c,d] and Miranda L. Rose[a,c]*

[a]*School of Allied Health, La Trobe University, Melbourne, Australia;* [b]*Speech Pathology, Cabrini Health, Melbourne, Australia;* [c]*Centre of Research Excellence in Aphasia Recovery and Rehabilitation, Australia;* [d]*Speech Pathology, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia*

John E Pierce. ORCID identifier: 0000-0001-5164-5106 Twitter:@johnpierce85 pierce.john.e@gmail.com

Robyn O'Halloran. ORCID identifier: 0000-0002-2772-2164 R.OHalloran@latrobe.edu.au

Maya Menahemi-Falkov. mayam10@gmail.com

Leanne Togher. ORCID identifier: 0000-0002-4518-6748 Leanne.togher@sydney.edu.au

Twitter: @LeanneTogher

Miranda L. Rose. ORCID identifier: 0000-0002-8892-0965 M.rose@latrobe.edu.au

Abstract

Optimising intensity for aphasia treatment is a high priority research issue for people with aphasia, their families and clinicians, and could result in healthcare cost savings. An important aspect of intensity is the frequency of intervention, or how regularly treatment should be provided each week. While principles of neuroplasticity endorse massed practice, cognitive psychology has established superiority of distributed practice within normal learning. Neither concept has been conclusively tested in aphasia. There have been many literature reviews of intensity in aphasia intervention, but most have not investigated treatment intensity whilst also ensuring that therapy dose and treatment type are identical between study groups. Some have also combined studies across acute, subacute and chronic aphasia. We searched systematically for studies directly comparing higher and lower weekly treatment frequency in chronic aphasia. Eight studies were retrieved and rated for methodological quality. Meta-analysis was completed for group and single case experimental designs. Results showed that there are few studies investigating treatment frequency in chronic aphasia and their quality is low-moderate. Meta-analyses were inconclusive due to limited data, but there was no indication of either schedule being superior. Further research directly comparing treatment schedules is needed.

Word Count: 6348 (inc. tables, figures and citations)

Keywords: Aphasia, Systematic Review, Intensity, Therapy, chronic

Introduction

The most recent Cochrane systematic review demonstrated the effectiveness of aphasia therapy after stroke, but concluded that establishing optimum intensity of treatment should be a key aim for future research (Brady, Kelly, Godwin, Enderby, & Campbell, 2016). Intensity is a crucial component of aphasia treatment, with one author describing it as, "possibly the biggest challenge facing speech-language pathologists today — that is, how much treatment is enough to be effective?" (Togher, 2012, p. 438). The Canadian Stroke Network's consensus study placed the optimal intensity of aphasia therapy as the third priority for general stroke research (Bayley et al., 2007).

There are two underlying theories for rehabilitation which provide conflicting guidance for treatment intensity. Results from cognitive psychology research assert that lower frequency training schedules (distributed practice) are best for long-term retention of new skills or knowledge. Rehabilitation can be considered as a form of learning and thus, distributed treatments should result in superior retention in aphasia rehabilitation (For a comprehensive overview, see Dignam, Rodriguez, & Copland, 2016b). In contrast, principles of neuroplasticity promoted within neuroscience, which are largely derived from animal models and studies of human motor and sensory rehabilitation, hold that higher intensity schedules/massed practice facilitate better recovery or learning than lower intensity schedules/distributed practice (Dignam, Rodriguez, & Copland, 2016b). However, neither theory has been conclusively evaluated in aphasia rehabilitation.

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Interpreting intensity

Intensity is a poorly defined term in speech pathology (Baker, 2012a), sometimes referring to the overall concept of treatment scheduling and duration, and at other times, to individual components. At a minimum, the variables that allow accurate reporting of intervention scheduling and intensity are:

1. The duration of each intervention session
2. The number of sessions in a given intervention period – the frequency of therapy sessions – most frequently described by the number of hours or sessions per week (e.g., Brady, Kelly, Godwin, & Enderby, 2012; Cherney, Patterson, & Raymer, 2011; Sage, Snell, & Lambon Ralph, 2011)
3. The duration of the whole intervention, typically number of weeks

In addition, some authors have called for treatment to be described more accurately by reporting treatment in terms of the number of "active ingredients" (e.g., one sentence production attempt) rather than by minutes or hours (Warren, Fey, & Yoder, 2007). In aphasia treatment, active ingredients have not yet been fully defined, but frameworks such as the Rehabilitation Treatment Taxonomy (Turkstra, Norman, Whyte, Dijkers & Hart, 2016) may provide a way to isolate and define such episodes. In this paper, we consider intensity as the amount of intervention provided in a given window, which can be considered at the level of *sessions* (active ingredients per session), *weeks* (active ingredients per week or hours per week), and the *total treatment* period (active ingredients or hours over the total duration of intervention). We will be investigating intensity at the level of each week rather than across the total treatment or within the session, though these are intrinsically linked, because weekly intensity is a commonly reported and practical measure (e.g. scheduling summaries in Cherney, Patterson & Raymer,

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

2011; Brady et al., 2016; Dignam, Rodriguez & Copland, 2016b). We will use the term 'weekly intensity' as we feel is intuitive and transparent to clinicians and researchers alike.

'High' versus 'low' intensity

Weekly intensity has been of great interest to clinicians and researchers in recent years, possibly due to increasing research and implementation of high intensity treatments such as Constraint Induced Aphasia Therapy (CIAT, Pulvermüller et al., 2001). In simple terms, clinicians need to know whether a lot of therapy in a short period of time (high intensity) does more than the same amount in a longer period of time (low intensity). Even within this concept, high and low intensity is described within research using multiple terms, as shown in Table 1.

Table 1

High intensity	Low intensity
Massed practice	Distributed practice
Intensive	Non-intensive
Intense	Non-intense

While most clinicians and researchers would agree that the 15 hours/week prescribed in CIAT is high intensity, there are no standard cut-offs for what constitutes 'low intensity'. Similarly, many papers discuss the efficacy of high intensity treatment for aphasia, yet 'high intensity' does not have a specific meaning or range. Previous research has created arbitrary boundaries for the comparison of high and low intensity (Cherney et al., 2011), with high and low intensity treatments grouped into:

- 3.5-10 versus 2 hours/week (Brady et al., 2016)
- 4-20 versus 1-4 hours/week (Brady et al., 2012)
- >8.8 versus <8.8 hours/week (Bhogal, Teasell, Foley, & Speechley, 2003a)
- 25 versus 4 hours/week (Hinckley & Carr, 2005)

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

- 5 versus 2 hours/week(Ramsberger & Marie, 2007)

Rather than considering weekly intensity in binary categories of high and low frequency, more accurate terminology might be 'high*er*' and 'low*er*' intensity as in reality, schedules exist on a continuum. At the upper end of the continuum lie intensive treatments such as CIAT which are typically 15 hours per week, with some Intensive Comprehensive Aphasia Programs providing up to 24 hours per week on average (Rose, Worrall & Cherney, 2013). Multiple studies have shown that outpatient aphasia treatment is rarely provided at more than 2-3 hours per week (Code & Heron, 2003; Mackenzie et al., 1993; Palmer, Witts, & Chater, 2018; Verna, Davidson, & Rose, 2009) which is the lower end of the continuum. With such a contrast between clinical practice and some high intensity aphasia treatments utilised within research, there will be significant implications for aphasia clinical practice worldwide if research does demonstrate superiority of higher intensity intervention.

Theoretical bases for massed and distributed schedules

Superiority of distributed practice has been demonstrated experimentally for learning and recall with a range of cognitive, verbal and motor activities (Dignam, Rodriguez, & Copland, 2016b). Distributed practice is thought to allow for more rehearsal between practice sessions and thus, deeper encoding (Moulton et al., 2006). In addition, longer intervals between sessions reduces the ability to rely on priming from the previous session and encourages true recall, promoting deeper changes to the underlying representations (Sage et al., 2011).

In contrast, it has also been shown that neural connections are created or strengthened when events occur simultaneously and the strength of such connections increases proportionally with the frequency of occurrences (Pulvermüller & Berthier, 2008). For aphasia, the concept has

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

been described as the Massed Practice Principle; that is, the hypothesis that more treatment and higher treatment intensity result in superior gains compared to less treatment and/or lower intensity (Pulvermüller & Berthier, 2008). The same reasoning leads to the proposal of a minimum threshold of treatment within a given timeframe that needs to be exceeded for the neural system to activate repair of connections or establish new pathways (Dignam et al., 2015; Harnish, Neils-Strunjas, Lamy, & Eliassen, 2008; Kleim & Jones, 2008).

