PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Remote Cognitive Assessment of Older Adults in Rural Areas by Telemedicine and Automatic Speech & Video Analysis: protocol for a cross-over feasibility study
AUTHORS	König, Alexandra; Zeghari, Radia; Guerchouche, Rachid; Duc Tran, Minh; Bremond, François; Linz, Nicklas; Lindsay, Hali; Langel, Kai; Ramakers, I.H.G.B; Lemoine, Pascale; Bultingaire, Vincent; Robert, Philippe

VERSION 1 – REVIEW

REVIEWER	Dham, Pallavi University of Toronto
REVIEW RETURNED	26-Jan-2021

GENERAL COMMENTS	Thank you for the opportunity to review this protocol which evaluates video-conference mode for cognitive assessment empowered by speech and video analysis. The authors should be commended for coming up with a comprehensive and detailed protocol that incorporates speech and video analysis. There are a few minor aspects which could be looked into: 1. Could the authors elaborate on the method for randomisation. 2. Could they specify if all tests they use have at least two different versions for f/F and video administration? 3. Authors talk about comprehensive neuro-cognitive battery. However, there are no tests for constructional tasks. It may help to mention this and specify the cognitive domains covered as well as how they make up for domains not covered such as constructional tasks. 4. Authors are encouraged to add its relevance to the current pandemic -this is crucial and an interesting aspect for future use. 5. The objectives mention specifically the assessment of acceptability by quantitative and qualitative methods. However anaylsis is focused on feasibility and reliability. There is no mention of how the feedback questionnaire and focus group data will be analysed and if there are other aspects of data collection which may add to this.

REVIEWER	Mars, Maurice University of Kwazulu-Natal, TeleHealth
REVIEW RETURNED	15-Feb-2021

GENERAL COMMENTS	The protocol describes a multi-component study investigating methods of assessing cognitive impairment in older people in rural areas. The study aims at evaluating the feasibility, reliability and

	users' experience and acceptance of assessment using a specifically developed 'videoconference system tool'. A concordance study will be undertaken to compare face to face testing with testing by videoconference using a battery of tests. The study's novelty appears to be the development and evaluation of automated analysis of video and audio recorded during the videoconference cognitive assessment interviews/assessments. This, it is hoped, will supplement assessment by videoconference and overcome perceived shortcomings of videoconferencing. The point is made in the introduction that clinicians have difficulty in extracting non-verbal cues of patients' emotional state during videoconferenced assessments. This was well-documented in the early telepsychiatry literature, and the solution has been to use the pan, tilt, and zoom features of videoconferencing unit cameras. The move to desktop type videoconferencing using Skype, Zoom, etc., and the built-in cameras of laptop computers, tablets, and smartphones removes this option. Insufficient information is provided on the features of the videoconference system tool to be used in the study other than what is shown in Figure 1. Is pan, tilt and zoom camera control possible, or is one of the unstated purposes of this study to overcome the absence of pan, tilt and zoom control by implementing audio and video analysis? The title of the paper refers to a randomised controlled study. According to the methods, 60 older adults will be recruited from a given region. No details of randomisation are provided, nor is there information about a control group. As set out in the protocol, the 60 participants will serve as their own controls in a concordance study. The only randomisation described relates to who is first assessed face to face who is first assessed by videoconference. There is no difference in the content of the interaction, only the mode of interaction. The title appears to be misleading. Does this study meet the definition of a randomised controlled study? User experience and acceptability of the videoconferenced evaluations will be assessed by questionnaire and focus group meetings. No mention is made of validation of the questionnaire. There is an aspect of data security and anonymity that needs further explanation. The video data will be used for the analysis of facial behaviours and activities. How will the video data be anonymised/de-identified such that the technicians involved will not be able to identify the subject? The paper requires minor English editing.

VERSION 1 – AUTHOR RESPONSE

I. Reply to Reviewer 1 :

We would like to thank the reviewer for his constructive feedback and helpful comments.

