Setting a Research Agenda to Inform Intensive Comprehensive Aphasia Programs

William D. Hula, PhD,[1,2] Leora R. Cherney, PhD,[3,4] and Linda E. Worrall, PhD[5]

[1]*Geriatric Research Education and Clinical Center, VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania;*
[2]*Department of Communication Science and Disorders, University of Pittsburgh, Pennsylvania;*
[3]*Center for Aphasia Research and Treatment, Rehabilitation Institute of Chicago, Illinois;*
[4]*Physical Medicine and Rehabilitation and Communication Sciences and Disorders, Northwestern University, Chicago, Illinois;*
[5]*School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Australia*

Research into intensive comprehensive aphasia programs (ICAPs) has yet to show that this service delivery model is efficacious, effective, has cost utility, or can be broadly implemented. This article describes a phased research approach to the study of ICAPs and sets out a research agenda that considers not only the specific issues surrounding ICAPs, but also the phase of the research. Current ICAP research is in the early phases, with dosing and outcome measurement as prime considerations as well as refinement of the best treatment protocol. Later phases of ICAP research are outlined, and the need for larger scale collaborative funded research is recognized. The need for more rapid translation into practice is also acknowledged, and the use of hybrid models of phased research is encouraged within the ICAP research agenda.
Key words: *aphasia, cost-benefit analysis, rehabilitation, treatment efficacy, treatment effectiveness, treatment outcome*

An intensive comprehensive aphasia program (ICAP) has been defined as a service delivery model that provides a minimum of 3 hours daily treatment over a period of at least 2 weeks; uses a variety of different formats including individual and group therapy, patient and family education, and technological advancements; targets directly the impairment and the activity and participation levels of language and communication functioning; and has a definable start and end date, with a cohort of participants entering and leaving the program at the same time.[1] A survey of ICAPs throughout the world found that the majority of the 12 programs have been formed relatively recently, with 1 program operating for 20 years.[1] Despite the growing number of ICAPs, there is little evidence about their efficacy, effectiveness, or cost-effectiveness. All stakeholders need this evidence. Funding agencies will require evidence to make decisions about their investments in aphasia rehabilitation. People with aphasia and their families should have evidence prior to investing their money and time into a program, and speech and language pathologists have an ethical obligation to provide evidence-based practices. This article summarizes current evidence for ICAPs and proposes a research agenda for taking this distinctive form of service delivery forward.

ICAPs are a complex intervention for a complex problem. Like any intervention, the components as well as the whole should be informed by evidence. According to definitions and descriptions of evidence-based practice, the sources of evidence include the scientific literature and the perspectives of the patients and clinicians within the context of current health and rehabilitation policies and practices.[2] This issue of *Topics in Stroke Rehabilitation* provides some clinician and consumer perspectives on ICAPs.[1,3,4] Although these perspectives are exceedingly valuable sources of information, this article will focus on the evaluation of ICAPs using empirical experimental methods that will ultimately help to determine whether ICAPs promote positive outcomes and are cost-effective in routine clinical practice.

Generally speaking, the highest level and most sought after empirical scientific evidence that is incorporated into influential knowledge syntheses such as systematic reviews and clinical guidelines is the randomized controlled trial

Top Stroke Rehabil 2013;20(5):409–420

www.strokejournal.com

doi: 10.1310/tsr2005-409

(RCT). The most recent summary of evidence of aphasia therapy[5] continues to highlight the mixed results and poor design of aphasia therapy RCTs. Robey and Schultz[6] argue that aphasia research cannot continue to rely on idiosyncratic research methods and that a universal phased approach should be used. The phased model calls for a series of experiments that use a variety of research designs but incorporate the RCT in the latter stages. Hence, incremental research questions are answered systematically prior to the full-scale RCT. This is often not the case in aphasia research, which likely contributes to the conflicting RCT results in the latest Cochrane review.[5] The phased approach has been applied to complex interventions in health care,[7] and it has more recently been adapted for general neurorehabilitation research.[8] In addition to arguing that randomized trials to investigate efficacy and (especially) effectiveness are appropriately conducted in later phases of research, Whyte and colleagues[8] forcefully advocate the funding and publishing of early-phase work as essential to development of the evidence base.

Although the authors writing on the topic of a phased approach to intervention evaluation have assigned different names and numbers to the phases, the various frameworks correspond closely to one another. We will take Whyte and colleagues' presentation as the outline for our discussion of a research program for ICAPs. The phases are Discovery; Phase I: proof-of-concept, safety, and feasibility; Phase II: efficacy; Phase III: effectiveness; and Phase IV: health services research – effectiveness beyond the clinic. In general, the number of participants in each phase increases over time so that the early phases include smaller samples whereas the latter phases include the largest numbers possible and seek to become more definitive. It is also worth noting that the phased approach can be iterative,[7] so if the results of a Phase II study are poor, a return to Phase I or even Discovery work may be indicated. Because progress is not necessarily constrained to be linear, work in multiple phases may proceed simultaneously. In the remainder of the article, we will discuss these phases as they apply to ICAPs.

