HAL
open science

Remote cognitive assessment of older adults in rural areas by telemedicine and automatic speech and video analysis: protocol for a cross-over feasibility study

Alexandra König, Radia Zeghari, Rachid Guerchouche, Minh Duc Tran, François Bremond, Nicklas Linz, Hali Lindsay, Kai Langel, Inez Ramakers, Pascale Lemoine, et al.

To cite this version:

Alexandra König, Radia Zeghari, Rachid Guerchouche, Minh Duc Tran, François Bremond, et al.. Remote cognitive assessment of older adults in rural areas by telemedicine and automatic speech and video analysis: protocol for a cross-over feasibility study. BMJ Open, 2021, 11 (9), pp.e047083. 10.1136/bmjopen-2020-047083 . hal-03550459

HAL Id: hal-03550459

https://hal.science/hal-03550459v1

Submitted on 7 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

BMJ Open

Remote cognitive assessment of older adults in rural areas by telemedicine and automatic speech and video analysis: protocol for a cross-over feasibility study

Alexandra König ,[1,2] Radia Zeghari,[2] Rachid Guerchouche,[1,2] Minh Duc Tran,[1] François Bremond,[1] Nicklas Linz,[3] Hali Lindsay,[4] Kai Langel,[5] Inez Ramakers,[6] Pascale Lemoine,[7] Vincent Bultingaire,[7] Philippe Robert[2]

To cite: König A, Zeghari R, Guerchouche R, *et al*. Remote cognitive assessment of older adults in rural areas by telemedicine and automatic speech and video analysis: protocol for a cross-over feasibility study. *BMJ Open* 2021;**11**:e047083. doi:10.1136/bmjopen-2020-047083

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online. (http://dx.doi.org/10.1136/bmjopen-2020-047083).

Received 19 November 2020
Accepted 27 July 2021

Check for updates

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

For numbered affiliations see end of article.

Correspondence to
Dr Alexandra König;
alexandra.konig@inria.fr

ABSTRACT

Introduction Early detection of cognitive impairments is crucial for the successful implementation of preventive strategies. However, in rural isolated areas or so-called 'medical deserts', access to diagnosis and care is very limited. With the current pandemic crisis, now even more than ever, remote solutions such as telemedicine platforms represent great potential and can help to overcome this barrier. Moreover, current advances made in voice and image analysis can help overcome the barrier of physical distance by providing additional information on a patients' emotional and cognitive state. Therefore, the aim of this study is to evaluate the feasibility and reliability of a videoconference system for remote cognitive testing empowered by automatic speech and video analysis.

Methods and analysis 60 participants (aged 55 and older) with and without cognitive impairment will be recruited. A complete neuropsychological assessment including a short clinical interview will be administered in two conditions, once by telemedicine and once by face-to-face. The order of administration procedure will be counterbalanced so half of the sample starts with the videoconference condition and the other half with the face-to-face condition. Acceptability and user experience will be assessed among participants and clinicians in a qualitative and quantitative manner. Speech and video features will be extracted and analysed to obtain additional information on mood and engagement levels. In a subgroup, measurements of stress indicators such as heart rate and skin conductance will be compared.

Ethics and dissemination The procedures are not invasive and there are no expected risks or burdens to participants. All participants will be informed that this is an observational study and their consent taken prior to the experiment. Demonstration of the effectiveness of such technology makes it possible to diffuse its use across all rural areas ('medical deserts') and thus, to improve the early diagnosis of neurodegenerative pathologies, while providing data crucial for basic research. Results from this study will be published in peer-reviewed journals.

Strengths and limitations of this study

- ► The study aims to evaluate the use of a specifically developed and tailored videoconference system for administering remotely a complete cognitive and psychological assessment.
- ► Additional audio and video features for emotion and engagement detection as well as objective stress measurements will be captured.
- ► A mobile unit will test the use of the system in an equipped van that goes to the participants' home.
- ► The sample size is relatively low since the study is limited in time.
- ► Selection bias may be that only participants open to the use of technology are willing to participate even if usage does not require any real interaction with the system.

INTRODUCTION

With the current COVID-19 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. Moreover, in regard to this, the journal *Lancet* published an article underlining the important negative consequences of social isolation on older adults[1] and that online technologies represent a promising solution to overcome these problems, by providing support and connection to healthcare professionals.

Today, many barriers, such as social isolation, still often hinder the early diagnosis of cognitive and affective disorders such as dementia which is crucial for timely treatment and management since lately the benefits of prevention strategies have been clearly demonstrated.[2]

BMJ Open: first published as 10.1136/bmjopen-2020-047083 on 2 September 2021. Downloaded from http://bmjopen.bmj.com/ on September 3, 2021 by guest. Protected by copyright.

For older adults living in rural isolated areas, it is particularly challenging due to the limited access and long travel times to specialised clinics, causing complete exclusion and thus, under-representation of this population in clinical trials.

