PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Intensive and standard group-based treatment for persons with social communication difficulties after an acquired brain injury: study protocol for a randomised controlled trial
AUTHORS	Hansen, Silje; Stubberud, Jan; Hjertstedt, Marianne; Kirmess, Melanie

VERSION 1 – REVIEW

REVIEWER	Sara da Silva Ramos The Disabilities Trust
REVIEW RETURNED	01-Mar-2019

GENERAL COMMENTS	Dear authors, Your paper proposes a valuable addition to the limited existing evidence around interventions for those with social communication difficulties following acquired brain injury (a term which I suggest should be included in the title) Understanding the differences between standard outpatient and intensive inpatient rehabilitation protocols could have important implications for practice and policy. It is clear that there was careful consideration of the study design, but upon careful review of the manuscript, I have a few concerns that require clarification and significant review. The specific objectives stated on page 9 are not fully consistent with the design of the study outlined in Figure 1. Objective 1 implies that outcomes of each group (standard and intensive) will be compared with the outcomes of a third waiting-list control group, but the schedule presented on Table 2 (p. 15) will not permit such comparisons. Figure 1 and the description of the procedure make it clear that there are only two groups in the study, and that the participants in the intensive GIST arm will be on a waiting list. However, the assessment schedule makes provision for a single baseline assessment, which is completed at the point of recruitment (p. 10). There will be no data on the performance of the waiting list group after the nine-month waiting period, and just before the intensive intervention begins. Thus, it will only be possible to compare the two groups at baseline, post-intervention, and at the two follow-up points (three and six months). The hypothesis that “there will be a greater reduction in social communication difficulties following GIST (both protocols) compared with the WL group” cannot be investigated with the described design. The waiting list condition seems to emerge for practical, rather than methodological reasons, and the description of the protocol should acknowledge this.

	There needs to be more clarity about how adequate stratification of the two groups will be ensured. What is considered “chronic phase”, and how will the authors ensure that the level of chronicity is comparable between the two groups? Other variables that could potentially introduce bias if they turn out not to be comparable between the groups are the time since injury at first admission into rehabilitation, the time since first rehabilitation, and the length of stay and areas of intervention at first rehabilitation. How will the authors ensure the two groups will not vastly differ in these areas? Why is there a need for an exclusion criterion around “unfitness for evaluation of outcome”? Would this not be addressed by criterion (iii) “cognitive , sensory, physical or language impairment affecting the capacity to complete the intervention”? Would the inclusion criterion of having “adequate Norwegian language proficiency to participate in the study” not ensure that individuals “lacking fluency in Norwegian prior to the ABI” would be naturally excluded? Caution is required in adding exclusion criteria that are too easily applied retrospectively. In addition, inclusion criterion (iv) “one close family member or close friend able to participate in the intervention” is concerning, as it can potentially preclude the inclusion of individuals with social communication difficulties who are more isolated, and thus would likely to need this treatment the most. How much information will participants receive about the two conditions prior to enrolment and randomisation? Could the fact that one condition requires an inpatient stay affect enrolment and adherence to that arm of the study? Could participants have a preference for one of the conditions, which could therefore affect their motivation to complete the study? On p. 14, it is stated that “The videos will be further analysed by a trained research assistant who is blinded to group allocation”. On the same page, it is also stated that “video recording [in some cases] is conducted by the participant and close family member or friend themselves at home”. Could this not potentially reveal the condition in which the participant has been enrolled? What precautions will be taken to prevent this? More detail is required about the differences between the two interventions. It is implied that the two programmes are exactly the same in contents and total contact hours, but that in one condition (standard) that contact occurs over 12 weeks, and in the other condition (intensive), it occurs over four weeks. However, from the descriptions provided in the checklist (pp. 34-35), this no longer seems to be the case. The standard group is said to receive 2.5 hours of treatment for 12 weeks (a total of 30 hours?), and the other group receives six hours of treatment a day, three days a week for four weeks (72 hours?). Please clarify how many hours of treatment in total will each group receive, and how are the 13 sessions of the programme distributed over time. The power calculations are not in line with the aims of the study. On page 9 the authors state “intensive GIST will be equally effective”, yet the power calculations are planned to uncover “moderate effects” on the LCQ. In addition, it is unclear whether the power calculations provided are for simple comparisons, or for the factorial design that will be employed (p. 16, ln. 52-60).

	In sum, this protocol needs clarification in some areas, including a clearer description of the recruitment process, more detail about the information provided to participants prior to recruitment, more detail about the contents and duration of the intervention, and a clearer description of the procedure across the two groups. The objectives of the study need to be revised to match the described methodology. In its current form, the study will only enable comparison between standard and intensive GIST, and the hypothesis outlined by the researchers is that there will be no differences between the two (p. 9). Power calculations and planned sample size will need to be adjusted to reflect this prediction. In order to compare both groups with a "waiting list control", the protocol would need to be modified to introduce a second baseline assessment at the end of the nine-month waiting period, just before the beginning of the intensive intervention. The authors also need to consider and discuss the potential bias introduced by having only participants in one of the arms of the study on a nine-month waiting list. This could affect adherence and drop-out rates more in one group than the other. These variables, and other missing data, must be included in the data analysis plan. Minor amendments p. 3, ln. 10 – Add comma after "isolated". p. 3., ln. 16-17 – Revise the aims of the study to reflect the type of paper (i. e. protocol)] p. 3, ln. 31 – Replace "subjects" with "participants". pp. 3-4 – Move the sentence about ethical approval to appear before the sentence beginning with "The trial will be conducted…". p. 5, ln. 6 – Delete "or hold back". p. 5, ln. 10 – Remove square brackets and replace with dash to facilitate reading, here and throughout the whole manuscript: (e. g., theory of mind – TOM). p. 5, ln. 38 – Comment on the evidence, or lack thereof, about persistence of social communication deficits in right hemisphere stroke. p. 8, ln. 38 – Replace "sic" with "six". p. 10, ln. 56 – Replace "3" with "three". p. 11, ln. 31 – If participants have not yet been randomised to a condition, strictly speaking the assessment is not blinded. p. 11, ln. 35 – Replace "60" with "Sixty". p. 13, ln. 57 – Please comment on the reliability and validity of the LCQ in patients with stroke. p. 15, ln. 17 – Please clarify to which raters the "sound-recorded to ensure interrater reliability" refers to. p. 18, ln. 36 – Replace "may offer the opportunity to access" with "may facilitate access to treatment".

REVIEWER	Natalie Gilmore, MS, CCC-SLP, PhD Candidate Boston University, USA
REVIEW RETURNED	28-Mar-2019

GENERAL COMMENTS	Major comments Abstract • Aims as stated on pg. 9, lines 9 – 18 are much clearer than when described in the abstract, perhaps rephrase pg. 3, lines 21 – 31 to match aims as stated on pg 9 Introduction • pg. 5, line 50 – It may be useful to specify that this is the most recent Cicerone review to distinguish from the 2000 and 2005 reviews. Of note, both prior Cicerone reviews (2000 & 2005) suggested pragmatic treatment for individuals with TBI as practice standard like the 2011 review, further strengthening the aim of this study. • pg. 5, line 51 – It may valuable to define class 1 and 1a studies, so the reader has a sense of what strength of evidence is available for social communication difficulty in this review. • Pg. 9, line 23 – It is not clear in this section why the authors believe that intensive GIST will be equally effective as standard GIST in terms of reduction of social communication difficulties and close family members and friends ratings. This section would be strengthened by providing a clear rationale for administering both standard and intensive GIST. Methods and Analysis • pg. 9, line 55 – This section would benefit from inclusion of a short rationale for use of the TiDier and SPIRIT checklist and guide. • pg. 10, line 19 – Please provide the end date for the study here as specified later on pg. 18, line 17 • pg. 11, line 15 – Please clarify this term “unfitness for evaluation of outcome” and how it will be determined. • pg. 15, line 10 – Please provide a rationale for why the assessor of ‘Mind in the Eyes’ will not be blinded to treatment allocation and how that limitation will be managed. • pg. 15, line 3 – It is not clear if both Trail Making Test Parts A and B be administered. Please clarify. • Page 18, line 8 – This section would benefit from definition and importance of eta – squared. Minor comments • Pg. 3, line 12 – consider adding commas – > holistic, multidisciplinary group treatment • Pg. 3, line 45 – consider adding commas – > international, peer – reviewed journals • Pg. 6, line 15 – rephrase follow – ups to follow – up assessments • Pg. 6, line 45 – no comma need after least effective • Pg. 7, line 40 – rephrase named to called • Pg. 7, line 42 – consider adding commas – > holistic, multidisciplinary group treatment • Pg. 7, line 5 – consider changing a to the • Pg. 8, line – no comma needed after weeks • Pg. 8, line 38 – correct spelling error (sic for six) • Pg. 9, line 55 – grammar – perhaps just delete “Based on” • Pg. 10, line 35 – correct traumatic injuries (TBI) to traumatic brain injuries (TBI) and cerebrovascular injuries to cerebrovascular accidents (CVA)

	• Pg. 13 and 14 – change present tense to future tense on these two pages as there is some inconsistency in this section of the document • Pg. 15, line 17 – change sound – recorded to audio – recorded • Pg. 15, line 27 – change WASI to WAIS • Pg. 18, line 33 – add the word assessment after follow – up • Pg. 20, line 12 – check reference 4 formatting – Translatin should be Translating. There is also an extra period and odd spacing in this reference • Pg. 21, line 23 – check reference 23 formatting – compentanse should be competence • Pg. 22, line 47 – check reference 49 formatting – all caps should be corrected

REVIEWER	Nicholas Behn City, University of London London, UK
REVIEW RETURNED	28-Mar-2019

GENERAL COMMENTS	Thank-you very much for allowing me to review this paper on a social communication intervention for people with ABI. It is really interesting to hear about more research happening in this field and the component of comparing standard with intensive treatment will be an important aspect of this study. I have provided comments to make this work stronger. I realise as this trial is underway there are particular things that can't be changed so further clarification and explanation may be needed in places throughout. Abstract Methods & Analysis: You mention that participants are seen four times, but is the WL group seen five times? (separated by how many weeks?) Article summary P3L19: Strong claims that I feel should be tempered. Especially as you go on to say that these aspects are unparalleled and address major limitations. McDonald et al (2008) had a sample size of 51, Dahlberg et al (2007) 52 participants, and Harrison-Felix 179(!) participants and they also had adequate outcome measures with good follow-up in the latter study. You may want to review this. P3L31: There are considerable issues and challenges with waitlist control designs e.g. participants in WL group may score themselves lower in an effort to get intervention earlier, or may seek out the intervention during the WL control period – I think these should be acknowledged and balanced. P3L38: This comes up later but the "active" involvement of close family member or friend is not clear and needs to be more clearly defined. Did they attend all the sessions as I would assume "active" involvement is more akin to Togher et al (2013) study on communication partner training. Introduction P4L13: Social communication disorders are seen as part of CCD as are pragmatic language disorders. I think this sentence needs to be better framed. P5L3: Omit "interesting" P5L3: I was intrigued why most of this paragraph was spent describing a SCED – I couldn't see the point? In fact, less attention was given to the RCT which was stronger methodologically. I realise the SCED was published after the review but I think this paragraph needs to be reworked.