Discovery Phase

In describing the Discovery phase, Whyte and colleagues[8] emphasized the role of basic science and theory in laying the foundation for a program of treatment research. In the context of clinical aphasiology, basic science may consist of research aimed at understanding the cognitive or neural substrates of normal and impaired communication. It may also consist of efforts to elucidate psychological factors or social systems that promote or hinder successful adaptation to aphasia. With respect to ICAPs in particular, the basic neuroscience literature on intensity and quantity of neurorehabilitation (reviewed in ref. 9) as well as the growing literature on intensity in aphasia treatment[5,10-12] are directly relevant.

Theory and basic research are important contributors to Discovery, but we believe that clinical experience can also be informative at this stage. Many prominent and well-studied approaches to the treatment of aphasia and related neurogenic communication disorders have their genesis in the experience of skilled and well-read clinicians as much or more so than in theoretical or basic empirical work (see, eg, refs. 13-15). Even in cases where there is a primary motivating theory, the theory does not necessarily exhaustively constrain or inform all the details of the approach (see, eg, ref. 16). Also, even though a useful theoretical explanation of all important aspects of a given treatment is certainly desirable, such understanding may be as much the result as it is the driver of treatment success, and an exhaustive theoretical account should not be mandatory for advancing treatment research to Phase I or even from Phase I to Phase II work.

Phase I: Proof-of-Concept, Safety, and Feasibility

The primary goal of Phase I is to apply the new treatment in its initial form and determine whether the expected therapeutic effect can be detected. According to Whyte and colleagues,[8] work in this phase has other objectives in addition to this primary goal, including demonstration of the safety and feasibility of the approach, selection of outcome measures, estimation of

effect size, and investigation of dosing. Robey[17] and Robey and Schultz[6] provide a somewhat more elaborate description of this work preliminary to efficacy trials, which they divide into 2 phases. The first of these phases is focused on selection of a treatment, detection of its effect, and initial estimation of its magnitude; and the second phase is concerned with refining effect size estimates and initial specification of the therapy protocol, target population, dosing, and outcomes. Given the interactive and highly iterative nature of these activities, we find it useful to consider them all as components of a single phase. Below, we discuss some of these components as they apply to ICAPs.

Proof-of-concept

Proof-of-concept requires a demonstration that a desired treatment effect can be obtained, at least potentially or under some conditions. Design considerations for research that is conducted to show proof-of-concept permit a fairly high degree of flexibility[6,8]; weaker designs such as case studies, small-group pre and post studies, and retrospective investigations may contribute along with stronger single-subject and group designs. Although recent systematic reviews have been guarded[5] or equivocal[10] on the benefits of intensive versus nonintensive treatment for aphasia, there is a small and growing number of reports documenting potential benefits of ICAPs.[18-21] Many questions remain, but we believe that the potential for ICAPs to show benefits has been demonstrated.

Safety

Animal models suggest that intensive motor rehabilitation applied very soon after stroke may interfere with recovery.[22,23] However, to our knowledge, such effects have never been demonstrated in humans nor in aphasia treatment specifically, and the limited evidence to date on the timing of aphasia treatment suggests no differences[24] or improved outcomes[25] for treatment begun in the first 6 months of recovery. Furthermore, as with neurorehabilitation in general,[8] the history of clinical practice in aphasiology suggests that research on ICAPs should not be concerned with safety as a primary issue. As a secondary issue, increased fatigue can create mobility safety issues, but in our experiences with 3 separate ICAPs, these have not posed such a risk to ICAPs that they require additional research.

Feasibility and acceptability

There is an evidence gap around the feasibility and acceptability of ICAPs. Brady and colleagues[5] suggest that the drop-out rate in intensive trials is much higher than in nonintensive trials. Certainly most existing ICAPs surveyed[1] have selection criteria about the endurance of potential participants. Most ICAPs are also being offered to participants who are at least 6 months post onset,[1] so the acceptability and feasibility of ICAPs in earlier stages has yet to be determined. In addition to the required endurance of participants with aphasia, therapists must have the capacity to provide an ICAP. Feasibility issues surrounding appropriate space requirements, workforce planning, and simultaneous access to multiple computers can be challenging to many facilities. Hence, the structural and organizational requirements of ICAPs have not been described to determine whether ICAPs are a feasible option for some service providers. Most existing ICAPs are associated with universities rather than health care providers, and organizational issues are likely to act as a barrier to further uptake by health care providers. Thus, demonstration of ICAP feasibility from the perspectives of providers and participants continues to be an important issue.

Outcome measurement

Selection of outcome measures in aphasia rehabilitation research continues to be a vexing question. Outcomes measured at all levels of the International Classification of Functioning, Disability and Health (ICF) and relevant to the patient journey may provide important information. Whyte and colleagues[8] differentiate between proximal outcomes that are closely related to the proposed treatment and measure immediate effects and distal outcomes that are more real-world measures and may assess new domains of function. They argue that proximal outcomes, which for many aphasia treatments

focus on language impairments as opposed to activities or participation, are appropriately used in Phase I studies where the purpose is to show proof-of-concept or to establish the underlying mechanism of the effect. For example, if naming treatments are applied as a part of an ICAP, tests of naming or measures of semantic and phonological processing may help to establish whether any observed changes have occurred for the reasons predicted by underlying theory. This information can help to constrain hypotheses about specific operative components of the treatment and in turn lead to refinement of the approach. When socially focused or participation-oriented aphasia treatments are applied, the proximal outcomes may be more appropriately measured at the activity or participation levels of the ICF.