New methods for remote screening and monitoring of people at risk are sorely needed. Additionally, the increasing risks caused by social isolation among older people in these areas have to be addressed rapidly. For this, over the past years, information and communication technologies have been employed in the field of dementia research with attempts of using computerised cognitive testing, sensors, automatic speech or image analysis for more objective and standardised evaluation of patients cognitive, behavioural and emotional states.[3] Furthermore, several studies have investigated the use of telemedicine for remote cognitive assessments of dementia disorders.[4 5]

The use of videoconference-mediated telemedicine is gradually increasing in diverse patient care settings including primary care, critical care, neurology, behavioural health and psychiatry, among other specialty areas. Lately, review studies show that a certain number of cognitive tests can be administered remotely with reliable results compared with face-to-face assessment with psychologists at a clinic.[6–8] Acceptability assessments show as well that this form of being assessed is relatively well tolerated by the users.[9]

However, most studies focused on just one or a few cognitive tests using already existing videoconference platforms such as Skype. Hence, it would be important to evaluate the feasibility of remote administration of a complete neuropsychological assessment including multiple tests of cognitive functions as well as a typical anamnesis interview. One hypothesis is that clinicians seem hesitant in adopting telemedicine technology due to the feeling of distance to their patients; they report the difficulty of extracting sufficient non-verbal cues on a patients' emotional state via this administration method. Today, with the advances in artificial intelligence empowered speech, language and image processing, it is possible to extract additional sophisticated and unobtrusive natural biomarkers from the videoconference-based consultations and provide it back to the clinician as feedback for a better understanding of a patients' behaviour and differential diagnosis. These sensor-based technologies can provide rich information about cognitive and emotional characteristics, such as, for instance, prosodic features from speech for mood detection,[10 11] or head and eye directions from video for indicating engagement levels.[12 13] This can be used to help with the process of clinical decision-making during consultations, improving diagnostic precision and overcoming the lack of communication cues usually provided in face-to-face interactions. New perceptual analysis (computer vision algorithms and natural language processing), can add more quantitative information to the interactions such as dynamics and intensity of behaviours.[14]

For this, a videoconference system tool was developed (figure 1) and specifically designed to remotely perform a full range of neuropsychological tests supporting early detection and monitoring of cognitive disorders. In opposition to already existing videoconference tools, it has two different adapted interfaces, one for the clinician (left image) with several clinical tests implemented and available in a platform with its visual content and scoring system and one very simplified for the patient (right image) showing mainly the clinician or the test content. For each test, speech and video can be recorded.

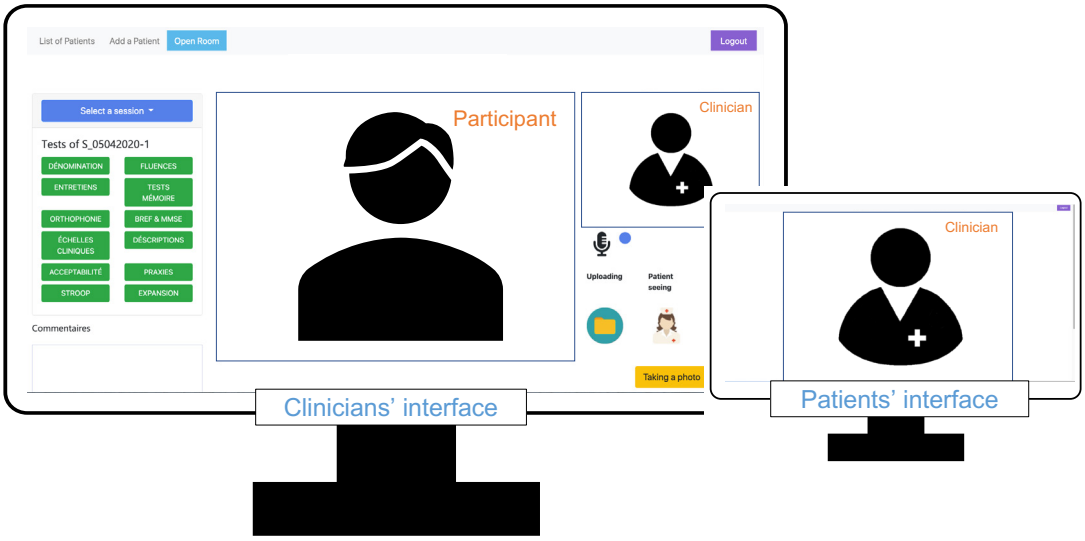

Figure 1 Clinician and patient interfaces of the videoconference tool; on the left side the clinician can access all different tests, control the visual content for the patient, score performances and record speech and video. The interface shows in the centre the video of the patient and on the right a feedback video of the clinician. On the right side, the patient's interface which displays mainly in the centre the video of the clinician or the test content.

König A, *et al*. *BMJ Open* 2021;**11**:e047083. doi:10.1136/bmjopen-2020-047083

BMJ Open: first published as 10.1136/bmjopen-2020-047083 on 2 September 2021. Downloaded from http://bmjopen.bmj.com/ on September 3, 2021 by guest. Protected by copyright.

This tool could allow older adults living in rural areas access to enrolment in future clinical dementia trials and thus, more effective interventions resulting in the reduction of overall costs associated with treatment and rehabilitation.[15] For patients, it could offer the comfort of flexible usage without physically visiting specialised clinicians.

Over the past years, in many of France's rural regions, a crisis of medical 'desertification' can be witnessed caused by a gradual steady decline in the number of local doctors. The region around Digne-les-Bains in the southeast of France is one of the most affected areas, urgently requiring novel solutions to address the lack of access to adequate specialised healthcare.

The overall aim of the current study is to evaluate the use of this system for administering remote cognitive assessments to isolated elderly. For this, a counterbalanced cross-over feasibility study will be performed in the rural areas in the southeast of France (Digne-les-Bains). The study will target (1) results from videoconference administration of a complete neuropsychological assessment will be compared with the classical face-to-face administration, (2) acceptability among the users, patients as clinicians, will be assessed and (3) speech features as well as information on facial expressions and posture will be additionally extracted and compared with classical assessment scales. Moreover, these features might help to indicate levels of engagement, stress, fatigue or mood states.

METHODS AND ANALYSIS

Objectives

We aim to deploy and test a specifically designed videoconference system for cognitive testing under real conditions of a clinical feasibility study.