1. Could the authors elaborate on the method for randomisation.

We thank the reviewer for pointing this out. We corrected the mistake of using the wrong term 'randomisation' and described in more detail the 'cross-over' design of our study. To reduce learning effect biases we tested both conditions with an alternate order across our sample and alternative versions of tests.

- title :

"Remote Cognitive Assessment of Older Adults in Rural Areas by Telemedicine and Automatic Speech & Video Analysis: protocol for a randomized controlled feasibility study "

"Remote Cognitive Assessment of Older Adults in Rural Areas by Telemedicine and Automatic Speech & Video Analysis: protocol for a cross-over feasibility study "

- Abstract Page 2:

"The administration procedure will be randomized. "

"The order of administration procedure will be counterbalanced so half of the sample starts with the videoconference condition and the other half the face-to-face condition."

- Page 8:

"For this, a randomized controlled feasibility study will be performed in the rural "

"For this, a counterbalanced cross-over controlled feasibility study will be performed in the rural areas "

- Page 10:

"This procedure will be randomized so that half of the participants will experience first the face-to-face and the other half the videoconference administration at first "

"To reduce learning effect biases and within rater variability, this procedure will be counterbalanced so that half of the participants will experience first the face-to-face and the other half the videoconference administration at first"

2. Could they specify if all tests they use have at least two different versions for f/F and video administration?

We specified that for the study we use different parallel versions of tests.

Page 10:

"For the following tests, we will use parallel versions in order to avoid a learning effect: MMSE, Free and Cued Selective Reminding Test, Digit span test, Denomination task, and Semantic and Phonemic fluencies."

3. Authors talk about comprehensive neuro-cognitive battery. However, there are no tests for constructional tasks. It may help to mention this and specify the cognitive domains covered as well as how they make up for domains not covered such as constructional tasks.

We thank the reviewer for pointing this out. The study is limited to use the video modality to verify test performances, which makes it complicated for constructional tasks that require drawing. However, there is actually a contructional task within the MMSE test. Subjects will be asked to perform the task on a white sheet and display it in front of their camera so the clinician can evaluate it remotely. We added an explanation on page 10.

We did not include other constructional tasks such as Rey's figure as we chose the most frequently used cognitive tests in a classical cognitive assessment for dementia. We covered memory (FCSRT, 5 words of Dubois, door test), executive functions (STROOP, Phonemic Verbal Fluency), working memory (Digit Span), Global functioning task (MMSE), language (naming tasks, Semantic Verbal Fluency), praxis. In a next step, we plan to implement constructional tasks such as Clock drawing to be performed either via a connected tablet or directly on the screen via a the computer mouse.

4. Authors are encouraged to add its relevance to the current pandemic -this is crucial and an interesting aspect for future use.

We appreciate the reviewer's comment on the relevance of this project to the current pandemic and added this in the conclusion section as well as in the introduction:

Page 5:
"With the current COVID19 pandemic crisis, now even more than ever, technical solutions such as telemedicine platforms are of great importance to provide isolated elderly people with timely and sufficient access to healthcare."
Page 21:
"Pushing the use of such remote solutions in the future is of particular relevance given the current context of the COVID19 pandemic."

5. The objectives mention specifically the assessment of acceptability by quantitative and qualitative methods. However anaylsis is focused on feasibility and reliability. There is no mention of how the feedback questionnaire and focus group data will be analysed and if there are other aspects of data collection which may add to this.

We agree with the reviewers comment and added a section on how feedback data from the questionnaires and the focus groups will be analyzed.