Another reason to choose proximal outcome measures is that they can lead to better understanding of the underlying treatment mechanism. When the underlying mechanism is understood, it supports inferences about whether the treatment has in fact caused the observed changes.[26] Conversely, when the results of an investigation of more distal outcomes are negative, knowledge of the treatment mechanism can increase the ability to interpret such negative findings. For example, if a treatment with established efficacy or effectiveness for improving language function fails to have an impact on a distal outcome such as quality of life or return to work, the negative finding may be due to variables other than simple inadequacy of the treatment.[8] Even in the early phases of ICAP research, both proximal and distal outcome measures may be required to capture the outcomes of the comprehensive program that targets all levels of the ICF.

It is also worth noting that in the context of a complex intervention like an ICAP, the distinction between proximal and distal outcome measures may be less clear. ICAPs, as they are defined here, target impairment, activity, and participation levels of functioning. Thus, whereas individual program components may have clearly delineated proximal and distal outcomes, it is easy to imagine that a distal outcome for one component (eg, improved self-reported participation in conversation as a function of a naming treatment) may be a proximal outcome of a more participation-oriented treatment. The intervention taken as a whole may have sets of proximal and distal outcomes that partially overlap. These considerations are further complicated by the possibility that there may be interactions between ICF levels such that activities and participation may impact impairments in addition to the reverse.

These considerations aside, it is likely that activity or participation outcomes will be of ultimate and primary interest to most stakeholders. Ideally, these outcomes should be apparent to and measurable from the perspectives of the person with aphasia and his or her family or social community. The last decade has seen an explosion of interest in the measurement of patient-reported health outcomes across the full range of health conditions,[27-32] and aphasia is no exception.[33-38] Although these efforts have produced some promising results and represent an appropriate focus on patients' perspectives, use of patient-reported outcomes as primary endpoints in RCTs does raise challenges.

One issue concerns the difficulty of blinding patients or surrogate outcome reporters to the experimental condition or time point. Blinding is included as an item in the Consolidated Standard of Reporting Trials (CONSORT) checklist,[39,40] and it is recognized as an important method of reducing bias, especially for measures that are less objective.[41] A potential solution might be to develop instruments and protocols that would use surrogate outcome reporters who could be blinded, recognizing that surrogate reports, although valuable, should not be directly substituted for patient reports.[34]

The second issue has been raised in the context of patient-reported outcomes but may apply more broadly to performance-based and clinician-reported outcomes. In cases where interventions are not expected to raise general functional ability (on whatever variable is being measured), but rather are intended to affect responses to specific test items, the size of the treatment effect will depend on the number of items included in the assessment that are targeted by the intervention.[42] For example, the Communication Outcome after Stroke Scale[33] includes items about how well the respondent can read and write in addition to questions about expression and comprehension

in conversation and the impact of speech and language problems on family and social life. If an intervention targeted at improving participation in family conversation is provided, and this intervention does not have goals related to reading and writing, the inclusion of reading and writing items in the outcome assessment may dilute the treatment effect. Conversely, if a treatment is expected to improve the underlying language impairment in such a way that it manifests as a general improvement in communication ability that impacts responses to all items equally (perhaps a heroic assumption), the treatment effect will be independent of the particular items administered and it may be beneficial to include as many items as possible to maximize reliability. Even with impairment-focused aphasia treatments, item-specific and domain-specific treatment effects are common, and researchers have traditionally dealt with this by assessing treated and untreated items and tasks separately. Users of patient-reported outcomes should consider the possibility of item-specific treatment effects and either demonstrate that they are absent or take steps to minimize or account for the potential bias on effect size estimates.

Specification of the treatment protocol

One of the central sources of tension in the evaluation of ICAPs will likely be specification of the protocol. On one hand, failure to define and understand the components of a complex health intervention in the early phases of work can lead to problems in later phases.[7] On the other hand, relying on traditional RCTs and delaying them until the research necessary to fully characterize the effects of the component treatments delivered in an ICAP (and their potential interactions) has been completed may deny participants, clinicians, policymakers, and other stakeholders needed evidence for decades to come.[43] We suggest that this conflict may be resolved at least in part by taking the view that the phases outlined here are iterative[7] and may at times proceed concurrently with one another. Our position is also informed by the distinction between explanatory and pragmatic trials,[44,45] discussed below under Phase III, and a recognition that both are necessary.

Clinical trial and error can provide initial conclusions about what components work best in an ICAP, and ICAPs have been evolving clinically for a number of years. Service providers have trialed different components and sought the views of stakeholders, particularly the service users, about the acceptability and perceived effectiveness of the individual components. To be defined as an ICAP, Rose, Cherney, and Worrall[1] stated that the program should involve the components of individual and group therapy, patient and family education, and technological advancements.[1] Furthermore, the targets of intervention should include impairments as well as the activity limitations and participation restrictions associated with aphasia. The survey identified considerable variability between ICAPs on these core components and also on additional components. For example, even though all ICAPs encouraged family involvement, some through observation and others through direct participation, not all programs included sessions specifically for family members. Also, the range of staff who are involved with the surveyed ICAPs (social work, music therapists, recreation officers, other allied health, psychology, volunteers) suggested that the programs may include other components as well.

