PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

| TITLE (PROVISIONAL) | Cognitive Training and Neuroplasticity in Mild Cognitive Impairment (COG-IT): Protocol for a two-site, blinded, randomized, controlled treatment trial. |
|---------------------|---|
| AUTHORS | D'Antonio, Jessica; Simon-Pearson, Laura; Goldberg, Terry; Sneed, Joel; Rushia, Sara; Kerner, Nancy; Andrews, Howard; Hellegers, Caroline; Tolbert, Sierra; Perea, Elena; Petrella, Jeffrey; Doraiswamy, Murali; Devanand, Davangere |

VERSION 1 - REVIEW

| REVIEWER | Jessica Peter |
|-----------------|---|
| | University Hospital of Old Age Psychiatry and Psychotherapy |
| REVIEW RETURNED | 28-Jan-2019 |

| GENERAL COMMENTS | I would say that the sample size is too small to apply all these statistical tests (particularly for moderation analyses). The figures are missing what about checking visus and motor functioning in all participants? what about computer skills? The authors use a very long list of neuropsychological tests. How do they prevent fatigue in the participants? What about missing data handling? |
|------------------|--|
| | - Do they apply parallel versions of all these tests at different time- points? |

| REVIEWER | Kee-Hong Choi Korea University |
|-----------------|-----------------------------------|
| REVIEW RETURNED | 23-Feb-2019 |

| GENERAL COMMENTS | The current NIMH funded study, entitled "Cognitive training and neuroplasticity in mild cognitive impairment: clinical rational and study design a two-site, blinded, randomized, controlled treatment trial" was well designed, and thoroughly described its rationale, methods, procedures, importance/significance and strengths and limitations. A few minor issues are provided for improving clarity. 1. Power analyses need more detailed information (e.g., effect sizes from referential studies) for justification of a total sample size |
|------------------|--|
| | sizes from referential studies) for justification of a total sample size of 100 participants. |

| Please describe expected drop-out rates over the 2.5 year of participation, and assumptions about missingness for proper statistical analyses and interpretation of results. Lumosity memory training appears largely non-verbal (e.g., memory matrix, tidal treasures, trouble brewing), except familiar faces. Would there be any limitations about limited verbal learning and memory training and measuring its domain without sufficient training? Please describe whether and how to track types of games/amount of times that each participant practices over the study. There is a possibility that some participants would avoid games targeting their lowered cognitive functioning, and instead engage more in games that they could do well. |
|--|
| 4. The crossword puzzle game appears to provide stimulation for some cognitive functioning (e.g., word generation related executive functioning, attention). This might also be considered when prediction of results. |
| |

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1

Reviewer Name: Jessica Peter

Institution and Country: University Hospital of Old Age Psychiatry and Psychotherapy

Please state any competing interests or state 'None declared': no conflicts of interest

Please leave your comments for the authors below

1. I would say that the sample size is too small to apply all these statistical tests (particularly for moderation analyses).

We have two primary outcome measures (i.e., multiple outcome measures), namely ADASCog and the UPSA. For multiple outcome measures, statistical significance on any one measure is meaningful and there is no need to correct for multiple comparisons (unlike coprimary outcome measures). All other outcome measures are secondary and exploratory.

Corresponding Manuscript Change:

The above language was placed verbatim in the manuscript on page 17.

2. The figures are missing.

Figure 1 is included in this submission for review.

3. What about checking visus and motor functioning in all participants?

We assume that visus is a typo meant to mean visual. During the screen, the study physician completes the Cumulative Illness Rating Scale for Geriatrics to document medical history. Additionally, participants with motor disorders, like Parkinson's disease, are excluded. Participants who have visual problems wear corrective lenses for the testing; this is documented.

Corresponding Manuscript Changes:

Page 8: "Patients who have a history of major psychiatric or neurological illness including motor disorders like Parkinson's disease, a dementia diagnosis of any type, contraindication to MRI scan, lack of English-speaking ability, or have been defined as regular online brain training or regular crossword puzzle users (\geq 2 times per week in the past year) will be excluded."

Page 14 addition: "When a participant wears corrective lenses during the testing battery, this is documented in the participant's research chart."

4. What about computer skills?

Participants who are screened are not required to have any particular level of computer skills for study inclusion. During the training session at the baseline visit, all participants are administered training on how to use the computer, how to successfully access the Lumosity website, and review basic trouble shooting techniques. The informant may also help the participant in using the computer at home, but not with the training sessions.

Corresponding Manuscript Change:

Page 10 addition: "Participants are not required to have any particular level of computer skills for study inclusion; however, at the initial baseline training, all participants will be trained on how to successfully access the training platform, and how they could obtain help both from research staff and their informant throughout the study."

5. The authors use a very long list of neuropsychological tests. How do they prevent fatigue in the participants? What about missing data handling?

