RESEARCH ARTICLE

REVISED Introducing the Tele-OCS: Preliminary evidence of validity for a remotely administered version of The Oxford Cognitive Screen

[version 3; peer review: 1 approved, 2 approved with reservations]
Previous Title 'Introducing the Tele-OCS: A validated remotely administered version of The Oxford Cognitive Screen'

Sam S. Webb[1*], Chloe Carrick[2*], Andrea Kusec[2], Nele Demeyere[2]

[1]Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, Oxfordshire, OX3 9DU, UK
[2]Experimental Psychology, University of Oxford, Oxford, Oxfordshire, OX2 6GG, UK

* Equal contributors

v3 **First published:** 13 Feb 2023, **5**:8
https://doi.org/10.12688/healthopenres.13291.1
Second version: 04 Apr 2024, **5**:8
https://doi.org/10.12688/healthopenres.13291.2
Latest published: 09 Jul 2025, **5**:8
https://doi.org/10.12688/healthopenres.13291.3

Abstract

Background

Remote cognitive assessments are increasingly used with the rising popularity of teleneuropsychology. Here, we evaluated the performance of the remotely administered Oxford Cognitive Screen (Tele-OCS) compared to in-person administration in adult stroke survivors.

Methods

40 stroke survivors (*M* age = 69.30, *SD* = 10.44; sex = 30% female) completed in-person and remote versions of the OCS on average 30 days apart, with different trained examiners. The order of administration was counterbalanced. Cohen's *d* estimates were used to compare performance between modalities.

Results

We found that the proportion of OCS subtasks impaired did not differ across modalities (*d* <.001). With regards to raw subtask scores, only

Open Peer Review

Approval Status ? ✓ ?

	1	2	3
version 3 (revision) 09 Jul 2025			
version 2 (revision) 04 Apr 2024		✓ view	? view
version 1 13 Feb 2023	? view	? view	

1. **Léon Franzen**, University of Lübeck, Lübeck, Germany
Frederike Schröpfer, University of Lübeck, Lübeck, Germany

2. **Dimosthenis Tsapekos**, King's College London, London, UK

3. **Hayley Wright**, Coventry University, Coventry, UK

the picture naming subtask and executive score from the trail making subtask were found to be statistically different across modalities, though raw differences were minimal (<1 point difference on average). These statistical differences did not affect impairment classifications.

Conclusions

The Tele-OCS classified cognitive impairments in a comparable way to the in-person version. The validation of the Tele-OCS allows for remote assessment to increase accessibility and pragmatically aid in addressing the clinical need for stroke-specific cognitive screening in a wider population.

Plain Language Summary

Background Commonly after a stroke, people have changes in specific thinking abilities (cognition). These may include difficulties with understanding, reading, writing, remembering, spatial awareness and planning activities. The Oxford Cognitive Screen (OCS) was made to briefly assess these stroke-specific changes in thinking abilities. The OCS is now well known in the stroke clinical community, with >1200 licensed stroke settings. During the pandemic the OCS was administered remotely (phone or video call etc.), but there was no evidence to support the use of this modality. Here, we evaluated the performance of the Oxford Cognitive Screen administered via phone and video call (Tele-OCS) compared to in-person administration in adult stroke survivors. Methods 40 stroke survivors (M age = 69.30, SD = 10.44; sex = 30% female) completed in-person and remote versions of the OCS approximately 30 days apart, with different trained examiners. The order of administration was counterbalanced. Statistical estimates were used to compare performance between modalities. Results We found that the proportion of OCS subtasks impaired did not differ across modalities, meaning individuals were not more impaired on one modality over the other. With regards to raw subtask scores, only the picture naming subtask and executive score from the trail making subtask were found to be statistically different across modalities, though raw differences were minimal (<1 point difference on average). These statistical differences did not affect impairment classifications. Conclusions The Tele-OCS classified mental ability impairments in a comparable way to the in-person version. The validation of the Tele-OCS allows for remote assessment to increase accessibility and pragmatically aid in addressing the clinical need for stroke-specific cognitive screening in a wider population.

Keywords

remote assessment, teleneuropsychology, cognitive screening, cognitive assessment

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Stroke Association gateway.

This article is included in the Neurology gateway.

Corresponding author: Nele Demeyere (nele.demeyere@psy.ox.ac.uk)

Author roles: Webb SS: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Carrick C**: Data Curation, Formal Analysis, Investigation, Project Administration, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Kusec A**: Formal Analysis, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Demeyere N**: Resources, Supervision, Writing – Review & Editing

Competing interests: ND is one of the creators of the original OCS. We report no other conflicts of interest.