The mix of components that make an ICAP "comprehensive" is potentially a response to documented unmet need in patient and family education[46,47] and group therapy.[48] Stroke rehabilitation guidelines in many countries have driven the inclusion of these components into practice. For example, the Australian Stroke Rehabilitation Guidelines[49] recommend that "all stroke survivors and their families/carers should be offered information tailored to meet their needs using relevant language and communication formats" (Level A evidence). There are similar statements regarding the other components: "Group therapy and conversation groups can be used for people with aphasia and should be available in the longer term for those with chronic and persisting aphasia" (Level C); "Interventions should be individually tailored but can include delivery of therapy programs via computer" (Level C). Although more recently published systematic reviews[48,50] have upgraded the level of evidence in some areas, there is

still little evidence for the superior efficacy or effectiveness of particular treatment approaches over others.[5,51,52] Clearly, the relative potency of different potential ICAP components in relation to patient characteristics and specific outcomes should be the object of continued empirical study. Is family and patient education the vital component for long-term maintenance of gains? Are gains in spoken output only achieved through impairment-focused interventions or is practice in real-life situations necessary? Are the greatest participation gains made through group therapy? What cognitive, linguistic, lesion, or psychosocial characteristics moderate treatment effects? The gaps in this phase include better evidence for the components themselves as well as comparative studies of the components typically included in an ICAP.

Dosage

There has been considerable attention about the intensity requirement for aphasia therapy.[5,10,11,53] This is one of the catalysts for the development of ICAPs and why "intensive" was included in the ICAP acronym. A research agenda for ICAPs includes research particularly addressing the intensity of treatment.

Evidence suggests that intensive therapy results in positive outcomes, although variations in patient characteristics such as aphasia type and severity, treatment type, and outcome measures complicate comparisons across studies. Some studies have examined a single therapy at a high intensity,[54,55] and others have compared 2 treatments given at the same high intensity.[56,57] Some treatment studies have directly compared conditions of higher and lower intensity treatment for aphasia.[11] Yet, optimal intensity has not been established for even one type of aphasia treatment.

Determining the optimal intensity of a behavioral intervention is complex and may need to be investigated iteratively.[8] For the ICAP, initial studies are needed that systematically contrast parameters such as session duration (eg, 6 hours a day), session frequency (eg, 5 days a week), and total duration (eg, 4 weeks), while keeping constant elements of the dose form (eg, type of treatment and treatment tasks).

To establish the optimal intensity of an intervention, clinicians must consider the number of times a teaching episode occurs per session. A teaching episode contains the unique combination of active ingredients that is essential to the therapeutic process.[58] For example, in a word retrieval task, a single teaching episode may include the sequence of steps given to cue a naming response; different outcomes may occur if the number of target words varies or if there are differences in the number of times each target is practiced (eg, 20 target words practiced 3 times vs 15 target words practiced 4 times). Later studies should investigate how these variations in dose (ie, teaching episodes) affect ICAP outcomes.

There are many factors that may influence the responsiveness of the person with aphasia to a specific dose of treatment; despite the selection of a homogeneous group of subjects, outcomes may differ. Therefore, for each participant in the ICAP, it may be useful to establish a dose-response relationship that describes the change in a specified outcome caused by differing levels of exposure (or doses). This would require frequent measurement of performance on specific "probes" or behaviors of interest. The dose-response curve is the graphic representation of the relationship between treatment amount or intensity (x axis) and the response or change in performance (y axis) on an outcome. It can illustrate the graded responses of the individual to varying amounts of treatment. Examination of the graph can show the threshold at which a minimal detectable response occurs and how much treatment is required to obtain a maximal response. The dose-response curves of participants could be analyzed statistically to provide information regarding responsiveness of subgroups of participants with aphasia to accumulating amounts and intensities of treatment.

Phase II: Efficacy

The goal of Phase II is to establish treatment efficacy. Robey and Schulz[6] provide a helpful definition of the term *efficacy* and point out that it has 3 main features. First, inferences regarding efficacy are made at the level of populations, not individuals. Second, efficacy research requires

that both the treatment and the population be focused and well defined. Third, the treatment must be applied under optimal conditions, that is, by well-trained clinicians, to carefully selected patients, with optimal facilities, and with outcomes evaluated by the most appropriate measures. Efficacy studies often use a parallel group design in which participants are randomly assigned to either the experimental intervention group or control group. Although the CONSORT statements are intended to provide guidance for the reporting rather than the design and conduct of clinical trials, the items included in the CONSORT checklists are quality markers and may thus help investigators to introduce as much rigor as possible into the work.[39,40] A well-designed and conducted RCT should determine whether the intervention works or not.