	P5L38: year of publication P5L42: what is meant by “standardised” social skills training – do you mean manualised? P6L15: “A few treatment studies” – how many are a few – what are these? Is the Westerhof-Evans the strongest of these? What was the treatment and what were the treatment aims? I wonder whether the Finch et al (2016) review and Togher et al (2014) INCOG recommendations are useful here. P6L31: You use the word “larger sample size” which you use to refer to the Westerhof-Evans paper but then you go on to have a smaller sample size at 60. I feel there needs to be some tempering here as the sample sizes are comparable with other RCT’s conducted in this field. P6L42: The references should only be those that refer to the studies that used GIST. P6L56: If this paper is not published, it shouldn’t be referenced. Your reader has no ability to access this if they would like (at least at this stage) P7L24: Improvement of social competence skills based on? P7L38: spelling error Methods & Analysis Study design: I would consider inclusion of a CONSORT diagram. Some of the content of Table 1 is already in the prose of the document and does not need to be duplicated e.g. outcomes Participants: are CVA the only non-traumatic ABI? If so, how are you addressing potential for aphasia? Can this occur alongside the social communication problems? P9L47: You are accepting people a minimum of 12 months post injury. What is your upper limit? Time post-injury and patient location (i.e. at home, or still in rehab) could have an effect on the outcome of the intervention. This needs to be clarified. P9L47: I’m not sure how you will assess the presence of social communication deficit, based on self-reports. Can you not have a practicing SLT make the diagnosis? How do you define moderate-to-sever difficulties on the LCQ? Someone could rate themselves as “usually or often” on an item but this could be largely representative of what they were like pre-injury (according to family report). P9L49: I would say “willing to participate” rather than motivated. P9L53: “participate” what does this mean – you want family members to attend all sessions? P9L58: How do you define minimum level on the AQ. You need to be specific here. P10L10: how do you define the impairments affecting capacity to complete the intervention as most people post-ABI have these issues. More specifics needed. P10L15: “unfitness” is not an appropriate word – what is meant by this? Procedures P10L21: How will the forms be sent to participants? E.g. mail, e-mail, in-person? As you have different recruitment procedures (through a hospital and SLPs) is the procedure for consent different? How will you assess a persons capacity to consent into the study? Overall, more information is needed on how you will obtain informed consent from participants (including those who may have some aphasia from their stroke) Randomisation: are you stratifying the randomisation on any factors? My concern is that you won’t have a matched group on some key demographics e.g. severity, time post-injury etc. Please consider.

	P12L19: is the group size based on any other publications or your pilot work? Measures I find it interesting that you have chosen the LCQ as your primary outcome measure given that the SCSQ-A had additional items that aligned most with GIST, and none of the published studies (Dahlberg 2007; Braden 2010; Harrison-Felix 2018) used the LCQ as a primary outcome measure but rather a secondary. I think the use of the LCQ as your primary needs to be more clearly justified. P13L13: Are you going to assess the feasibility of getting video recordings? I suspect you may encounter problems here. This calls into question the involvement of the family member in the treatment if they can't attend to have this done. P14L10: Why did you choose the mind in the eyes test? This test is being recommended less as a social cognition assessment. I am also intrigued to understand how you think GIST will specifically make a change on this measure? You could assess the feasibility of doing all these measures as this large assessment battery (10 questionnaires and a video recorded conversation) will entail considerable participant burden. I would perhaps put the outcomes as primary, secondary and exploratory. Blinding – are all questionnaires completed by a blinded assessor? How will you ensure the person remains blinded. I would recommend you record instances of unblinding or near misses. Intervention I noticed that the standard GIST is in an outpatient setting and intensive is an inpatient setting. I would therefore assume that the intensive participants are earlier post-injury than those in the standard GIST. As a result factors such as time post-injury and location (setting of treatment) could have an effect on outcome. I feel you need to be very cautious in your interpretation of the end results. How will you prospectively measure fidelity? – I note this is not in your TIDIER checklist. Sample size What outcome was the power analysis based on? I assume the LCQ as this is your primary outcome but this should be clarified. Data analysis How will you deal with attrition – what amount of data do you expect to lose to follow-up? How will you deal with missing data? Do you intend to do per protocol analyses? Any analyses that compared standard with intensive GIST need to be controlled for (for reasons stated above). My concern is how misleading the results could be given location of treatment and time post-injury. I would want you to compare baseline differences before you start comparing intensive with standard GIST. What is the purpose of this study in the long-term – how would this study inform future research work? Having a discussion and/or conclusions section would be recommended. TIDIER checklist (see paper on use of TIDIER by Isaksen et al 2018 in stroke and aphasia – it will provide some clarification of some of the items and detail needed) 2. This item should be about the "essential elements" of the intervention. Your description does not adequately address this item – is it group work, family involvement 4. More information is needed about the procedures and activities 11. You need to mention what the planned fidelity procedures are.

REVIEWER	Rachael Rietdijk The University of Sydney Australia
REVIEW RETURNED	29-Mar-2019

GENERAL COMMENTS	Thank you for the opportunity to review this interesting study protocol. This protocol reports on an innovative study comparing standard social communication skills training with intensive (inpatient-based) training for people with acquired brain injuries. Inpatient-based social communication skills training is not commonly used within the field of ABI, and so this study will represent an important contribution to the literature. I will be interested to read about the outcomes of this study when they are available. I have made the following comments related to improving the clarity of this protocol: COMMENTS: Title: The key focus of your study (intensive versus standard treatment) is not clear from the title of this paper. I would recommend adding the word "intensive" to the title. Abstract: In the introduction section, the innovation of your study (comparing intensive with standard treatment) is not clear. It is stated that GIST has received empirical support, so from reading this section alone, it is not clear how this study will add to the evidence base that already exists around GIST. In the methods and analysis section, it is stated that "The WL group will receive intensive GIST". Could you clarify that the WL group will receive the intensive GIST after a waiting period? The first mention of "intensive GIST (inpatient)" here confused me, as it sounded like this was a different group of inpatients (separate from the 60 adults recruited and randomised). I would suggest moving the section where you define the differences between standard GIST and intensive GIST earlier in this section. Also, rather than stating that standard GIST will be delivered to outpatients, and intensive GIST will be delivered to inpatients (which sounds like you have recruited participants from two different contexts), it would be clearer to state, "Standard GIST will be delivered to participants in 2.5 hour outpatient sessions..." and "For intensive GIST, participants will be admitted as inpatients for four weeks and receive three full-day sessions each week." Introduction: Page 4 line 5 - what do you mean by "hold back information"? Does this mean the same thing as the next part, i.e. "avoid saying inappropriate things"? Page 6 line 38 - I was expecting to see in this paragraph some more detailed information about the previous studies using GIST to back up the claim that this is "one of the best-validated group interventions", but I note that this content follows in the next paragraph. Restructuring these two paragraphs may assist in developing your argument, starting from the broad evidence base for GIST, and narrowing to the reasons for trialling the more intensive version.

	Methods and Analysis: Page 8 Line 46 - The figure 1 provided relates to randomisation of participants, rather than showing the assessment time points. Is there another figure that should be referenced here? Page 9 line 45 - "No concomitant diseases" sounds very broad - would this mean, for example, that a person with diabetes would be excluded? Would it be more correct to state "no concomitant neurological diseases"? Page 9 line 56 - Could you clarify this criteria relating to the LCQ? Were you basing this on the self-report or other-report version? By "aspects of communication", do you mean three questionnaire items? How were moderate to severe difficulties defined (e.g. problem happening either often, or usually/always - rating of 3/4)? Page 10 line 15 - What is meant by "unfitness for evaluation of outcome"? Table 1 - I am unclear on the timing of the T2 assessment for the WL/intensive GIST group? Do they complete this at the end of the waiting period, or after their intensive GIST treatment? It is unclear from this table. Page 12 line 45 - I assume the LCQ was translated into Norwegian for the purposes of this study - if so, this should be mentioned. Was any support provided to people with ABI who may have difficulty with comprehending the items? This could similarly be reported for the SCQ-A. Page 14 line 12 - It is not really correct to state that these outcome assessments are blinded, as the participants and family members are aware whether intervention has been received or not. Page 14 line 17 - How does sound recording the assessments ensure inter-rater reliability? Will a different assessor listen to the sound recordings to score some assessments to ensure the original assessor noted the scores correctly? Page 15 line 15 - How are the follow-up treatment sessions scheduled in relation to the follow-up assessment sessions i.e. do the participants complete their three month follow-up session and then their T3 assessment? If so, the "booster" effect of the extra training session should be considered in evaluating long-term effects. Page 15 - What are the arrangements if a participants misses a session due to illness etc? Page 15 line 31 - The number of hours of training the intensive GIST participants receive is unclear. If the standard GIST participants are receiving 12 x 2.5 hr sessions = 25 hours, how many hours are the intensive GIST participants receiving? It sounds like the intensive GIST intervention differs from the standard GIST in both intensity (i.e. how concentrated the intervention is) and dose (i.e. how many total hours of intervention are received). This should be clarified, and will be an important issue to consider in interpreting your findings.

	Page 16 - As per previous comments, it is difficult to follow the timing of the assessments and how these relate to the interventions being provided. I assume that the comparison between groups at T2 will provide information about standard GIST vs waitlist. Which data will be relevant to comparing the standard GIST vs intensive GIST, and intensive GIST vs waitlist? How will analyses of data from T3 and T4 be informative about your research questions? Further note - it would be relevant to add somewhere some comments about the potential limitations of this research. For example, I wonder whether participants will find the intensive GIST training acceptable, or whether the expense of intensive GIST training may make this delivery format difficult to implement in practice. Additional Files: Tidier checklist - The information provided in the table does not always match up with the checklist item. For "Why?", I expected to read about the rationale/theory that is behind the GIST intervention, rather than why this study is being conducted. For "Materials", are participants provided with any physical materials? For "Planned", it would be relevant to provide some information about how you plan to measure intervention fidelity e.g. attendance record, record of completion of home practice, data collection record etc.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Comments to the authors
This paper proposes a valuable addition to the limited existing evidence around interventions for those with social communication difficulties following acquired brain injury. Understanding the differences between standard outpatient and intensive inpatient protocols could have important implications for practice and policy.
Thank you for all the helpful feedback on our protocol article. Questions and comments are answered below. The minor comments are changed in the manuscript and marked in the document “main document- marked copy”, and the major comments are addressed here.

1. It is clear that there was careful consideration of the design which, at first glance, seems appropriate to address the research question. However, there are a few points of concern that require clarification. The specific objectives stated on page 9 are not fully consistent with the design of the study outlined in Figure 1. Objective 1 implies that outcomes of each group (standard and intensive) will be compared with the outcomes of a third waiting-list control group, but the schedule presented on Table 2 (p. 15) will not permit such comparisons. Figure 1 and the description of the procedure make it clear that there are only two groups in the study, and that the participants in the intensive GIST arm will be on a waiting list. However, the assessment schedule makes provision for a single baseline assessment, which is completed at the point of recruitment (p. 10). There will be no data on the performance of the waiting list group after the nine-month waiting period, and just before the intensive intervention begins. Thus, it will only be possible to compare the two groups at baseline, post-intervention, and at the two follow-up points (three and six months). The hypothesis that “there will be a greater reduction in social communication difficulties following GIST (both protocols) compared with the WL group” cannot be investigated with the described design. The waiting list condition seems to

emerge for practical, rather than methodological reasons, and the description of the protocol should acknowledge this.

Thank you for valuable feedback. We hope that by adding some clarification that you see how we are planning to compare the groups. The baseline for intensive GIST will be test point 4 (T4). See clarifications in the test points and WL group on page 8, lines 14-18 and page 11, lines 14-16 and in Figure 1 and Table 1.

2. There needs to be more clarity about how adequate stratification of the two groups will be ensured. What is considered “chronic phase”, and how will the authors ensure that the level of chronicity is comparable between the two groups? Other variables that could potentially introduce bias if they turn out not to be comparable between the groups are the time since injury at first admission into rehabilitation, the time since first rehabilitation, and the length of stay and areas of intervention at first rehabilitation. How will the authors ensure the two groups will not vastly differ in these areas?

Thank you for addressing this. In this study a simple randomisation is performed, without stratification. We have decided not to use stratification since all the participants included in the study are in the chronic phase with a minimum of 12 months since the ABI. Hence, no cognitive changes are expected on a group level. Importantly, most of the participants will be recruited from the same hospital (same section/department). As described in the manuscript, participants are randomized to either standard GIST or WL/intensive GIST in a 1:1 ratio. However, data regarding prior rehabilitation is collected during the screening interview and inclusion assessment.

3. Why is there a need for an exclusion criterion around “unfitness for evaluation of outcome”? Would this not be addressed by criterion (iii) “cognitive , sensory, physical or language impairment affecting the capacity to complete the intervention”? Would the inclusion criterion of having “adequate Norwegian language proficiency to participate in the study” not ensure that individuals “lacking fluency in Norwegian prior to the ABI” would be naturally excluded? Caution is required in adding exclusion criteria that are too easily applied retrospectively.

Thank you for pointing this out. We agree, and have as such deleted these from the exclusion criteria.

4. In addition, inclusion criterion (iv) “one close family member or close friend able to participate in the intervention” is concerning, as it can potentially preclude the inclusion of individuals with social communication difficulties who are more isolated, and thus would likely to need this treatment the most.