1. Evaluate the new telemedical service in terms of clinical relevance and usability and acceptability.
   a. Evaluate the reliability of the different cognitive tests administered through the videoconference system to isolated older adults living in rural areas.
   b. Assess acceptability of videoconference modality among the users qualitatively and quantitatively.
   c. Identify automatically extracted speech, language and video features, which correlate with cognitive performance and neuropsychiatric/psychological symptoms (eg, depression, apathy, anxiety), as measured by standard clinical cognitive and behavioural assessments.
   d. Support the proposed speech analysis by a complementary computer-vision-based analysis. Such analysis exploits advanced methods related to automated face analysis, tracking, detection and recognition, as well as human behaviour analysis (capturing mood, levels of engagement in tasks and levels of arousal).

Participants

For this non-interventional observational study, 60 older adults will be recruited (age ≥55) from the region of Digne-les-Bains, France over an inclusion period of 12 months. Participants will be referred by the local Hospital Center's memory clinic, general practitioners and/or the ADMR (*Aide à domicile en milieu rural*/home services in rural areas) Federation. Those interested in the study will receive an information sheet. On attendance, a member of the research team will address any potential queries and take informed consent, prior to the experiment.

The main inclusion criteria are:

- Lives in the region of Digne-les-Bains.
- ≥55 years old.
- Speaks French as a first language.
- Can independently understand the informed consent form, and voluntarily consents to participate, OR has an alternate decision-maker who can provide consent on their behalf while the participant provides assent.

Exclusion criteria are:

- Has significant vision problems which would impact ability to perceive visual stimuli.
- Has significant auditory problems which would impact ability to understand verbal prompts.

Patient and public involvement

Participants were not directly involved in the development of the protocol and research questions. However, the association ADMR is working very closely with isolated elderly and represents an important partner in this project. Its members were involved from the early beginning of the planning and their input shaped the overall design of the study. Focus groups with participants will be organised to gather feedback on the experience of the videoconference-based administration procedure.

Study protocol

To assess the usability of the system in several conditions, we propose two scenarios: the first one is to validate the telemedicine tool in a clinical setup when patients can move to close-by places, where they have access to the needed infrastructures (internet connection, device with a webcam, etc); the second one is to move close-by the homes of isolated patients (eg, in rural areas) with an equipped mobile unit.

A comprehensive neuropsychological assessment (see box 1) consisting of a clinical interview followed by a set of cognitive tests (memory, attention, etc) will be administered face-to-face and 2 weeks later remotely via a videoconference system by two different psychologists. The system is installed in a room in the local Hospital Center's memory clinic. To reduce learning effect biases and within rater variability, this procedure will be counterbalanced so that half of the participants first experience the face-to-face interaction and the other half will initially receive the videoconference administration (see figure 2).

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. The study

BMJ Open: first published as 10.1136/bmjopen-2020-047083 on 2 September 2021. Downloaded from http://bmjopen.bmj.com/ on September 3, 2021 by guest. Protected by copyright.

Box 1 Protocol overview

Description

1. Psychologist will call the patient through the teleconference system.
2. Short introduction about aim and procedures.
3. Informed consent.
4. Clinical interview: demographic information, medical history, subjective memory complaint; scales: Subjective Cognitive Functioning (SCF-4 items),[25] Mood (GDS-15 item),[26] Apathy Inventory.[27]
5. Test of global cognition (Mini Mental State Examination (MMSE)[28]).
6. Visual episodic memory test (free and cued selective reminding test[29] if MMSE <20; 5-word test of Dubois *et al*[30]).
7. Praxis test.[31]
8. Visual recognition test (doors and people test).[32]
9. Working memory task (digit span test).[33]
10. Open questions (positive/negative storytelling).[18]
11. Verbal episodic memory recall.
12. Stroop test.[34]
13. Semantic (animals/fruits) and phonemic ('p'/'r') fluency tasks.[34]
14. Picture description (Cookie Theft picture).[35]
15. Denomination task (Lexis[36]/DO80[37]).
16. Completion of conversation with patient, including providing information about the procedure of the study (face-to-face assessment).
17. User experience questionnaire (copy under online supplemental file).

is limited to use the video modality to verify test performances, which makes it complicated for constructional tasks that require drawing. 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.

For those participants, who live further away, it is planned to perform the videoconference administration nearby their home with the help of a van equipped with a computer connected to 4G internet. The participants will sit in the van and connect to the psychologist via the system on a dedicated laptop. In this way, the scenario of remotely testing those who are living very isolated will be evaluated.

After the inclusion of all participants, results obtained at the videoconference administration will be compared with the classical face-to-face method to evaluate their reliability. Evaluation reports of the neuropsychological assessments obtained in this study will be sent to the referring clinicians of the hospital in Digne-les-Bains.

Regarding acceptability, participants of the study will be asked to complete a questionnaire on their experience of the videoconference-administered assessment (compared with the classical). In addition, a subgroup of study participants, as well as other stakeholders, will be invited to participate in a focus group with semi-structured qualitative interviews in to assess, in more depth, the ease and usability of the system.

A corpus of video and speech samples will be created for further analysis. Features, potentially relevant for early detection of cognitive disorders and/or behavioural and psychological symptoms, will be extracted. Speech, language and video features extracted during the videoconference administration will be compared partly to manual annotations of the psychologists (information regarding engagement, mood and arousal), and partly to information extracted during the clinical interview and scales/questionnaires on the presence of behavioural and psychological symptoms (depression, apathy, etc).