There are no RCTs of ICAPs as yet. An RCT would require a selection of "ideal" participants who are randomly allocated to an ICAP or a control group. Control groups are often difficult to design in ICAPs, because the participants are usually aware of the therapy they are receiving and it is difficult to find a control condition of similar dosage and intensity that does not have an impact on aphasia outcomes.[59] Given that ICAPs are of relatively short duration, wait-list control groups, in which participants eventually receive the intervention, may be an ethical option.[59] A no-treatment condition also does not control for the effect of the contact time between client and therapist or the Hawthorne effect that occurs when a client is expecting change. A placebo analogue or attention control group that controls for a similar dosage of contact would be desirable.[59] Such attention control groups have typically undertaken activities that should not have an impact on the outcome of interest, for example, a hand task when lower limb mobility is the focus of intervention. Communication is a feature of most interpersonal contact and is often thought to have an impact on outcomes generally, so it is challenging to design attention control groups that do not involve communication. An additional challenge is making the control task as interesting as the experimental task so that participants do not drop out of the study. An efficacy RCT is only possible in a research environment, not a user self-pay environment, where funds are available to pay for each participant's treatment as well as for the administration of the research protocol. Many ICAPs are user self-pay, hence there is a need to obtain large-scale research funding to take ICAP evaluation to the stage of an efficacy trial.

Another issue to be considered in designing an RCT for an ICAP is the potential need for cluster randomization. Given that ICAPs typically involve cohorts of patients who begin and end treatment at the same time, it may be appropriate to randomize participants at the level of these cohorts rather than at the level of the individual participant. Cluster randomization has often been used to guard against experimental contamination caused by contact between participants in the active treatment and control arms of an RCT.[60-62] Such contact would be likely to occur in the context of an ICAP if participants in both the active and control groups were being seen at the same facility concurrently. On a related but distinct issue, the typical structure of ICAPs also suggests that multilevel statistical models in which participants are nested within cohorts should be considered for data analysis.[63]

Phase III: Effectiveness

In contrast to efficacy, effectiveness refers to inferences made when treatments are studied under conditions consistent with typical clinical practice.[6,8,64] Effectiveness studies are concerned with the generalizability of the treatment and are usually a step closer to the real world of health care. The treatment is administered to larger numbers of heterogeneous patients in different treatment centers by different clinicians who vary in their level of expertise and must deal with the time constraints and competing demands of the clinic.

As in Phase II, RCTs are the preferred design. However, different clinical environments may influence the structure of the ICAP (eg, number of hours per day, number of weeks) or the availability of some of its components (eg, computer-based treatment, family education). Furthermore, for a Phase III effectiveness trial, specification of treatment procedures and thus fidelity of treatment

pose difficulties because of the heterogeneity of the subjects. This contrasts with a Phase II efficacy trial, where fidelity of treatment is accomplished, in part, by the standardization and manualization of the treatment procedures together with extensive training of the clinicians and periodic monitoring throughout the trial to ensure that all participants are receiving the same treatment. There are many aphasia treatments to choose from, and it may not be feasible to standardize the components of the ICAP and the exact treatments so that the ICAP can be replicated across sites for a heterogeneous group of subjects with different strengths and deficits. Conducting a meta-analysis of the outcomes from different ICAPs may be one solution to assessing effectiveness when there is lack of uniformity of ICAPs across different settings.[8]

According to Whyte and colleagues,[8] improvement in Phase III is measured by outcomes that are closely related to the target of treatment as well as outcomes that assess important but more distal functional areas. As previously discussed, the focus on 2 distinct outcome domains may be less relevant to ICAPs given the comprehensive nature of the ICAP and its inclusion of treatments that target both impairment and activity/participation. Both outcome domains would have been addressed in Phase II.

The literature on phased models for clinical outcome research in aphasia has held that efficacy must be established before effectiveness can be investigated.[6,17] Robey,[17] on the assumption that subjects in effectiveness trials are likely to be clients of clinical service providers, suggested that establishing efficacy before studying effectiveness satisfies the ethical obligation to provide services only when benefit can reasonably be expected. However, it has been observed that the transition between efficacy trials and effectiveness trials lacks a clear boundary,[8,65] and the view that efficacy must be established first in all cases is not universally held.[44,65,66] The idea of an explanatory-to-pragmatic continuum in clinical outcomes research is helpful in considering these questions.[44,45,66] Briefly, explanatory trials emphasize internal validity and provide inferences about efficacy, whereas pragmatic trials emphasize external validity and generate answers about effectiveness.[5,67] In some cases, explanatory trials are appropriately conducted before pragmatic trials, but in others the reverse may be true.[66]

To the ethical issue raised by Robey,[17] we would add the following observations. First, services may be provided with reasonable expectation of benefit based on evidence other than efficacy established by one or more RCTs. Second, delays in research that is relevant to the needs of patients and practitioners can be considered a risk to research participants, patients, and to the public.[43] Thus, we agree that the negative findings with limited interpretability arising from an expensive effectiveness trial are a risk that should be mitigated against. However, we also believe that it may be appropriate in some cases to address questions of effectiveness without first definitively establishing efficacy or to consider ways of integrating questions of efficacy, effectiveness, and implementation into the same trial.[68,69]

Phase IV: Health services research

The final phase of research aims to determine whether change is required in health service delivery or policy. Questions surrounding the cost-effectiveness and cost benefit of the intervention as well as comparison trials to determine whether other current interventions have better or worse cost-effectiveness are studied. There is an increasing demand for implementation studies to determine the most effective ways to implement the intervention into practice.