Under the section 'Study protocol' on Page 11:
"Regarding acceptability, participants of the study will be asked to complete a questionnaire on their experience and acceptability of the videoconference-administered assessment (compared to the classical) and a subgroup, as well as other stakeholders, will be invited to participate in a focus group and semi-structured qualitative interviews in order to assess, in more depth, their user experiencee feasibility and the usability of the system

Under the section 'Data collection, management and analysis" on page 14

"Acceptability evaluation
All participants will be asked to answer a questionnaire (copy under supplementary file) on the acceptance of the videoconference as well as of the face-to-face modality for cognitive testing including 7 questions with a response ranging from 1 to 7, where 1=I strongly disagree and 7=I strongly agree. This questionnaire assesses the user experience, including an overall evaluation, if participants are satisfied, if they want to repeat the experience, attitudes and clarity of instructions as well as what type of method is preferred and why., and what could be improved. After each question, participants have the option to add a comment.

The three following open questions are included at the end of the questionnaire : What was missing or disappointing in your experience? What do you like most/least about this procedure? What is the one thing we could do to make it better?
Descriptive statistics will be performed on the obtained scores. Qualitative analysis will be applied on the comments and written answers to the questionnaires as well as on feedback provided during informal focus group discussions with participants in order to define encountered problems and points of improvement of the system."

II. Reply to Reviewer 2 :

We appreciated very much the feedback of the reviewer and tried to answer all his comments in the revised version of the manuscript.

1. The point is made in the introduction that clinicians have difficulty in extracting non-verbal cues of patients' emotional state during videoconferenced assessments. This was well-documented in the early telepsychiatry literature, and the solution has been to use the pan, tilt, and zoom features of videoconferencing unit cameras. The move to desktop type videoconferencing using Skype, Zoom, etc., and the built-in cameras of laptop computers, tablets, and smartphones removes this option. Insufficient information is provided on the features of the videoconference system tool to be used in the study other than what is shown in Figure 1.

Is pan, tilt and zoom camera control possible, or is one of the unstated purposes of this study to overcome the absence of pan, tilt and zoom control by implementing audio and video analysis?
We understand the reviewer's comment and will explain our preference for integrated cameras rather than the Pan-tilt-zoom.The implemented telemedical tool is web-based and both clinician and patient connect to the platform through existing web browsers (chrome, firefox, safari, edge) or any web browser supporting the webRTC standard. This web-based system runs under any operating system supporting webRTC. Thus we could only use the digital zoom and not the PTZ cameras. The goal is to provide a simple and easy to use platform without the necessity of complex infrastructure and heavy-to-use materials and devices.

Since the communication between the clinician and the patient is made through the web by connecting to an interface on a web browser, the built-in webcam is used when a laptop or a tablet are employed, or a webcam plugged to acomputer. The telemedicine tool was completely developed by us, using common web development technologies (mainly javascript, webRTC, html) and some web development libraries. No skype, zoom or any other existing videoconferencing systems are involved.

We added a detailed description of the developed system to the manuscript under the section 'Technical Description of the videoconference system ' on page 12:

"Technical Description of the videoconference system
The videoconference system (or telemedicine) tool was internally developed as a web-based platform, using common web-development technologies and libraries (JavaScript, Node.JS, HTML…). No skype, zoom or any other existing known videoconferencing systems are involved. A secured server allows connecting two clients (clinician and patient) through the two interfaces described in Figure 1. Both clinician and patient connect to the platform through existing web browsers (chrome, firefox, safari, edge) or any web browser, under any Operating System, supporting the webRTC standard. WebRTC[1] (Web Real-Time Communication) is a free, open-source project providing web browsers and mobile applications with real-time communication (RTC) via simple application programming interfaces (APIs). It allows audio and video communication to work inside web pages by allowing direct peer-to-peer communication, eliminating the need to install plugins or download native apps. Since the communication between the clinician and the patient is directly made through the web, the used devices (camera, microphone and speakers) are either the built-in devices when a laptop is used, or external plugged ones to the PC (in the case of a desktop computer for example). When a tablet or a smartphone is used the integrated devices are used.