The presence or absence of an activation threshold for aphasia treatment is a crucial piece of knowledge for clinicians. Therapy provided below this proposed minimum threshold would provide suboptimal results, or at worst, be completely ineffective (Baker, 2012a). Current low rates of treatment intensity (<3 hours/week) in clinical practice may sit below such a threshold. The risk of ineffective treatment is one reason treatment intensity is regarded as such a fundamental and pressing question for speech pathology (Baker, 2012a) and some have argued that if true, lower intensity would be unethical (Togher, 2012). However, neuroplasticity theory is largely based on research from motor actions and animal models to date (Kleim & Jones, 2008), and it remains unclear whether neuroplasticity or cognitive psychology models are more suitable for language recovery.

As a counterpart to a minimum threshold of activation, there might also be a ceiling level of treatment per week. In a synthesis of paediatric treatment for phonological awareness and print concepts, Schmitt and Justice (2012) noted that increasing the total dose does not indefinitely result in superior outcomes. They predicted diminishing returns after a point, where more treatment would not necessarily be better. Assuming a parallel within aphasia intervention, it is probable that there is an upper limit of effectiveness in terms of weekly dose, perhaps due to redundancy and patient fatigue. In this case speech pathology intervention provided *more*

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

frequently than required would also be a waste of resources (Baker, 2012a), potentially stressful or even harmful.

Practical considerations

Conflicting cognitive psychology and neuroscience models notwithstanding, there are practical considerations for both higher and lower intensity of treatment. The higher treatment intensity of 15 hours per week found in some research studies is unlikely to be feasible within current healthcare models. In a survey of clinicians in the USA, 60% reported CIAT would be *very difficult* or *extremely difficult* to administer in their facility and 90% felt they would be unable to implement CIAT at all (Page & Wallace, 2014). Although this survey specifically enquired about CIAT, responses centred around the challenges of high intensity rather than specific treatment components of CIAT and the results are therefore likely to apply to other intensive treatments. Reimbursement from health insurers was also a common concern. Some authors predict that the financial constraints of healthcare are unlikely to change significantly in the future (Code & Petheram, 2011) which would preclude uptake of more expensive treatments.

While feasibility does require consideration when developing treatments, the architecture of current service models should not solely dictate what treatments are developed and researched (the "tail wagging the dog"). The other counter-argument to higher intensity treatment being too challenging to implement is that, should higher intensity treatment prove to be superior, redesigning services to provide the *same* total dose of treatment in a shorter span of time could be an economically rational measure, as there would be greater recovery for the same hours. Given significant funding and resource limitations for aphasia rehabilitation worldwide (Rose, Ferguson, Power, Togher, & Worrall, 2013), funding could be best spent on shorter but more

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

frequent intervention (Harnish et al., 2008). Such comparative effectiveness and economic data is not yet available.

Treatment adherence is another potential barrier to higher intensity treatments. Within physiotherapy, a survey found that the majority of patients were unwilling to participate in Constraint Induced Movement Therapy (CIMT) for post-stroke hemiparesis (Page, Levine, Sisto, Bond, & Johnston, 2002). The intensity of CIMT, which involves 6 hours/day over 2 weeks, was a major concern for these respondents. Surveyed clinicians also predicted their patients would be unlikely to adhere to CIMT. In a parallel survey for CIAT, more than 60% of clinicians believed their patients would be *unlikely* or *very unlikely* to adhere to the treatment protocol (Page & Wallace, 2014). Data on dropouts from a systematic review of aphasia treatment supports this view, finding significantly higher dropouts from more intensive treatments ($p = .03$) in acute and subacute aphasia (Brady et al., 2016). However, the analysis showed no significant difference in dropouts between higher and lower intensity treatments in chronic aphasia. This result might mean that people with chronic aphasia can tolerate intensive treatments more readily than those in acute or sub-acute phases of recovery. Another explanation is that people in acute and subacute phases are in the process of adjusting to the stroke and learning about their capacity to participate in a range of life activities including research activities. This may mean they are more likely to consent initially without fully appreciating the demands of the research and later drop out. People with chronic aphasia may be better able to predict the feasibility of attending intensive treatments at the time of consenting. Additionally, people with chronic aphasia generally cannot be routinely approached to participate in research and so researchers would likely recruit enthusiastic participants who were motivated to approach the researcher. Hence, the

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

acceptability of higher intensity treatments across the population with chronic aphasia is not yet clear.

Another concern reported by clinicians when asked about higher intensity treatment was burden on caregivers in transporting patients to and from each session (Page & Wallace, 2014). New models of care such as telehealth offer a possible solution to travel and distance. However, in two studies of clinician perceptions after administration of higher intensity treatments, other negatives reported were patient and clinician fatigue (Gunning et al., 2016), frustration in patients who make limited progress (Gunning et al., 2016) and unrealistic patient expectations of progress (Babbitt, Worrall, & Cherney, 2013). The time required to plan and provide treatments was also a concern (Babbitt et al., 2013). However, clinicians also identified that the progress seen in patients was highly motivating and helped them appreciate the gains that are possible (Babbitt et al., 2013). They described increased patient confidence (Gunning et al., 2016) and improved relationships between patients, family and clinicians (Babbitt et al., 2013; Gunning et al., 2016). For the clinicians themselves, reported rewards included better teamwork and support and learning new techniques and clinical skills (Babbitt et al., 2013; Gunning et al., 2016), though some of these benefits might be attributable to the group aspect of therapy rather than the higher intensity. Finally, in one of the studies, clinicians reported that returning to typical (non-intense) clinical practice was difficult and even "depressing," as they felt that the designs of their services presented fewer opportunities to offer meaningful treatment gains and high quality therapy (Babbitt et al., 2013).

Previous research on intensity

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

So far, theoretical and practical arguments have been discussed for higher versus lower intensity treatments, but in clinical research, is higher or lower weekly intensity more efficacious for patient outcomes? There have been multiple previous reviews on intensity, both systematic and narrative, with conflicting results. However, the methods in these reviews do not allow conclusions to be made about treatment frequency in chronic aphasia.

Most past reviews have examined studies that compared different treatment frequencies, but also different total durations of treatment (Brady et al., 2016; Cherney, Patterson, Raymer, Frymark, & Schooling, 2008). For example, the most recent Cochrane review concluded that there is some tentative evidence for higher intensity being more efficacious, yet the higher intensity arms of these studies provided a mean of 84 total hours per participant, while lower intensity arms provided 44 hours (Brady et al., 2016). Thus, with higher intensity treatments providing nearly double the treatment hours, the effects of *more* therapy overall and the *frequency* of therapy are conflated.

Previous reviews have also included papers that provide a *different* therapy in each arm (Brady et al., 2016; Cherney et al., 2008). For example, Pulvermüller et al. (2001) is often included as evidence in favour of high intensity treatment (as in Bhogal, Teasell, & Speechley, 2003b), yet that research compared two different treatment approaches in each arm and thus the contribution of the intensity and the type of treatment cannot be separated. To isolate the effect of treatment intensity, the same treatment should be offered at each intensity.

Finally, some reviews of intensity aggregated data from all phases of aphasia recovery, including acute, subacute and chronic (Bhogal et al., 2003b). Based on previous analyses of treatment effect sizes, early and chronic aphasia should not be expected to improve by the same magnitude (Brady et al., 2016; Robey, 1998). The specific response to higher or lower intensity

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

of treatment could also differ between chronic and acute/subacute patients (Cherney et al., 2011). For example, the principles of neuroplasticity that promote higher intensity therapy might be more relevant to a brain which has recently been injured, and the cognitive psychology findings related to enhanced learning in a lower intensity might apply more to the more 'stable' neurophysiology of a brain with chronic aphasia.

In order to accurately determine whether 'a lot in a little time does more than a lot in a longer time,' treatment schedules need to be altered between treatment arms with all other variables controlled, including the type and total dose of therapy and the participant characteristics. One review to date has been careful to include only studies with these criteria (Dignam, Rodriguez, & Copland, 2016b). This paper was a narrative review of research up to November 2014 and identified four papers. The authors concluded that there is some preliminary suggestion of superiority of lower intensity treatment when considering the longer-term maintenance timepoints. As a narrative review, and no doubt due to the low yield of papers, Dignam et al. did not attempt a meta-analysis. As a high priority topic, it is likely that further relevant evidence has been published in the past five years. The aim of this paper is to systematically review papers that directly compare higher and lower weekly intensity treatments while controlling other variables and, if feasible, meta-analyse results of high-quality papers.