Rehabilitation is expensive, and an intensive specialized program such as an ICAP incurs substantial costs. Do these translate into benefits for the participants and society in general? An economic analysis will evaluate the relative resource use and costs and, thereby, the cost-effectiveness of an ICAP. Generally, cost data (cost of all key services and resources used during the ICAP) are collected during the RCT, and costs are then displayed against their benefits as measured by units of improvement on the relevant outcome measure. For example, this calculation may

allow researchers to indicate the dollar amount required per unit of improvement on the chosen outcome measure. This is useful to persons who have aphasia, but the cost-effectiveness of ICAPs cannot be compared to health care interventions for other conditions. Hence, the broader measure of quality-adjusted life years (QALYs) is used. This introduces the concept of cost utility analyses. Even though an ICAP may offer value for money, unless people value the benefit that it may bring (and are willing to pay for it), then health service providers may be reluctant to choose an ICAP, for example, over another health or rehabilitation program.

Comparative studies may occur at this stage or earlier. Comparisons between different interventions may be studied, because they may either be less expensive or convey more benefit, thereby altering the cost-effectiveness ratio. A potential comparison treatment may be a noncomprehensive treatment such as intensive impairment-based therapy (eg, constraint-induced aphasia therapy [CIAT]) that has high levels of evidence and is known to be effective. However, a comparable dosage of noncomprehensive therapy, such as 4 hours a day of CIAT, may not be tolerated as well by both clients and therapists. Another comparison would be with a nonintensive but comprehensive treatment. Participants with aphasia who are receiving standard care may also be an appropriate comparison group, because a primary question about ICAPs must be whether they produce better outcomes to the standard care provided in typical clinical practice.

Finally, there is a need to translate strong evidence into practice. A growing concern is the gap between research and practice and the slow rate at which research results are translated into clinical practice.[43,69] For an RCT, it may take more than 7 years from grant submission to publication and as much as 17 years from initiation of a concept to implementation into practice.[43] Indeed, there is a relative paucity of effectiveness trials and health services research in relation to the number of efficacy trials of behavioral interventions,[69] a situation that is also clearly present in the area of aphasia research. To facilitate more rapid integration of ICAP therapy into clinical practice, consideration of the implementation process is critical. Implementation research identifies ways in which the context or environment affects implementation, provides information on how to refine or adapt the intervention, and provides critical information for future replication and dissemination. Hybrid designs that simultaneously assess the effectiveness of an intervention and its potential for implementation have been proposed and may be appropriate for future investigation of the ICAP.[68]

Conclusion

A considerable amount of research on the effects and value of ICAPs remains to be done. Greater attention to earlier phases of research may mean that subsequent RCTs are more likely to show benefit. Hence, there is a need to refine the intervention, the responders, and the outcome measures so that efficacy and effectiveness trials are as well controlled as possible. There is a clear need for more Phase I research on ICAPs. Also, multisite studies and large-scale funding will likely be required for all subsequent phases, making collaboration across ICAP sites and researchers essential.

ICAP researchers may need to consider hybrid models to speed up the process of phased research. There is now a need to collect data about costs and begin to measure the value of ICAPs in terms of willingness to pay per unit of improvement. There is also need to examine the barriers and facilitators to implementing ICAPs. If at the end of this research effort, ICAPs are shown to be cost-effective and have cost utility, then they can be implemented so that all people with aphasia can access an ICAP within their local health service.

Acknowledgments

We acknowledge the Veterans Affairs Rehabilitation Research and Development Career Development Award C6210M and NHMRC grant 569935 to the Centre for Clinical Research Excellence in Aphasia Rehabilitation.