Yes, this is a concern. However, since the discussion of the home assignments has a central role in GIST we have decided that it is necessary to have a support person to discuss the home assignments with. Nevertheless, the participants select their support person themselves and are as such free to choose a neighbour or a friend (if the support persons consent to this).

5. How much information will participants receive about the two conditions prior to enrolment and randomisation? Could the fact that one condition requires an inpatient stay affect enrolment and adherence to that arm of the study? Could participants have a preference for one of the conditions, which could therefore affect their motivation to complete the study?

Yes, all participants receive detailed information about the two treatment conditions prior to the enrolment. Some of the participants express a desire to one of the treatments. However, the assessor clarifies that group allocation is decided through randomisation. This issue has been clarified on page 10, line 12-20.

6. On p. 14, it is stated that “The videos will be further analysed by a trained research assistant who is blinded to group allocation”. On the same page, it is also stated that “video recording [in some cases] is conducted by the participant and close family member or friend themselves at home”. Could this not potentially reveal the condition in which the participant has been enrolled? What precautions will be taken to prevent this?

Video recordings at home can be used when the close family member does not have the opportunity to join the assessment point. Since the close family member does not participate in the treatment session, the video recordings will be random to group allocation. There will be video recordings

conducted at home in both groups (WL/intensive GIST and in standard GIST). This is clarified on page 12 line 5-10.

7. More detail is required about the differences between the two interventions. It is implied that the two programmes are exactly the same in contents and total contact hours, but that in one condition (standard) that contact occurs over 12 weeks, and in the other condition (intensive), it occurs over four weeks. However, from the descriptions provided in the checklist (pp. 34-35), this no longer seems to be the case. The standard group is said to receive 2.5 hours of treatment for 12 weeks (a total of 30 hours?), and the other group receives six hours of treatment a day, three days a week for four weeks (72 hours?). Please clarify how many hours of treatment in total will each group receive, and how are the 13 sessions of the programme distributed over time.

Thank you for these comments. We have specified the amount of treatment hours in Figure 1, and elaborated on the differences between the two treatments on page 14, section 1-3.

8. The power calculations are not in line with the aims of the study. On page 9 the authors state "intensive GIST will be equally effective", yet the power calculations are planned to uncover "moderate effects" on the LCQ. In addition, it is unclear whether the power calculations provided are for simple comparisons, or for the factorial design that will be employed (p. 16, ln. 52-60).

Thank you for pointing this out to us. We have clarified this section on page 15 (line 16-18) by adding that we expect moderate effects in both treatment groups independent of group allocation.

9. In sum, this protocol needs clarification in some areas, including a clearer description of the recruitment process, more detail about the information provided to participants prior to recruitment, more detail about the contents and duration of the intervention, and a clearer description of the procedure across the two groups.

Yes, we agree on this. A clarification of the recruitment procedure is added on page 9 line 15-20 and 10 line 12-20. The patient journal at SRH will be screened for patients who have been to rehabilitation in the period 2014–2017 as a result of either TBI or non-traumatic brain injuries and who have documented cognitive deficits, such as difficulties with memory, attention and/or executive function. Participants will also be recruited from advertising in user organisations and through local speech and language pathologists (SLP). Regarding the content and duration of the intervention see comment number 7. A clearer description of the procedure is explained in figure 1.

10. The objectives of the study need to be revised to match the described methodology. In its current form, the study will only enable comparison between standard and intensive GIST, and the hypothesis outlined by the researchers is that there will be no differences between the two (p. 9). Power calculations and planned sample size will need to be adjusted to reflect this prediction. In order to compare both groups with a "waiting list control", the protocol would need to be modified to introduce a second baseline assessment at the end of the nine-month waiting period, just before the beginning of the intensive intervention.

Thank you for pointing out the need for clarification. By clarifying all the test points and procedure for the two groups we hope that we have showed how we are going to compare the treatments with the WL group. Assessment point 4 (T4) for the WL group will also be the baseline and pre-test before intensive GIST. See Figure 1.

11. The authors also need to consider and discuss the potential bias introduced by having only participants in one of the arms of the study on a nine-month waiting list. This could affect adherence and drop-out rates more in one group than the other. These variables, and other missing data, must be included in the data analysis plan.

Yes, this is an important issue. We have addressed this on page 18 section 1, and on page 15 line 21-25.

12. Minor amendments

The minor amendments have been addressed in the text and marked in the document "main document- marked copy".

Reviewer: 2
Thank you for all the helpful feedback on our protocol article. Questions and comments are answered below. The minor comments are changed in the manuscript and marked in the document "main document- marked copy", and the major comments are addressed here.

Major comments
Abstract
Aims as stated on pg. 9, lines 9 – 18 are much clearer than when described in the abstract, perhaps rephrase pg. 3, lines 21 – 31 to match aims as stated on pg 9
Thank you for your comments, the main aim in the abstract is now clarified. The aim of this study is to determine the efficacy of two different GIST protocols, standard GIST and intensive GIST, with efficacy of both conditions also being compared to results of subjects in a waiting-list control group (WL). See page 7 line 23-24 and page 8 line 1-5.
Introduction
- pg. 5, line 50 – It may be useful to specify that this is the most recent Cicerone review to distinguish from the 2000 and 2005 reviews. Of note, both prior Cicerone reviews (2000& 2005) suggested pragmatic treatment for individuals with TBI as practice standard like the 2011 review, further strengthening the aim of this study.

Yes, thank you for pointing this out. This has now been specified on page 4 line 17.
- pg. 5, line 51 – It may valuable to define class 1 and 1a studies, so the reader has a sense of what strength of evidence is available for social communication difficulty in this review.

Yes, I agree. A definition on class 1 and 1a studies has now been added to page 4 page 17-19.
- Pg. 9, line 23 – It is not clear in this section why the authors believe that intensive GIST will be equally effective as standard GIST in terms of reduction of social communication difficulties and close family members and friends ratings. This section would be strengthened by providing a clear rationale for administering both standard and intensive GIST.

We have added a clarification and strengthened our rationale for the use of both intensive GIST and standard GIST. Based in previous research (e.g. Harrison-Felix et al., 2018, Dahlberg et al., 2007) on page 6 section 3 we have argued that there is still a need for more research determining the best format for GIST. Rationale for the intensive GIST is also expressed on page 7 Line 2-7.
Methods and Analysis
- pg. 9, line 55 – This section would benefit from inclusion of a short rationale for use of the TiDier and SPIRIT checklist and guide.

We have chosen to remove the TiDier checklist, and the SPIRIT checklist is used as recommended by BMJ open, as the protocol closely mirrors the CONSORT statement and ethical considerations of the trial. A short rationale is added to page 8 line 21-24.
- pg. 10, line 19 – Please provide the end date for the study here as specified later on pg. 18, line 17

Yes, we have highlighted the anticipated end date of the recruitment that is planned to be completed by autumn 2019. See page 9 line 10-11.
- pg. 11, line 15 – Please clarify this term "unfitness for evaluation of outcome" and how it will be determined.

Thank you for pointing this out. We have decided to omit this criterion as the criterion (iii) "cognitive, sensory, physical or language impairment affecting the capacity to complete the intervention" is addressing the same issue.
- pg. 15, line 10 – Please provide a rationale for why the assessor of 'Mind in the Eyes' will not be blinded to treatment allocation and how that limitation will be managed.

Yes, thank you for pointing this out. We have added some clarifications addressing this issue on page 13 line 4-8.
- pg. 15, line 3 – It is not clear if both Trail Making Test Parts A and B be administered. Please clarify.

We use TMT from D-KEFS. This has now been clarified in Table 1.

- Page 18, line 8 – This section would benefit from definition and importance of eta squared.

A clarification is added to page 16 on line 12-13.

Minor comments

Most of the minor comments have been addressed in the text and marked in the document "main document- marked copy". Subtests from both Wechsler Abbreviated Scale of Intelligence (WASI) and Wechsler Adult Intelligence Scale IV (WAIS-IV) are included. See Table 1.

Reviewer: 3

Thank-you very much for allowing me to review this paper on a social communication intervention for people with ABI. It is really interesting to hear about more research happening in this field and the component of comparing standard with intensive treatment will be an important aspect of this study. I have provided comments to make this work stronger. I realise as this trial is underway there are particular things that can't be changed so further clarification and explanation may be needed in places throughout.

Thank you for all the helpful feedback on our protocol article. Questions and comments are answered below. The minor comments are changed in the manuscript and marked in the document "main document- marked copy", and the major comments are addressed here.

Abstract

Methods & Analysis: You mention that participants are seen four times, but is the WL group seen five times? (separated by how many weeks?)

Thank you for this comment. We have now clarified this in the Abstract and in the article. All the assessment points for the WL/intensive group have now been described on page 2 line 13-15, and on page 8 line 14-20 and in figure 1.

Article summary

P3L19: Strong claims that I feel should be tempered. Especially as you go on to say that these aspects are unparalleled and address major limitations. McDonald et al (2008) had a sample size of 51, Dahlberg et al (2007) 52 participants, and Harrison-Felix 179(!) participants and they also had adequate outcome measures with good follow-up in the latter study. You may want to review this.

Yes, thank you for commenting on this. We agree, and have reviewed and modified the statement in the summary on page 3 line 5- 18.

P3L31: There are considerable issues and challenges with waitlist control designs e.g. participants in WL group may score themselves lower in an effort to get intervention earlier, or may seek out the intervention during the WL control period – I think these should be acknowledged and balanced.

Thank you for pointing this out, we agree on this. We have added the limitation of the use of WL group on page 3, as well as addressed this issue in the discussion at the end of the paper. on page 18 line 20-24.

P3L38: This comes up later but the "active" involvement of close family member or friend is not clear and needs to be more clearly defined. Did they attend all the sessions as I would assume "active" involvement is more akin to Togher et al (2013) study on communication partner training.

Yes, we agree. We have clarified that the active involvement is related to home assignments specifically. See page 3 line 13, page 6 line 13-15, page 9 line 25-26.

Introduction

P4L13: Social communication disorders are seen as part of CCD as are pragmatic language disorders. I think this sentence needs to be better framed.

This sentence has been reframed to Social communication deficits following ABI are often referred to as social communication difficulties (SCD), cognitive communication disorders8 or pragmatic language disorders. on page 4 line 1-3.

P5L3: I was intrigued why most of this paragraph was spent describing a SCED – I couldn't see the point? In fact, less attention was given to the RCT which was stronger methodologically. I realise the SCED was published after the review but I think this paragraph needs to be reworked.
Thank you for pointing this out and we totally agree with you on this. We have shortened the focus on the SCED, and added more information about the outcome measures from the RCT. See page 4 line 21-25 and page 5 line 1-7.

P5L42: what is meant by "standardised" social skills training – do you mean manualised?
Standardized social skills training is the term Togher er al (2014)., used, and is not specifically defined in their article. But as I interpret this, they use the term standardized social skills training about standardized or manualized training that not involve training of the family members or communication partners. I have now tried to make this clearer on page 5 line 13-19.

P6L15: "A few treatment studies" – how many are a few – what are these? Is the Westerhof-Evans the strongest of these? What was the treatment and what were the treatment aims? I wonder whether the Finch et al (2016) review and Togher et al (2014) INCOG recommendations are useful here.
Finch et al (2016) and Togher et al (2014) only included studies through 2013, and to our knowledge, this is the most recent systematic review describing SCD interventions. We did a search in Ovid MEDLINE in April 2017 which showed only 4 SCD intervention studies since 2013, in the field of ABI. This has been clarified on page 6 line 1-2.

P6L31: You use the word "larger sample size" which you use to refer to the Westerhof-Evans paper but then you go on to have a smaller sample size at 60. I feel there needs to be some tempering here as the sample sizes are comparable with other RCT's conducted in this field.
Yes, we agree and we have deleted this statement from the paper.

P6L56: If this paper is not published, it shouldn't be referenced. Your reader has no ability to access this if they would like (at least at this stage)
According to the BMJ Opens guidelines we have now changed this reference (and removed it from the reference list). The unpublished data is now cited in parentheses in the text with the names of the authors with the year of the submission. See changes on page 7 line 1 and page 14 line 23.