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Methods

Four major databases were searched in June 2018: Medline (OVID, 1946-present), CINAHL (Ebscohost), PubMed and Psycinfo (OVID, 1987-present). The search strategy combined two key concepts, **aphasia** and **treatment intensity**, and a variety of search terms to represent these were used. An iterative approach using known results was used to ensure all relevant terms were included. An example of the search strategy employed is found in Table 2. Medical Subject Headings (MeSH) and keyword searches were used. Searches were limited to English only and terms were exploded where databases allowed. Results were imported into citation management software to identify duplicates manually and automatically, and non-English papers were excluded. Results were then exported into Rayyan (Ouzzani, Hammady, Fedorowicz, & Elmagarmid, 2016), online systematic review software, for screening.

Table 2 – Search strategy example, Medline (OVID, 1946-present)

Keywords:	aphasi* dysphasi* anomic anomia NOT "progressive aphasia"	**AND**	duration total hours total therapy total treatment total intervention dose cumulative intervention intensity intensive amount of therapy amount of treatment amount of intervention intense
	OR		
Subject headings:	*exp: Aphasia/* *exp: anomia/* *NOT exp: Aphasia, primary progressive/*		

Note. All subject headings were exploded, and all subheadings were included.

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

A search was also conducted in Google Scholar in August 2018 using a variety of search phrases (see Table 3). As recommended in systematic review research (Haddaway, Collins, Coughlin, & Kirk, 2015), the first 300 results were collected, and then the search was repeated using the operator "AllInTitle:" which only returns results with search terms in the titles.

The first and third author independently screened remaining articles on Rayyan using the following inclusion criteria:

1. English language

2. Adults with chronic stroke-induced aphasia (≥ 6months)

3. Original research data

4. Two schedule (intensity) conditions with the same total dose of therapy (e.g. 30 hrs at each intensity)

After screening, the two authors discussed and resolved discrepancies in inclusion and exclusion decisions. Full texts of the remaining articles were then obtained and further examined according to the inclusion criteria. Reference lists of resulting articles were inspected by the first author to identify any additional studies.

Table 3 – Search strategy, Google Scholar

Search 1:	"aphasia duration" OR "total hours" OR "total therapy" OR "total treatment" OR "total intervention" OR dose OR "cumulative intervention" OR intensity OR intensive OR "amount of therapy" OR "amount of treatment" OR "amount of intervention" OR intense"
Search 2:	Allintitle: "aphasia duration" OR "total hours" OR "total therapy" OR "total treatment" OR "total intervention" OR dose OR "cumulative intervention" OR intensity OR intensive OR "amount of therapy" OR "amount of treatment" OR "amount of intervention" OR intense"

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Included articles were categorised by study type using a modified version of the Oxford Centre for Evidence Based Medicine levels of evidence (OCEBM Levels of Evidence Working Group, 2011). An additional level was included for Single Case Experimental Designs (SCEDs) outside of the n-of-1 design, including multiple baseline, changing criterion, alternating treatment and withdrawal designs. Rather than considering mean changes within or between groups, SCEDs repeatedly assess individuals over time and determine whether manipulation of treatment (e.g. commencing and withdrawing) results in clear changes to outcomes. When conducted rigorously, SCEDs are recognised alongside n-of-1 designs as comprising a high level of experimental control (Tate et al., 2013). The first and third authors categorised papers and reached agreement.

Randomised and non-randomised controlled trials were rated for methodological rigour with the PEDro-P scale (Murray et al., 2013). SCEDs were rated for methodological rigour with the RoBiNT scale (Tate et al., 2015). PEDro-P was rated by the first and third authors and SCEDs by the first and last authors. Discrepancies were resolved by discussion. Remaining study types, including pre/post case series and non-experimental single case designs such as single phase or biphase designs, were not rated as these typically contain no experimental control and form a low level of evidence (Tate et al., 2013).

Data from each paper was extracted to a spreadsheet including participant characteristics, treatment type, outcome measures, therapy schedule, and results.

Effect size calculation - group studies

The primary outcome measures at the post-intervention timepoint were meta-analysed as all were measures of expressive language. The mean and standard deviation of change scores per arm were calculated and analysed using RevMan (The Cochrane Collaboration, 2014). A random

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

effects model was applied with a standardised mean difference, given the different outcome measures across studies. Only data from the first phase was considered in crossover studies.

For the maintenance timepoint, studies containing data at one month post intervention were meta-analysed. For Mozeiko et al. (2015), the one month follow-up scores and mean baseline scores were used to calculate a change score as a percentage of baseline within each participant. This approach was more conservative than using Tau-U which would have resulted in ceiling (1.0) or floor (-1.0) scores due to there being only one data point at maintenance. A forest plot was generated using Revman.

Effect size calculation - single case experimental designs

Data points were measured graphically using software image measurements when not obvious directly from the chart. Baselines were corrected for trend if the baseline phase Tau was >0.40 *and* a trend was apparent on visual analysis. Tau was consistent with visual analysis on each occasion. Tau-U was calculated using the software at www.singlecaseresearch.org. To meta-analyse the maintenance timepoint, only SCEDs that included "withdrawal" phases (Ramsberger & Marie, 2007; Raymer, Kohen, & Saffell, 2006) had Tau-U applied, and this was conducted as baseline compared to maintenance data. This comparison was made because language behaviour is unlikely to rapidly react to withdrawal of treatments and the improvement compared to pre-treatment is the most relevant outcome. A forest plot was generated using the DistillerSR Forest Plot Generator from Evidence Partners (Evidence Partners, 2019) and overall effect sizes were calculated weighted by the number of pairs.

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Results

Results of the search and screening process are displayed in Figure 1. The initial yield was high (2462) but after screening, eight papers remained which represented seven studies (two papers reported on different aspects of the same study). Hand searching reference lists revealed two additional references. One was added to the review while the other was a conference abstract and the authors stated the data was not ready to be released (Rochon et al., 2016). The final yield therefore included nine articles reporting on eight studies. It should be noted that Stahl et al. (2018) was a comparison of different schedules *and* total dose, but published data allowed comparison of the two groups at a timepoint where each had received 24 hours of treatment.

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

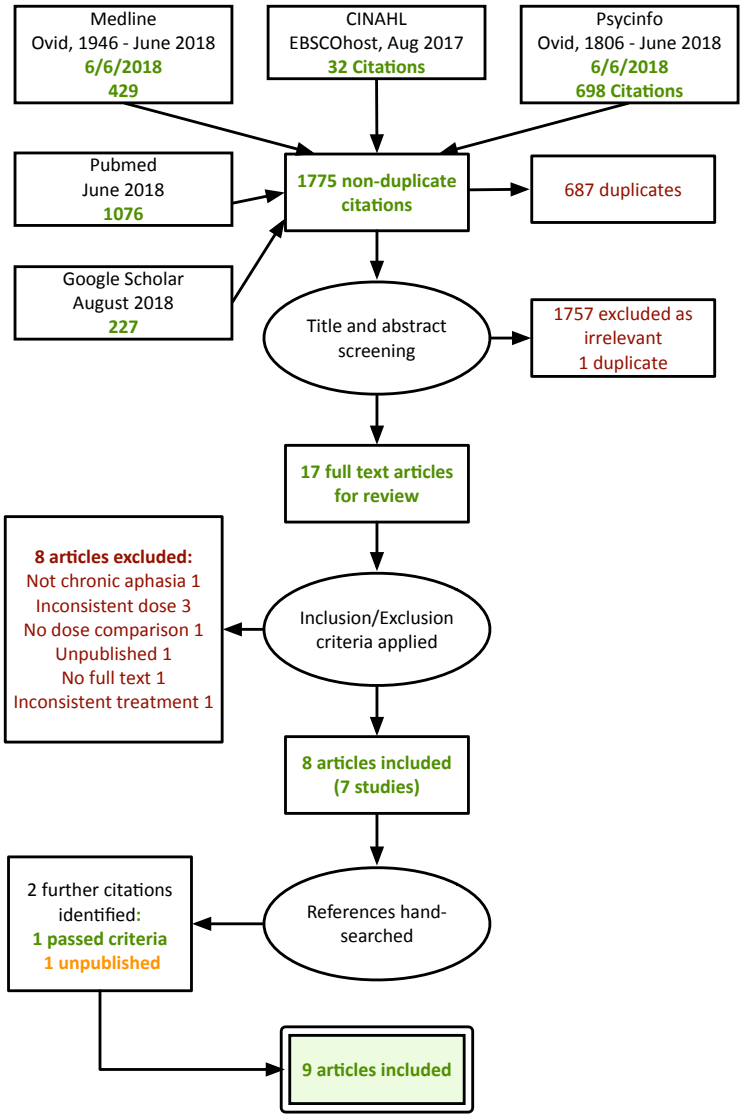

Figure 1 – Search strategy results

Level of evidence

Results included one randomised controlled trial (Stahl et al., 2018) and one non-randomised controlled trial represented by two papers (Dignam et al., 2015; Dignam, Copland, Rawlings, OBrien, et al., 2016a). Sage, Snell and Lambon Ralph (2011) employed a crossover