For practical reasons, we will use a laptop or a desktop with a minimum of 17 inch screen wide. This will make the clinical tests' contents (images and text) visible for the patients. In addition, we will use a wide-angle camera, especially for the patient's side, in order to allow the clinician to see the upper body of the patient to be able to observe gestures. This is particularly important for some clinical tests, in which seeing the patient's hands is required.
The videoconferencing communication requires a dedicated server to allow the transmission of different information between the two clients (clinician and patient). Servers to store different data

(database, patients' information, videos, speech recordings, scores …); and to run the different services allowing the videoconferencing communication are mandatory. All the servers could be hosted in dedicated and regulated infrastructure such as the ones of the clinical partners, and thus respect the legislation related to health data and privacy.
The developed system is linked to third parties cloud infrastructures allowing speech and video analysis. For the planned clinical study, audio and video data will be stored on secured servers. The processing of this data is done according to the procedure explained under Data processing.
The use of speech and video analysis, in addition to providing potential digital biomarkers, helps in overcoming both the physical-absence of the patient and the non-use of sophisticated and complex observation devices (such as Pan-Tilt-Zoom "PTZ" cameras). By providing the clinician with meaningful information about the patient's behaviours and state (comfort, fatigue, stress), the physical distance can be potentially compensated for."

2.The title of the paper refers to a randomised controlled study. According to the methods, 60 older adults will be recruited from a given region. No details of randomisation are provided, nor is there information about a control group. As set out in the protocol, the 60 participants will serve as their own controls in a concordance study. The only randomisation described relates to who is first assessed face to face who is first assessed by videoconference. There is no difference in the content of the interaction, only the mode of interaction. The title appears to be misleading. Does this study meet the definition of a randomised controlled study?

We thank the reviewer for pointing this out. We corrected the mistake of using the wrong term 'randomisation' and described in more detail the 'cross-over' design of our study. To reduce learning effect biases we tested both conditions with an alternate order across our sample and alternative versions of tests. We corrected the title, the abstract and the manuscript.

- Abstract Page 2:

"The administration procedure will be randomized. "
"The order of administration procedure will be counterbalanced so half of the sample starts with the videoconference condition and the other half the face-to-face condition."

- Page 8:

"For this, a randomized controlled feasibility study will be performed in the rural "
"For this, a counterbalanced cross-over controlled feasibility study will be performed in the rural areas "

- Page 10:

"This procedure will be randomized so that half of the participants will experience first the face-to-face and the other half the videoconference administration at first "
"To reduce learning effect biases and within rater variability, this procedure will be counterbalanced so that half of the participants will experience first the face-to-face and the other half the videoconference administration at first"

3.User experience and acceptability of the videoconferenced evaluations will be assessed by questionnaire and focus group meetings. No mention is made of validation of the questionnaire.

We added a more detailed description of the questionnaire, its origin and the planned analysis on page page 11 :
"Regarding acceptability, participants of the study will be asked to complete a questionnaire on their experience and acceptability of the videoconference-administered assessment (compared to the classical) and a subgroup, as well as other stakeholders, will be invited to participate in a focus group and semi-structured qualitative interviews in order to assess, in more depth, the feasibility and the eas ease and usability of the system."

Page 14:
“Acceptability evaluation
All participants will be asked to answer a questionnaire (copy under supplementary file) on the acceptance of the videoconference as well as of the face-to-face modality for cognitive testing including 7 questions with a response ranging from 1 to 7, where 1=I strongly disagree and 7=I strongly agree. This questionnaire is based on the ‘System Usability Scale’[31] and assesses the user experience, including an overall evaluation, if participants are satisfied, if they want to repeat the experience, attitudes and clarity of instructions as well as what type of method is preferred and why., and what could be improved. After each question, participants have the option to add a comment.

The three following open questions are included at the end of the questionnaire : What was missing or disappointing in your experience? What do you like most/least about this procedure? What is the one thing we could do to make it better?
Descriptive statistics will be performed on the obtained scores. Qualitative analysis will be applied on the comments and written answers to the questionnaires as well as on feedback provided during informal focus group discussions with participants in order to define encountered problems and points of improvement of the system.”