Grant information: This work was supported by the Stroke Association (SA PPA 18/100032). SSW was further supported by SA PGF 21100015. The views expressed in the submitted article are the authors own views and not an official position of the institution or funder.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Webb SS, Carrick C, Kusec A and Demeyere N. **Introducing the Tele-OCS: Preliminary evidence of validity for a remotely administered version of The Oxford Cognitive Screen [version 3; peer review: 1 approved, 2 approved with reservations]** Health Open Research 2025, **5**:8 https://doi.org/10.12688/healthopenres.13291.3

First published: 13 Feb 2023, **5**:8 https://doi.org/10.12688/healthopenres.13291.1

REVISED **Amendments from Version 2**

Minor updates include rephrasing of text to enable an easier reading experience. Other updates include justification for not examining sex differences, on the differences between those with capacity and no capacity to consent to research, and on older age/stroke use of video conferencing.

Any further responses from the reviewers can be found at the end of the article

Introduction

Whilst improved stroke care in hospital has led to a decrease in stroke-related mortality, the number of individuals living with the long-term complications of stroke has risen (Amy & Kapral, 2019). There are greater than 100,000 individuals affected by stroke each year (King *et al.*, 2020). The number of strokes in Europe is projected to rise by 34% by 2035 (King *et al.*, 2020). Post-stroke cognitive impairment is common, with cognitive deficits experienced by almost all stroke survivors early after stroke (Demeyere *et al.*, 2016; Esmael *et al.*, 2021; Merriman *et al.*, 2019; Sexton *et al.*, 2019), and can negatively impact rehabilitation outcomes and post-stroke quality of life (Mole & Demeyere, 2020; Nys *et al.*, 2006). Clinical guidelines have called for the early screening of post-stroke cognition using valid and reliable tools (Intercollegiate Stroke Working Party, 2016; Intercollegiate Stroke Working Party, 2023; Quinn *et al.*, 2021; Rudd *et al.*, 2017).

Teleneuropsychology, which includes the remote delivery of neuropsychological assessments alongside therapy interventions, has long existed alongside the technology to support it (Munro Cullum *et al.*, 2014). The use of remotely administered cognitive tests dramatically increased following the onset of the COVID-19 pandemic, with many healthcare professionals shifting the modality of assessment to telephone or video conferencing (Chapman *et al.*, 2020; Fox-Fuller *et al.*, 2022; Hammers *et al.*, 2020; Webb *et al.*, 2022). This expansion in remote neuropsychological assessment led to the adaptation of widely used cognitive assessment tools for remote use. However, research on the validity of these adaptations has lagged behind. Widely-accessible remote cognitive assessment tools were underdeveloped at the onset of COVID-19, with low numbers of freely available standardized and validated versions of cognitive screening tools (Webb *et al.*, 2022). Tools such as the telephone Montreal Cognitive Assessment, (MoCA; Benge & Kiselica, 2021; Katz *et al.*, 2021; Pendlebury *et al.*, 2013), the Telephone Interview for Cognitive Status (Brandt *et al.*, 1988), and the 26-point telephone version of the Mini Mental State Examination (MMSE; Newkirk *et al.*, 2004) were developed for screening for general dementia, not for early stroke-specific cognitive domains typically affected by stroke (e.g., visuospatial neglect, apraxia, aphasia etc.; Demeyere *et al.*, 2015; Demeyere *et al.*, 2016). Most of these tools also require a paid license, or paid training certification to gain access and use the materials.

Whilst there has been an increase in remote cognitive assessments, no stroke specific screen was validated for remote use. Domain-specific screening allows health professionals to identify specific cognitive profiles, with both cognitive impairments, as well as domain-specific strengths. This allows for targeted and personalized rehabilitation and discharge planning, instead of providing a broad brush overall cognitive picture in binary terms (impaired/un-impaired classification). The Oxford Cognitive Screen (OCS) is a stroke-specific, in-person screen of cognitive impairment, consisting of 11 normed tasks which assess functioning in the domains of memory, language, praxis, executive functioning, and attention (Demeyere *et al.*, 2015). The OCS provides both raw continuous performance scores per subtask as well as subtask-specific impairment classification compared normative data, and a total proportion of subtasks impaired score. The visual snapshot reports the performance across five cognitive domains. The OCS has been shown to be more sensitive to detecting post-stroke cognitive impairment than both the MoCA and MMSE (Demeyere *et al.*, 2016; Mancuso *et al.*, 2018). Since its original English publication with UK norms, the OCS has seen several cultural and language adaptations and translations, normed and validated for use in different countries (Hong *et al.*, 2018; Huygelier *et al.*, 2020; Kong *et al.*, 2016; Mancuso *et al.*, 2016; Ramos *et al.*, 2018; Robotham *et al.*, 2020; Valera-Gran *et al.*, 2019; Valério *et al.*, 2022) and is now widely used across the globe.

