













BMJ Open Supporting wellbeing through PEEr-Befriending (SUPERB) feasibility trial: fidelity of peer-befriending for people with aphasia

Nicholas Behn ¹, Becky Moss ¹, Sally McVicker ¹, Abi Roper ¹, Sarah Northcott ¹, Jane Marshall ¹, Shirley Thomas ², Alan Simpson ³, Chris Flood ⁴, Kirsty James ⁵, Kimberley Goldsmith ⁵, Katerina Hilari ¹

To cite: Behn N, Moss B, McVicker S, *et al*. Supporting wellbeing through PEEr-Befriending (SUPERB) feasibility trial: fidelity of peer-befriending for people with aphasia. *BMJ Open* 2021;**11**:e047994. doi:10.1136/bmjopen-2020-047994

► Prepublication history and supplemental material for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-047994>).

Received 15 December 2020
Accepted 12 July 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Professor Katerina Hilari;
k.hilari@city.ac.uk

ABSTRACT

Objective To evaluate systematically the fidelity of a peer-befriending intervention for people with aphasia.

Design Supporting wellbeing through Peer-befriending (SUPERB) was a feasibility randomised controlled trial comparing usual care to usual care +peer-befriending. This paper reports on the fidelity of all intervention aspects (training and supervision of providers/befrienders; intervention visits) which was evaluated across all areas of the Behaviour Change Consortium framework.

Setting Community.

Participants People with aphasia early poststroke and low levels of distress, randomised to the intervention arm of the trial (n=28); 10 peer-befrienders at least 1-year poststroke.

Intervention Peer-befrienders were trained (4–6 hours); and received regular supervision (monthly group while actively befriending, and one-to-one as and when needed) in order to provide six 1-hour peer-befriending visits over 3 months.

Main measures Metrics included number and length of training, supervision sessions and visits. All training and supervision sessions and one (of six) visits per pair were rated against fidelity checklists and evaluated for inter-rater and intrarater reliability (Gwets AC1 agreement coefficient). Per-cent adherence to protocol was evaluated.

Results All peer-befrienders received 4–6 hours training over 2–3 days as intended. There were 25 group supervision sessions with a median number attended of 14 (IQR=8–18). Twenty-six participants agreed (92.8%) to the intervention and 21 (80.8%) received all six visits (median visit length 60 min). Adherence was high for training (91.7%–100%) and supervision (83%–100%) and moderate-to-high for befriending visits (66.7%–100%). Where calculable, inter-rater and intrarater reliability was high for training and supervision (Gwets AC1 >0.90) and moderate-to-high for intervention visits (Gwets AC1 0.44–1.0).

Conclusion Planning of fidelity processes at the outset of the trial and monitoring throughout was feasible and ensured good-to-high fidelity for this peer-befriending intervention. The results permit confidence in other findings from the SUPERB trial.

Trial registration number NCT02947776.

Strengths and limitations of this study

- Based on the National Institutes of Health Behaviour Change Consortium framework, this study systematically evaluated the fidelity of all aspects (training, supervision, and intervention visits) of the Supporting wellbeing through PEEr-Befriending (SUPERB) intervention for people with aphasia.
- A range of strategies and metrics were planned and used to ensure fidelity across study design, training providers, delivery of treatment, receipt of treatment and treatment enactment.
- A support programme informed by user involvement and including standardised, manualised training and regular supervision was in place to ensure the treatment providers (befrienders with aphasia) were able to deliver the intervention as intended.
- Videoing all intervention visits rather than a sample would have provided a larger pool of data on which to base treatment adherence for visits.
- Data presented in this report need to be supplemented with qualitative interview data in order to capture treatment enactment as comprehensively as other fidelity areas.

INTRODUCTION

In stroke rehabilitation research, evaluation of fidelity is a core recommendation and an integral component of intervention trial design.¹ Fidelity refers to the ‘methodological strategies used to monitor and enhance the reliability and validity of behavioural interventions’ (p.443)² and is needed to ensure that an intervention is delivered as intended.^{3–5} Assessing fidelity can assist researchers to identify the essential elements that help to make a treatment work.⁴ Fidelity can also help explain nonsignificant or ambiguous findings^{6,7} as the lack of an effect may be due to poor implementation of a treatment or problems with the treatment itself.^{4,8} Fidelity is essential in making decisions about an

intervention's efficacy and replication. The treatment Fidelity Workgroup of the National Institutes of Health Behaviour Change Consortium reviewed existing definitions and practices of treatment fidelity and developed recommendations for consistent reporting across five key areas: study design, training providers, delivery of treatment, receipt of treatment and enactment of treatment skills.² Over the years, this workgroup has further described and refined methods for the assessment, monitoring and enhancement of fidelity.^{9 10}

Despite its importance, fidelity is rarely reported in stroke and aphasia studies. In a review of stroke rehabilitation trials published in 2015 (n=182 papers), fidelity was reported in less than 10% of the studies.¹ In a review of 149 aphasia intervention studies (2002–2011), fidelity was reported in 14% of studies.¹¹ In a more recent review of aphasia randomised controlled trials (RCTs) (2012–2017), 21% of studies explicitly reported fidelity.¹² Similarly, fidelity is rarely reported in befriending studies. In 34 intervention studies across two systematic reviews,^{13 14} only six studies (18%) referred to fidelity.

Peer-befriending is an intervention where social and emotional support is provided by people with experience of a condition to others sharing a similar condition to bring about change.¹⁵ Peer-befriending fits within the stepped care model for offering psychological care,¹⁶ which offers simpler interventions first (level 1) before progressing to more complex interventions. Peer-befriending may be beneficial for people with aphasia post-stroke with no or low mood problems. It may avert some long-term psychological consequences such as depression¹⁷ and reduce the need for more costly psychological therapies. The SUPporting wellbeing through PEER-Befriending (SUPERB) study is a feasibility RCT of peer-befriending for people with aphasia poststroke with low levels of psychological distress.¹⁸ Peer-befriending is a complex intervention requiring active engagement of both the provider and recipient of treatment, with scope for low compliance. Therefore, a comprehensive assessment of fidelity was prespecified and built into the protocol,¹⁸ to monitor the intervention and help interpret the trial results.

The aim of this study was to systematically evaluate fidelity within the SUPERB trial. We describe the development, assessment and monitoring of fidelity strategies and report on the fidelity of all aspects of the intervention, including training and supervision of providers.