4.There is an aspect of data security and anonymity that needs further explanation. The video data will be used for the analysis of facial behaviours and activities. How will the video data be anonymised/de-identified such that the technicians involved will not be able to identify the subject?

We added the following paragraph to the paper to explain how security and the pseudo-anonymity of the speech and video data will be handled.

“Concretely, an IT technician (or engineer) working for the Institut Claude Pompidou, and authorised to access to all patients data as part of his duties (creating patients records, correcting informations about patients…) will be the only non-clinical person who can access to the identifying data of the participants (speech and video recordings). This IT technician will be trained by the technical partners (using similar type of non-confidential data and not belonging to the participants), on how to use the different softwares -provided by the technical partners, to generate pseudo-anonymous metadata : low level features extracted form speech and videos such as signal’s intensities, acoustic characteristics, 2D points positions, head/eyes positions in different images…. These pseudo-anonymised metadata do not contain the identity of the participants. For each participant, a random code will be assigned, known only by the IT technician and the clinicians involved in the study. All the metadata extracted by the different softwares and executed only by the IT technicien, will be then transmitted in a pseudo-anonymous manner : the technical partners do not know the identity of the participants, span style="font-family:Arial; font-style:italic; color:#2f5496">and have access to only a set of metadata per à random and unidentifiable code.
In order to perform the processing of the data, other pseudo-anonymous data could be transmitted to the technical partners such as tests’ scores and values of different clinical scales. .”

5.The paper requires minor English editing.

We had the paper proof read by a native English speaker.

VERSION 2 – REVIEW

REVIEWER	Mars, Maurice University of Kwazulu-Natal, TeleHealth
REVIEW RETURNED	17-May-2021

GENERAL COMMENTS	The previous comments have been addressed. Their resolution raises further questions. It would appear that most subjects in the videoconferenced component of the study will come to a central site and will not be using a computer or tablet device in their home setting. This is not clearly stated in the protocol. For those in rural settings, a van equipped with the necessary computer and connectivity will be used. The authors note that the patients will have access to a 17 inch screen and that a wide angle camera will be used. This will then limit the interpretation of the worth of the method as 17 inch screens and wide angle cameras are not the norm for laptop or tablet computers. The paper does not describe in sufficient detail how the information gained in the focus group discussions will be gathered and analysed. Will the discussions be recorded and transcribed? Will thematic analysis be used? An IT technician has access to all patient data. Was this person required to sign a confidentiality agreement as is generally required when non-medical personnel are present during a videoconference consultation? Was there a technician in the mobile van when the consultations were performed? Did he/she sign a confidentiality agreement? I note that the instructions for reviewers of study protocols states that, 'If data collection is complete, we will not consider the manuscript." According to the information provided in the "dates of the study" section, the study commenced in September 2019 and data collection has been completed. This renders review and correction of the protocol meaningless. If, for example, the technicians involved in the data analysis and the consultation in the mobile van did not sign a confidentiality agreement, will the authors repeat the study?

VERSION 2 – AUTHOR RESPONSE

I. Reply to Reviewer 1 :

We would like to thank the reviewer for his constructive feedback and helpful comments.

1.It would appear that most subjects in the videoconferenced component of the study will come to a central site and will not be using a computer or tablet device in their home setting. This is not clearly stated in the protocol.

For those in rural settings, a van equipped with the necessary computer and connectivity will be used. The authors note that the patients will have access to a 17 inch screen and that a wide angle camera will be used. This will then limit the interpretation of the worth of the method as 17 inch screens and wide angle cameras are not the norm for laptop or tablet computers.

We thank the reviewer for this comment. The first goal of our clinical study is to investigate the feasibility of remote assessments. Since we developed the telemedicine tool from scratch and implemented several clinical tests, the first step is to validate these tests and to assess the acceptability among the users. The implemented clinical tests are normalised tests used by neuropsychologists and related medical professionals. We have 3 types of clinical tests :

(1) tests which do not require sharing any visual contents (such as verbal fluency). For these tests, any device can be used (smartphone, tablet, PC, laptops with the integrated webcams, microphones and speakers).