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

design with randomisation. When considering only the data from the first condition for each participant, this design can be considered a randomised controlled trial. There were three SCEDs: two crossover multiple baseline designs (Ramsberger & Marie, 2007; Raymer et al., 2006) and a biphasic design (Mozeiko, Coelho, & Myers, 2015). Two papers were pre/post case series (Harnish et al., 2008; Marcotte et al., 2018) and were not considered in meta-analysis.

Descriptive characteristics of data

The Appendix contains extracted data for retrieved studies. Overall, there were 92 participants across the studies, with a mean reported age of 59 (SD 13). There were 60 males and 32 females, a higher proportion of males (1.875 : 1) than overall stroke populations (1.4 : 1) (Appelros, Stegmayr, & Terént, 2009). The mean number of months post stroke in papers reporting this data was 48 (range 4 months to 21 years), though these were not normally distributed, with 65/83 (78%) participants less than six years post stroke. Data from participants who were less than six months post stroke were excluded from meta-analysis.

Four studies investigated traditional, cueing hierarchy-based naming treatments (two using software), two Constraint Induced Aphasia Therapy, one an Intensive Comprehensive Aphasia Program and one Phonological Components Analysis. The total dose ranged from 10-48 hours, assuming sessions were one hour, as some reported in sessions per week rather than hours. The weekly intensity for the higher intensity schedules (as labelled by the authors) was 3-16 hours per week and for the lower schedules, 1-6 hours. Within each study, the weekly dose for higher intensity schedules was at least double the lower intensity schedule, with one paper providing five times more per week (Mozeiko et al., 2015).

The six studies used a range of impairment-based measures as their primary outcome, including picture naming of treated items (3), a naming test (1), subtests from a language battery

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

(1) and a discourse measure (1). Only two studies investigated activity/participation outcomes (Dignam et al., 2015; Mozeiko et al., 2015), both as secondary outcomes, while only Dignam et al. investigated quality of life.

Methodological quality

Methodological ratings for papers are displayed in Table 4. For group trials, Stahl et al. (2018) was rated as high quality methodology (7/10) while Dignam et al. (2015) and Sage et al. (2011) were rated as fair (4/10), as per commonly accepted benchmarks for the Pedro scale (Stroke Engine, n.d.). For the RobinT, a recently published algorithm indicates the risk of bias for internal validity (Perdices, Tate & Rosenkoetter, 2019). The algorithm gives higher weighting to items considered more important for internal validity. Using this algorithm, the SCEDs in this review had low (Ramsberger & Marie, 2007; Raymer et al., 2006) and very low (Mozeiko et al., 2015) methodological rigour for internal validity. The RobinT does not yet have consensus on a cut-off or benchmark of scores for overall methodology that includes external validity.

Table 4 – Methodological ratings

PEDro-P	**Total**	Eligibility criteria + source	Random Allocation	Allocation concealed	Baseline Similarity	Participant Blinding	Therapist blinding	Assessor blinding	85% retention	ITT Analysis	b/w group Comparisons	Point estimates and variability
Sage et al., 2011	**4**	✓	✓	×	×	×	×	×	✓	×	✓	✓
Stahl et al., 2018	**7**	✓	✓	✓	✓	×	×	✓	✓	×	✓	✓
Dignam et al., 2015	**4**	×	×	×	✓	×	×	×	✓	×	✓	✓

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

RobinT	Total	Design with control	Randomisation	Sampling behaviour	Participant Blinding	Assessor blinding	Interrater agreement	Treatment Adherence	Baseline characteristics	Setting	Dependent variable defined	Independent variable defined	Raw data record	Data analysis	Replication	Generalisation
Mozeiko et al., 2015	**15**	×	×	1	×	×	1	×	2	1	2	1	×	2	2	1
Ramsberger & Marie, 2007	**17**	1	1	×	×	×	1	×	2	1	2	2	2	2	2	1
Raymer et al., 2006	**17**	1	1	1	×	×	1	×	2	×	1	2	×	2	2	2

Meta-analysis - immediately post intervention

The forest plot for outcomes immediately post intervention for the three group trials is shown in Figure 2. The total overall estimate indicates that a distributed, lower weekly intensity schedule is superior (p = 0.02), though this result is based on a small data set (n=70). The Tau-U meta-analysis chart for three SCEDs is shown in Figure 3 (see note on interpretation). Visual comparison of higher and lower intensity effect sizes reveals that higher intensity results had a marginally higher weighted effect size, but the confidence intervals of both schedules overlap, indicating no significant differences.