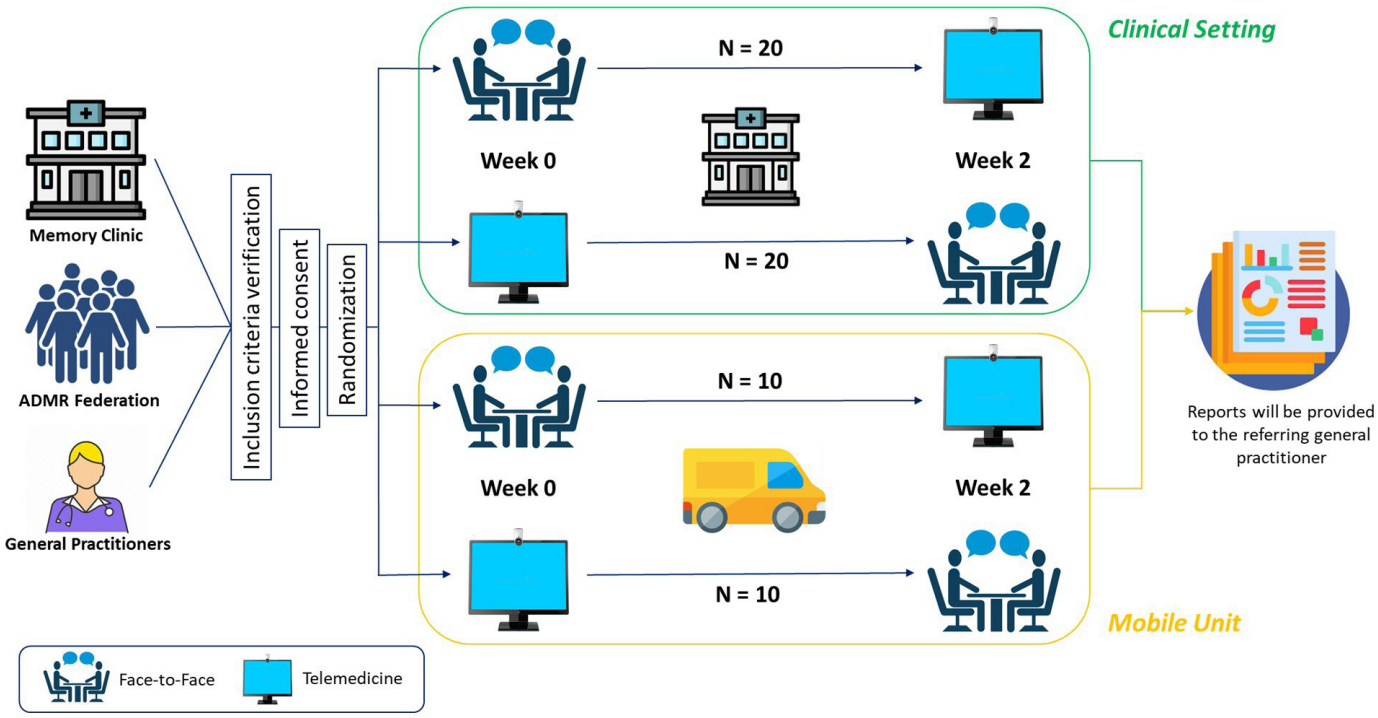

Figure 2 Study protocol design. ADMR, *Aide à domicile en milieu rural*/home services in rural areas.

BMJ Open: first published as 10.1136/bmjopen-2020-047083 on 2 September 2021. Downloaded from http://bmjopen.bmj.com/ on September 3, 2021 by guest. Protected by copyright.

In a subgroup of participants, stress levels will be measured during the face-to-face and remote assessment administration both subjectively via a questionnaire and quantitatively, with a bracelet (Empatica E4[16]) which measures physiological data in real-time.

PROTOCOL OF THE ASSESSMENT

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, etc). No Skype, Zoom or other existing videoconferencing systems were 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 any existing web browsers (Chrome, Firefox, Safari, Edge) under any operating system that supports webRTC standard. WebRTC (https://en.wikipedia.org/wiki/WebRTC) (Web Real-Time Communication) is a free, open-source project providing web browsers and mobile applications with real-time communication via simple application programming interfaces. 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 made directly through the web, the used devices (camera, microphone and speakers) are either the built-in devices when a laptop is used (in the van), or externally attached to the PC (in the case of a desktop computer, for example/in the clinic). For a tablet or a smartphone, the integrated hardware devices can be used.

The implemented clinical tests are normalised tests used by neuropsychologists and related medical professionals. The following three types of clinical tests will be used:

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, etc). 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.

For practical reasons, we will use a laptop or a desktop with a minimum of 17-inch screen width. 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, etc); and to run the different services allowing the videoconferencing communication are mandatory. All the servers will 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 these 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 lack 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.

DATA COLLECTION, MANAGEMENT AND ANALYSIS

Data collection

Data will be collected at the Hospital Center in Digne-les-Bains; the assessments will be performed remotely and face-to-face by clinicians from the Memory Clinic in Nice. For remote assessments, the videoconference software will record test scores, video and speech and store them on a dedicated secured server complying with healthcare data hosting regulations. Other clinical and neuropsychological data collected by the clinicians will follow the standard practice of the centre.

Prior to initial participation in the study, each participant will consent on paper, or an alternate decision-maker will consent and the participant will provide assent. Each individual (or dyad) will be given as much time as they need to review the consent form and decide whether they want to participate. The consent only needs to be provided once, before initial participation in the study.