The current study aimed to validate a remote version of the Oxford Cognitive Screen (Demeyere *et al.*, 2015) that can be administered via telephone or videoconferencing. Interim provisional guidance for administration of a remote version of the OCS was released in May 2020 (Raymer, 2020), following frequent requests to the authors. Indeed, there is evidence to suggest that remote versions of the OCS were administered throughout the COVID-19 pandemic (Webb *et al.*, 2022). However, no formal validation of a remote delivery format of OCS had been conducted until now.

Study purpose

The purpose of this study is to compare classification of impairment on the Oxford Cognitive Screen when administered in-person versus remotely (via telephone or videoconferencing) with adults at least 6-months post-stroke. We do not expect any sex differences to affect results. Our pre-registered hypotheses (Webb *et al.*, 2022) were as follows and any deviations from the pre-registration are reported transparently:

1. Each OCS subtask continuous performance score will have moderate to strong associations across in-person and remote versions (ICC $\geq$ 0.50);

2. The sensitivity and specificity values of the in-person and remote version of the OCS will be approximately equitable.

Methods

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study

(Simmons *et al.*, 2012). Approval for the study was granted by the Medical Sciences Interdivisional Research Ethics Committee (First approved July 2021, amendment approval Jan 2022, REC REF: R58224/RE001). We adhere to the STROBE cross-sectional study checklist (von Elm *et al.*, 2007).

Participant identification

Eligibility criteria were: 1) history of confirmed stroke; 2) 18 years of age or older at time of stroke; 3) capacity to consent to research; 4) able to remain alert for at least 20 minutes; and 5) spoke and understood sufficient English. Participants were excluded if they had hearing, language, or visual impairments (not including visual neglect) that would not allow for remote or in-person assessment outside of reasonable adjustments. Any hearing difficulties that became evident whilst testing were overcome by repeating a single instruction twice.

Study design and sample size

We used a prospective cross-sectional within-subjects design. Due to the standard practice of inpatient stroke care within Oxfordshire, most stroke survivors admitted to a hospital for stroke will complete, or partially complete, the OCS. We recruited chronic stroke survivors in the UK between November 2021 to August 2022. All participants completed the OCS when originally admitted to hospital, separate to this study. We contacted long-term stroke survivors from the OXCHRONIC (Kusec *et al.*, 2023) study who had recently (within 30 days at time of contact) completed the tele-OCS as part of the OXCHRONIC protocol. The OXCHRONIC population were at minimum 2 years post-stroke with an average of 4.5 years. Additionally, we asked all participants from the Oxford Cognitive Screening programme, who had just (at time of being contacted) completed an in person 6-month follow up paper OCS (Demeyere *et al.*, 2015; Milosevich *et al.*, 2023), to take part. We aimed to get both the OCS and Tele-OCS completed within 30 days of each other, but this was not always possible. Inclusion criteria for both the Oxford Cognitive Screening programme and OXCHRONIC are very inclusive and aligned with the current inclusion criteria. This means the sampling approaches originally used have minimal bias (except bias of those who are willing to participate), and the underlying samples are representative of the UK stroke population. We acknowledge that by study design, participants will be exposed to the same test twice and no matter the order of administration (remote or in-person) they may improve on the second test. The study design is typical and is shared with multiple other comparison studies of modality equivalence (Brown *et al.*, 2024; Duran *et al.*, 2024; Morrissey *et al.*, 2024). We used two methods to address possible practice effects: We ensured that different sets of the sample receive opposite administration order (Tele-OCS vs. OCS) and different versions (A or b). We further used group statistics to examine if participants improve on their second administered test.

Sample size was pre-determined in our pre-registration at a minimum of 30 participants to detect a two-way random effects single unit agreement based intraclass correlation coefficient (ICC) of greater than .50, with 90% power, and an alpha of 0.05. Recruitment was conducted irrespective of participant demographics, including sex, and how these may relate to OCS subtasks, as this was not a main aim of the study. Sex was determined via medical records of sex-assigned at birth when recruiting participants.