METHODS

The SUPERB trial was a single-blind, parallel-group feasibility RCT comparing usual care to usual care +peer-befriending; full methods are reported elsewhere.¹⁸ SUPERB was funded by the Stroke Association and sponsored by City, University of London. Recruitment commenced in April 2017; data collection was completed in August 2019. This paper focuses on the methods related to checking and reporting of fidelity (table 1). Fidelity

strategies were developed, assessed and monitored across all five areas proposed by Bellg *et al.* study design, training providers, delivery of treatment, treatment receipt and treatment enactment.²

Patient and public involvement

People with aphasia were engaged at every stage of the SUPERB trial, with full details reported in the protocol paper.¹⁸ At the proposal stage, a group with experience of befriending from the organisation Aphasia Re-Connect (n=6) influenced our plans, including, the nature and timing of intervention and the use of mixed methods (quantitative and qualitative). At the development phase of the trial, another group of six consultants with aphasia participated in a series of workshops and advised on aspects, such as trial materials including) intervention manuals, outcome measures and selection and matching of peer-befrienders. They also contributed to the development of fidelity checklists. During and after the trial, a user group comprising five people with aphasia and one significant other advised on management issues, helped us explore the implications of the findings, and informed our dissemination to the stroke community (regular trial newsletters, social media promotion, website, videos after the trial).

Participants

Fifty-six participants were recruited and randomised to either usual care (n=28) or usual care +peer-befriending (n=28), soon after discharge from hospital and when active care was withdrawn. Participant characteristics are reported in detail elsewhere.¹⁹ The sample in this report are those in the intervention arm: their mean age was 70.5 years; the majority were white (53.6%), retired (67.9%) with a fairly even split between male (53.6%) and female (46.4%). Most of the group had suffered an ischaemic stroke (82.1%), and left hemisphere stroke (96.4%). Participants were a median (IQR) 48 (21.5–86.5) days poststroke. Aphasia severity ranged from mild (64%), moderate (11%), to severe and very severe (25%), based on the Western Aphasia Battery-Revised.²⁰ Cognition ranged from within normal limits (18%) to mild (32%), moderate (25%) and severe (25%) cognitive impairment (Cognitive Linguistic Quick Test²¹).

Peer-befrienders were people with mild-moderate aphasia, ≥1 year post-stroke, nominated by community services and screened by the trial manager. Ten befrienders were recruited, trained and delivered the befriending intervention. Their mean age was 54.2 years; the majority were female (80%), white (60%) and in volunteer work (50%). Most befrienders had suffered an ischaemic stroke (60%) in the left hemisphere (70%). All befrienders had mild-to-moderate aphasia based on the Frenchay Aphasia Screening Test.²²

Intervention

Full details are provided in the SUPERB protocol paper.¹⁸ Befrienders attended training and participated in

Table 1 Fidelity strategies for the SUPERB trial

Goal ²	Strategies used in SUPERB	Metrics and criteria (primary metrics in bold)
Study design	<p>Ensure same treatment dose within conditions</p> <ul style="list-style-type: none"> ▶ Feasibility RCT. Theoretical model for intervention identified and active ingredients operationalised in a treatment manual and befriender handbook to ensure consistency of training. ▶ Befrienders trained in 4–6 hours over 2–3 days covering a range of topics related to peer-befriending, health and safety, and dealing with adverse events. ▶ Treatment dose set a priori. Six 1-hour visits over 3 months (two further optional visits over 6 months). ▶ Supervision of befrienders (monthly group while they are befriending and additional individual contact as needed) to monitor deviations from the protocol (eg, in no and length of visits). ▶ Befrienders completed an aphasia friendly record form after each participant contact indicating length and content of visit and submitted to supervisor. 	<ul style="list-style-type: none"> ▶ No (and length) of training and supervision sessions recorded ▶ No of group supervision sessions attended by each befriender ▶ Count of visits ▶ Median length of visits
Ensure equivalent dose across conditions	<ul style="list-style-type: none"> ▶ The control condition in SUPERB received usual care only as opposed to the intervention condition who received usual care +peer-befriending. Uptake of services was monitored by documenting the no and frequency of contacts with therapy services, GP appointments and total social and healthcare costs using the Adapted Client Service Receipt Inventory (CSRI)⁴⁴ (for both conditions). Information was gathered: at 4 months (following the intervention) and at 10 months. 	<ul style="list-style-type: none"> ▶ Group comparisons on contacts and NHS services (reported in health economic paper), stroke and befriending services
Plan for implementation setbacks	<ul style="list-style-type: none"> ▶ Training to befrienders was staggered to accommodate a range of geographical areas. ▶ Monthly group supervision for befrienders to monitor implementation and give additional information and resources as and when needed. ▶ Individual supervision (face to face or remote) as and when needed. ▶ All training and supervision of befrienders videotaped to monitor consistency of training and supervision. 	<ul style="list-style-type: none"> ▶ No of group supervision sessions attended by each befriender ▶ Amount of individual supervision provided to each befriender ▶ Topics discussed for each group supervision session recorded
Training providers (ie, befrienders)	<ul style="list-style-type: none"> ▶ All befrienders working on the project were people with mild-moderate aphasia, at least 1 year poststroke and had to meet further predetermined inclusion and exclusion criteria as described in protocol. ▶ Befrienders trained by the same two trainers (a qualified speech and language therapist and a clinical linguist) using set training materials which included a study-adapted version of the treatment manual⁴⁵ which covered a range of topics (eg, the role of a befriender, hopes and fears, dealing with challenging situations) and a befriender handbook that had key information in an aphasia-accessible format. ▶ Befrienders trained in groups of up to five. ▶ All training sessions videotaped. 	<ul style="list-style-type: none"> ▶ No (and length) of training sessions recorded ▶ Training adherence rated against fidelity checklists and inter-rater and intrarater reliability calculated. Acceptable adherence: >80% of checklist behaviours rated as present.²⁷ Reliability: >0.60 is considered good.²⁸
Ensure provider skill acquisition	<ul style="list-style-type: none"> ▶ Trainers checked understanding and used variety of training techniques to ensure befriender engagement and skills acquisition (eg, group discussion, pair work, raised questions and issues, and invited befrienders to express their own opinions about befriending). 	<ul style="list-style-type: none"> ▶ 12 (of 24 items) of training fidelity checklist considered befriender involvement in discussing content or raising questions or issues. ▶ Training adherence rated and inter-rater and intrarater reliability calculated (see criteria above).