P7L24: Improvement of social competence skills based on?
We have now tried to clarify this sentence on page 6 line 19-22.
Methods & Analysis
Study design: I would consider inclusion of a CONSORT diagram. Some of the content of Table 1 is already in the prose of the document and does not need to be duplicated e.g. outcomes
Thank you for this suggestion. The study design has been clarified by the use of a CONSORT diagram in Figure 1.

Participants: are CVA the only non-traumatic ABI? If so, how are you addressing potential for aphasia? Can this occur alongside the social communication problems?
The non-traumatic ABI that is included covers also brain tumours and anoxic incidents. This has now been clarified on page 9 line 16-18. In some cases social communication difficulties and aphasia do occur together. However, aphasia is an exclusion criterion as we want to focus solely on the social communication difficulties. Aphasia is assessed by the local speech and language therapist, or during the WASI (Vocabulary test) conducted before randomisation. Clarified on p. 10 line 9.

P9L47: You are accepting people a minimum of 12 months post injury. What is your upper limit? Time post-injury and patient location (i.e. at home, or still in rehab) could have an effect on the outcome of the intervention. This needs to be clarified.
We do not have an upper limit regarding post injury. In Norway these persons have a limited rehabilitation offer, and a lot of people live with their social communication difficulties undetected for many years before treatment. All the participants included in the present study are in the chronic phase with a minimum of 12 months post ABI. Changes are not expected on a group level. However, data regarding prior rehabilitation is collected during the demographic interview conducted during the screening and inclusion assessment. Clarified on page 9 line 22.

P9L47: I'm not sure how you will assess the presence of social communication deficit, based on self-reports. Can you not have a practicing SLT make the diagnosis? How do you define moderate-to-sever difficulties on the LCQ? Someone could rate themselves as "usually or often" on an item but this could be largely representative of what they were like pre-injury (according to family report).
A clarification of the inclusion criteria (vi) has been added on page 10. A general issue regarding assessing social communication difficulties in Norway is the lack of standardised assessment tools. However, the demographic interview is conducted by an SLT. We have chosen LCQ in this assessment, in addition to an interview and a screening by an SLP with experience in assessing social communication difficulties. The LCQ form also include a "change" item which the participants and close family member score if the communication skills have changed (+= happens more, 0= no change, - happens less) since the injury.

P9L53: "participate" what does this mean – you want family members to attend all sessions?
This has now been clarified in the inclusion criteria on page 9 line 25.

P9L58: How do you define minimum level on the AQ. You need to be specific here.
The total score is estimated between 17-85, and a score of >20 in the discrepancy score has been suggested as indicating clinically significant impairment of self-awareness (Evans et al., 2005). This is clarified on page 10 line 4-5.

P10L10: how do you define the impairments affecting capacity to complete the intervention as most people post-ABI have these issues. More specifics needed.
Yes, we agree, specifics are added in the inclusion criteria on page 10 section 1.
Procedures
P10L21: How will the forms be sent to participants? E.g. mail, e-mail, in-person? As you have different recruitment procedures (through a hospital and SLPs) is the procedure for consent different? How will you assess a person's capacity to consent into the study? Overall, more information is needed on how you will obtain informed consent from participants (including those who may have some aphasia from their stroke)
Yes, we use different recruitment procedures. However, all participants recruited from medical records, advertising or local SLPs receive an information letter about the study in the mail. The participants is also screened during a telephone interview. Persons that after this interview still is interested in participating in the study is summoned for the first assessment meeting together with their close family member or friend. In this screening and inclusion meeting the participants and their close family member or friend both receive detailed information about the two treatment conditions prior to the enrolment so that informed consent is ensured. This has been clarified under procedures p. 10 line 11-20.

Randomisation: are you stratifying the randomisation on any factors? My concern is that you won't have a matched group on some key demographics e.g. severity, time post-injury etc. Please consider.
Thank you for addressing this. A simple randomisation without stratification is conducted with regard to feasibility in a small population. All the participants included in the study are in the chronic phase

with a minimum of 12 months since the ABI. All participants included in the study are in the chronic phase with a minimum of 12 months post ABI and changes are not expected on a group level. However, data regarding severity or prior rehabilitation is collected during the demographic interview conducted during the screening and inclusion assessment.

P12L19: is the group size based on any other publications or your pilot work?
The group size is based on the recommendations of the developers of GIST (Hawley and Newman, 2012). This is now specified on page 11 line 9.
Measures
I find it interesting that you have chosen the LCQ as your primary outcome measure given that the SCSQ-A had additional items that aligned most with GIST, and none of the published studies (Dahlberg 2007; Braden 2010; Harrison-Felix 2018) used the LCQ as a primary outcome measure but rather a secondary. I think the use of the LCQ as your primary needs to be more clearly justified. Braden and colleagues found the LCQ to be sensitive to change in social communication skills, and comparable to changes found in SCSQ-A used as primary outcome measures in other GIST trials. The authors also reported that they plan to use LCQ in future studies. This is also seen in the Harrison-Felix study from 2018 as they use PPIC as the primary outcome, and no longer use SCSQ-A but LCQ as secondary outcome measure. Based on this, and that the LCQ has been psychometrically evaluated with excellent results in the TBI population, the LCQ was translated into Norwegian and employed as primary outcome measure in the present RCT. We have included PPIC as secondary outcome measure. We have added this justification on page 11 line 16-23.
P13L13: Are you going to assess the feasibility of getting video recordings? I suspect you may encounter problems here. This calls into question the involvement of the family member in the treatment if they can't attend to have this done.
Yes, the feasibility of getting video recordings was assessed in the beginning of recruitment. Most of the participant and close family members/friends have until now been able to attend to the assessment meetings. However, we have opened up for the participants and family members to do the video recordings at home in those cases when the family member is not able to attend the assessment meeting.
P14L10: Why did you choose the mind in the eyes test? This test is being recommended less as a social cognition assessment. I am also intrigued to understand how you think GIST will specifically make a change on this measure?
The capacity to interpret other people's behaviour and mental states is a vital part of social communication. Reading the Mind in the Eyes Test, is widely used and a simple test for emotion recognition, involving the first stage of attribution of Theory of mind: attribution of the relevant mental state (Baron-Cohen et al., 2001). Emotion recognition is also addressed in the GIST groups during the treatment, and participants may also have individual goals related to these kinds of problems.
You could assess the feasibility of doing all these measures as this large assessment battery (10 questionnaires and a video recorded conversation) will entail considerable participant burden. I would perhaps put the outcomes as primary, secondary and exploratory.
Thank you for pointing this out. We have chosen primary and secondary outcome measure since this assessment battery is not unusually large compared to other studies. Nevertheless, a large outcome battery may also be considered a participant burden, and this is added in the discussion on page 19 line 1-2.

Blinding – are all questionnaires completed by a blinded assessor? How will you ensure the person remains blinded. I would recommend you record instances of unblinding or near misses.
All of the questionnaires are self- and informant forms and are conducted by the participants themselves. The majority of these surveys are completed at home without any assessor present. This issue is now clarified on page 13 line 4-10.
Intervention

I noticed that the standard GIST is in an outpatient setting and intensive is an inpatient setting. I would therefore assume that the intensive participants are earlier post-injury than those in the standard GIST. As a result factors such as time post-injury and location (setting of treatment) could have an effect on outcome. I feel you need to be very cautious in your interpretation of the end results.
All the participants (both inpatients and outpatients) are minimum 12 months post-injury. Hence, the intensive participants are not earlier post-injury than those in the standard GIST. We have clarified this issue on the inclusion criteria's on page 9 line 22.

How will you prospectively measure fidelity? – I note this is not in your TIDIER checklist.
Both training protocols are manualized, ensuring the consistency in which intervention content is delivered. Regular meetings with external supervisors are also conducted, with monitoring of intervention delivery, to further ensure intervention fidelity. Specified on page 15 line 3-4.
Sample size
What outcome was the power analysis based on? I assume the LCQ as this is your primary outcome but this should be clarified.
Yes, this has now been clarified in the manuscript on page15 line 14.
Data analysis
How will you deal with attrition – what amount of data do you expect to lose to follow-up? How will you deal with missing data?
Throughout the trial all participants have regularly contact with the therapists/research team. Based on the Dahlberg et al. (2007) study, we expect to lose some, but not many, participants during the treatments, WL period and follow up. However since we have chosen the intention to treat principle every patient randomized to the clinical study should enter the primary analysis. However, participants are excluded if no treatment has been applied at all or if there are no data available after randomization. Page 15 line 21-23.
Do you intend to do per protocol analyses?
We intend to use the intention-to-treat principle as described in the previous comment.

Any analyses that compared standard with intensive GIST need to be controlled for (for reasons stated above). My concern is how misleading the results could be given location of treatment and time post-injury. I would want you to compare baseline differences before you start comparing intensive with standard GIST.
We hope our answers and changes in the manuscript help to clarify how we plan to compare and control for the factors you are addressing.

What is the purpose of this study in the long-term – how would this study inform future research work? Having a discussion and/or conclusions section would be recommended.
Yes, this is an important comment. We have made changes accordingly by adding a discussion on page 18 section 1 addressing this.

Thank you for all your helpful comments.

Reviewer: 4
Thank you for the opportunity to review this interesting study protocol. This protocol reports on an innovative study comparing standard social communication skills training with intensive (inpatient-based) training for people with acquired brain injuries. Inpatient-based social communication skills training is not commonly used within the field of ABI, and so this study will represent an important contribution to the literature. I will be interested to read about the outcomes of this study when they are available.
I have made the following comments related to improving the clarity of this protocol:

Thank you for all the helpful feedback on our protocol article. Questions and comments are answered below. The minor comments are changed in the manuscript and marked in the document “main document- marked copy”, and the major comments are addressed here.

COMMENTS:
Title: The key focus of your study (intensive versus standard treatment) is not clear from the title of this paper. I would recommend adding the word "intensive" to the title.
Yes, this is a good point. We agree on this. The title has now been changed to: Intensive and standard group-based treatment for persons with social communication difficulties after acquired brain injury: study protocol for a randomised controlled trial

Abstract:
In the introduction section, the innovation of your study (comparing intensive with standard treatment) is not clear. It is stated that GIST has received empirical support, so from reading this section alone, it is not clear how this study will add to the evidence base that already exists around GIST.
Thank you for pointing this out. We have made this clearer in the aim of the study by stating that intensive GIST is a newly developed GIST protocol.
We have also added a clarification and strengthened our rationale for the use of both intensive GIST and standard GIST. Based in previous research (e.g. Harrison-Felix et al., 2018, Dahlberg et al., 2007) on page 6 section 2 and 3 we have argued that there is still a need for more research determining the best format for GIST. Rationale for the intensive GIST is also expressed on page 7 Line 8-9.

In the methods and analysis section, it is stated that "The WL group will receive intensive GIST". Could you clarify that the WL group will receive the intensive GIST after a waiting period?
Yes, this has been clarified in the abstract. This is also clarified in figure 1 and on page 8 line 14-18 and page 14 section line 12-22.

The first mention of "intensive GIST (inpatient)" here confused me, as it sounded like this was a different group of inpatients (separate from the 60 adults recruited and randomised). I would suggest moving the section where you define the differences between standard GIST and intensive GIST earlier in this section. Also, rather than stating that standard GIST will be delivered to outpatients, and intensive GIST will be delivered to inpatients (which sounds like you have recruited participants from two different contexts), it would be clearer to state, "Standard GIST will be delivered to participants in 2.5 hour outpatient sessions..." and "For intensive GIST, participants will be admitted as inpatients for four weeks and receive three full-day sessions each week."
Thank you for pointing this out to us. We have made some changes in the method section of the abstract to clarify this issue.

Introduction:
Page 4 line 5 - what do you mean by "hold back information"? Does this mean the same thing as the next part, i.e. "avoid saying inappropriate things"?
We have omitted “hold back information” from this section.

Page 6 line 38 - I was expecting to see in this paragraph some more detailed information about the previous studies using GIST to back up the claim that this is "one of the best-validated group interventions", but I note that this content follows in the next paragraph. Restructuring these two paragraphs may assist in developing your argument, starting from the broad evidence base for GIST, and narrowing to the reasons for trialling the more intensive version.
Thank you for pointing this out to us. We have followed you suggestion and have as such restructured this paragraph accordingly. Detailed information about the previous GIST studies is now found on page 6 section 3 and on page 7 section 1.