Changes from pre-registration

This project was originally pre-registered at the start of data collection (3 participants had taken part). The preregistered analysis plan was inspired by ICCs between OCS parallel versions A and B in the original OCS normative article (Demeyere *et al.*, 2015), though once data was collected, it became clear that this was not the best approach for the current data that was observed. The small variation of scores observed in the subtasks meant the raw continuous scores ICCs between in-person and remote subtest comparisons were not appropriate. We instead opted to shift towards test of difference between modalities, to retain consistent statistical power. The sample size calculations remained the same as we were similarly powered for equivalent comparisons via difference based *t*-tests. We present ICC analyses in extended data.

In line with the clinical relevance of cognitive screening, we instead used the binary impairment classifications rather than continuous performance scores to determine if deviations between in-person and remote assessments would change interpretation of overall impairment status and hence interpretation of the assessment. We re-ran the analysis after collecting further participants (n = 9) who completed only OCS version A both remotely and in-person due to the timeline in which we changed our analysis plan. We included these participants in the final dataset. Power to detect effects was unchanged when adapting only impairment classification analysis over continuous scores.

The Oxford Cognitive Screen (OCS)

The OCS is usually administered in-person by health care professionals using pen and paper and takes around 15 to 20 minutes to administer (Demeyere *et al.*, 2015). The task can be administered at bedside but is best administered at a table. The OCS requires a testing booklet, examiner forms, and participant pack (see www.ocs-test.org).

The OCS tasks include:

- a picture naming task
- a semantics task (pointing to images based on verbal instruction)
- an orientation task (verbal and multiple choice question response permitted)
- a basic visual field assessment (the examiner holds hands up in upper and lower left and right visual quadrants)
- a sentence to read
- numbers to write
- calculations to complete verbally (or via multiple choice if the participant has expressive problems)

- a spatial inattention task
- a gesture imitation task (the examiner presents gestures with hands and fingers and the participants is instructed to mirror these)
- delayed verbal and episodic recognition of the sentence and previously presented images
- a symbol trail making test (two shape based lines and a mixed trail).

All subtasks have continuous performance scores and are classified as impaired/not impaired based on normed cut-offs (Demeyere *et al.*, 2015). A proportion of tasks impaired score can also be calculated (the summed number of subtasks impaired divided by the number of subtasks attempted). Subtasks belong to five separate cognitive domains; language (picture naming, semantics, sentence reading), number (number writing and number calculation), praxis (gesture imitation), memory (orientation, episodic memory, delayed recall and recognition, and sentence recall), and executive attention (trail making test and spatial inattention test).

The Tele-OCS: a remotely administered version of the Oxford Cognitive Screen

The remote version of the OCS is an adaptation of version A of the OCS. Version A is currently the only version used here, as this was the form used in OxCHRONIC which did not repeat testing with version B. A version B will be available in the future, but is not used in the current study. Administration of the Tele-OCS can be conducted via telephone or video-call. Before the appointment, participants are sent a single testing booklet through the post containing the necessary stimuli to complete the OCS remotely. Participants are instructed not to open the pack before the phone or videocall appointment and are asked to ensure they sit at a table for the duration of the session and have a pen or pencil to hand. The assessor keeps a copy of the remotely sent materials in front of them to jointly follow along with the participant through each subtask. During the session, participants are given specific instructions (provided in the user manual for the examiner) on how to complete each subtask of the Tele-OCS, including the relevant page to turn to in their testing booklet, when to pick up and put down their pen or pencil, and which way to position their booklet on the table. If completing the Tele-OCS via telephone call, participants were instructed to put the telephone on speaker mode, allowing participants freedom to use both hands to complete the subtasks. Upon completion of the assessment session, participants were asked to place the testing booklet into a pre-paid addressed envelope and return it to the assessors via the post. Where necessary, facilitation of page turning and reinforcing assessor instructions was supported by a carer or family member.

Several differences exist between the in-person and remote version of the OCS. Given that the praxis and visual field subtasks require the participant to be in full view of the assessor, these are not assessed as part of the Tele-OCS. To assess constructional praxis, the Tele-OCS instead contains a figure-copy task, requiring participants to copy, and immediately recall, a complex figure. The figure was taken from the OCS-Plus assessment, under the same copyright as the OCS (Demeyere *et al.*, 2015). The spatial attention task was adapted for remote administration by giving two practice rounds, instead of one, to complete before the main task. This was put in place to allow the examiner to gauge whether the participant has understood the instructions, without being able to see the practice trials. On the first round, participants practice a short version of the spatial attention subtask, where they are presented with six line-drawings of hearts, three of which contain a gap on either the left or right-hand side. Participants are instructed to only cross out the complete hearts on the page. On the second round, participants verbally state which of these hearts have a gap and which ones are complete. This additional practice round allows the assessor to evaluate whether the participant has understood the instructions before they proceed to the main task.