Subsequent completion of additional sessions will not require additional consent. Each participant will be associated with a unique, randomly generated, non-personally

BMJ Open: first published as 10.1136/bmjopen-2020-047083 on 2 September 2021. Downloaded from http://bmjopen.bmj.com/ on September 3, 2021 by guest. Protected by copyright.

identifying number ('participant ID'). During their first session, on providing consent to participate, the participant will be asked to provide general demographic data (eg, month and year of birth, sex, number of years of education, spoken languages, history of dementia, etc); responses will be associated with the appropriate participant ID. If consent is provided by an alternate decision-maker, the participant's data will be collected from the designated proxy. This information will be used to control for confounding variables when conducting analyses of the collected data and will only be reported anonymously or in aggregate.

Acceptability evaluation

All participants will be asked to answer a questionnaire (copy under online supplemental file) on the acceptance of the videoconference as well as of the face-to-face modality for cognitive testing including seven 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'[17] and assesses the user experience, including an overall evaluation, participant satisfaction, if they would repeat the experience, attitudes and clarity of instructions as well as what type of method is preferred and why. 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. Thematic qualitative analysis will be applied to the comments and written answers to the questionnaires.

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 report and based on thematic analysis 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.

Speech data

The speech of the different participants will be recorded as audio files for different cognitive tests. Depending on the test, the recorded speech will be either free speech or direct answers to questions, or verbal or visual stimuli.

We will directly record the speech of the patients. From the audio files, we will use automatic speech recognition to obtain textual transcripts of the recordings. A subset of the data will be manually transcribed to compare results between automated and manual transcriptions.

We will use either the internal microphone of the device (PC or tablet) or an external microphone for better recording quality. The recorded data will be automatically stored on the secured server.

Video data

Towards our general goal to detect and remotely monitor cognitive decline in the context of dementia in a partially automated way, we intend to support the proposed speech analysis by a complementary computer-vision-based analysis. Such analysis exploits advanced methods related to automated face analysis, tracking, detection and recognition, as well as human behaviour analysis.

First, we plan to record two-dimensional video data from all participants for the computer vision-based analysis. Then, we intend to analyse these data with a focus on finding facial/gestural behaviours and facial/gestural activities that are representative for cognitive performance and neuropsychiatric symptoms of dementia in different situations, such as free speech interviews, and in some cognitive tests (eg, patients are encouraged to describe the content of a series of images).

We intend to acquire the video data using the integrated web camera of the device (PC or tablet) used for the telemedicine session. Possible use of an external web camera connected to a PC is also considered. In both cases, the recorded data will be automatically stored on the secured server.

Stress measure data

We would like to explore the use of additional objective markers of stress levels within this study. For this, for only a subgroup of participants (during the mobile home unit phase), we will extract through the Empatica E4 device physiological data during both administration methods:

- Electro-dermal activity: measures sympathetic nervous system activity manifested through the skin, by measuring the constantly fluctuating changes in certain electrical properties of the skin.
- Heart rate variability: derived from measuring blood volume pulse.
- Peripheral skin temperature: measured using an integrated infrared thermopile.
- Three-axis accelerometer: captures motion-based activity, which identifies intensity and frequency of movements that could be a seizure.

We aim to compare the different time point measures with each other to assess the stress levels of the participants during the face-to-face and the remote assessment administration.

Data security

A collection of data will be made with the videoconferencing web-based platform (telemedicine tool). Digital data (audio, video, recorded scores, answers to questionnaires), as well as written records will be collected.

Concerning the paper data, these will be stored in armoires with access limited to clinicians participating in

BMJ Open: first published as 10.1136/bmjopen-2020-047083 on 2 September 2021. Downloaded from http://bmjopen.bmj.com/ on September 3, 2021 by guest. Protected by copyright.

the study (a key is needed for access). Each involved clinical partner will store these papers in their clinic. The data will be digitised through the web-platform to conduct the research work. In the following paragraphs, we provide details about the security of all digital data.

To perform the study, a secured connection to the web-platform is required by both clinicians and participant subjects. The secured and encrypted connection (ie, HTTPS) requires authentication with an email address and an encrypted password. Only limited email addresses' domains will be allowed to connect to the web-platform. The clinician should have a professional email that is provided by their organisation (such as '@chu-nice.fr', '@ch-digne.fr', etc). Only the domains of clinical partners involved in the clinical study will be allowed.

The enrolled subjects will have authorised access to the web-platform with their email and a generated temporary password valid only for the time of their participation in the study. The password will be entered by the person accompanying the subject (a clinician, part of the clinical study) to connect the subject to the platform. For the subjects who have no email, a unique identifier (temporary email) will be provided.

The data collected during the study will be stored in a secured encrypted database. The database will be hosted by the Institute Claude Pompidou (ICP) servers. These servers are owned and part of the University Hospital of Nice (CHU NICE). These servers are secure and follow all the required security and healthcare data hosting regulations. The security of the servers is managed by the IT personnel of CHU NICE who has to sign a confidentiality agreement and only has access to anonymised clinical data.

An IT technician from the ICP will maintain the anonymised database. The technician will have all the rights on the database:

- Create/delete/edit the list of clinicians who can access the web-platform.
- Create/delete/edit the list of subjects participating in the clinical study.

Regarding the mobile van, the technician is only there before the consultation to ensure the technical equipment is working correctly, then he will leave the van.

The authorised clinicians will have the right to access the web-interface, which will allow them to perform the study. The collected data will automatically be stored in the database.

Access to the stored data is strictly reserved to the clinicians with a secured account. The type of database is MySQL database. The security of the database is ensured by Advanced Encryption Standard techniques. The following rules will be applied:

- The database will only be accessible from trusted hosts.
- No use of data from input without filtering.
- All types of data will be protected.
- The administrator of the database will not be a user (the IT technician) of the web platform.