Continued

Table 1 Continued	
Goal ²	Strategies used in SUPERB
Minimise therapist (ie, befriender) drift	<p>Monthly group supervision of bidders while actively befriending (and attendance recorded), providing opportunity to share experiences and discuss difficulties that had arisen from intervention visits. Supervision provided by supervisor who was the qualified speech and language therapist involved in the training.</p> <p>Individual supervision (face to face or remote) as and when needed.</p> <p>All group supervision sessions (except first and last) videotaped and independently rated for adherence against fidelity checklist with timely feedback given to supervisor.</p> <p>A single intervention visit (of 6) was videotaped and independently rated for adherence against fidelity checklist with timely feedback given to the befriender by the supervisor.</p>
Accommodate provider (ie, befriender) differences	<p>Two trainers delivered each training session despite the small no of bidders trained (≤ 5) to ensure individual needs were accommodated (eg, by communication support, checking, understanding).</p> <p>Bidders were trained in communication skills for different people with aphasia (mild, moderate, severe).</p> <p>Bidders were trained to deal with different types of participants (eg, quiet, talkative).</p> <p>Trainers accommodated differences between bidders in monthly group and individual supervision sessions.</p>
Delivery of treatment	<p>Matching information about bidders was made available to bidders.</p> <p>Bidders recorded content of visits in aphasia-friendly record forms which were checked by supervisor for deviations from protocol regularly and feedback given.</p> <p>Group and individual supervision sessions to discuss bidder differences and deviations minimised through supervisor feedback.</p> <p>Participant complaints monitored and matching breakdowns recorded.</p> <p>All bidders interviewed about their experiences of delivering the intervention.</p> <p>A purposive sample of participants (n=10) interviewed to explore their perceptions of the intervention and bidder ability to provide the intervention.</p>
Reduce differences within treatment	<p>One (of six) intervention visits and all (but first and last) supervision sessions were videotaped and independently rated for adherence against a checklist.</p> <p>Adherence ratings fed back to supervisor to give information to bidders either individually, or through group supervision sessions.</p> <p>Monthly group supervision sessions conducted to monitor delivery of intervention.</p>
Ensure adherence to protocol	<p>Supervisor monitored bidders correct no of sessions were given to participants.</p> <p>Adherence to the protocol for all (but first and last) supervision sessions and one (of six) intervention visits videotaped and independently rated for adherence against a fidelity checklist.</p> <p>All bidders completed aphasia friendly record forms for each intervention visit, which were monitored by supervisor.</p>
	<p>Metrics and criteria (primary metrics in bold)</p> <ul style="list-style-type: none"> ▶ No (and length) of group supervision sessions. ▶ No of group supervision sessions attended by each befriender. ▶ No of bidders in each group supervision session. ▶ Amount of individual supervision provided to each befriender. ▶ Topics discussed for each group supervision session recorded ▶ Supervision adherence rated against fidelity checklist and inter-rater and intrarater reliability calculated (see criteria above). ▶ Intervention visit adherence rated and inter-rater and intrarater reliability calculated (see criteria above). <p>12 (of 24 items) of training fidelity checklist considered bidder involvement in discussing content or raising questions or issues.</p> <ul style="list-style-type: none"> ▶ Training adherence rated and inter-rater and intrarater reliability calculated (see criteria above). <p>Matching criteria for each bidder-bidder recorded.</p> <ul style="list-style-type: none"> ▶ No adverse events in relation to matches recorded. ▶ No of visits rescheduled and reasons recorded. ▶ Qualitative data (reported elsewhere). <p>Intervention visit and supervision adherence rate and reliability calculated (see criteria above).</p> <ul style="list-style-type: none"> ▶ Topics discussed for each group supervision session recorded. ▶ No of group supervision sessions attended ▶ Intervention visit and supervision adherence rate and reliability calculated (see criteria above). ▶ No of intervention visits was recorded with any protocol deviations or problems recorded. ▶ Completion rates for aphasia friendly record form recorded (>80—good) ▶ Bidder participation in supervision recorded (details above).

Continued

Table 1 Continued	
Goal ²	Strategies used in SUPERB
Minimise contamination between conditions	<p>► Participants in control arm were not aware of what intervention was tested. Two-stage consent process where all participants first consented to a study about adjustment poststroke and aphasia and only participants randomised to the intervention arm were given information about befriending intervention and asked to consent to this arm.</p> <p>► Recruitment sites selected on the basis that no befriender intervention was available in that setting.</p> <p>► Befrienders were trained to only give a single intervention.</p> <p>► CSRI information were collected to check control group did not receive a different befriending intervention.</p>
Receipt of treatment	
Ensure participant comprehension	<p>► Befrienders were trained on giving information in an accessible way to participants (see handbook) and this was discussed on an ongoing basis within group supervision sessions.</p> <p>► One (of six) intervention visits from each match were videotaped and independently rated for adherence against a fidelity checklist.</p> <p>► Purposive sample of participants interviewed postintervention about their experiences.</p>
Ensure participant ability to use cognitive skills	<p>► One (of six) intervention visits from each match was videotaped and independently rated for adherence against a fidelity checklist.</p>
Ensure participants ability to perform behavioural skills	<p>► Not directly assessed.</p>
Enactment of treatment skills	
Ensure participant use of cognitive skills	<p>► Purposive sample of participants interviewed post-intervention and 6 months later to explore their experiences and impact of intervention.</p>
Ensure participant use of behavioural skills	<p>► Purposive sample of participants interviewed postintervention and 6 months later to explore their experiences and impact of intervention.</p>
	<p>GP, general practitioner; NHS, National Health Service; RCT, randomised controlled trial; SUPERB, Supporting wellbeing through Peer-befriending.</p>
	<p>Metrics and criteria (primary metrics in bold)</p> <p>► Group comparisons on contacts and NHS services (reported in full in health economic paper), stroke and befriending services, based on CSRI.</p>
	<p>► Intervention visit adherence checklist includes six items on acknowledging the competence of the person with aphasia (four essential; two optional).</p> <p>► Intervention visit adherence rate and reliability calculated (see criteria above)</p> <p>► No of times accessible communication was discussed in group supervision.</p> <p>► Qualitative data (reported elsewhere).</p>
	<p>► Visit adherence checklist includes nine (out of 16) items on acknowledging and revealing the competence of the person with aphasia.</p> <p>► Intervention visit adherence rate and reliability calculated (see criteria above).</p>

regular supervision (monthly group, one-to-one as and when needed) to deliver the intervention as intended. Befrienders were matched with and introduced to their participant with aphasia soon after and within 1 month of randomisation. Matching criteria included geographical location and preferences around sex, cultural factors (eg, ethnicity, religion, languages spoken), age, and personal interests. After they met, both parties had the option to request a rematch, but this did not happen. Befrienders worked with 2–4 participants during the project and no more than two at any one time.

Befrienders visited participants in their own homes six times over a period of 3 months (each visit ~1 hour). Another two optional visits within the next 6 months could also be offered for a gradual transition to the end of peer-befriending. The schedule and nature of visits was agreed between the pair at their first meeting, as well as possible goals for the intervention (eg, participants might highlight concerns that they would like to discuss or activities that they would like to pursue). Subsequent visits included: conversation, problem solving, trips out and joint activities.