Testing fatigue is mitigated by allowing participants to take breaks during the testing. If there is missing data from one time point, the study team will attempt to bring the participant back to the clinic within the allowed window to complete the missing measures. Missing data is managed statistically through use of mixed model repeated measures analyses.

Corresponding Manuscript Changes:

Page 11 addition: "Testing fatigue is mitigated by allowing participants to take breaks during the testing. If there is missing data from one time point, the study team will attempt to bring the participant back to the clinic within the allowed window to complete missing measures."

Page 17 addition: "Missing data is managed statistically through use of mixed model repeated measures analyses."

6. Do they apply parallel versions of all these tests at different time-points?

Yes, for word learning lists, the neuropsychological testing materials provide different but parallel word lists, so as to avoid practice effects in MMSE and ADAS-Cog, but not for AVLT. With respect to the latter we did not adopt this approach because we were concerned that different forms have not been established as equivalent in difficulty level.

Corresponding Manuscript Change:

Page 14: "For word learning lists, the neuropsychological testing materials provide different but parallel word lists, so as to avoid practice effects in MMSE and ADAS-Cog, but not for AVLT. With respect to the latter we did not adopt this approach because we were concerned that different forms have not been established as equivalent in difficulty level."

Reviewer: 2

Reviewer Name: Kee-Hong Choi

Institution and Country: Korea University

Please state any competing interests or state 'None declared': None declared.

Please leave your comments for the authors below

The current NIMH funded study, entitled "Cognitive training and neuroplasticity in mild cognitive impairment: clinical rational and study design a two-site, blinded, randomized, controlled treatment trial" was well designed, and thoroughly described its rationale, methods, procedures, importance/significance and strengths and limitations. A few minor issues are provided for improving clarity.

1. Power analyses need more detailed information (e.g., effect sizes from referential studies) for justification of a total sample size of 100 participants.

We powered our trial to detect an effect size at 18 months of d=.58 (80% power). This effect size is more conservative than published treatment changes associated with CCT (for instance, Buschert et. al, 2011).

Corresponding Manuscript Changes:

Page 16 addition: "We powered our trial to detect an effect size at 18 months of d=.58 (80% power). This effect size is more conservative than published treatment changes associated with CCT (for instance, see [39]).

2. Please describe expected drop-out rates over the 2.5 year of participation, and assumptions about missingness for proper statistical analyses and interpretation of results.

We assume that dropout is distributed uniformly across waves of follow-up assessments (with 5% attrition between each consecutive pair of the 5 major time-points, i.e. 20% by 18 months).

Corresponding Manuscript Changes:

Page 16: "We assume that dropout is distributed uniformly across waves of follow-up assessments (with 5% attrition between each consecutive pair of the 5 major time-points, i.e. 20% by 18 months)."

3. Lumosity memory training appears largely non-verbal (e.g., memory matrix, tidal treasures, trouble brewing), except familiar faces. Would there be any limitations about limited verbal learning and memory training and measuring its domain without sufficient training?

In the Lumos Games battery, we included some verbal tasks, like Word Bubbles (Verbal Fluency) Word Snatchers (Vocabulary proficiency) and Editor's Choice (Vocabulary proficiency).

Corresponding Manuscript Change:

Page 9: "Verbal Fluency and Vocabulary Proficiency tasks were included to promote verbal learning in the CCT group."

4. Please describe whether and how to track types of games/amount of times that each participant practices over the study. There is a possibility that some participants would avoid games targeting their lowered cognitive functioning, and instead engage more in games that they could do well.

Participants assigned to the Lumosity games receive a suite of 6 games randomly from the prefixed 18 set of games chosen by the study team. Once they begin a session, they do not choose the games, and are not give the opportunity to skip over or change the suite of games.

To track type of games/crossword puzzles and amount of time that the subject spends doing the games/crossword puzzles, only unblinded study coordinators receive reports from Lumosity each week. If the Lumosity reports of computer games/crosswords access do not match the subject's assigned instructions, the unblinded coordinator then contacts the subject to guide and ensure adherence to the protocol.

Corresponding Manuscript Changes:

Page 9: "Participants in the CCT condition are not allowed to choose the games, and are not allowed to skip over or change the suite of games."

Page 12: "To track type of games/crossword puzzles and amount of time that the subject spends doing the games/crossword puzzles, only unblinded study coordinators receive reports from Lumosity each week. If the Lumosity reports of computer games/crosswords access do not match the subject's assigned instructions, the unblinded coordinator then contacts the subject to guide and ensure adherence to the protocol."

5. The crossword puzzle game appears to provide stimulation for some cognitive functioning (e.g., word generation related executive functioning, attention). This might also be considered when prediction of results.

We agree with this point and that is why we consider crossword puzzles to be an active control condition. Nevertheless, while crosswords engage primarily verbal abilities and perhaps executive and attentional mechanisms, the Lumosity games target different cognitive domains, such as speed of processing and memory, in addition to the verbal games mentioned above. The effect that these different trainings have on the so-called far transfer problem will of course be of major interest.