Security of transmission between clinicians and participants during study sessions: After the clinician and the participant are connected to the platform, video streams are circulating in both directions. Video streams are protected by integrating the following rules:

- The sessions of the clinical study will be scheduled by the involved clinicians and will not be publicly known.
- We will add a signalling protocol that will provide encryption of signalling traffic.
- The connection between the clinician and the patient will be a P2P connection, the media contents (audio and video channels) will be transmitted between peers directly in full duplex. Thus, as the signalling server maintains the number of peers in communication, we will monitor the connection for additional suspicious peers in a call session. If the number of peers actually present on signalling server is more than the number of peers interacting on the connection, then it could mean that someone is eavesdropping secretly and should be terminated from session access by force.
- Request permission from both sides to use the camera and the microphone.

The dissemination of the research results will be based on the analysis of generated pseudo-anonymous data, which does not include any information that would allow any reference to the identity of participants.

Data processing

The collected data will be processed to generate pseudo-anonymous metadata, which will be used by the different technical and clinical partners of the project. These technical partners (who are not the clinicians involved in passing the tests and authorised to access the participant data) will never have access to the raw data: identity of the patient, personnel information, videos and audio or any other information which allows identifying the subject. The only processed data will be: scores of tests, videos and audio files. The processing will be done on the servers of the Institut Claude Pompidou, a processing which will generate anonymous metadata which will be transmitted by the IT technician to the partners with no information about the identity of the participants.

Concretely, an IT technician (or engineer) working for the Institut Claude Pompidou, and authorised to access all patient's data as part of his duties (creating patients records, correcting information about patient information, etc) will be the only non-clinical person who can access the identifying data of the participants (speech and video recordings). This IT technician will be trained by the technical partners (using similar non-confidential data that does not belong to the participants), for training on relevant software-provided by the technical partners-to generate pseudo-anonymous metadata: low-level features extracted from speech and video such as signal intensities, acoustic characteristics, two-dimensional points positions and head/eyes positions in different images. This pseudo-anonymised metadata will not contain the identity of the participants. For each participant, a random

BMJ Open: first published as 10.1136/bmjopen-2020-047083 on 2 September 2021. Downloaded from http://bmjopen.bmj.com/ on September 3, 2021 by guest. Protected by copyright.

code will be assigned, known only by the IT technician and the clinicians involved in the study. All the metadata extracted by the different software and executed only by the IT technician will then be transmitted in a pseudo-anonymous manner: the technical partners will not know the identity of the participants, and they will only have access to a set of metadata matched to an 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.

The data will be stored for a maximum period of 3 years. This time period will allow us to do the research work, analyse the data and publish results.

Analysis

In this study, we will mainly work on the analysis of three types of data: speech, video and tests scores, including multimodal analysis combining the three data types.

To verify the agreement between both administration procedures, we will compare the face-to-face test results with the videoconference-based test results. For this, the mean and SD for each test score will be calculated. Intraclass correlation coefficients will be used to assess agreement between the two testing formats. The Analysis of variance (ANOVA) test will be used to assess if the administration modality (independent variable) was associated with any difference in total scores of the different tests (dependent variables).

Speech analysis

The audio and transcribed text data will be processed to obtain acoustic, lexico-syntactic and semantic features. First, acoustic features will be extracted from the audio samples. The audio data will be passed through frequency domain transformations to obtain standard acoustic measures used in language processing, such as Mel frequency cepstral coefficients based on the cepstral representation of the acoustic signal which detects periodicity in its spectrum, jitter and shimmer measures based on irregularities in signal periodicity, recurrence period density entropy and pitch period entropy based on measures of aperiodicity and noise in the signal, measures based on the frequency domain harmonics, signal to noise ratios and filled and unfilled pause features, among others. Similar work was previously performed for the detection of signs of apathy.[18]

Second, syntactic measures will be obtained from transcriptions of the verbal and written responses. The text transcripts will be parsed with a standard probabilistic context-free grammar parser, and the resulting parse trees will be searched for the presence of specific syntactic constructions. Extracted syntactic features include the depth of syntactic parse trees, use of subordinate and coordinate clauses, use of passive voice, use of different types of syntactic constructions (eg, noun phrases, auxiliary verb phrases, etc) and mean length of utterance, among others. Lexical features will be extracted from part of speech tagged transcriptions and written responses; these include relative word-class usage, lexical richness evaluated using standard measures such as type-to-token ratio and Honoré's statistic, average age of acquisition, familiarity and imageability, among others.

Finally, semantic features will be extracted from the transcriptions based on the criteria of the corresponding language task, for example, qualitative features from the semantic verbal fluency tasks[19–21] or recall and precision-based measures of semantic content units present in a picture description task.[22] Other relevant acoustic and linguistic features may also be computed. Standard statistical testing techniques (eg, t-test, ANOVA) will be used to determine the significance of any trends, which are detected in the data.

Video analysis

For the video analysis, we will first annotate the data in an initial step to find similar video sequences (eg, such as depicted patients talking, as well as such showing smiles or neutral states). From these annotated sequences, features that target behavioural analysis[23 24] will be extracted. Specifically, we will use algorithms for face detection and then extract dense trajectories from the video sequence and represent these with facial spatio-temporal features, which incorporate motion of sampled points in the video sequence. Using the extracted features, emotions and engagement of the patients will be computed. Combining the features from video and speech analysis, we will perform an analysis to correlate the emotions expressed by a patient and their engagement during the cognitive tests with the cognitive disorders they could be affected with. In particular, the engagement information of a patient during a test will be used as an indicator of the success and the efficiency of a cognitive test. Our approach will be validated by the acquired data, as well as by benchmark datasets, which are publicly available. The recorded data will only be used to develop, test and validate algorithms (analysis, classification, deep-learning, etc) and will be deleted after the study.