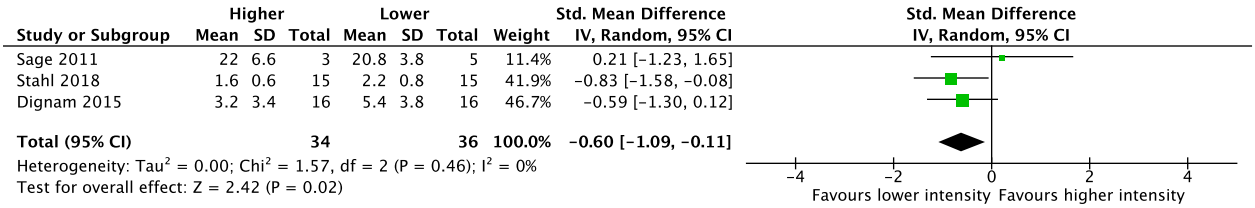

Figure 2 – Forest plot for group studies, immediately post treatment

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

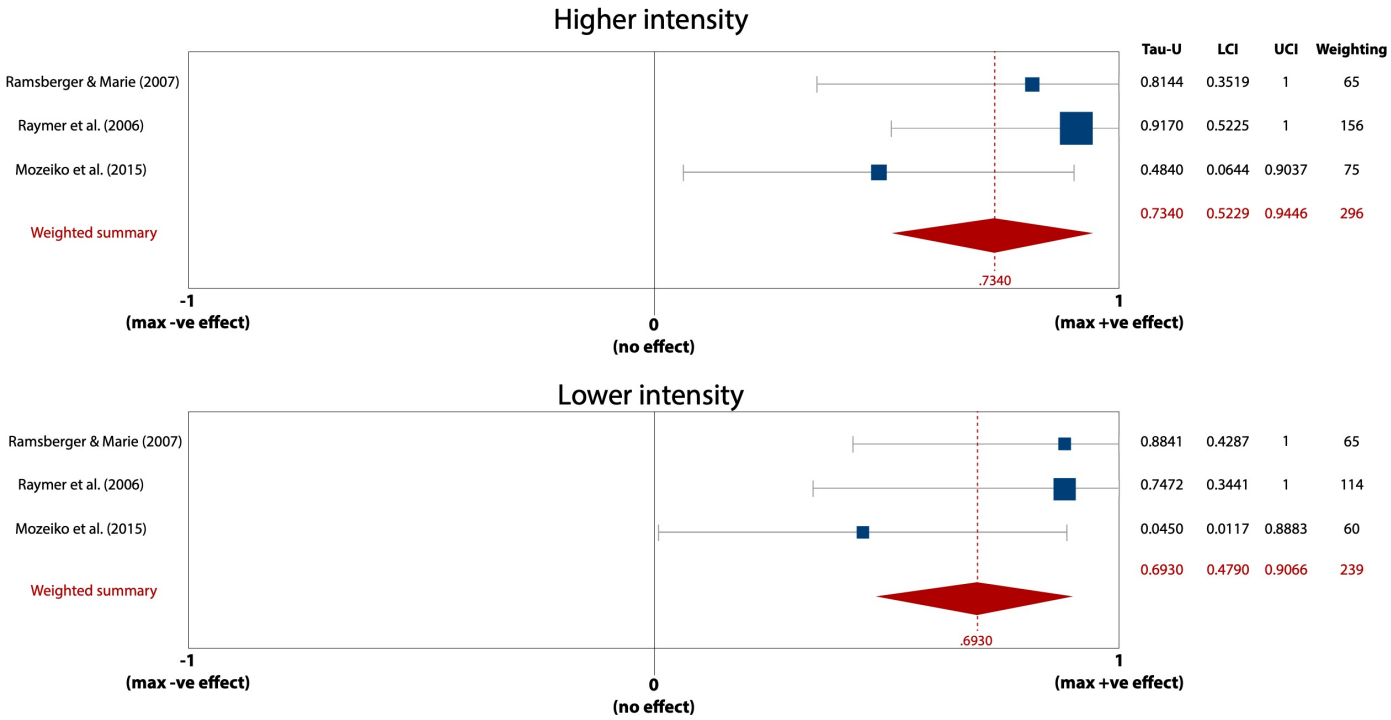

Figure 3 – Tau-U plot for single case experimental designs, immediately post treatment
Note: Upper plot is higher intensity phases, lower plot is lower intensity. Lines −1.0 and 1.0 indicate the possible limits of Tau-U scores; positive values demonstrate improvement relative to baseline. Dotted lines indicate the weighted average and diamond width indicates the 95% confidence interval. Square sizes indicate relative weighting for each study.

Meta-analysis - one month post intervention

The total overall estimate in the forest plot (Figure 4) shows no significant difference between intensities at this timepoint, again based on limited data (n=47). In the Tau-U chart for two SCEDs (Figure 5), both charts show a large effect size with no appreciable difference between intensities.

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

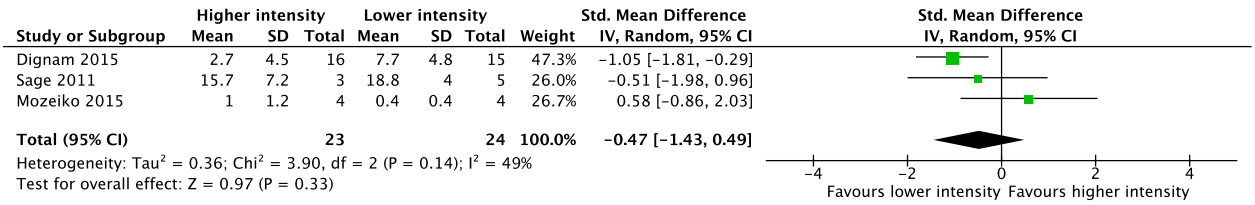

Figure 4 – Forest plot one month post treatment

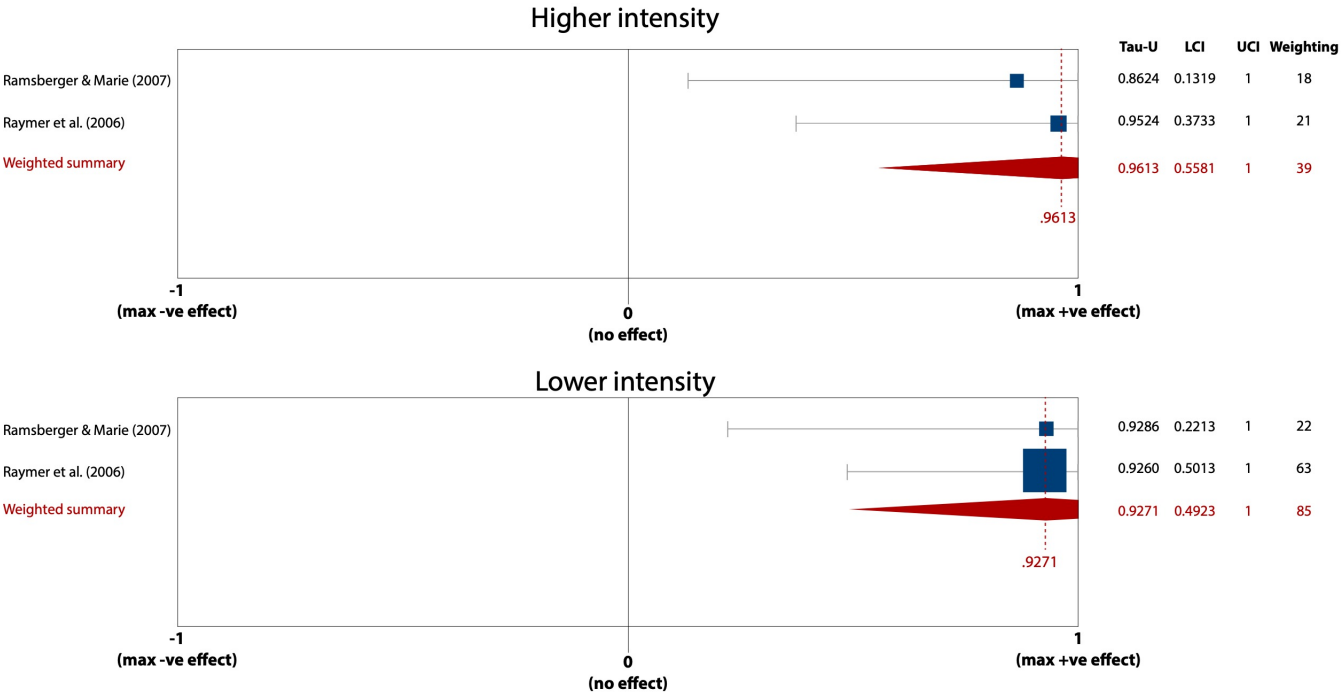

Figure 5 – Tau-U plot for single case experimental designs, one month post treatment

Secondary Outcomes

In the two studies employing activity/participation outcomes (Dignam et al., 2015; Mozeiko et al., 2015), results do not clearly favour either intensity. Mozeiko et al. (2015) used the Communication Activities of Daily Living-2 (CADL-2) which assesses functional communication in any modality across a range of simulated scenarios. Participants in the higher intensity arm were unchanged or improved (0, +6, +11, +25 point changes) and mixed in the lower intensity arm (-14, -12, +3, +38 point changes). Dignam et al. (2015) found no significant

between group differences on the Communication Effectiveness Index (P=0.05 post, P=0.21 followup) or the Communication Confidence Rating Scale for Aphasia (P=0.79 post, P=0.48 followup), although both improved compared to baseline. In the one study that employed a quality of life outcome measure (Dignam et al. 2015), no difference was found between higher and lower intensity groups (P=0.37 post, P=0.75 followup).

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Discussion

Considering the importance of the question of weekly intensity for clinicians, patients and health service funding bodies, this systematic review found a surprisingly low number of studies examining the question directly in chronic aphasia. However, the eight studies retrieved in this review includes four new studies (Dignam et al., 2015; Marcotte et al., 2018; Mozeiko et al., 2015; Stahl et al., 2018) since the recent narrative review by Dignam, Rodrizuez and Copland (2016b), particularly considering we limited our search to chronic aphasia whereas they included all phases. More may have been published in acute or subacute aphasia since this time. It is unclear why there have not been larger scale studies regarding this question. Perhaps the relative efficacy of different approaches – another important topic for aphasia – is considered by researchers to be a higher priority than scheduling of any single treatment. Whatever the reason for the low number of studies, the demand for guidance on intensity demonstrates a disconnect between research priorities and practice.

In our review results, both the SCEDs and the group trials had high risk of bias. To more stringently examine the influence of weekly intensity on outcomes, group studies should randomise participants (2/3 group studies were randomised) and use blinded assessors (1/3). An intention to treat approach (0/3) is also recommended, as the acceptability and feasibility of different treatment schedules has not been clearly established for chronic aphasia. SCEDs addressing this question should employ more stringent designs such as Multiple Baseline Designs, collect a minimum of five data points per phase, and randomise the point at which treatment commences.