Corresponding Manuscript Changes:

Page 10 addition: "Crosswords engage primarily verbal abilities and perhaps, executive and attentional mechanisms. The Lumosity games target different cognitive domains, such as speed of processing and memory, as well as verbal abilities. The effect that these different trainings have on the so-called far transfer problem will of course be of major interest."

Figures 2, 3, and 4 in the prior version were partly redundant with the description in the text. Further, these figures did not add significant or valuable information to the manuscript, so they have been deleted in this revision.

VERSION 2 – REVIEW

| REVIEWER | Jessica Peter University Clinic of Old Age Psychiatry and Psychotherapy |
|-----------------|--|
| REVIEW RETURNED | 19-May-2019 |

| GENERAL COMMENTS | The authors have adequately addressed my concerns. |
|------------------|--|
|------------------|--|

| REVIEWER | Kee-Hong Choi Korea University, Seoul, Korea; Wesleyan University, Middletown, CT |
|-----------------|---|
| REVIEW RETURNED | 02-May-2019 |

| GENERAL COMMENTS | The authors properly addressed and clarified issues raised. One thing that has not been fully addressed is # 3: Lumosity memory training appears largely non-verbal (e.g., memory matrix, tidal treasures, trouble brewing), except familiar faces. Would there be any limitations about limited 'verbal learning and memory' training and measuring its domain without sufficient training? |
|------------------|---|
| | The authors responded that they included some verbal tasks, like Word Bubbles (Verbal Fluency) Word Snatchers (Vocabulary proficiency) and Editor's Choice (Vocabulary proficiency). However, those games seem not targeting on verbal memory and learning (but verbal fluency). This issue should be re-addressed. |

VERSION 2 – AUTHOR RESPONSE

The authors properly addressed and clarified issues raised. One thing that has not been fully addressed is # 3: Lumosity memory training appears largely non-verbal (e.g., memory matrix, tidal treasures, trouble brewing), except familiar faces. Would there be any limitations about limited 'verbal learning and memory' training and measuring its domain without sufficient training?

The authors responded that they included some verbal tasks, like Word Bubbles (Verbal Fluency) Word Snatchers (Vocabulary proficiency) and Editor's Choice (Vocabulary proficiency). However, those games seem not targeting on verbal memory and learning (but verbal fluency). This issue should be re-addressed.

While there are games that engage verbal fluency or vocabulary proficiency (Word Bubbles, Word Snatchers, Editor's Choice, Continuum), the key episodic memory task, Familiar Faces, involves recall of newly acquired verbal information. There is an episodic memory task involving visuospatial information as well, which is entitled Memory Matrix. Nevertheless, we now acknowledge in the Treatment Regimen section of the manuscript that episodic memory training may be somewhat limited.

Corresponding changes in text:

Page 10: "Verbal Fluency and Vocabulary Proficiency tasks were included to promote verbal fluency in the CCT group. Further, the episodic memory task, Familiar Faces, targets verbal memory and learning. With this, it is acknowledged that episodic memory training may be somewhat limited in the selected battery of modules provided by Lumosity.

We have edited Table 2 to include "Verbal Memory and Learning" in the targeted cognitive domain column for Familiar Faces. Please see below:

Table 2. Complete list of CCT game battery and associated cognitive domains (provided by Lumos Labs)

| Game Name | Cognitive Domain |
|------------------|--|
| Tidal Treasures | Working Memory i.e., delayed, non-matching to sample; self-ordered pointing |
| | i.e., delayed, non-matching to sample, sen-ordered pointing |
| Speed Match | Processing Speed |
| Color Match | Response Inhibition |
| Word Bubbles | Verbal Fluency |
| Train of Thought | Planning |
| | Divided Attention |
| | Multiple attractions |
| Familiar Faces | Episodic Memory; verbal memory and learning |
| Memory Matrix | Episodic Memory; Visuospatial memory |

| Lost in Migration | Visual Interference |
|-------------------|--|
| Brain Shift | Task Switching |
| Trouble Brewing | Multitasking, divided attention, sustained attention, planning, working memory |
| Ebb and Flow | Task switching, semantic and visual interference |
| Masterpiece | Mental rotation; visualization; spatial reasoning |
| River Ranger | Identification |
| Word Snatchers | Vocabulary proficiency |
| Speed Pack | Visualization |
| Disillusion | Task Switching |
| Editor's Choice | Vocabulary Proficiency |
| Continuum | Vocabulary Proficiency |

VERSION 3 - REVIEW

| REVIEWER | Kee-Hong Choi |
|-----------------|-------------------------------|
| | Korea University, South Korea |
| REVIEW RETURNED | 19-Jun-2019 |

| GENERAL COMMENTS | The revision properly addressed the issue raised. |
|------------------|---|