REFERENCES

1. Rose ML, Cherney LR, Worrall LE. Intensive comprehensive aphasia programs: An international survey of practice. *Top Stroke Rehabil.* 2013;20(5):379-387.
2. Dollaghan CA. Evidence-based practice in communication disorders: What do we know, and when do we know it? *J Commun Disord.* 2004;37(5):391-400
3. Persad C, Wozniak L, Kostopoulos E. Retrospective analysis of outcomes from two intensive comprehensive aphasia programs. *Top Stroke Rehabil.* 2013;20(5):388-397.
4. Babbitt EM, Worrall LE, Cherney LR. Clinician perspectives of an intensive comprehensive aphasia program. *Top Stroke Rehabil.* 2013;20(5):398-408.
5. Brady MC, Kelly H, Godwin J, Enderby P. Speech and language therapy for aphasia following stroke. *Cochrane Database Syst Rev.* 2012; 5:1-228.
6. Robey RR, Schultz MC. Review: A model for conducting clinical-outcome research: An adaptation of the standard protocol for use in aphasiology. *Aphasiology.* 1998;12(9):787-810.
7. Campbell M, Fitzpatrick R, Haines A, et al. Framework for design and evaluation of complex interventions to improve health. *Br Med J.* 2000;321(7262):694.
8. Whyte J, Gordon W, Rothi LJ. A phased developmental approach to neurorehabilitation research: The science of knowledge building. *Arch Phys Med Rehabil.* 2009; 90(11):S3-S10.
9. Raymer AM, Beeson P, Holland A, et al. Translational research in aphasia: From neuroscience to neurorehabilitation. *J Speech Lang Hear Res.* 2008;51:S259-S275.
10. Cherney LR, Patterson JP, Raymer AM. Intensity of aphasia therapy: Evidence and efficacy. 2011;11(6):560-569.
11. Cherney LR. Aphasia treatment: Intensity, dose parameters, and script training. *Int J Speech Lang Pathol.* 2012;14(5):424-431.
12. Bhogal SK, Teasell R, Speechley M. Intensity of aphasia therapy, impact on recovery. *Stroke.* 2003;34(4):987-993.
13. Albert ML, Sparks RW, Helm NA. Melodic intonation therapy for aphasia. *Arch Neurol.* 1973;29(2):130-131.
14. Kearns KP. Response elaboration training for patient initiated utterances. *Clin Aphasiol.* 1985;15: 196-204.
15. Rosenbek JC, Lemme ML, Ahern MB, Harris EH, Wertz RT. A treatment for apraxia of speech in adults. *J Speech Hear Disord.* 1973;38(4):462.
16. Boyle M, Coelho CA. Application of semantic feature analysis as a treatment for aphasic dysnomia. *Am J Speech Lang Pathol.* 1995; 4(4): 94-98.
17. Robey RR. A five-phase model for clinical-outcome research. *J Commun Disord.* 2004; 37(5):401-411.
18. Hinckley JJ, Craig HK. Influence of rate of treatment on the naming abilities of adults with chronic aphasia. *Aphasiology.* 1998; 12(11):989-1006.
19. Code C, Torney A, Gildea-Howardine E, Willmes K. Outcome of a one-month therapy intensive for chronic aphasia: Variable individual responses. *Semin Speech Lang.* 2010;31(1):21-33.
20. Winans-Mitrik R, Schumacher JG, Hula WD, Dickey MW, Doyle PJ. Description of an intensive residential aphasia treatment program: Rationale, clinical processes, and outcomes. Presented at: Clinical Aphasiology Conference; May 2013; Tucson, AZ.
21. Rodriguez A, Worrall L, McKinnon E, et al. Therapeutic effect of an intensive, comprehensive aphasia program: Aphasia LIFT. Presented at: Clinical Aphasiology Conference; May 2013; Tucson, AZ.
22. Kleim JA, Jones TA, Schallert T. Motor enrichment and the induction of plasticity before or after brain injury. *Neurochem Res.* 2003;28(11):1757-1769.
23. Woodlee MT, Schallert T. The interplay between behavior and neurodegeneration in rat models of Parkinson's disease and stroke. *Restor Neurol Neurosci.* 2004;22(3):153-161.
24. Wertz RT, Weiss DG, Aten JL, et al. Comparison of clinic, home, and deferred language treatment for aphasia. A Veterans Administration Cooperative Study. *Arch Neurol.* 1986;43(7):653-658.
25. Robey RR. A meta-analysis of clinical outcomes in the treatment of aphasia. *J Speech Lang Hear Res.* 1998;41(1):172-187.
26. Russo F, Williamson J. Interpreting causality in the health sciences. *Int Stud Philos Sci.* 2007; 21(2):157-170.
27. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol.* 2010;63(11): 1179-1194.
28. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH roadmap cooperative group during its first two years. *Med Care.* 2007; 45(5 Suppl 1):S3-S11.
29. Coster WJ, Haley SM, Jette AM. Measuring patient-reported outcomes after discharge from inpatient rehabilitation settings. *J Rehabil Med.* 2006;38(4):237-242.
30. Massof RW, Ahmadian L, Grover LL, et al. The activity inventory: An adaptive visual function questionnaire. *Optom Vis Sci.* 2007; 84(8): 763-774.
31. Doyle PJ, McNeil MR, Bost JE, et al. The Burden of Stroke Scale (BOSS) provided valid, reliable, and responsive score estimates of functioning and well-being during the first year of recovery from stroke. *Qual Life Res.* 2007;16(8):1389-1398.
32. Duncan PW, Bode R, Lai SM, Perera S. Rasch analysis of a new stroke-specific outcome scale: The Stroke Impact Scale. *Arch Phys Med Rehabil.* 2003;84:950-963.