Impairment based outcome measures were employed with all retrieved studies, however, only two used activity/participation level outcomes and one used a quality of life measure; all as

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

secondary outcomes. People with aphasia and their families have identified that impairment, activity/participation and wellbeing levels are important priorities for treatment outcomes (Wallace et al., 2016). There were no clear differences in the data retrieved in this review, but it is possible that higher and lower intensity schedules have different effects on these outcome domains compared to impairment, and thus more investigation of such outcomes with contrasting weekly intensities is required. Specifically, a Core Outcome Set was recently developed based on patient, family and clinician input and use of these outcome measures will ensure that future research is consistent, comparable and covers key domains (Wallace et al., 2016).

Although studies employed a disparate range of weekly schedules, within each study the schedules were sufficiently varied to allow comparison, with higher intensity arms at least double the weekly dose as the lower intensity arms. In addition, 6/8 studies used <3 hours per week for the lower intensity arm which is a schedule similar to most clinical practice worldwide (Mackenzie et al., 1993; Palmer et al., 2018; Verna et al., 2009). Comparison of typical clinical schedules to higher intensity means that results can be used to aid decision making with regards to increasing intensity from current clinical levels.

Importantly, for most papers *both* treatment intensities showed a positive mean change for the primary outcome measure. Overall, results tentatively suggested that no schedule was ineffective and thus current clinical practice, although considered low intensity, is likely superior to no treatment – a finding consistent with the most recent Cochrane review of aphasia treatment (Brady et al., 2016).

The meta-analysis conducted in this review is preliminary and limited by a low yield of papers, low participant numbers and insufficiently rigorous designs. The immediate post-

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

treatment meta-analysis for group studies was drawn from a very small pool of data, but it is interesting that the lower weekly intensity had a stronger effect at this timepoint. If this finding was replicated in future research, this would contradict neuroplasticity principles which predict better immediate performance with intensive practice. However, the forest plot does show wide confidence intervals which indicates uncertainty of the result, and the results shown on the SCED figure did not favour either intensity at the same timepoint.

For the meta-analysis of the maintenance timepoint (1 month), there was no apparent difference between schedules on either meta-analysis, yet cognitive psychology predicts that lower intensity would be more favourable in the longer term. If intensity had a simple linear correlation with outcomes, some sort of trend in results might be expected given the significant difference in schedules between participants, even considering the limitations of this meta-analysis. Thus, the inconclusive result of this review is an interesting outcome in itself.

Higher versus lower intensity is a crucial issue, but it should not be taken as the only important treatment variable. Many authors have suggested that there are multiple interacting factors that moderate response to intensity (Baker, 2012b; Cherney et al., 2011; Dignam, Rodriguez, & Copland, 2016b). As one example, Baker (2012b) described a finding in paediatric treatment where the mean response to 1/week versus 5/week schedules was non-significant, but those without Down Syndrome and those who played with many objects did respond more to higher intensity (Yoder et al. 2012, as cited in Baker, 2012b). Within this review, Ramsberger and Marie (2007) had one participant respond more favourably to higher intensity while the remaining three showed no difference between schedules. Sage, Snell and Lambon-Ralph (2011) also reported one participant who demonstrated better naming at both timepoints for the wordlist

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

treated at higher intensity, while the other participants showed no significant overall difference. There were no obvious factors in this participant to explain their stronger response.

Kiran and Thompson (2019) outlined a list of variables that influence recovery from stroke related aphasia. These included aphasia severity, cognitive reserve, psychosocial environment, age at time of stroke, lesion volume, lesion location and impact on connectivity, perfusion characteristics and treatment variables. The authors point out that these factors are not linearly correlated with recovery and form a complex picture. Any one of these factors, and probably many of them, may interact with intensity so that particular patient profiles respond better to higher intensity and others to lower intensity (Brady et al., 2016). These questions require large datasets to solve.

Further to moderating factors, intensity is not simply a matter of "counting the hours" (Togher, 2012). For example, the active ingredient is important to report as it provides a more exact estimate and description of the treatment provided (Baker, 2012b). In addition, the number of active ingredients per session provided to participants is likely to vary across therapist and patient, depending on factors such as talkativeness and severity. As an extreme, one hour of therapy might contain, for example, several hundred naming attempts or only a handful. In comparative studies, active ingredients may therefore be an uncontrolled variable where not reported. While all studies retrieved in this review reported on the number of hours or sessions provided to participants, the active ingredient was not described with the exception of Dignam et al. (2015) who showed that the session dose and total dose were not significantly different between groups. Randomised trials of sufficiently large sample size are likely to control for this variance.

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Limitations

This review considered only studies which compared different weekly intensities in a binary manner. Other reviews have also included papers in which treatment hours varied between each participant, allowing a correlational analysis of intensity (Cherney et al., 2011). In our case we were looking for maximally contrasting schedules in order to find any clear indications of higher or lower intensity being superior to one another. For some retrieved papers, the duration of each session was not reported and we had to assume these were one hour. Although this assumption might be imprecise, our analysis was based on the comparison *within each study*, thus the units involved are not crucial. Assuming a simple linear relationship (e.g. higher intensity = better outcomes), each paper should have found better outcomes for the higher intensity treatment regardless of units. As higher and lower intensities overlapped between retrieved studies, this review is unable to provide guidance on *optimal* weekly intensity but instead allowed us to investigate the simple linear predictions of neuroplasticity and cognitive learning.

Conclusions

Results of this review are preliminary, given the limited number of studies retrieved, yet they did not support the predictions of neuroplasticity or cognitive psychology. If anything, lower intensity was marginally more favourable immediately after treatment while no difference was demonstrated at one month follow up. This is in contrast to many current perspectives – higher intensity is often described as superior as a foregone conclusion, yet this review demonstrates that there is no strong evidence for this in chronic aphasia.

As described above, many authors have concluded that treatment response is likely to emerge as a highly complicated system of interacting factors. Speech pathologists seeking an

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

answer to the question of how often to schedule therapy may find this frustrating. Nonetheless, even lower intensity treatments had favourable outcomes and therefore clinicians who are limited to these treatment schedules due to funding and policy constraints can be confident in providing lower intensity services. It will be many years before large scale models are produced which provide accurate predictions of treatment response based on patient and treatment factors – an example of such work is in progress from the RELEASE collaboration (Brady et al., 2019). However, despite the complexity of treatment scheduling, weekly intensity remains an important piece of the treatment puzzle. The call for larger RCTs directly comparing aspects of intensity, put forward by many before (Cherney et al., 2011; Dignam, Rodriguez, & Copland, 2016b), remains. Larger sample sizes with low risk of bias will control for the variance between clinicians, patient demographics and aphasia and stroke profiles and ultimately add data to this question.

Acknowledgements

The work was supported in part by a La Trobe University Postgraduate Research Scholarship and the NHMRC Centre for Research Excellence in Aphasia Recovery and Rehabilitation (#1153236).

Declaration of Interest

The authors report no conflicts of interest

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

References

Appelros, P., Stegmayr, B., & Terént, A. (2009). Sex Differences in Stroke Epidemiology. *Stroke*, *40*(4), 1082–1090.

Babbitt, E. M., Worrall, L. E., & Cherney, L. R. (2013). Clinician perspectives of an intensive comprehensive aphasia program. *Topics in Stroke Rehabilitation*, *20*(5), 398–408. https://doi.org/10.1310/tsr2005-398

Baker, E. (2012a). Optimal intervention intensity. *International Journal of Speech-Language Pathology*, *14*(5), 401–409. https://doi.org/10.3109/17549507.2012.700323

Baker, E. (2012b). Optimal intervention intensity in speech-language pathology: Discoveries, challenges, and unchartered territories. *International Journal of Speech-Language Pathology*, *14*(5), 478–485. https://doi.org/10.3109/17549507.2012.717967

Bayley, M. T., Hurdowar, A., Teasell, R., Wood-Dauphinee, S., Korner-Bitensky, N., Richards, C. L., et al. (2007). Priorities for stroke rehabilitation and research: results of a 2003 Canadian Stroke Network consensus conference. *Archives of Physical Medicine and Rehabilitation*, *88*(4), 526–528. https://doi.org/10.1016/j.apmr.2007.01.005

Bhogal, S. K., Teasell, R. W., Foley, N. C., & Speechley, M. R. (2003a). Rehabilitation of aphasia: more is better. *Topics in Stroke Rehabilitation*, *10*(2), 66–76. https://doi.org/10.1310/RCM8-5TUL-NC5D-BX58