Methods and Analysis:
Page 8 Line 46 - The figure 1 provided relates to randomisation of participants, rather than showing the assessment time points. Is there another figure that should be referenced here?
We have now added a new figure: a flow diagram over the trial and assessment points to make this clearer. See figure 1.

Page 9 line 45 - "No concomitant diseases" sounds very broad - would this mean, for example, that a person with diabetes would be excluded? Would it be more correct to state "no concomitant neurological diseases"?
Yes we agree, and changes have been made on page 10 line 7.

Page 9 line 56 - Could you clarify this criteria relating to the LCQ? Were you basing this on the self-report or other-report version? By "aspects of communication", do you mean three questionnaire items? How were moderate to severe difficulties defined (e.g. problem happening either often, or usually/always - rating of 3/4)?
A clarification of the inclusion criterion (vi) has now been added: communication difficulties reported on minimum three questions (i.e., “often” or “always”) on the LCQ. See page 10 line 2-3.

Page 10 line 15 - What is meant by "unfitness for evaluation of outcome"?
We agree that this is unclear, and have decided to delete this criterion.

Table 1 - I am unclear on the timing of the T2 assessment for the WL/intensive GIST group? Do they complete this at the end of the waiting period, or after their intensive GIST treatment? It is unclear from this table.
Yes, based on your comments we have decided to remove table 1 and make a flow diagram to make the procedure more clear and visual. See figure 1.

Page 12 line 45 - I assume the LCQ was translated into Norwegian for the purposes of this study - if so, this should be mentioned. Was any support provided to people with ABI who may have difficulty with comprehending the items? This could similarly be reported for the SCQ-A.
Yes, we have added this on page 11 section 2. Braden and colleagues found the LCQ to be sensitive to change in social communication skills, and comparable to changes found in SCQ-A used as primary outcome measures in other GIST trials. The authors also reported that they plan to use LCQ in future studies. Based on this, and that the LCQ has been psychometrically evaluated with excellent results in the TBI population, the LCQ was translated into Norwegian and employed as primary outcome measure in the present RCT. Also, based on the multiple causes of SCD the developer of LCQ anticipates LCQ to provide useful information across ABI-populations.

Page 14 line 12 - It is not really correct to state that these outcome assessments are blinded, as the participants and family members are aware whether intervention has been received or not.
We have now described these procedures in a different way to make this more correct on page 13 line 1-10.

Page 14 line 17 - How does sound recording the assessments ensure inter-rater reliability? Will a different assessor listen to the sound recordings to score some assessments to ensure the original assessor noted the scores correctly?
The tests are scored by the help of sound recordings by a trained research assistant who is blinded to group allocation to ensure interrater reliability. This has been clarified on page 13 8-10.

Page 15 line 15 - How are the follow-up treatment sessions scheduled in relation to the follow-up assessment sessions i.e. do the participants complete their three month follow-up session and then their T3 assessment? If so, the "booster" effect of the extra training session should be considered in evaluating long-term effects.
The participants complete their three and six month follow up assessments together with their close family member or friend before the follow up sessions to reduce the “booster” effect of the extra training session.

Page 15 - What are the arrangements if a participants misses a session due to illness etc?
We have added a clarification on this issue on page 15 line 1-2. If a participant misses a session, the amount of repetition in GIST allows him/her to continue direct to the next session. However, if one misses more than one session consecutively, an individual session will be allowed to bring them up to speed.

Page 15 line 31 - The number of hours of training the intensive GIST participants receive is unclear. If the standard GIST participants are receiving 12 x 2.5 hr sessions = 25 hours, how many hours are the intensive GIST participants receiving? It sounds like the intensive GIST intervention differs from the standard GIST in both intensity (i.e. how concentrated the intervention is) and dose (i.e. how many total hours of intervention are received). This should be clarified, and will be an important issue to consider in interpreting your findings.
We have now specified the number of treatment hours in Figure 1, and elaborated on the differences between the two treatments on page 14 sections 1-2 and on page 15 section 1.

Page 16 - As per previous comments, it is difficult to follow the timing of the assessments and how these relate to the interventions being provided. I assume that the comparison between groups at T2 will provide information about standard GIST vs waitlist. Which data will be relevant to comparing the standard GIST vs intensive GIST, and intensive GIST vs waitlist? How will analyses of data from T3 and T4 be informative about your research questions?
We have addressed this by clarifying the interventions on page 14 sections 1-2, and in the data analyses plan section on page 15-16. We have also added a new figure/flow diagram over the trial and assessment points to make this clearer. See figure 1.

Further note - it would be relevant to add somewhere some comments about the potential limitations of this research. For example, I wonder whether participants will find the intensive GIST training acceptable, or whether the expense of intensive GIST training may make this delivery format difficult to implement in practice.
Yes, we agree on this and have addressed limitations on page 3 under the sections strengths and limitations of this study, and we have also added a discussion section on page 18 section 1.

VERSION 2 – REVIEW

REVIEWER	Sara da Silva Ramos The Disabilities Trust United Kingdom
REVIEW RETURNED	09-May-2019

GENERAL COMMENTS	Thank you to the authors for their extensive review of the manuscript. Having looked at their response and at the revised manuscript, I believe that there are still some points that require clarification. These are outlined below. I've also noticed a few typos throughout the manuscript, which will need to be corrected (e. g. missing spaces between words). In my opinion, the manuscript would be acceptable for publication, provided that the clarifications outlined below, and those of other Reviewers, are addressed in the final version manuscript. In addition, the revised manuscript should also include further comment and recommendations for future implementations of this protocol, with respect to the points that could not be mitigated by a change in the present protocol (e. g. absence of sample stratification). 1. Despite the authors' answers, this point is still unclear and further revision is required I was not able to locate Figure 1. Furthermore, while now it seems that data will be collected at enough time points to make all the required comparisons, the sentence "The proposed study is an RCT comparing two different GIST versions—standard GIST (n=30) and intensive GIST (n=30)—to results of persons in a WL group using a repeated-measures design" implies that there is a third group of people in a waiting list who might receive either standard treatment, intensive treatment, or some other treatment not included in the study. However, further down in the same sentence "seven time points for the WL/intensive GIST group (see Figure 1)" it becomes clear that the waiting list group is also the group who receives intensive GIST. In other words, there is repeated measures data to allow comparisons pre-and post- for the standard treatment and for the intensive treatment, and to compare a control condition (no treatment) with intensive treatment. However, no comparable data are available for the standard condition. Please make it clear in the manuscript that this is no third group, and that the only repeated-measures comparisons possible involving the waiting list controls, are those between WL and intensive condition. 2. From the authors' response, it seems that there are no aspects of the procedure to ensure that the groups will be comparable. The limitations introduced by this (which go beyond the consideration of whether or not cognitive changes occur after 12 months) will need to be discussed. In particular, the extent and nature of rehabilitation received previously could introduce significant differences between the two groups and undermine the strength of any between group comparisons. I suggest that the authors acknowledge this and provide some practical recommendations on how this limitation could be addressed by introducing minor changes in future studies using this protocol. 4. I take the authors' point, and think that their response is acceptable, however there is some value in expanding on this

	issue further, from a more general standpoint (e. g. in the discussion). How might we make social and environmental interventions more accessible to those who are more socially isolated? 5. The authors' response clarifies the question, but I can't find these specific details in the indicated portion of manuscript (i. e. clear explanation that all participants receive detailed information about the two treatment conditions prior to enrolment, and that although some of the participants may express a preference for one of the treatments, the assessor clarifies that group allocation is decided through randomisation). As an aside comment, it would be useful to include, as part of the protocol / analysis, a record of when participants express a preference for one of the conditions and make a note of whether they were randomised to the preferred treatment or not. This can give us some general insight about patient preferences and also help us evaluate whether a mismatch in preferences influence the achieved outcomes. 7. Information about extending sessions to 2.5 hours now appears twice. The following sentence implies there are more than 12 modules, or perhaps some modules require more than one session "The intensive GIST treatment will consist of the same 12 modules as arm 1 (i.e. manualised) in an inpatient setting (2 x three days/week, 2 x four days/week)". Please clarify. In addition, the information provided in the abstract and that in the body manuscript is not consistent. The abstract states that "Standard GIST (n=30) will be delivered to outpatients in 2.5-hour sessions once per week for 12 weeks", which would make a total of 30 hours. However, on page 16, it is stated "this has been extended to 2.5 hours (32.5 contact/treatment hours in total), allowing for more breaks". Which is correct? 8. The power calculations provided seem to have been made on the basis of separate pre- post- comparisons within each group, rather than on a mixed factorial design with time (pre-post) as within-subjects factor and group (standard versus intensive) as between-subjects factor. However, on p. 18, the authors recommend a factorial analysis. The power calculations should match the recommended analyses. Please also explain in more detail the comparisons involving the waiting list data, including considering the impact that within- and between-subjects' comparisons may have on power. 11. Can the authors expand on the proposed revision to fully address this point? We agree that "the long waiting period for the WL group may affect adherence and drop-out rates", but how so? (Possibly higher drop-out rates on the intensive group during the waiting list period. Maybe more motivation to adhere to the treatment, in comparison to the standard group, once it is started?) If these differences in drop-out rates and adherence do turn out to be observed, how might they affect the study conclusions? Is the waiting list condition really going provide data that will allow robust conclusions about the effectiveness of the treatment? All other points (3, 6, 9 and 10) have been adequately addressed.

REVIEWER	Natalie Gilmore Boston University, US
REVIEW RETURNED	23-May-2019

GENERAL COMMENTS	The authors have addressed my comments adequately.

REVIEWER	Nicholas Behn City, University of London UK
REVIEW RETURNED	15-May-2019

GENERAL COMMENTS	There are a lot of acronyms used throughout the paper and I wonder if all are needed? There are some sentences where 3 acronyms are used, which I think detracts. Do you need the SCD acronym? Also check them as there are many inconsistencies throughout the paper where sometimes you use the acronym and sometimes you don't and occasions where you use different acronyms e.g. SCSQ-A vs. SCQ-A. Abstract Methods: You should mention the assessment points for the Standard GIST group. Strengths and limitations P3L17. Reword as this is a repetition. Perhaps say "two follow-up sessions (at three and six months)" P3L31. Is the involvement of close family to help with homework assignments that big a strength. I read this as them having a low dose of input. Introduction P4L26. "And denied access to" Rephrase. P4L27. Consider omitting sentence "loneliness and social isolation" as this does not directly link to communication. P5.L17. Again, I am unsure why you have mentioned the coping intervention as this didn't come up in the Cicerone review and is something that was published after Cicerone, ACRM, Togher and Finch reviews. I'm struggling to understand the point of why you mention these two studies from Cicerone but don't describe studies from any of the other reviews. If you to include the Douglas paper there needs to be a clearer justification and it needs to be moved further down the manuscript. P7. The mention of your pilot study should come further down as a lead in to your aims and hypotheses. P8L10. Similar patterns of improvement in what? I wonder if mentioning each outcome detracts from the point you are trying to say and in your methods you can more clearly explain why you chose the outcomes you did. Methods and analysis P9L16. I would consider omitting the description of the time points as its odd you doing it for the standard but not the intensive group. I noticed you omitted the TIDieR table which is a shame as this is considered a gold standard for intervention description and reporting according to CONSORT. I would reconsider. Study setting – as you are recruiting through social media and user organisations do you think you will recruit from hospitals other than SRH? Eligibility criteria – you are excluding those with aphasia but as you suggested in your response people can have social communication difficulties alongside aphasia which is entirely correct. I would have thought you would exclude severe aphasia so mild-moderate may still be included?