After each visit, to control for differences between befrienders and ensure adherence to protocol, befrienders completed an aphasia friendly, 12-question record form (including whether a visit was cancelled and reason why, topics discussed, activities undertaken and date and time of next visit). The supervisor helped if necessary and collated the forms during monthly supervision sessions.

Adherence rating: development of fidelity checklists for training, supervision and delivery of treatment

All training sessions, all but first and last supervision sessions, and one of six visits per pair were recorded to be rated for adherence. The visit chosen was not selected randomly but rather for practical reasons: as befrienders experienced difficulties with recording, a mutually convenient visit was chosen for a research assistant to set up equipment, leave for visit to take place and return to collect equipment. Three checklists were developed for the training, supervision and intervention visits (see online supplemental files 1–3). In line with recent recommendations, the development process involved identifying intervention components, reviewing related measures and obtaining user feedback on the checklists' content.²³ Training checklist items were based on training manual content. Supervision and intervention visit items were based on the Health Behaviour Change Competency Framework²⁴ which describes the competencies required to develop and deliver interventions; and the Measure of Participation in Conversation and Measure of Support in Conversation²⁵ which described behaviours integral to communicating with people with aphasia.

The content of each checklist was reviewed through an iterative process involving four authors (KH, SM, NB and BM) with experience of both the befriending intervention and fidelity. Items were either essential or optional.

Essential items (rated 'present' or 'absent') were behaviours considered critical to the delivery of the intervention. Optional items ('present', 'absent' or 'not applicable') were behaviours which would enhance aspects of the intervention but were not critical in every session (eg, discussion of risk and safeguarding in supervision, presence of humour). The training checklist comprised 24 essential items. The supervision checklist comprised 14 items (13 essential). The intervention visit checklist contained 16 items (13 essential, three optional).

The checklists were piloted by an MSc student who rated two training sessions, four supervision sessions and one visit. The ratings were discussed with the two lead authors (NB, KH) and minor changes were made to make item descriptions clearer.

Fidelity rating process

Raters comprised eight MSc Speech and Language Therapy students and one research assistant with a degree in Speech and Language Therapy. No raters were involved in the intervention. Depending on their skill and experience, raters attended a training session of 2–4 hours which involved familiarising them with the checklists, watching and rating session videos using the fidelity checklists and discussing rating discrepancies.

Each session was rated for adherence by two raters, which allowed testing of inter-rater reliability. For intra-rater reliability, sessions were rerated by the same rater after a period of at least 2 weeks. Interim fidelity results were fed back to the supervisor to use during befriender supervision to ensure protocol adherence.

Analysis of adherence

The adherence score for each session was calculated by dividing the number of checklist items present by the number of items expected (excluding items recorded as not applicable).²⁶ These scores were then converted to a percentage score. A score of 80% and above reflected a high level of treatment adherence.^{26 27} To calculate inter-rater and intrarater reliability, Gwets AC1 was used.^{28 29} This statistic is less susceptible to skew due to prevalence compared with Cohen's kappa.³⁰ A reliability coefficient of 0.81–1.00 is considered very good, 0.61–0.80 good, 0.41–0.60 moderate, 0.21–0.40 fair and below 0.20 poor.²⁸ Reliability was calculated for essential items only. Statistical analyses were carried out using STATA V.13.

RESULTS

Study design

Treatment dose within condition and implementation setbacks: training and supervision

Three training programmes were completed, attended by five, three and two befrienders, respectively. Each befriender attended one training programme. All training sessions were videotaped and completed as intended in 4–6 hours (not including rest periods) over 2–3 days.

Table 2 Topics discussed in group and individual supervision sessions (from most to least discussed)

Content	% occurrence over all supervision
Administration (eg, update on matches, befriender record forms, arranging visits)	63.6
Endings (eg, how to say goodbye, managing only six visits)	36.8
Other (eg, study information, disability resources)	36.8
Communication (eg, using total communication strategies, identifying topics to discuss)	34.9
Boundaries (eg, managing distractions, personal boundaries, understanding your limitations)	27.3
Managing relatives	19.6
Emotions (eg, managing own and other's emotions, providing emotional support)	19.1
Risks and safeguarding (eg, reminders of process, check in-out)	19.1
Coping with people living in difficult circumstances	16.3
Getting started (eg, managing first visit, practical considerations)	13.9

There were 25 group supervision sessions, and 77 individual sessions (6 face to face, 67 telephone and 4 email). Each group session lasted a median length of 90 min (IQR=90–90) and the median number of befrienders attending each group session was 5 (IQR=4–6). Befrienders were expected to attend group supervision while they were actively befriending during the trial and the median number of group supervision sessions attended per befriender was 14 (IQR=8–18). Befrienders attended 73% (IQR=52–91) of sessions while they were actively befriending. They also attended between one and 13 sessions when they were not actively befriending (IQR=4.5–10). The topics discussed in each group supervision session are detailed in [table 2](#). Topics discussed in the earlier group supervision sessions related to administration, other topics (eg, disability resources available to people post-stroke), risk and safeguarding, boundaries and managing relatives. In later sessions, how to communicate with people with aphasia (and of different severities) and endings were commonly discussed. To minimise implementation setbacks (and befriender drift) supervision sessions consistently covered administration issues (63.6%). The seventy-seven individual supervision sessions also helped minimise drift and accommodate befriender differences. Befrienders attended a median of

Table 3 Number of intervention visits completed by participants (n=26)

Befriending visits	N (%)
Visit 1	24 (92.3)
Visit 2	24 (92.3)
Visit 3	22 (84.6)
Visit 4	21 (80.8)
Visit 5	21 (80.8)
Visit 6	21 (80.8)
Visit 7 (optional)	6 (23.1)
Visit 8 (optional)	4 (15.4)
Total visits completed	143

7.5 (IQR=7–8) individual sessions, lasting a median of 10 min (IQR=10–15).

Treatment dose within condition: intervention visits

The number ([table 3](#)) and length of intervention visits was recorded. Of the 28 people randomised to the intervention arm, 26 gave consent to participate in the befriending intervention. The remaining two were either not interested or were moving overseas for the duration of the intervention. As per protocol, the median number of visits attended was 6 (IQR=6–6.5) with a median length of 60 min (IQR=60–70 min).