ETHICS AND DISSEMINATION

The procedures are not invasive and there are no expected risks or burdens to participants. All participants will be informed that this is an observational study and their consent taken prior to the experiment. Participants with cognitive impairment may experience feelings of confusion, anxiety, distress, embarrassment or sadness during participation. We mitigate against this by explaining sufficiently beforehand the procedure, answering possible questions and keeping the session relatively short. In situations where a participant is significantly distressed as a result of participating in the study, the clinician will stop the session. Participants may withdraw from the study completely at any time. No other known risks to the participants exist. After completion of the tests, each

BMJ Open: first published as 10.1136/bmjopen-2020-047083 on 2 September 2021. Downloaded from http://bmjopen.bmj.com/ on September 3, 2021 by guest. Protected by copyright.

participant will receive feedback on her/his performance as well as a summary report will be sent to the referring practitioner.

Confidentiality aspects such as data encryption and storage comply with the General Data Protection Regulation (GDPR) and the requirements from sponsoring bodies and ethical committees. Results from this study will be published in peer-reviewed journals. However, all communications will only include results on analyses undertaken after preprocessing the recordings, ensuring that audio-visual data will never be published or disseminated.

The project will enable the validation of technology that allows the implementation of sophisticated and unobtrusive remote neuropsychological assessment, eventually at home or at easily reachable locations to facilitate access to clinical experts. The solution will furthermore allow for easier recruitment and onboarding of people living isolated in rural areas into clinical trials who, are until today, under-represented due to the lack of access to clinical sites.

CONCLUSIONS

Demonstration of the effectiveness of this technology may later allow its distribution across rural areas ('medical deserts') in France and thus improve early diagnosis of neurodegenerative pathologies while providing data crucial for basic research. Ultimately, it will lead to an improvement of healthcare access and care of isolated seniors in these regions. Furthermore, recruitment, onboarding and monitoring of potential candidates in these regions for clinical trials will be facilitated. Pushing the use of such remote solutions in the future is of particular relevance given the current context of the COVID-19 pandemic.

Author affiliations
[1]STARS Team, Institut National de Recherche en Informatique et en Automatique Centre de Recherche Sophia Antipolis Méditerranée, Sophia Antipolis, France
[2]Cobtek (Cognition-Behaviour-Technology) Lab, FRIS, Universite Cote d'Azur, Nice, France
[3]German Research Centre for Artificial Intelligence Saarbrucken Branch, Saarbrucken, Germany
[4]Deutsches Forschungszentrum fur Kunstliche Intelligenz GmbH Standort Saarbrucken, Saarbrucken, Germany
[5]Janssen Healthcare Innovation, Beerse, Belgium
[6]Maastricht University, Maastricht, The Netherlands
[7]Centre Hospitalier de Digne-les-Bains, Digne-les-Bains, France

Acknowledgements This research is funded by the EIT Health Activity 19249, DeepSPA, with great support from Modis Belgium, Janssen Clinical Innovation, the Association IA, ADMR 04, and the Centre Hospitalier and the city of Digne-les-bains, France. Special thanks to Valérie DEPRAD President, Francis Kuhn and Emmanuelle Martin Vice-President of ADMR 04.

Contributors Conceptualisation: AK, RZ, RG, KL, IHGBR, PR; methodology: RZ, AK, IHGBR; software: RG, MDT, FB; validation: RZ, AK, PL, VB; formal analysis: NL, HL, RZ; investigation: RZ, AK, PL, VB; resources: PR, PL, VB; data curation: RG, FB, MDT; writing—original draft preparation: AK, RZ, RG; visualisation: RZ, RG; supervision: PR, FB; project administration: AK, FB; funding acquisition: FB, PR. All authors have read and agreed to the published version of the manuscript.

Funding This work is supported by the European Institute for Innovation and Technology (EIT)—Health, grant number 19249.

Competing interests We have read and understood BMJ policy on declaration of interests. NL is an employee and shareholder of ki elements UG.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

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

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

ORCID iD
Alexandra König http://orcid.org/0000-0001-9960-9657

REFERENCES

1 Armitage R, Nellums LB. COVID-19 and the consequences of isolating the elderly. *Lancet Public Health* 2020;5:e256.
2 Rakesh G, Szabo ST, Alexopoulos GS, *et al*. Strategies for dementia prevention: latest evidence and implications. *Ther Adv Chronic Dis* 2017;8:121–36.
3 Robert PH, Konig A, Andrieu S, *et al*. Recommendations for ICT use in Alzheimer's disease assessment: Monaco CTAD expert meeting. *J Nutr Health Aging* 2013;17:653–60.
4 Kim H, Jhoo JH, Jang J-W. The effect of telemedicine on cognitive decline in patients with dementia. *J Telemed Telecare* 2017;23:149–54.
5 Timpano F, Pirrotta F, Bonanno L, *et al*. Videoconference-based mini mental state examination: a validation study. *Telemed J E Health* 2013;19:931–7.
6 Haralambous B, Subramaniam S, Hwang K, *et al*. A narrative review of the evidence regarding the use of telemedicine to deliver video-interpreting during dementia assessments for older people. *Asia Pac Psychiatry* 2019;11:e12355.
7 Barth J, Nickel F, Kolominsky-Rabas PL. Diagnosis of cognitive decline and dementia in rural areas - A scoping review. *Int J Geriatr Psychiatry* 2018;33:459–74.
8 Brearly TW, Shura RD, Martindale SL, *et al*. Neuropsychological test administration by Videoconference: a systematic review and meta-analysis. *Neuropsychol Rev* 2017;27:174–86.
9 Dario C, Luisotto E, Dal Pozzo E, *et al*. Assessment of patients' perception of telemedicine services using the service user technology acceptability questionnaire. *Int J Integr Care* 2016;16:13.
10 Cummins N, Scherer S, Krajewski J, *et al*. A review of depression and suicide risk assessment using speech analysis. *Speech Commun* 2015;71:10–49.
11 Valstar M, Pantic M, Gratch J. *AVEC 2016: depression, mood, and emotion recognition workshop and challenge*, 2016: 3–10.
12 Dewan MAA, Murshed M, Lin F. Engagement detection in online learning: a review. *Smart Learn. Environ.* 2019;6:1.
13 Asteriadis S, Karpouzis K, Kollias S. *The importance of eye gaze and head pose to estimating levels of attention the importance of eye gaze and head pose to estimating levels of attention, (June)*, 2011.
14 Scherer S, Stratou G, Lucas G, *et al*. Automatic audiovisual behavior descriptors for psychological disorder analysis. *Image Vis Comput* 2014;32:648–58.
15 Armaignac DL, Saxena A, Rubens M, *et al*. Impact of telemedicine on mortality, length of stay, and cost among patients in progressive care units: experience from a large healthcare system. *Crit Care Med* 2018;46:728–35.
16 Real-Time physiological data streaming and visualization. Available: https://www.empatica.com/en-eu/research/e4/