	P11L10. Verbal consent rather than oral consent. How will this be obtained? I assume you expect participants to make first contact with you? Can I also check – is the inclusion assessment face-to-face? P13L40. Sentence not needed. I would omit. The LCQ rates perceived communicative ability. Higher scores do not reflect "reduced communication skills" but rather "greater perceived communicative difficulty" or "less perceived communicative ability". You have a sub-heading for secondary outcome measures but not primary. You might want to amend. P14L10 What is meant by the "assessment meeting"? – this is the first time this term was used. Is this different to the conversation? P16L20 Problems with wording. Seems there is some repetition here. Interventions – this is why you need a TIDieR description or something similar as there is a lack of detail e.g. who does the sessions? Setting? Tasks done within session? What are the lists of topics discussed? As you are adding time for practicing goals (i.e. 2 hours/week) does this mean there is an additional 8 hours in total for the four weeks so 44 + 8 = 52 hours? This needs clarifying. Sample size – I would remove the statement about sample size of other studies as you should be led by your power analysis on the primary outcome measure. Data analysis – you plan to do many t-tests. How will you control for this? I wonder whether you should be doing this? My concern is if you don't find a result in the repeated ANOVA you will be looking for something to come through on the t-tests – I would at least apply a Bonferroni adjustment or something along those lines. In your aims you mention that you will determine the efficacy of the two protocols to those in the WL group but your analyses seem to be comparing standard with intensive GIST. Some clarification would be useful here. Your analyses don't seem to match your aims. Discussion You provide reasons for why you may get better outcomes in either the standard or intensive GIST but your hypotheses suggested you don't expect that. Your conclusion might need some more work as it feels odd to end on issues with WL group and participant burden from a large assessment battery Other I would avoid starting sentences with "because" – its awkward wording.

REVIEWER	Rachael Rietdijk The University of Sydney, Australia
REVIEW RETURNED	16-May-2019

GENERAL COMMENTS	Thank you for the opportunity to review this revision. I appreciate that this manuscript has been significantly improved in clarity. The document, "Response to editor and reviewers" did not appear to be available for me to review, so I have based this review on the content in the manuscript only. There are a few minor comments regarding further edits to improve clarity.

	ABSTRACT The abstract is now much easier to follow. The introduction now indicates the innovation of this study and the method section reports the study design more clearly. However, it still is possible for the reader to get the impression that the two intervention groups were recruited from different contexts, since “Standard GIST will be delivered to outpatients”, and “Intensive GIST will be delivered to inpatients”. To make it clearer that the whole sample are outpatients before being recruited to the study, I would suggest editing this to say something like, “Standard GIST will be delivered via outpatient sessions” and “For Intensive GIST, patients will be admitted to an inpatient rehabilitation unit for four weeks and receive (etc).” INTRODUCTION I appreciate the authors adding further detail here to explain the context of research studies on social communication disorders after TBI. Some further editing would assist to develop the clarity of the argument for studying GIST intervention, as the context and links. For example: • If starting with the Cicerone et al systematic review, it is important to note that this review focuses on literature published in 2003 to 2008 (and therefore does not reflect what has been published in the last decade). • In the same paragraph as Cicerone et al, the study by Douglas et al (reference 24) (published in 2015) is introduced. The transition to the Douglas paper is unclear. • You mention that searching in 2017 found 4 additional papers published since 2013, but then only mention one of them. • In transitioning to focussing on GIST intervention, it is unclear whether these GIST intervention papers were included in the Cicerone or Finch systematic reviews. • The argument would be more developed logically if you finish summarising the evidence from GIST intervention papers, then introduce your own pilot data. The transition from your pilot data to the Braden paper does not assist you to build your argument for this protocol. If you need to cut down on the word length, you may consider reducing or deleting the content about the Cicerone et al systematic review, and instead use the more recent and relevant systematic review by Finch et al as a starting point for developing your argument. METHODS I queried during my last review whether “no concomitant diseases” really meant “no concomitant neurological diseases”. I note this has been edited in the exclusion criteria but not in the inclusion criteria. Also, if the exclusion criterion essentially means the same thing as the inclusion criterion, you do not need to have both. La Trobe Communication Questionnaire – was the self-report or other-report score considered the primary outcome measure? I note that at the end of the secondary outcome measures section, you have noted that people will complete this measure at home independently, but it would be helpful to state this at the start of the Measures section, so that the reader has this information while reading about the LCQ. The follow-up treatment sessions for Standard GIST at three and six months should be marked on Figure 1 so the reader can see how these treatment sessions relate to the assessment timepoints.

	DISCUSSION The additional consideration of strengths and limitations of the research included in this section is helpful and appropriate.

VERSION 2 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1

Thank you to the authors for their extensive review of the manuscript. Having looked at their response and at the revised manuscript, I believe that there are still some points that require clarification. These are outlined below. I've also noticed a few typos throughout the manuscript, which will need to be corrected (e. g. missing spaces between words). In my opinion, the manuscript would be acceptable for publication, provided that the clarifications outlined below, and those of other Reviewers, are addressed in the final version manuscript. In addition, the revised manuscript should also include further comment and recommendations for future implementations of this protocol, with respect to the points that could not be mitigated by a change in the present protocol (e. g. absence of sample stratification).

Thank you for all the helpful feedback on our protocol article. We have now included further limitations and recommendations for future implementations to the discussion section on page17, in addition to correction of typos. Questions and comments are answered below.

1. Despite the authors' answers, this point is still unclear and further revision is required I was not able to locate Figure 1. Furthermore, while now it seems that data will be collected at enough time points to make all the required comparisons, the sentence "The proposed study is an RCT comparing two different GIST versions—standard GIST (n=30) and intensive GIST (n=30)—to results of persons in a WL group using a repeated-measures design" implies that there is a third group of people in a waiting list who might receive either standard treatment, intensive treatment, or some other treatment not included in the study. However, further down in the same sentence "seven time points for the WL/intensive GIST group (see Figure 1)" it becomes clear that the waiting list group is also the group who receives intensive GIST. In other words, there is repeated measures data to allow comparisons pre-and post- for the standard treatment and for the intensive treatment, and to compare a control condition (no treatment) with intensive treatment. However, no comparable data are available for the standard condition. Please make it clear in the manuscript that this is no third group, and that the only repeated-measures comparisons possible involving the waiting list controls, are those between WL and intensive condition.

Thank you for pointing this out. It was unfortunate to hear that you did not get access to Figure 1 attached to this protocol. Nevertheless, we have adjusted the aims in the abstract and protocol to clarify. We have also reformulated the study setting section and the sentence you point out to clarify that there are two groups of participants and that the WL and intensive GIST group is the same group. The WL design we have chosen is based on Dahlberg et al.´s 2007 design, comparing results from treatment with deferred treatment in their RCT. As this is a common design to use, it is our understanding that the WL group will provide us with data that we can compare with the results of treatment. This will allow us to make comparisons between subjects for standard GIST and WL, and for standard GIST and intensive GIST, and within subjects for WL and Intensive GIST. This clarification is added to page 7 line 9-13.

2. From the authors' response, it seems that there are no aspects of the procedure to ensure that the groups will be comparable. The limitations introduced by this (which go beyond the consideration of whether or not cognitive changes occur after 12 months) will need to be discussed. In particular, the extent and nature of rehabilitation received previously could introduce significant differences between

the two groups and undermine the strength of any between group comparisons. I suggest that the authors acknowledge this and provide some practical recommendations on how this limitation could be addressed by introducing minor changes in future studies using this protocol.

Thank you for pointing this out. The use of a randomisation is used to equalize group differences. However, a simple randomisation does not rule out factors such as time since injury or amount of rehabilitation received prior to participation as you point out. Nevertheless, this type of data will be collected during the demographic interviews and will as such be controlled for in the data analyses.

On the other hand, the use of stratification and a larger sample size would have been valuable to our results, and for future studies, using this protocol, we highly recommend to explore design with stratification to ensure a higher degree of group homogeneity. We have as such added this to the discussion section on page 19 line 5-8.

4. I take the authors' point, and think that their response is acceptable, however there is some value in expanding on this issue further, from a more general standpoint (e. g. in the discussion). How might we make social and environmental interventions more accessible to those who are more socially isolated?

This is a interesting question, and there would be important for future studies to include this aspect. In our protocol we are not able to investigate this aspect in detail as this paper is a study protocol. Further discussion of this is more relevant when presenting results from the RCT. However, we have included this challenge in the discussion and highlighted the importance of future studies to include this perspective. See page. 18 line 2-7.

5. The authors' response clarifies the question, but I can't find these specific details in the indicated portion of manuscript (i. e. clear explanation that all participants receive detailed information about the two treatment conditions prior to enrolment, and that although some of the participants may express a preference for one of the treatments, the assessor clarifies that group allocation is decided through randomisation). As an aside comment, it would be useful to include, as part of the protocol / analysis, a record of when participants express a preference for one of the conditions and make a note of whether they were randomised to the preferred treatment or not. This can give us some general insight about patient preferences and also help us evaluate whether a mismatch in preferences influence the achieved outcomes.

Thank you for addressing this. As such, we have now included a clarification of this in the manuscript on page 9 line 15-17. To comment on the aside comment, this is an interesting point. We will include a record of when participants express a preference for one of the conditions prior to inclusion in the cases that this is registered. This has now been included in the procedure section on page 9 line 17-19.

7. Information about extending sessions to 2.5 hours now appears twice. The following sentence implies there are more than 12 modules, or perhaps some modules require more than one session "The intensive GIST treatment will consist of the same 12 modules as arm 1 (i.e. manualised) in an inpatient setting (2 x three days/week, 2 x four days/week)". Please clarify. In addition, the information provided in the abstract and that in the body manuscript is not consistent. The abstract states that "Standard GIST (n=30) will be delivered to outpatients in 2.5-hour sessions once per week for 12 weeks", which would make a total of 30 hours. However, on page 16, it is stated "this has been extended to 2.5 hours (32.5 contact/treatment hours in total), allowing for more breaks". Which is correct?

Thank you for commenting on this. 32.5 contact hours is correct for the standard GIST group and is including the 12 treatment sessions and the initial group orientation session. This has now been clarified on page 13 line 3-8. Regarding the Intensive GIST group, we agree that this was an unclear description. The Intensive GIST treatment weeks include 44 contact hours with GIST curriculum. Because of the reduced time to practise on their individual goals (4 weeks compared to 12 weeks for the standard GIST group) the time spent on each session is expanded to allow the participants to practise, discuss and interact further within the treatment sessions. In addition, time for practicing social communication goals and informal group activities is added to the participant's weekly schedule

(2 hours/week), and the participants is encouraged to participate in social activities in the evenings offered at the hospital or initiated by the participants themselves. The four weeks also include time for assessments pre-and post-treatment. This has now been clarified with examples on page 13-14 line 13-22, 1-8.

8. The power calculations provided seem to have been made on the basis of separate pre- post-comparisons within each group, rather than on a mixed factorial design with time (pre-post) as within-subjects factor and group (standard versus intensive) as between-subjects factor. However, on p. 18, the authors recommend a factorial analysis. The power calculations should match the recommended analyses. Please also explain in more detail the comparisons involving the waiting list data, including considering the impact that within- and between-subjects' comparisons may have on power.

Thank you for pointing this out. Preparing our study, the sample size was in line with previous studies available Dahlberg et al, Douglas et al., and Braden et al., however, in the process of publishing this protocol, the multi-centre RCT from Harrison-Felix et al. was published with a greater sample size than any other studies (n=179). Based on the small population in Norway it is not possible to increase the sample size in the present RCT. However, in future studier, multi-centre designs should be considered in order to increase the sample size. In our study we calculated power based on two treatment groups and are planning to run two between subject analyses comparing; I) standard GIST and intensive GIST: the between-subjects factor and time (baseline [T1/T4], post-intervention [T2/T5], three-month follow-up [T3/T6] and six-month follow-up [T4/T7]) and II) standard GIST and WL: the between-subjects factor and time (baseline [T1/T1], post-intervention/WL [T2/T2], three-month follow-up/WL [T3/T3] and six-month follow-up/WL [T4/T4]). In addition will intensive GIST and WL be compared using t-tests to explore the within-subjects factors change of scores (baseline -post-intervention [T4-T5], baseline- three-month follow-up [T4-T6], baseline - six-month follow-up [T4-T7]) and accordingly also for the WL ([T1–T2], [T1–T3], [T1–T4]). To adjust for multiple comparisons the Bonferroni procedure will be applied.

11. Can the authors expand on the proposed revision to fully address this point? We agree that "the long waiting period for the WL group may affect adherence and drop-out rates", but how so? (Possibly higher drop-out rates on the intensive group during the waiting list period. Maybe more motivation to adhere to the treatment, in comparison to the standard group, once it is started?) If these differences in drop-out rates and adherence do turn out to be observed, how might they affect the study conclusions? Is the waiting list condition really going provide data that will allow robust conclusions about the effectiveness of the treatment?