Treatment dose across conditions

Usual care was received in both arms of the trial and equivalence of dose assessed. There were no significant differences between the usual care and peer-befriending groups based on the Client Service Receipt Inventory for residential and nursing home care; hospital overnight stays; day patient treatment; accident and emergency visits; general practitioner appointments; physiotherapy, occupational therapy, speech and language therapy sessions; community-based healthcare professionals; dentists; social services; individual travel and out of pocket expenses. Importantly, there were no significant differences between the two arms on total social and healthcare costs and overall total costs. There was a higher number of outpatient appointments for the control group at 10 months ($p=0.04$).³¹ To explore contamination between conditions, we looked in more detail at support services such as befriending, stroke groups and other support services (eg, church support, gym groups, charity stroke co-ordinator visits). One participant from the intervention condition at the 4-month assessment point reported receiving four visits by a Red Cross volunteer providing instrumental support with community mobility. No participant reported receiving peer-befriending (outside that delivered in the intervention arm). An equal number of participants from both conditions were going to stroke groups at 4 (n=16) and 10 months (n=10). A small number of

participants received other support services across arms (n=4 in intervention vs n=3 in usual care at 4 months; n=2 in intervention at 10 months).

Training providers

Standardisation of training and provider skill acquisition

Adherence for training sessions ranged from 91.7% to 100%. For the 12 items in the checklists that more closely tapped into befriender skill acquisition and accommodating befriender differences, 10/12 (83%) were rated as present by both raters, at all trainings. Only 1–2/24 items were rated as absent in 4/9 training ratings. Training inter-rater reliability was very good (Gwets AC1 0.91–0.96, $p<0.001$). Intra-rater reliability was very good for one training session (Gwets AC1=0.95 (0.85–1.0), $p<0.001$), with no variability between ratings for the remaining sessions (rated 100%).

Minimising befriender drift and accommodating differences

Adherence for supervision sessions ranged from 83% to 100%, based on 22/25 sessions (first and last excluded as planned; fifth session not recorded due to human error). Only two items were rated as absent. Inter-rater reliability could not be calculated (no variability in scores, that is, all essential items rated present by both raters). Intra-rater reliability was calculable and very good for one supervision session (Gwets AC1=0.90 (95% CI 0.66 to 1.00), $p<0.001$), with perfect agreement for remaining sessions.

Delivery of treatment

Aspects of the fidelity of the treatment delivery that relate to supervision (in terms of controlling for befriender differences, reducing differences in treatment, ensuring adherence to protocol) and group comparisons of services received (to minimise contamination effects) go across categories and have been covered under study design and training of providers above. In this section, the focus is primarily on the intervention visits.

Table 4 Proportion of participants matched to befrienders on the following criteria

Matching criteria	N (%)
Geographical location	25 (96.2)
Similar interests	22 (84.6)
Similar hobbies	15 (57.7)
Ethnic/cultural background	17 (65.4)
Religious background	16 (61.5)
Similar age	6 (23)
Same sex	16 (61.5)
Environment for example, smokers/pets	24 (92.3)
Other (eg, same language, both have children, befriender experience in working with people with severe aphasia)	5 (19.2)

Controlling for befriender differences and adherence to protocol: intervention visits

Participants were matched with befrienders according to set criteria (table 4). Given that befrienders had to travel to the participant's home, geography was the major deciding factor. To control for differences between befrienders, researchers consistently considered practicalities (eg, smokers, pets), interests and hobbies, and commonalities in terms of ethnic/cultural background, religion and sex. Most participants did not mind what the ethnic or religious background, age or sex of the befriender was (16/26; 61.5%). Where participants did indicate a preference, it was related to a befriender's sex (6/26; 23.1%), ethnic background (3/26, 11.5%), age (2/26, 8%) and religious background (1/26, 3.8%).

No adverse events were recorded for any of the 26 matches. Two participants who initially gave consent to participate in peer-befriending withdrew before the first visit as they did not feel they had the time to participate or felt uncomfortable with having visitors in their home. For the remaining 24 participants, four befrienders were matched with three participants, three with two participants; two with one participant; and one befriender was matched with four participants. Six visits were completed by 21 (87.5%) of befrienders. Of the three participants who did not complete the required number of visits, one could not be contacted after the third visit, one stopped as they felt they did not need more visits, and for the last one, their significant other did not feel the visits were beneficial.

Further evidence of controlling for befriender differences and ensuring adherence to protocol come from visit statistics and the befriender record forms. Of the 143 visits completed, 116 (79.5%) happened as planned, 27 (18.5%) were rescheduled and only three (2.1%) were cancelled. Of those rescheduled, 22 (81.5%) were rescheduled only once. Reasons for first rescheduled or cancelled visits were related to person with aphasia (50%), peer-befriender (40%) or other reasons. Across the 143 visits, befrienders completed the aphasia-friendly record form for 139 visits (97.2%) with a median of 11 (of 12) responses recorded on each form (IQR=10–11). These forms were checked by the supervisor to ensure intervention was delivered as intended.

Reducing differences within treatment

In terms of adherence of intervention visits, of the 24 participants that had befriending visits, 20 videotaped visits were obtained; four participants completed fewer than the required six visits and were unable to be videotaped during the intervention period for practical reasons. Adherence for intervention visits is reported in table 5 and ranged from 66.7% to 100%, with only one visit rated below 80% by one rater. Across the 20 visits, nine different checklist items were rated as absent on at least one occasion. The three items most often reported as absent were: ability to manage the physical and social

Table 5 Treatment adherence for intervention visits

Intervention visit (visit no out of six)	Rater 1% (items implemented/items expected)	Rater 2 time 1* % (items implemented/items expected)	Rater 2 time 2* % (items implemented/items expected)
1 (V4)	92.9 (13/14)	100 (16/16)	100 (16/16)
2 (V4)	100 (15/15)	100 (16/16)	100 (16/16)
3 (V2)	92.9 (13/14)	100 (14/14)	100 (15/15)
4 (V2)	93.3 (14/15)	100 (14/14)	100 (15/15)
5 (V3)	100 (14/14)	100 (14/14)	100 (16/16)
6 (V4)	100 (16/16)	87.5 (14/16)	85.7 (12/14)
7 (V2)	100 (15/15)	93.3 (14/15)	87.5 (14/16)
8 (V2)	100 (16/16)	100 (15/15)	100 (15/15)
9 (V4)	100 (14/14)	100 (15/15)	100 (15/15)
10 (V5)	66.7 (10/15)	100 (15/15)	100 (15/15)
11 (V5)	92.3 (12/13)	92.9 (13/14)	92.9 (13/14)
12 (V5)	100 (16/16)	100 (16/16)	100 (16/16)
13 (V4)	100 (16/16)	100 (16/16)	100 (16/16)
14 (V5)	100 (15/15)	100 (14/14)	100 (15/15)
15 (V4)	100 (16/16)	100 (15/15)	100 (15/15)
16 (V3)	100 (15/15)	100 (14/14)	100 (15/15)
17 (V4)	100 (16/16)	100 (13/13)	100 (13/13)
18 (V5)	100 (16/16)	100 (15/15)	100 (15/15)
19 (V5)	100 (15/15)	84.6 (11/13)	85.7 (12/14)
20 (V6)	100 (14/14)	100 (13/13)	100 (13/13)

*Time 1 and 2 refer to rating time points by the second rater to calculate intrarater reliability.

environment, give relevant support to the befriender and have a natural and non-patronising conversation.