Bhogal, S. K., Teasell, R., & Speechley, M. (2003b). Intensity of aphasia therapy, impact on recovery. *Stroke; a Journal of Cerebral Circulation*, *34*(4), 987–993. https://doi.org/10.1161/01.STR.0000062343.64383.D0

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Brady, M. C., Kelly, H., Godwin, J., & Enderby, P. (2012). Speech and language therapy for aphasia following stroke. *The Cochrane Database of Systematic Reviews*, *5*(5), CD000425. https://doi.org/10.1002/14651858.CD000425.pub3

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *The Cochrane Database of Systematic Reviews*, *6*(6), CD000425–CD000425. https://doi.org/10.1002/14651858.CD000425.pub4

Cherney, L. R., Patterson, J. P., & Raymer, A. M. (2011). Intensity of aphasia therapy: evidence and efficacy. *Current Neurology and Neuroscience Reports*, *11*(6), 560–569. https://doi.org/10.1007/s11910-011-0227-6

Cherney, L. R., Patterson, J. P., Raymer, A. M., Frymark, T., & Schooling, T. (2008). Evidence-Based Systematic Review: Effects of Intensity of Treatment and Constraint-Induced Language Therapy for Individuals With Stroke-Induced Aphasia. *Journal of Speech, Language, and Hearing Research*, *51*(5), 1282–1299. https://doi.org/10.1044/1092-4388(2008/07-0206)

Code, C., & Heron, C. (2003). Services for aphasia, other acquired adult neurogenic communication and swallowing disorders in the United Kingdom, 2000. *Disability and Rehabilitation*, *25*(21), 1231–1237. https://doi.org/10.1080/0963828031000159996l

Code, C., & Petheram, B. (2011). Delivering for aphasia. *International Journal of Speech-Language Pathology*, *13*(1), 3–10. https://doi.org/10.3109/17549507.2010.520090

Dignam, J. K., Copland, D., McKinnon, E., Burfein, P., OBrien, K., Farrell, A., & Rodriguez, A. (2015). Intensive Versus Distributed Aphasia Therapy A Nonrandomized, Parallel-Group, Dosage-Controlled Study. *Stroke*, *46*(8), 2206–2211. https://doi.org/10.1161/STROKEAHA.115.009522/-/DC1

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Dignam, J. K., Copland, D., Rawlings, A., OBrien, K., Burfein, P., & Rodriguez, A. D. (2016a). The relationship between novel word learning and anomia treatment success in adults with chronic aphasia. *Neuropsychologia*, *81*, 186–197. https://doi.org/10.1016/j.neuropsychologia.2015.12.026

Dignam, J. K., Rodriguez, A. D., & Copland, D. (2016b). Evidence for Intensive Aphasia Therapy: Consideration of Theories from Neuroscience and Cognitive Psychology. *American Academy of Physical Medicine and Rehabilitation*, *8*(3), 254–267. https://doi.org/10.1016/j.pmrj.2015.06.010

Evidence Partners. (2019). DistillerSR Forest Plot Generator. Retrieved from https://www.evidencepartners.com/resources/forest-plot-generator/

Gunning, D., Wenke, R., Ward, E. C., Chalk, S., Lawrie, M., Romano, M., et al. (2016). Clinicians' perceptions of delivering new models of high intensity aphasia treatment. *Aphasiology*, *31*(4), 406–426. https://doi.org/10.1080/02687038.2016.1236359

Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching. *PloS One*, *10*(9), e0138237. https://doi.org/10.1371/journal.pone.0138237

Harnish, S. M., Neils-Strunjas, J., Lamy, M., & Eliassen, J. C. (2008). Use of fMRI in the study of chronic aphasia recovery after therapy: a case study. *Topics in Stroke Rehabilitation*, *15*(5), 468–483. https://doi.org/10.1310/tsr1505-468

Hinckley, J., & Carr, T. (2005). Comparing the outcomes of intensive and non-intensive context-based aphasia treatment. *Aphasiology*, *19*(10-11), 965–974. https://doi.org/10.1080/02687030544000173

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Kiran, S., & Thompson, C. K. (2019). Neuroplasticity of Language Networks in Aphasia: Advances, Updates, and Future Challenges. *Frontiers in Neurology*, *10*, S225. https://doi.org/10.3389/fneur.2019.00295

Kleim, J. A., & Jones, T. A. (2008). Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. (Vol. 51, pp. S225–39). Presented at the Journal of speech, language, and hearing research : JSLHR. https://doi.org/10.1044/1092-4388(2008/018)

Mackenzie, C., Le May, M., Lendrem, W., McGuirk, E., Marshall, J., & Rossiter, D. (1993). A survey of aphasia services in the United Kingdom. *European Journal of Disorders of Communication : the Journal of the College of Speech and Language Therapists, London*, *28*(1), 43–61.

Marcotte, K., Laird, L., Bitan, T., Meltzer, J. A., Graham, S. J., Leonard, C., & Rochon, E. (2018). Therapy-Induced Neuroplasticity in Chronic Aphasia After Phonological Component Analysis: A Matter of Intensity. *Frontiers in Neurology*, *9*, 225–7. https://doi.org/10.3389/fneur.2018.00225

Moulton, C.-A. E., Dubrowski, A., Macrae, H., Graham, B., Grober, E., & Reznick, R. (2006). Teaching surgical skills: what kind of practice makes perfect?: a randomized, controlled trial. *Annals of Surgery*, *244*(3), 400–409. https://doi.org/10.1097/01.sla.0000234808.85789.6a

Mozeiko, J., Coelho, C. A., & Myers, E. B. (2015). The role of intensity in constraint-induced language therapy for people with chronic aphasia. *Aphasiology*, *30*(4), 339–363. https://doi.org/10.1080/02687038.2015.1070949

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Murray, E., Power, E., Togher, L., McCabe, P., Munro, N., & Smith, K. (2013). The reliability of methodological ratings for speechBITE using the PEDro-P scale. *International Journal of Language \& Communication Disorders / Royal College of Speech \& Language Therapists*, *48*(3), 297–306. https://doi.org/10.1111/1460-6984.12007

OCEBM Levels of Evidence Working Group. (2011). *The Oxford Levels of Evidence 2*. Oxford Centre for Evidence-Based Medicine. Retrieved from http://www.cebm.net/index.aspx?o=5653

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Page, S. J., & Wallace, S. E. (2014). Speech language pathologists' opinions of constraint-induced language therapy. *Topics in Stroke Rehabilitation*, *21*(4), 332–338. https://doi.org/10.1310/tsr2104-332

Page, S. J., Levine, P., Sisto, S., Bond, Q., & Johnston, M. V. (2002). Stroke patients“ and therapists” opinions of constraint-induced movement therapy. *Clinical Rehabilitation*, *16*(1), 55–60. https://doi.org/10.1191/0269215502cr473oa

Palmer, R., Witts, H., & Chater, T. (2018). What speech and language therapy do community dwelling stroke survivors with aphasia receive in the UK? *PloS One*, *13*(7), e0200096. https://doi.org/10.1371/journal.pone.0200096

Pulvermüller, F., & Berthier, M. L. (2008). Aphasia therapy on a neuroscience basis. *Aphasiology*, *22*(6), 563–599. https://doi.org/10.1080/02687030701612213

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-induced therapy of chronic aphasia after stroke. *Stroke*, *32*(7), 1621–1626. https://doi.org/10.1161/01.STR.32.7.1621

Ramsberger, G., & Marie, B. (2007). Self-administered cued naming therapy: a single-participant investigation of a computer-based therapy program replicated in four cases. *American Journal of Speech-Language Pathology*, *16*(4), 343–358. https://doi.org/10.1044/1058-0360(2007/038)

Raymer, A. M., Kohen, F. P., & Saffell, D. (2006). Computerised training for impairments of word comprehension and retrieval in aphasia. *Aphasiology*, *20*(2-4), 257–268. https://doi.org/10.1080/02687030500473312

Robey, R. R. (1998). A Meta-Analysis of Clinical Outcomes in the Treatment of Aphasia. *Journal of Speech, Language, and Hearing Research*, *41*(1), 172–16. https://doi.org/10.1044/jslhr.4101.172

Rochon, E., Laird, L., Simic, T., Höbler, F., Marcotte, K., Bitan, T., et al. (2016). The Effect of Treatment Intensity on the Efficacy of Therapy for Anomia in Aphasia. Presented at the Canadian Stroke Congress, Quebec City, Canada. https://doi.org/10.1177/1747493016659793

Rose, M. L., Ferguson, A., Power, E., Togher, L., & Worrall, L. E. (2013). Aphasia rehabilitation in Australia: Current practices, challenges and future directions. *International Journal of Speech-Language Pathology*, *16*(2), 1–12. https://doi.org/10.3109/17549507.2013.794474

Sage, K., Snell, C., & Lambon Ralph, M. A. (2011). How intensive does anomia therapy for people with aphasia need to be? *Neuropsychological Rehabilitation*, *21*(1), 26–41. https://doi.org/10.1080/09602011.2010.528966

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Schmitt, M. B., & Justice, L. M. (2012). Optimal intervention intensity for emergent literacy: What we know and need to learn. *International Journal of Speech-Language Pathology*, *14*(5), 451–455.