33. Long A, Hesketh A, Paszek G, Booth M, Bowen A. Development of a reliable self-report outcome measure for pragmatic trials of communication therapy following stroke: The Communication Outcome after Stroke (COAST) scale. *Clin Rehabil.* 2008;22(12):1083-1094.
34. Doyle PJ, Hula WD, Hula SNA, et al. Self-and surrogate-reported communication functioning in aphasia. *Qual Life Res.* 2012; 22(5):957-967.
35. Kagan A, Simmons-Mackie N, Rowland A, et al. Counting what counts: A framework for capturing real-life outcomes of aphasia intervention. *Aphasiology.* 2008;22(3):258-280.
36. Chue WL, Rose ML. The reliability of the Communication Disability Profile: A patient-reported outcome measure for aphasia. *Aphasiology.* 2010;64(6-8):940-956.
37. Babbitt EM, Cherney LR. Communication confidence in persons with aphasia. *Top Stroke Rehabil.* 2010;17(3):214-223.
38. Hilari K, Byng S, Lamping DL, Smith SC. Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Evaluation of acceptability, reliability, and validity. *Stroke.* 2003;34(8):1944-1950.
39. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMC Med.* 2010;8(1):18.
40. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: Explanation and elaboration. *Ann Intern Med.* 2008;148(4):295-309.
41. Schulz KF, Grimes DA. Blinding in randomised trials: Hiding who got what. *Lancet (Lond).* 2002;359(9307):696-700.
42. Massof RW. A general theoretical framework for interpreting patient-reported outcomes estimated from ordinally scaled item responses [published online ahead of print February 19, 2013]. *Stat Methods Med Res.*
43. Riley WT, Glasgow RE, Etheredge L, Abernethy AP. Rapid, responsive, relevant (R3) research: A call for a rapid learning health research enterprise. *Clin Transl Med.* 2013;2:10.
44. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chroicn Dis.* 1967;20(8):637-648.
45. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials. *J Am Med Assoc.* 2003; 290(12):1624-1632.
46. Rose T, Worrall L, Hickson L, Hoffmann T. Do people with aphasia want written information? A verbal survey exploring preferences for when and how to provide stroke and aphasia information. *Top Stroke Rehabil.* 2010;17(2):79-98.
47. Rose TA, Worrall LE, McKenna KT, Hickson LM, Hoffmann TC. Do people with aphasia receive written stroke and aphasia information? *Aphasiology.* 2009;23(3):364-392.
48. Lanyon LE, Rose ML, Worrall L. The efficacy of outpatient and community-based aphasia group interventions: A systematic review [published online ahead of print January 22, 2013). *Int J Speech Lang Pathol.*
49. National Stroke Foundation (Australia). *Clinical guidelines for stroke rehabilitation and recovery*. National Stroke Foundation. http://strokefoundation.com.au/health-professionals/. Accessed July 29, 2013.
50. Salter K, Teasell R, Bhogal S, Zettler L, Foley N. *Aphasia. Evidence-Based Systematic Review of Stroke Rehabilitation.* http://www.ebrsr.com. Accessed July 29, 2013.
51. Allen L, Mehta S, McClure JA, Teasell R. Therapeutic interventions for aphasia initiated more than six months post stroke: A review of the evidence. *Top Stroke Rehabil.* 2012;19(6):523-535.
52. Wisenburn B, Mahoney K. A meta-analysis of word-finding treatments for aphasia. *Aphasiology.* 2009;23(11):1338-1352.
53. Cherney LR, Patterson JP, Raymer A, Frymark T, Schooling T. Evidence-based systematic review: Effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia. *J Speech Lang Hear Res.* 2008;51(5):1282-1299.
54. Schlaug G, Marchina S, Norton A. Evidence for plasticity in white-matter tracts of patients with chronic Broca's aphasia undergoing intense intonation-based speech therapy. *Ann N Y Acad Sci.* 2009;1169:385-394.
55. Schlaug G, Marchina S, Norton A. From singing to speaking: Why singing may lead to recovery of expressive language function in patients with Broca's aphasia. *Music Percept.* 2008;25(4):315.
56. Barthel G, Meinzer M, Djundja D, Rockstroh B. Intensive language therapy in chronic aphasia: Which aspects contribute most? *Aphasiology.* 2008;22(4):408-421.
57. Maher LM, Kendall D, Swearengin JA, Rodriguez A, Gonzalez-Rothi LJ. A pilot study of use-dependent learning in the context of constraint induced language therapy. *J Int Neuropsychol Soc.* 2006;12:843-852.
58. Baker E. Optimal intervention intensity. *Int J Speech Lang Pathol.* 2012;14(5):401-409.
59. Hart T, Fann JR, Novack TA. The dilemma of the control condition in experience-based cognitive and behavioural treatment research. *Neuropsychol Rehabil.* 2008;18(1):1-21.
60. Donner A, Klar N. Pitfalls of and controversies in cluster randomization trials. *Am J Public Health.* 2004;94(3):416-422.
61. Torgerson DJ. Contamination in trials: Is cluster randomisation the answer? *Br Med J.* 2001;322(7282):355.
62. Campbell MK, Elbourne DR, Altman DG. CONSORT statement: Extension to cluster randomised trials. *Br Med J.* 2004;328(7441):702-708.
63. Gelman A. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press; 2007.
64. Johnston MV, Granger CV. Outcomes research in medical rehabilitation: A primer and introduction to a series. *Am J Phys Med Rehabil.* 1994;73(4): 296-303.
65. Hoagwood K, Hibbs E, Brent D, Jensen P. Introduction to the special section: Efficacy

and effectiveness in studies of child and adolescent psychotherapy. *J Consult Clin Psych.* 1995;63(5):683.
66. Treweek S, Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability. *Trials.* 2009;10(37):9.
67. Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic–explanatory continuum indicator summary (PRECIS): A tool to help trial designers. *J Clin Epidemiol.* 2009;62(5):464-475.
68. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: Combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care.* 2012;50(3):217-226.
69. Glasgow RE, Lichtenstein E, Marcus AC. Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *Am J Public Health.* 2003;93(8):1261-1267.