For inter-rater reliability, there was a moderate (Gwets ACI=0.44 (95% CI -0.15 to 1.00), $p=0.13$) to very good level of agreement (Gwets ACI=0.82–1.00, $p<0.001$) between the two raters for eight intervention visits, with perfect agreement for the remaining sessions ($n=12$). Intrarater reliability was good to very good for four visits (Gwets ACI=0.76–1.00, $p<0.01$), with perfect agreement for other sessions ($n=16$).

Treatment receipt

Some fidelity strategies span across categories. Aspects of fidelity of treatment receipt that related to training (eg, communication skills training) and supervision (eg, number of times communication was discussed) have been covered above.

Ensuring participant comprehension of the intervention

Checklist items that rated acknowledging the competence of the person with aphasia were reviewed. The essential items ($n=4$) related to how the befriender created a natural, non-patronising and sensitive conversation, that demonstrated active listening of the thoughts and concerns of the participant and responded in a way that was sensitive and respectful. These four items were

rated as present for 18/20 visits. The optional items ($n=2$) related to how a befriender used humour to help participants (eg, to reduce tension) and how they responded to a participant's humour. These two items were frequently rated as not applicable in 10/20 and 9/20 visits.

Ensuring participant use of cognitive skills

Checklist items that rated the ability of the befriender to acknowledge and reveal the competence of the person with aphasia were reviewed (three additional essential items). All items were rated as present with the exception of one item that was rated as absent in 2/20 visits by a single rater.

Treatment enactment

Treatment enactment of befriending is complex and relates not only to participant activities but also their feelings, identity and coping. Qualitative data from the SUPERB trial are rich and extensive and capture treatment enactment; they are reported elsewhere.³² Here, we provide some illustrative comments, using pseudonyms, that demonstrate a degree of treatment enactment. There were instances where participants with aphasia undertook activities that they would not have done if it were not for their befriender. One participant, Betsy did not like to go

out but said that seeing her befriender, who she felt was like her, going out helped her start going out.

Other participants reflected on taking on the advice of the bendifenders. They valued that bendifenders gave advice from personal experience. Tips, such as recommending local stroke support groups and a taxi-card service, were perceived as helpful and had been acted on. James said a key realisation for him was that his befriender was still able to drive, and that he could aspire to that too. Ivy continued to meet her befriender at a local stroke group, which he had first invited her to.

DISCUSSION

We planned and employed a range of strategies to ensure the fidelity of the peer-befriending intervention for people with aphasia in the SUPERB trial. Fidelity across a range of categories was high. All bendifenders attended the training and a high number of group supervision sessions. Over 80% of participants received the required six 1-hour intervention visits. Adherence was high for training and supervision sessions and moderate-to-high for intervention visits. Fidelity ratings were reliable: very good for the training and supervision sessions; and moderate to very good for the intervention visits.

High adherence to fidelity checklists is consistent with the few other studies in aphasia.^{26 27 33} High adherence to training and supervision is likely due to using the same set of materials, structured agenda and the same experienced facilitators across all sessions.¹⁰ Strategies to support the bendifenders included monthly group sessions and individual supervision as required, which were key to preventing drift in their skills and ensured they completed visits. Monthly group meetings have previously been offered for other befriending studies.¹⁴ Bendifenders were highly supported to attend group supervision particularly during the time they were completing intervention visits. Supervision provided an opportunity to minimise implementation setbacks (eg, managing communication, boundaries) and to refamiliarise bendifenders with relevant topics from the training (eg, risk and safeguarding, ending visits).

There was excellent adherence for a large proportion of intervention visits. To maximise protocol adherence, monitoring fidelity prospectively ensured that differences within the intervention arm (and between bendifenders) were minimised.^{2 9} Deviations could be fed forward to the facilitator of the supervision sessions to review and discuss visit related topics. Bendifenders were able to promote supported conversation for participants³⁴ where they listened to the opinions and concerns of people with aphasia and gave them the opportunity to speak. Even those bendifenders (n=4) who were matched with participants with severe aphasia could accommodate their skills to communicate effectively with participants. Nevertheless, some variability was noted in adherence for the intervention visits. The nature of the sessions may help to explain this. More highly structured sessions

(eg, training and supervision) tend to show greater fidelity.³⁵ Befriending visits are less structured and inherently require a more subjective judgement, for example, whether a befriender has spoken in a natural and non-patronising manner or used and responded appropriately to humour. Greater rater training to clearly explain vague or ambiguous terms and establish a predetermined level of agreement before rating study samples may help.³⁶ However, more variability may also reflect lack of/developing skills for some bendifenders, during their initial pairings before fidelity checks had occurred. Subsequent feedback ensured bendifenders displayed the necessary skills for future matches.

Behaviours related to managing the environment, giving support to the participant and having a natural conversation were the ones most commonly rated as absent. However, they were mostly rated as absent by a single rather than both raters. This finding may suggest they were missed, too subtle to have been observed or difficult to assess, with training inadequate for helping raters accurately recognise them. Therefore, clearer specification of behaviours may be needed. In half the evaluated sessions both raters judged the use and response to humour as 'non-applicable', suggesting these items should remain optional. Humour is considered integral to interactions with friends but there may be fewer opportunities for humour with less familiar people,³⁷ and not all interactions require humour.

This study advances our understanding of fidelity beyond simply adherence to an intervention protocol, as in previous studies.^{1 3 4 38} While a focus on implementation and factors that may moderate fidelity⁴ could be appropriate for pragmatic trials,³⁹ the results of this study demonstrate the importance of planning a comprehensive set of fidelity practices at the outset of an early explanatory trial, like SUPERB. Through consideration of strategies proposed by Bellg *et al*,² our understanding of fidelity moves beyond treatment delivery alone. Insight is gleaned of the planned strategies that led to the fidelity results, such as, same skilled facilitators, standardised training with set materials, regular and flexible supervision, feeding fidelity findings into supervision. This latter strategy did not happen with one befriender early in the trial who had low adherence; in contrast, overall, discussing fidelity of visits in supervision before giving a befriender a second match ensured differences were considered, engagement encouraged and understanding of important components of intervention maximised.