Thank you for your comment. The use of WL has both strengths and limitations. However, we see it as an ethical strength that all participants will receive a treatment. The WL design we have chosen is based on Dahlberg et al. 2007 design comparing results from treatment with deferred treatment in their RCT. McDonald et al. 2012 did also use a WL group in their RCT and compared the treatment group results with WL results. In addition, they compared the treatment group results to the results available from treatment after WL. It is our understanding that the WL group will provide us with data that we can compare with the results of treatment. Regarding your comment on adherence, the adherence in both groups (standard GIST and WL/intensive GIST) will be registered and included in the data analyses, and we see differences in adherence in the two groups. It will be important to report and discuss this further in the efficacy article, when presenting the results.

All other points (3, 6, 9 and 10) have been adequately addressed.
Thank you, and again thank you for your comments in the previous round.

Reviewer: 2

The authors have addressed my comments adequately.
Thank you so much for your comments in the

Reviewer: 3

There are a lot of acronyms used throughout the paper and I wonder if all are needed? There are some sentences where 3 acronyms are used, which I think detracts. Do you need the SCD acronym? Also check them as there are many inconsistencies throughout the paper where sometimes you use the acronym and sometimes you don't and occasions where you use different acronyms e.g. SCSQ-A vs. SCQ-A.

Thank you for pointing this out. We added the SCD because of the word limit, but the inconsistencies of acronyms throughout the paper has now been corrected.

Abstract
Methods: You should mention the assessment points for the Standard GIST group.

Yes, the assessment points for the standard GIST group have now been added to the abstract. Thank you for pointing this out.

Strengths and limitations
P3L17. Reword as this is a repetition. Perhaps say "two follow-up sessions (at three and six months)"
Yes, we agree and have now changed the sentence accordingly.

P3L31. Is the involvement of close family to help with homework assignments that big a strength. I read this as them having a low dose of input.
Research on context-sensitive approaches and GIST suggest that the involvement of a close family member is a strength (e.g., Dahlberg et al., 2007; Ylvisaker, 2003). Before treatment the close family member is instructed on their role as a support person during treatment. Also, as part of the home assignments they instructed to give feedback to the participants on how they are working on their social communication skills goals and the different sessions in general. To have a support person knowing what the participants are aiming to work on, and by allowing the support person to give feedback and talk about how one communicates may have a positive impact on how they communicate with each other, including generalization effects. Of course, there is no guarantee that the close family member or support person is active in their role as instructed. However, we ask the participant and close family member to describe the family members role as support person during treatment and the cooperation regarding home assignments at the end of the treatment.

Introduction
P4L26. "And denied access to" Rephrase.
The sentence has now been rephrased to "A lack of successful social skills can lead to conflicts, isolation, and limited access to social and vocational opportunities"

P4L27. Consider omitting sentence "loneliness and social isolation" as this does not directly link to communication.
The sentence has now been omitted.

P5.L17. Again, I am unsure why you have mentioned the coping intervention as this didn't come up in the Cicerone review and is something that was published after Cicerone, ACRM, Togher and Finch reviews. I'm struggling to understand the point of why you mention these two studies from Cicerone but don't describe studies from any of the other reviews. If you to include the Douglas paper there needs to be a clearer justification and it needs to be moved further down the manuscript.
Yes, we agree with you on this. Based on your feedback we have now chosen to remove the Douglas study. Additionally, we have also removed the Cicerone et al., review, and decided to focus on the more recent systematic reviews Finch et al., and Togher et al. In addition of adding the most recent Cicerone et al. see page 4 line17-24, page 5 line 1-5

P7. The mention of your pilot study should come further down as a lead in to your aims and hypotheses.
Thank you for pointing this out. We have moved our pilot study at the end of this section so that this leads to our aims and hypotheses. See page 6 line 2-12.

P8L10. Similar patterns of improvement in what? I wonder if mentioning each outcome detracts from the point you are trying to say and in your methods you can more clearly explain why you chose the outcomes you did.
Thank you for pointing this out to us. We have now removed the assessment tools in this section, as this is referred to in the method section.

Methods and analysis
P9L16. I would consider omitting the description of the time points as its odd you doing it for the standard but not the intensive group.
Yes, the time points (T1-T4) is now omitted from this section.

I noticed you omitted the TIDieR table which is a shame as this is considered a gold standard for intervention description and reporting according to CONSORT. I would reconsider.
Thank you for pointing this out. Hence, we have decided to include the TIDieR table again. See the attached file.

Study setting – as you are recruiting through social media and user organisations do you think you will recruit from hospitals other than SRH?
Yes, the participants will be recruited during several channels in order to cover health institutions, rehabilitation settings and arenas where eligible participants in the chronic stage are typically found. This has now been clarified on page 8 line 1-3.

Eligibility criteria – you are excluding those with aphasia but as you suggested in your response people can have social communication difficulties alongside aphasia which is entirely correct. I would have thought you would exclude severe aphasia so mild-moderate may still be included?
Yes, this something we discussed a lot when developing this protocol. We have set the criteria as we want to have a focus on social communication in the present study, to be able to compare the results with similar studies. This mean that the participants included in this study do not have conspicuously aphasia symptoms evaluated by an SLP, but they experience difficulties with pragmatic and social communication similar to the symptoms often observed in patients with TBI or right hemisphere stroke. However, it is possible that we include some participants with mild aphasia that is not detected by the tests. As such, we have now reformulated the exclusion criteria to: (iv) if the communication difficulties mainly is associated with aphasia assessed by an SLP. See page 9 line 6-7.

P11L10. Verbal consent rather than oral consent. How will this be obtained? I assume you expect participants to make first contact with you? Can I also check – is the inclusion assessment face-to-face?

Oral consent has now been changed to verbal consent. In the information letter it is stated that we will call them after 7-10 days to ask if they are interested in hearing more about the project or not. In the information letter it is also stated that participating in this study is voluntary and that a refusal will not affect further treatment at the hospital. The inclusion assessment is face to face at an assessment meeting at the hospital. A clarification about this matter has now been added to page 9 line 10-21.

P13L40. Sentence not needed. I would omit.
We are very sorry, but cannot find the page and line referred to here in the manuscript, and hence, we are unable to make any correction.

The LCQ rates perceived communicative ability. Higher scores do not reflect “reduced communication skills” but rather “greater perceived communicative difficulty” or “less perceived communicative ability”
Yes agree, this sentence has now been changed.

You have a sub-heading for secondary outcome measures but not primary. You might want to amend.
We have added a sub-heading for the primary outcome measure. Thank you for pointing this out to us.

P14L10 What is meant by the "assessment meeting"? – this is the first time this term was used. Is this different to the conversation?
We have now changed the assessment meeting to just the assessments. All assessments are conducted at the hospital's outpatient clinic. However, sometimes these are conducted over the telephone if the close family/ friend do not have the opportunity to join in.

P16L20 Problems with wording. Seems there is some repetition here.
We was unfortunately not able to find the page and line referred to in the manuscript, and was unable to make any correction.

Interventions – this is why you need a TIDieR description or something similar as there is a lack of detail e.g. who does the sessions? Setting? Tasks done within session? What are the lists of topics discussed? As you are adding time for practicing goals (i.e. 2 hours/week) does this mean there is an additional 8 hours in total for the four weeks so 44 + 8 = 52 hours? This needs clarifying.
Yes, I see your point and we have decided to include TIDieR again to make this clearer. Due to the reduced time to practice on their individual social communication goals (4 weeks compared to12 weeks for the standard GIST group) the time spent on each GIST-session is expanded to allow the participants to practice, discuss and interact further within the treatment sessions. When compared to standard GIST, the intervention is more intensive and extensive with a total of 44 contact/treatment hours over four weeks. In addition, informal group activities (e.g. cooking or garden groups) are added to the participant's weekly schedule (2 hours/week, total of 8 hours). In addition, the participants is encouraged to participate in social activities in the evenings offered at the hospital (e.g. quiz, ceramics or morning walks) or initiated by the participants themselves (e.g. shuffle-board or cafè/restaurant visits). The four weeks also include time for assessments pre- and post-treatment. These clarifications have now been added on page 13 and 14 line 13-22, 1-8.

Sample size – I would remove the statement about sample size of other studies as you should be led by your power analysis on the primary outcome measure.
Thank you for pointing this out to us. We have rephrased this sentence on page. 14 line 23-24.

Data analysis – you plan to do many t-tests. How will you control for this? I wonder whether you should be doing this? My concern is if you don't find a result in the repeated ANOVA you will be looking for something to come through on the t-tests – I would at least apply a Bonferroni adjustment or something along those lines.
Thank you for commenting on this. We agree with your suggestion, and have now added this to our data analysis. See page 16 line 1-2.

In your aims you mention that you will determine the efficacy of the two protocols to those in the WL group but your analyses seem to be comparing standard with intensive GIST. Some clarification would be useful here. Your analyses don't seem to match your aims.

Thank you for commenting on this. The aim of the study is now clarified in the abstract, and on page 6. We are planning to run 3 analyses comparing (1.) standard GIST and intensive GIST, (2.) standard GIST and WL, and (3.) intensive GIST and WL. This is clarified on page in the analysis section on page 15-16 lines 18-25, 1-2.

Discussion
You provide reasons for why you may get better outcomes in either the standard or intensive GIST but your hypotheses suggested you don't expect that.
Your conclusion might need some more work as it feels odd to end on issues with WL group and participant burden from a large assessment battery.
Thank you for commenting on this. We have now made some adjustments to the discussion to make our point clearer. See discussion on page 18-19, line 18-25, 1-7.

Other
I would avoid starting sentences with "because" – its awkward wording.

Thank you for pointing this out, and this is now changed throughout the manuscript. Thank you for all your helpful feedback and comments.

Reviewer: 4
Thank you for the opportunity to review this revision. I appreciate that this manuscript has been significantly improved in clarity. The document, “Response to editor and reviewers” did not appear to be available for me to review, so I have based this review on the content in the manuscript only. There are a few minor comments regarding further edits to improve clarity.

Thank you for all the helpful feedback on our protocol article. Questions and comments are answered below.

ABSTRACT
The abstract is now much easier to follow. The introduction now indicates the innovation of this study and the method section reports the study design more clearly. However, it still is possible for the reader to get the impression that the two intervention groups were recruited from different contexts, since “Standard GIST will be delivered to outpatients”, and “Intensive GIST will be delivered to inpatients”. To make it clearer that the whole sample are outpatients before being recruited to the study, I would suggest editing this to say something like, “Standard GIST will be delivered via outpatient sessions” and “For Intensive GIST, patients will be admitted to an inpatient rehabilitation unit for four weeks and receive (etc).”

Thank you for pointing this out to us and thank you so much for your suggestion on how to address this. We have now changed to wording to: “Standard GIST (n=30) will be delivered via outpatient sessions for 2.5-hour once per week for 12 weeks and one additional initial group orientation session. Participants will be assessed at pre- and post-intervention and at three- and six-month follow-up (T1-T4). For Intensive GIST (n=30), participants will be admitted to an inpatient rehabilitation unit for four weeks and receive three full-day sessions each week, following a waiting period.” See page 2 section 3.

INTRODUCTION
I appreciate the authors adding further detail here to explain the context of research studies on social communication disorders after TBI. Some further editing would assist to develop the clarity of the argument for studying GIST intervention, as the context and links. For example:
- If starting with the Cicerone et al systematic review, it is important to note that this review focuses on literature published in 2003 to 2008 (and therefore does not reflect what has been published in the last decade).
- In the same paragraph as Cicerone et al, the study by Douglas et al (reference 24) (published in 2015) is introduced. The transition to the Douglas paper is unclear.
- You mention that searching in 2017 found 4 additional papers published since 2013, but then only mention one of them.
- In transitioning to focussing on GIST intervention, it is unclear whether these GIST intervention papers were included in the Cicerone or Finch systematic reviews.
- The argument would be more developed logically if you finish summarising the evidence from GIST intervention papers, then introduce your own pilot data. The transition from your pilot data to the Braden paper does not assist you to build your argument for this protocol.

If you need to cut down on the word length, you may consider reducing or deleting the content about the Cicerone et al systematic review, and instead use the more recent and relevant systematic review by Finch et al as a starting point for developing your argument.