(2) tests which require sharing visual contents (images, words, pictures …). For these tests, the content should be visible to the patient, and using small devices such as smartphones can lead to bad perception of the visual contents. In the face-to-face clinical assessments, we use pen and paper, and the size of the visual contents are normed with a minimum size (size of a picture, size of a word or a figure). For this reason we recommend a minimum size of 10 inch for the screen of the used device, so some tablets can be used, in addition to laptops and PCs; however for such tests, a smartphone is not recommended.

(3) for few tests, the medical professional needs to see the hands of the patient and generally the upper body (such as psychometric tests); for such tests a wide-angle camera is needed.

We did add these details to the paper (page 9&12).

One of the main goals of developing this telemedicine tool is to fight against medical deserts in rural areas, and reach subjects living in such isolated regions when it comes to clinical trials. This telemedicine tool allows to bring firstline screening of cognitive impairment to these areas, with dedicated materials and infrastructures.

We just want to point out that the telemedicine tool is developed for cognitive disorders and validated, in this study, for neurocognitive assessments in elderly. As with any new telemedicine tool, some constraints related to the pathology and the functioning of the consultations has to be respected. To make our system usable in several conditions, we proposed 2 scenarios for the clinical study: the first one is to validate the telemedicine tool in a clinical setup when patients/subjects can move to close-by places, where they have access to the needed infrastructures (internet connexion, device with a webcam ...); the second one is to move close-by the homes of isolated patients/subjects (e.g., in rural areas) with an equipped mobile unit.

2. The paper does not describe in sufficient detail how the information gained in the focus group discussions will be gathered and analysed. Will the discussions be recorded and transcribed? Will thematic analysis be used?

We added a more detailed description of the analysis we aim to perform of the focus group discussions. The discussions will be recorded and parts of it transcribed (mainly the different responses). We will then perform thematic analysis on the transcripts of the different responses.

Page 14

“Focus group discussions with some participants will be recorded and transcribed. We will then rearrange the comments so that answers are together for each interview question. For each question we will note the main ideas that occur in the answers. Recurring main ideas will be used to identify themes which in turn will be illustrated by quotations. The analysis results will be described in a narrative rrepop will be thematic analysis on the transcripts of the different responses as well as on feedback provided during informal focus group discussions with participants to presenting the user experiences and define encountered problems and points of improvement of the system.”

3. An IT technician has access to all patient data. Was this person required to sign a confidentiality agreement as is generally required when non-medical personnel are present during a videoconference consultation? Was there a technician in the mobile van when the consultations were performed? Did he/she sign a confidentiality agreement?

The technician has only access to anonymised clinical data, he is never present during the video consultations which take place only with the psychologist. Data analysis will be performed by the psychologists. Regarding the mobile van, the technician is only there before the consultation to ensure the technical equipment is working correctly, then he leaves the van. We added this clarification to the manuscript (page 18). The technician did sign a confidentiality agreement additionally.

4. I note that the instructions for reviewers of study protocols states that, ``If data collection is complete, we will not consider the manuscript.” According to the information provided in the “dates of the study” section, the study commenced in September 2019 and data collection has been completed. This renders review and correction of the protocol meaningless. If, for example, the technicians involved in the data analysis and the consultation in the mobile van did not sign a confidentiality agreement, will the authors repeat the study?

We thank the reviewer for pointing this out. Due to the COVID pandemic the study had an important delay in the recruitment. Participants could not come for a long period to the hospital for the face-to-face assessments, and additionally we wanted to avoid all unnecessary risk since the participants represent a high risk population. This means the study is actually still ongoing and we are planning a second round for the mobile van part in autumn 2021. We corrected this on page 2. The agreement has been signed and as explained under #3, the technician has never had direct contact with the participants. Data analysis will be performed by the psychologists.