Stahl, B., Mohr, B., Büscher, V., Dreyer, F. R., Lucchese, G., & Pulvermüller, F. (2018). Efficacy of intensive aphasia therapy in patients with chronic stroke: a randomised controlled trial. *Journal of Neurology, Neurosurgery & Psychiatry*, *89*(6), 586–592. https://doi.org/10.1136/jnnp-2017-315962

Stroke Engine. (n.d.). PEDro score - Stroke Engine. Retrieved August 7, 2019, from https://www.strokengine.ca/en/glossary/pedro-score/

Tate, R. L., Perdices, M., Rosenkoetter, U., Wakim, D., Godbee, K., Togher, L., & McDonald, S. (2013). Revision of a method quality rating scale for single-case experimental designs and n-of-1 trials: the 15-item Risk of Bias in N-of-1 Trials (RoBiNT) Scale. *Neuropsychological Rehabilitation*, *23*(5), 619–638. https://doi.org/10.1080/09602011.2013.824383

Tate, R. L., Rosenkoetter, U., Wakim, D., Sigmundsdottir, L., Doubleday, J., Togher, L., et al. (2015). The Risk of Bias in N-of-1 Trials (RoBiNT) Scale: An expanded manual for the critical appraisal of single-case reports. Sydney, Australia: Author.

The Cochrane Collaboration. (2014). Review Manager (RevMan) [Computer Software]. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration.

Togher, L. (2012). Challenges inherent in optimizing speech-language pathology outcomes: It's not just about counting the hours. *International Journal of Speech-Language Pathology*, *14*(5), 438–442. https://doi.org/10.3109/17549507.2012.689334

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Verna, A., Davidson, B., & Rose, T. (2009). Speech-language pathology services for people with aphasia: A survey of current practice in Australia. *International Journal of Speech-Language Pathology*, *11*(3), 191–205. https://doi.org/10.1080/17549500902726059

Wallace, S. J., Worrall, L. E., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., et al. (2016). Which outcomes are most important to people with aphasia and their families? an international nominal group technique study framed within the ICF. *Disability and Rehabilitation*, *91*, 1–16. https://doi.org/10.1080/09638288.2016.1194899

Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: a missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*(1), 70–77. https://doi.org/10.1002/mrdd.20139

This is an original preprint of an article published at
https://doi.org/10.1080/09602011.2020.1768127

Appendix

Study	Design	n *(Intensity allocations)*	MPO	Aphasia type	Severity	Age	Sex	Treatment	**Outcome measures** (bold = primary outcome) I	A/P	QoL	Total dose per arm	Schedule - higher	Schedule – lower	High:low ratio‡
Marcotte et al., 2018	Case series	2 *(1 higher, 1 lower)*	12, 36	2 Broca's	WAB AQ 66 WAB AQ 62	59, 58	1M 1F	Phonological Component Analysis	**BOLD signal changes on fMRI during picture naming** Picture naming – treated items Picture naming – untreated items	-	-	30 hrs	12 hrs/wk x 2.5 wks	3hrs/wk x 10 wks	4:1
Harnish, Neils-Strunjas, Kamy & Eliassen, 2008	Crossover pre-post case series	1 *(Higher then lower)*	≈96	1 Conduction	WAB AQ 71.6	52	1F	Traditional – naming, picture description, cueing hierarchy, writing to dictation	**BOLD signal changes on fMRI during letter decision task** BNT	-	-	15 hrs	7.5 hrs/wk x 2 wks	2 hrs/wk x 7.5 wks	3.75:1
Sage, Snell & Lambon Ralph, 2011	Crossover design *(phase 1 only ≈ RCT)*	8 *(3 higher, 5 lower)*	Range not available. Mean 58.3 (SD 41.1)	*5 Fluent* *3 Non-fluent*	6 Mild 1 Moderate 1 Severe*	Mean 61.2 (SD 8)	6M 2F	Picture naming with cueing hierarchy	**Picture naming – treated items** Untreated items	-	-	10 sessions (?hrs)	5/wk x 2 wks	2/wk x 5 wks	2.5:1
Stahl et al., 2018	RCT	30 *(15 higher, 15 lower)*	12-243	23 Broca's 5 Global 1 Anomic 1 Wernicke's	AAT mean 50.5 (SD 2.9)	33-84	17M 13F	Constraint Induced Language Therapy (no cueing provided)	**AAT (four subscales)** Action Communication Test	-	-	24 hrs†	12hrs/wk x 2 wks†	6hrs/wk x 4 wks†	2:1
Dignam et al., 2015	NonRCT	34 *(16 higher, 18 lower)*	4-225	Not reported	CAT Severity scores 51.6 (higher), 52.3 (lower)	35-77	28M 6F	Language Impairment and Functioning Therapy – impairment, computer, functional & group	**BNT**	CETI CCRSA	ALA	48 hrs	16hrs/wk x 3 wks	6hrs/wk x 8 wks	2.67:1
Mozeiko, Coelho & Myers, 2015	SCED – AB	8 *(4 higher, 4 lower)*	13-134	3 Broca's 1 Global 1 Anomic 1 Conduction 2 Unclassifiable	WAB AQ Mean 48.5, range 24-84	26-77	5M 3F	Constraint Induced Language Therapy	**CIU count** CIUs/word count CIUs/min *WAB AQ (Pre-post)*	*CADL (pre-post)*	-	30 hrs	15hrs/wk x 2 wks	3hrs/wk x 10 wks	5:1
Ramsberger & Marie, 2007	SCED – Multiple baseline across conditions	4 *(Higher & lower phases for each participant)*	6-72	1 Broca's 1 Wernicke's 1 Conduction 1 Anomic	WAB AQ 69 WAB AQ 53 ADP SS 25th %ile ADP SS 68th %ile	63-74	3F 1M	Cued Naming module of MossTalk Words software – cues customised per participant and across sessions	**Picture naming – treated items**	-	-	**P1/P2:** 15 sessions **P3/P4:** 20 sessions	**P1/P2:** 5/wk x 3 wks **P3/P4:** 5wk x 4 wks	**P1/P2:** 2/wk x 7.5 wks **P3/P4:** 2/wk x 10 wks	2.5:1
Raymer, Kohen & Saffell, 2006	SCED – multiple baseline across conditions	5 *(Higher & lower phases for each participant)*	4-42	2 Broca's 2 Conduction 1 Mixed transcortical	WAB AQ Mean 53.3, range 33-76	51-82	3F 2M	Three matching modules of MossTalk Words software	**Picture naming – treated items** Semantic word/picture decisions – treated items (C01, C02) *WAB (Pre-post)* *BNT (Pre-post)*	-	-	12 hrs	3-4hrs/wk x 3-4wks	1-2hrs/wk x 6-12wks	2:1 - 4:1

*Our interpretation of reported language scores for multiple assessments
† Reported schedule for our pre-endpoint sub-analysis; complete study investigated more therapy hours
‡ Ratio of weekly treatment in higher intensity to low intensity arms
Note. RCT – Randomised Controlled Trial; SCED – Single Case Experimental Design, MPO – Months Post Onset from stroke; WAB AQ – Western Aphasia Battery Aphasia Quotient; BNT – Boston Naming Test; ADP SS – Aphasia Diagnostic Profiles standard score; I – Impairment level; A/P – Activity/Participation level; QoL – Quality of life level; CCRSA - Communication Confidence Rating Scale for Aphasia ; AAT – Aachen Aphasia Test; ALA – Assessment of Living with Aphasia; CIU – Correct Information Units