BMJ Open: first published as 10.1136/bmjopen-2020-047083 on 2 September 2021. Downloaded from http://bmjopen.bmj.com/ on September 3, 2021 by guest. Protected by copyright.

17 Brooke J. SUS: A “quick and dirty” usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, *et al*, eds. *Usability evaluation in industry*. London: Taylor & Francis, 1996: 189–94.
18 König A, Linz N, Zeghari R, *et al*. Detecting apathy in older adults with cognitive disorders using automatic speech analysis. *J Alzheimers Dis* 2019;69:1183–93.
19 König A, Linz N, Tröger J, *et al*. Fully automatic Speech-Based analysis of the semantic verbal fluency task. *Dement Geriatr Cogn Disord* 2018;45:198–209.
20 et alLinz N, Tröger J, Alexandersson J. Predicting dementia screening and staging scores from semantic verbal fluency performance. *Proceedings of the IEEE International Conference on Data Mining (ICDM) - Workshop on Data Mining for Aging, Rehabilitation and Independent Assisted Living*, 2017:18–21.
21 Linz N, Tröger J, Alexandersson J. Using Neural Word Embeddings in the Analysis of the Clinical Semantic Verbal Fluency Task. In: *Proceedings of the 12th International Conference on computational semantics (IWCS)*. Montpellier (France), 2017: 19–22.
22 Fraser K, Meltzer J, Rudzicz F. Linguistic Features Identify Alzheimer’s Disease in Narrative Speech. *JAD* 2015;49.
23 Cohn JF, Kruez TS, Matthews I. Detecting depression from facial actions and vocal prosody. *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, 2009:1–7.
24 Joshi J, Goecke R, Parker G. Can body expressions contribute to automatic depression analysis? *Proceedings of IEEE Conference on Automatic Face and Gesture Recognition*, 2013.
25 Aalten P, Ramakers IHGB, Biessels GJ, *et al*. The Dutch Parelsnoer Institute--Neurodegenerative diseases; methods, design and baseline results. *BMC Neurol* 2014;14:254.
26 Almeida OP, Almeida SA. Short versions of the geriatric depression scale: a study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. *Int J Geriatr Psychiatry* 1999;14:858–65.
27 Robert PH, Clairet S, Benoit M, *et al*. The apathy inventory: assessment of apathy and awareness in Alzheimer's disease, Parkinson's disease and mild cognitive impairment. *Int J Geriatr Psychiatry* 2002;17:1099–105.
28 Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–98.
29 Grober E, Merling A, Heimlich T, *et al*. Free and cued selective reminding and selective reminding in the elderly. *J Clin Exp Neuropsychol* 1997;19:643–54.
30 Dubois B, Touchon J, Portet F. The "5 words": A simple and sensitive test for the diagnosis of Alzheimer’s disease. *Presse médicale* 2002;31:1696–9.
31 Peigneux P, Van der Linden M. « Présentation d’une batterie neuropsychologique et cognitive pour l’évaluation de l'apraxie gestuelle ». *Revue de neuropsychologie* 2000;10:2000:311–62.
32 Baddeley AD, Emslie H, Nimmo-Smith I. *The doors and people test: a test of visual and verbal recall and recognition*. Bury-St-Edmunds, UK: Thames Valley Test Company, 1994.
33 Botwinick J, Storandt M. *Memory, related functions and age*. Springfield, IL: Thomas, 1974.
34 Godefroy O, Gibbons L, Diouf M, *et al*. Validation of an integrated method for determining cognitive ability: implications for routine assessments and clinical trials. *Cortex* 2014;54:51–62.
35 Goodglass H, Kaplan E, Bar-resi B. *Boston Diagnostic Aphasia Examina- tion. Lippincott Williams & Wilkins*, 2000.
36 De Partz MP, Bilocq V, Wilde D V, *et al*. *LEXIS, test pour Le diagnostic des troubles lexicaux CheZ Le patient aphasique*. Solal: Marseille, 2001.
37 Deloche G, Hannequin D. *Test de Dénomination Orale d'Images. Paris : les Éditions du Centre de Psychologie Appliquée*, 1997.

BMJ Open: first published as 10.1136/bmjopen-2020-047083 on 2 September 2021. Downloaded from http://bmjopen.bmj.com/ on September 3, 2021 by guest. Protected by copyright.