Another strength of this study was the broad range of information gathered about the dosage of training, supervision and intervention; adherence and reliability. In SUPERB, 100% of training, 88% of supervision and 14% of intervention sessions were checked by two independent raters for adherence and all of these were rated for reliability. In most studies, a proportion of sessions (usually 10%–33%) is rated for adherence with a proportion of those second rated for reliability.^{11 26 27 40} Furthermore, reliability was calculated in multiple ways which is

rarely done in aphasia trials. Many studies use per cent agreement to determine reliability, which does not take chance agreement into account.⁴¹ This study provided more reliability evidence by using Gwets AC1 that can be calculated if there is variability in the scores of one rater, instead of the commonly reported kappa, which requires variability in the scores of both raters. The study also provided a description of services received by those participants in the usual care arm to monitor for contamination effects.

Limitations included not having a videotaped intervention visit for all 26 pairings, and only rating for adherence 1/6 visits for each match. Visits to record were not chosen randomly but on practical grounds (availability of researcher to visit session) which may have led to some selection bias. It is also possible that the highly selective videoing of sessions influenced the behaviours of befrienders. Videoing all sessions might reduce such an impact and allow random selection of sessions for rating. Despite the restricted sampling, 14% visits were rated for adherence which is consistent with other studies.¹¹

In terms of clinical implications, peer-befriending for people with aphasia can be implemented with good-high fidelity. Careful planning, thorough training, regular (group) and flexible (individual) befriender supervision minimise drift and ensure the intervention is delivered as intended. Supervision should be provided by a skilled facilitator with experience in working with people with aphasia. Information gathered from supervision suggested that helping befrienders to plan and organise visits was a vital part of the practical support they needed. In addition, refamiliarising befrienders throughout supervision with topics initially trained and checking fidelity early and feeding back to befrienders is needed to ensure the intervention is delivered as intended.

CONCLUSION

Peer-befriending for people with aphasia post-stroke was delivered with good-high fidelity in the SUPERB RCT. The study was novel as one of still only a handful of aphasia trials reporting fidelity in a systematic way, based on a theoretical framework.^{33 39 42 43} The study demonstrated the importance of comprehensive planning for fidelity at the outset of a trial and the need for ongoing rigorous monitoring throughout the trial to ensure an intervention is implemented as intended.

Author affiliations

¹Centre for Language and Communication Science Research, City University of London, London, UK

²Division of Rehabilitation & Ageing, University of Nottingham, Nottingham, UK

³Health Services and Population Research, King's College London, London, UK

⁴School of Health and Social Care, London South Bank University, London, UK

⁵Biostatistics, Institute of Psychiatry Psychology and Neuroscience, King's College London, London, UK

Twitter Nicholas Behn @NicholasBehn, Becky Moss @BeckMoss, Sally McVicker @sallymckvicker, Abi Roper @abracadabger, Sarah Northcott @sarahnorthcott8,

Shirley Thomas @DrShirleyThomas, Alan Simpson @cityalan, Kimberley Goldsmith @KimberleyGol and Katerina Hilari @KaterinaHilari

Acknowledgements We thank the recruitment sites and the participants in the study. We also thank Chloe Chalmers, Lauren Cook, Lauren Cooper, Bethanie Hall, Rachael Lee, Nicole Palanyandi, Tamsin Sharp and Georgia Whittle, who were student SLTs at the time of the study and contributed to fidelity ratings. KG contributions to this paper represent independent research part funded by the NIHR Biomedical Research Centre (South London and Maudsley NHS Foundation Trust and KCL) and the NIHR Applied Research Collaboration South London (King's College Hospital NHS Foundation Trust).

Contributors KH designed the study with contributions from SN, JM, ST, AS, CF, SM and KG. NB and KH were responsible for monitoring fidelity checks by raters throughout the trial. NB, KH, BM and SM developed the fidelity checklists and AR was one of the fidelity raters. NB analysed the data with contributions from KJ. NB wrote early drafts of the paper and KH contributed and finalised. All authors reviewed final versions of the manuscript.

Funding SUPERB was funded by The Stroke Association, Psychological Consequences of Stroke – Priority Programme Award (PPA2015-03).

Disclaimer The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care or The Stroke Association.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval SUPERB received full ethics approval from the London Bloomsbury National Health Service ethics committee (ref 16/LO/2187) and all participants provided written consent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. Data reported in this paper comprise de-identified participant data. Data will be available on reasonable request from the corresponding author, <https://orcid.org/0000-0003-2091-4849>

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Nicholas Behn <http://orcid.org/0000-0001-9356-9957>

Becky Moss <http://orcid.org/0000-0003-0340-533X>

Sally McVicker <http://orcid.org/0000-0002-4939-4210>

Abi Roper <http://orcid.org/0000-0001-6950-6294>

Sarah Northcott <http://orcid.org/0000-0001-8229-5452>

Jane Marshall <http://orcid.org/0000-0002-6589-221X>

Shirley Thomas <http://orcid.org/0000-0003-0704-9387>

Alan Simpson <http://orcid.org/0000-0003-3286-9846>

Chris Flood <http://orcid.org/0000-0001-5170-7792>

Kirsty James <http://orcid.org/0000-0002-2657-8148>

Kimberley Goldsmith <http://orcid.org/0000-0002-0620-7868>

Katerina Hilari <http://orcid.org/0000-0003-2091-4849>

REFERENCES

- 1 Walker MF, Hoffmann TC, Brady MC, *et al*. Improving the development, monitoring and reporting of stroke rehabilitation research: consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. *Int J Stroke* 2017;12:472–9.