Thank you for all your comments on this section. We have decided to omit the content about Cicerone et al., (2011) systematic review, and focus on the more recent systematic reviews from Finch et al. and Togher et al. We have also included the latest Cicerone et al. systematic review (2019; including studies from 2009-2014). In addition, we have removed the Douglas et al., study because this was not that relevant for building our argument. The search we did was actually in done in 2019, and not 2017 as it says in the manuscript, so this has now been corrected. We have also added more information about the other three studies we found on our search. As you suggest, we added a clarification regarding GIST studies reported in the systematic reviews, and restructured the section regarding our

pilot study, and moved this part to the end of the section and building up to our aims for this RCT study.
See page from page 4 line 17 to page 6 line 12.

METHODS
I queried during my last review whether “no concomitant diseases” really meant “no concomitant neurological diseases”. I note this has been edited in the exclusion criteria but not in the inclusion criteria. Also, if the exclusion criterion essentially means the same thing as the inclusion criterion, you do not need to have both.

Thank you for pointing this out. We have now omitted this from the inclusion criteria. As you say, there is no need to mention this twice.

La Trobe Communication Questionnaire – was the self-report or other-report score considered the primary outcome measure? I note that at the end of the secondary outcome measures section, you have noted that people will complete this measure at home independently, but it would be helpful to state this at the start of the Measures section, so that the reader has this information while reading about the LCQ.
Thank you for pointing this out. Yes, the LCQ is the primary outcome measure. We agree that there is a need for clarification. Unlike the other self- and other report forms (secondary outcome measures) the LCQ is completed at the assessment points with the assessor present and available for questions. However, this self-report test does not require the assessor to interpret the answers. This is now clarified on page 11, line 4-6.

The follow-up treatment sessions for Standard GIST at three and six months should be marked on Figure 1 so the reader can see how these treatment sessions relate to the assessment timepoints.
Thank you for commenting on this. We have now marked the follow-up treatment sessions after 3 and 6 months in Figure 1.

DISCUSSION
The additional consideration of strengths and limitations of the research included in this section is helpful and appropriate.

Thank you for all the helpful comments and feedback on our protocol article.

VERSION 3 – REVIEW

REVIEWER	Sara D S Ramos The Disabilities Trust
REVIEW RETURNED	20-Jul-2019

GENERAL COMMENTS	I believe the authors have addressed all of the major concerns raised in previous reviews. The design of the study is now described much more clearly, and the authors discuss the risks of bias introduced by lack of sample stratification in the discussion. However, I would sugest that this is also mentioned in the summary of the study's strengths and limitations. On the section about power, I would suggest that the justification of the sample size in the current study should focus more on whether or not there will be enough power to detect a small to moderate effect with the proposed sample size, instead of simply directly comparing sample numbers with previous studies. For example, when discussing the Harrison-Felix study, the authors could comment on the whether effect size observed in that study

	would be detected with a sample of 60, and whether it was larger or smaller than the expected d = 0.65 for the current study. The authors should also check and clarify whether the following statement applies to the standard versus intensive GIST comparison: “Moderate effects are expected for both treatment groups independent of group allocation” (p. 15). The reflection on the difficulties of making these types of treatment more accessible to those with limited social networks is welcome, but the authors could also give some consideration as to the role of intensive GIST might play in offering rehabilitation to these individuals as well as, more generally, to how different formats of rehabilitation might be more indicated for different individuals and situations. The attached document highlights a small number of typos and unclear wording throughout the document. In my view, the points raised are minor issues that I believe the authors will be able to address in a final version of the manuscript, and I would now recommend the protocol for publication. The reviewer provided a marked copy with additional comments. Please contact the publisher for full details.

REVIEWER	Nicholas Behn City, University of London United Kingdom
REVIEW RETURNED	22-Jul-2019

GENERAL COMMENTS	Thank-you for the changes you have made to the manuscript and the attention to detail you have employed throughout this version. The changes I suggest below are very minor. Abstract: Perhaps mention that the primary outcome is the LCQ (Self). Introduction: Where you have mentioned the four studies you identified since 2013, you could cut your words by simply saying that these four studies provided further evidence for the efficacy of SCD interventions on X, Y, Z. The amount of detail on these four studies is disproportionate to the other reviews and their findings. Methods: No need to use the word “key” for primary and secondary outcome measures unless you are suggesting there are other outcomes you have not described, in which case they should be mentioned (as a protocol suggests). Data analysis: Can you specify the Bonferroni you will apply? Discussion: Another limitation is that you are not doing a health economic analysis that would give you information about the cost-benefit of one intervention over another. This doesn’t need to be mentioned in this paper but something to consider.

	Two small changes to the TIDieR table: (1) On item 2 just a little more information on the essential elements (those areas critical to the delivery of GIST e.g. group, goal setting, involvement of a family member); (2) On item 7 a little more information on the location e.g. where in the outpatient clinic (?small room within a clinic based in a hospital) and same for inpatient group. Grammar: Just double-check as there are punctuation errors (e.g. two full stop marks, missing spaces between words).

REVIEWER	Rachael Rietdijk The University of Sydney, Australia.
REVIEW RETURNED	05-Jul-2019

GENERAL COMMENTS	Thank you for your work in providing this revised version of your manuscript. Most of my queries raised in my last review have now been addressed. I have made a few minor comments relating to increasing clarity: Page 5 - The transition from the conclusions of the systematic reviews and more recent literature, to the specific focus on GIST feels somewhat awkward. It might help to add a sentence on page 5 line 8 noting something like, "Group Interactive Structured Treatment (GIST) is one of the key approaches that incorporates these components (and add the key references)" and then on page 5 line 42, add a sentence which helps link to your next paragraph, stating something like, "This study provides evidence for the efficacy of GIST across different delivery formats". It may also assist the direction of your argument to provide less detail about the other three recently published studies (references 21-23). Page 9 line 31 - "desire preferable" may be replaced with "preference" Page 9 line 35 - "a written informed consent" - "a" not necessary here Page 11 line 3-5 - These two sentences are somewhat confusing. Would it be correct to say something like, "This LCQ is self-completed at each assessment point by participants with ABI, with the assessor present to clarify items if needed"? Page 18 line 5 - "with socially isolated life" perhaps could be replace by "experiencing social isolation"

VERSION 3 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 4

Thank you for your work in providing this revised version of your manuscript. Most of my queries raised in my last review have now been addressed. I have made a few minor comments relating to increasing clarity:
Thank you so much for all your helpful comments.

Page 5 - The transition from the conclusions of the systematic reviews and more recent literature, to the specific focus on GIST feels somewhat awkward. It might help to add a sentence on page 5 line 8 noting something like, "Group Interactive Structured Treatment (GIST) is one of the key approaches that incorporates these components (and add the key references)" and then on page 5 line 42, add a sentence which helps link to your next paragraph, stating something like, "This study provides evidence for the efficacy of GIST across different delivery formats". It may also assist the direction of your argument to provide less detail about the other three recently published studies (references 21-23).
Thank you for your suggestions. We agree that the transition from the literature review to GIST needed to be clarified. In order to meet both the suggestions from you and reviewer no 3, we have rewritten and removed the details from the other three recently published studies, and added the clarifying sentences you suggested on page 5 lines 9-15.

Page 9 line 31 - "desire preferable" may be replaced with "preference"
Page 9 line 35 - "a written informed consent" - "a" not necessary here
Page 11 line 3-5 - These two sentences are somewhat confusing. Would it be correct to say something like, "This LCQ is self-completed at each assessment point by participants with ABI, with the assessor present to clarify items if needed"?
Page 18 line 5 - "with socially isolated life" perhaps could be replace by "experiencing social isolation"
Thank you for pointing these sentences out and for you suggestions. We have changed the sentences accordingly in our manuscript.

Reviewer: 1
I believe the authors have addressed all of the major concerns raised in previous reviews.
The design of the study is now described much more clearly, and the authors discuss the risks of bias introduced by lack of sample stratification in the discussion. However, I would suggest that this is also mentioned in the summary of the study's strengths and limitations.
Thank you so much for all the helpful feedback you have provided regarding our manuscript. We have added the lack of sample stratification to the strengths and limitations in the summary.

On the section about power, I would suggest that the justification of the sample size in the current study should focus more on whether or not there will be enough power to detect a small to moderate effect with the proposed sample size, instead of simply directly comparing sample numbers with previous studies. For example, when discussing the Harrison-Felix study, the authors could comment on the wether effect size observed in that study would be detected with a sample of 60, and whether it was larger or smaller than the expected d = 0.65 for the current study.
Thank you for your comment. In accordance with your suggestion we have now focused the justification of the sample size on the power analysis conducted, i.e., ”sample size of 30 participants in each group, with an expected effect size of d=.65 provides a statistical power of .80.” see page 14 lines 14-21.

The authors should also check and clarify whether the following statement applies to the standard versus intensive GIST comparison: “Moderate effects are expected for both treatment groups independent of group allocation” (p. 15).
In response to this comment we have now omitted this statement. Sufficient information is already provided in the section about power, i.e., prior research on the effect of GIST in ABI has reported moderate to large effect sizes. To detect an effect size of d=0.65 with a power=0.80, 30 patients are needed in each group.

The reflection on the difficulties of making these types of treatment more accessible to those with limited social networks is welcome, but the authors could also give some consideration as to the role of intensive GIST might play in offering rehabilitation to these individuals as well as, more generally, to how different formats of rehabilitation might be more indicated for different individuals and situations.
Thank you for your comment. Whether or not intensive GIST is proven acceptable and equally effective as standard GIST, this will be interesting to discuss both in relation to the potential role intensive GIST might have in offering rehabilitation to individuals with limited social networks, in addition to how the different formats of rehabilitation might be more indicated for different individuals and situations. Intensive GIST might give people with limited social networks the opportunity to work on their social communication skills in a social environment. Nevertheless, we will explore these issues in more detail in later publications. In the present paper, we have commented on this topic in the Discussion on page 18 lines 15-17.

The attached document highlights a small number of typos and unclear wording throughout the document.
Thank you for highlighting typos and unclear wording. We really appreciate this and have made changes accordingly.

In my view, the points raised are minor issues that I believe the authors will be able to address in a final version of the manuscript, and I would now recommend the protocol for publication.
Thank you so much for all your helpful comments.

Reviewer: 3
Thank-you for the changes you have made to the manuscript and the attention to detail you have employed throughout this version. The changes I suggest below are very minor.

Abstract: Perhaps mention that the primary outcome is the LCQ (Self).
Thank you for pointing this out. We have now clarified that the primary outcome is LCQ self-report in the abstract.

Introduction: Where you have mentioned the four studies you identified since 2013, you could cut your words by simply saying that these four studies provided further evidence for the efficacy of SCD interventions on X, Y, Z. The amount of detail on these four studies is disproportionate to the other reviews and their findings.
Thank you for commenting on this. We have decided to follow your suggestion, and cut the details of the other studies. However, we have kept the section about the Harrison-Felix et al. study, as this is important for our protocol. We have further added a sentence to make the transition to further GIST studies clearer. See page 5 lines 9-15.

Methods: No need to use the word “key” for primary and secondary outcome measures unless you are suggesting there are other outcomes you have not described, in which case they should be mentioned (as a protocol suggests).
Agree. We have now omitted the word “key” from the headings.

Data analysis: Can you specify the Bonferroni you will apply?
To adjust for multiple comparisons the Holm-Bonferroni procedure will be applied. See page 15 line 20.

Discussion: Another limitation is that you are not doing a health economic analysis that would give you information about the cost-benefit of one intervention over another. This doesn’t need to be mentioned in this paper but something to consider.
Thank you for commenting on this. This is an interesting point of view that we will consider when discussing the results from this study in later publications.

Two small changes to the TIDieR table: (1) On item 2 just a little more information on the essential elements (those areas critical to the delivery of GIST e.g. group, goal setting, involvement of a family member); (2) On item 7 a little more information on the location e.g. where in the outpatient clinic (?small room within a clinic based in a hospital) and same for inpatient group.
Thank you for commenting on this. I have now added more details to the TIDieR table as you suggest. See item 2 and 7.

Grammar: Just double-check as there are punctuation errors (e.g. two full stop marks, missing spaces between words).
Thank you for pointing this out. We have detected that some of the errors are a result from transforming the manuscript to a plain text document without endnote links. We have now double-checked for punctuation errors in the plain text version.