- 2 Bellg AJ, Borrelli B, Resnick B, *et al.* Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH behavior change Consortium. *Health Psychol* 2004;23:443–51.
- 3 Mowbray CT, Holter MC, Teague GB, *et al.* Fidelity criteria: development, measurement, and validation. *Am J Eval* 2003;24:315–40.
- 4 Carroll C, Patterson M, Wood S, *et al.* A conceptual framework for implementation fidelity. *Implement Sci* 2007;2:40.
- 5 Moncher FJ, Prinz RJ. Treatment fidelity in outcome studies. *Clin Psychol Rev* 1991;11:247–66.
- 6 Resnick B, Inguito P, Orwig D, *et al.* Treatment fidelity in behavior change research: a case example. *Nurs Res* 2005;54:139–43.
- 7 Hohmann AA, Shear MK. Community-based intervention research: coping with the "noise" of real life in study design. *Am J Psychiatry* 2002;159:201–7.
- 8 Perepletchikova F, Treat TA, Kazdin AE. Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. *J Consult Clin Psychol* 2007;75:829–41.
- 9 Borrelli B, Sepinwall D, Ernst D, *et al.* A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *J Consult Clin Psychol* 2005;73:852–60.
- 10 Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent* 2011;71:S52–63.
- 11 Hinckley JJ, Douglas NF. Treatment fidelity: its importance and reported frequency in aphasia treatment studies. *Am J Speech Lang Pathol* 2013;22:S279–84.
- 12 Brogan E, Ciccone N, Godecke E. Treatment fidelity in aphasia randomised controlled trials. *Aphasiology* 2019;33:759–79.
- 13 Mead N, Lester H, Chew-Graham C, *et al.* Effects of befriending on depressive symptoms and distress: systematic review and meta-analysis. *Br J Psychiatry* 2010;196:96–101.
- 14 Siette J, Cassidy M, Priebe S. Effectiveness of befriending interventions: a systematic review and meta-analysis. *BMJ Open* 2017;7:e014304.
- 15 Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. *Psychiatr Rehabil J* 2004;27:392–401.
- 16 Richards DA, Bower P, Pagel C, *et al.* Delivering stepped care: an analysis of implementation in routine practice. *Implement Sci* 2012;7:3.
- 17 Kauhanen M, Korpelainen JT, Hiltunen P, *et al.* Poststroke depression correlates with cognitive impairment and neurological deficits. *Stroke* 1999;30:1875–80.
- 18 Hilari K, Behn N, Marshall J, *et al.* Adjustment with aphasia after stroke: study protocol for a pilot feasibility randomised controlled trial for supporting wellbeing through peer befriending (SUPERB). *Pilot Feasibility Stud* 2019;5:14.
- 19 Hilari K, Behn N, James K. Supporting wellbeing through peer-befriending (SUPERB) for people with aphasia: a feasibility randomised controlled trial. *Clin Rehabil* 2021;35:1151–63.
- 20 Kertesz A. Western aphasia battery - revised (WAB-R), 2006
- 21 Helm-Estabrooks N. Cognitive linguistic quick test (CLQT), 2001
- 22 Enderby P, Wood V, Wade D. *Frenchay aphasia screening test (FAST)*. West Sussex: Wiley, 1986.
- 23 Walton H, Spector A, Williamson M, *et al.* Developing quality fidelity and engagement measures for complex health interventions. *Br J Health Psychol* 2020;25:39–60.
- 24 Dixon D, Johnston M. Health behaviour change competency framework: Competences to deliver interventions to change lifestyle behaviours that affect health 2010. Available: <http://www.healthscotland.com/documents/4877.aspx> [Accessed 17 Feb 2017].
- 25 Kagan A, Winckel J, Black S, *et al.* A set of observational measures for rating support and participation in conversation between adults with aphasia and their conversation partners. *Top Stroke Rehabil* 2004;11:67–83.
- 26 Kladochou V, Papathanasiou I, Efstathiadou EA, *et al.* Treatment integrity of elaborated semantic feature analysis aphasia therapy delivered in individual and group settings. *Int J Lang Commun Disord* 2017;52:733–49.
- 27 Heilemann C, Best W, Johnson F. Investigating treatment fidelity in a conversation-based therapy. *Aphasie und Verwandte Gebiete* 2014;2:14–26.
- 28 Gwet KL. Handbook of inter-rater reliability: the definitive guide to measuring the extent of agreement among raters.. In: 4th. Gaithersburg: Advanced Analytics, LLC, 2014.
- 29 Gwet KL. *Kappa statistic is not satisfactory for assessing the extent of agreement between raters social capital and health.* , 2002: 1, 1–6.
- 30 Wongpakaran N, Wongpakaran T, Wedding D, *et al.* A comparison of Cohen's kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Med Res Methodol* 2013;13:61.
- 31 Flood C, Behn N, Marshall J. A pilot economic evaluation of a feasibility trial for supporting wellbeing through PEeR-Befriending for post stroke aphasia (SUPERB). Clinical rehabilitation. submitted.
- 32 Moss B, Behn N, Northcott S. "Loneliness can also kill:" a qualitative exploration of outcomes and experiences of the SUPERB peer-befriending scheme for people with aphasia and their significant others. *Disabil Rehabil* 2021;231:1–10.
- 33 Conlon EL, Braun EJ, Babbitt EM, *et al.* Treatment fidelity procedures for an aphasia intervention within a randomized controlled trial: design, feasibility, and results. *Am J Speech Lang Pathol* 2020;29:412–24.
- 34 Kagan A, Black SE, Duchan FJ, *et al.* Training volunteers as conversation partners using "Supported Conversation for Adults with Aphasia" (SCA): a controlled trial. *J Speech Lang Hear Res* 2001;44:624–38.
- 35 Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol* 2008;41:327–50.
- 36 Hart T, Brockway JA, Whyte J, *et al.* Analyzing the ingredients of a telephone counseling intervention for traumatic brain injury. *Disabil Rehabil* 2013;35:1668–75.
- 37 Davidson B, Howe T, Worrall L, *et al.* Social participation for older people with aphasia: the impact of communication disability on friendships. *Top Stroke Rehabil* 2008;15:325–40.
- 38 Orwin RG. Assessing program fidelity in substance abuse health services research. *Addiction* 2000;95(Suppl 3):S309–27.
- 39 Palmer R, Dimairo M, Latimer N, *et al.* Computerised speech and language therapy or attention control added to usual care for people with long-term post-stroke aphasia: the big cactus three-arm RCT. *Health Technol Assess* 2020;24:1–176.
- 40 Jackson HJ, McGorry PD, Killackey E, *et al.* Acute-Phase and 1-year follow-up results of a randomized controlled trial of CBT versus Befriending for first-episode psychosis: the ACE project. *Psychol Med* 2008;38:725–35.
- 41 Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol* 2012;8:23–34.
- 42 Carragher M, Ryan B, Worrall L, *et al.* Fidelity protocol for the action success knowledge (ASK) trial: a psychosocial intervention administered by speech and language therapists to prevent depression in people with post-stroke aphasia. *BMJ Open* 2019;9:e023560.
- 43 Thomas SA, Russell C, Seed R, *et al.* An evaluation of treatment integrity in a randomized trial of behavioural therapy for low mood in stroke patients with aphasia. *Clin Rehabil* 2013;27:1097–106.
- 44 Forster A, Dickerson J, Young J, *et al.* A cluster randomised controlled trial and economic evaluation of a structured training programme for caregivers of inpatients after stroke: the TRACS trial. *Health Technol Assess* 2013;17:1–216.
- 45 McVicker S, Swinburn K. *How to set up a befriending scheme: a guide for scheme organisers*. UK: Connect Press, 2009.