Finally, adaptations were made to the remote version of the OCS trail making task. When this subtask is administered in-person, participants are shown a completed example by the assessor which they can refer to throughout the duration of the task. As this is not possible through remote administration, instructions have been printed at the top of the practice task and the main task in the remote testing booklet. The materials for the Tele-OCS and the manual for administration are freely available for publicly funded clinical and research use in the same way as the in-person OCS. The Tele-OCS is licensed through Oxford University Innovations Health Outcomes (https://innovation.ox.ac.uk/outcome-measures/the-oxford-cognitive-screen-ocs/).

We timed the duration of the Tele-OCS administration for a portion of participants only (n = 18). We additionally recorded any pragmatic information regarding the optimal assessment of the Tele-OCS. This involved recording unsolicited spontaneous feedback from participants on the Tele-OCS after administration.

Procedure

Once participants from the Oxford Cognitive Screening programme were identified as having just (within 30 days) completed an in-person OCS in their homes, we asked if they would complete the Tele-OCS. For those who agreed, we posted the Tele-OCS pack at least 1 week before the appointment and reminded them not to open the pack until instructed. In the session, informed consent was taken and the Tele-OCS was administered either by phone or videoconferencing depending on the preference of the participant. Shifting modality during a session was not permitted, nor did this come up during data collection. We provided a pre-paid envelope for participants to post their packs back. If participants were recruited from the OXCHRONIC study, we contacted them within 30 days of having completed the Tele-OCS for the OXCHRONIC protocol, and administered the OCS in person at their homes.

Data analysis

Prior to formal analyses we examined sample characteristics, including demographics and prior OCS performance taken within 20 days of stroke. Acute OCS performance is presented in Table S2 in the extended data.

First, we summarized the administration time of participants and comments given by participants about the Tele-OCS. Cut offs for impairment per OCS subtask were created using the original OCS normative cut offs from healthy aging data (Demeyere *et al.*, 2015). Then we visually depicted proportion of impaired subtasks per modality and impairment classification per sub-task for both modalities. We further explored continuous performance scores of participants who were classed as unimpaired on the in-person OCS but impaired on the Tele-OCS, and vice versa. We compared raw continuous performance scores for statistically significant differences uses *t*-tests and Cohen's *d* effect size.

All statistical analysis and data sorting was computed in R Studio (Posit team, 2024) using R (R Core Team, 2021) version 4.0.4. We used the following packages for the production of the RMarkdown manuscript and analysis: *bookdown* version 0.26 (Xie, 2021); *readxl* version 1.3.1 (Wickham & Bryan, 2019); *cowplot* version 1.1.1 (Wilke, 2020); *ggplot2* version 3.3.5 (Wickham, 2016); *kableExtra* version 1.3.4 (Zhu, 2021); *httr* version 1.4.2 (Wickham, 2022); and *tidyr* 1.2.0 (Wickham, 2021). Data and analysis scripts to recreate the manuscript are openly available in CC-BY 4.0 license (Webb *et al.*, 2023).

Results

Sample

40 chronic stroke survivors took part in this study, with demographics of participants presented in Table 1. The median time post-stroke was 228 days (*IQR=189.25 to 1500.50).* One participant did not return their Tele-OCS pack, as such we were unable to generate scores for the number writing, broken hearts, or trails accuracies subtasks. We include this participant in the majority analyses where complete data were available, including proportion of subtasks and domains impaired as proportion accounts for missing subtasks. We present OCS data taken within 20 days of stroke in the extended data for context on generalisability of the current sample and impairment levels.

Tele-OCS overview

The time in days between in-person and Tele-OCS administration ranged from 6 to 72, with an average of *M*=27.70 days (*SD*=12.29). 23 participants were administered the Tele-OCS first (57.5%). Time interval was not correlated with any differences in subtasks impaired on either modality ($r(38)$= -0.03, p=.85). Administration of the Tele-OCS took an average time of 17.17 minutes (*SD*=2.73, range = 14 - 22). The Tele-OCS was administered using phone to the majority of participants (*n*=36, 90%).

14 participants volunteered feedback after the Tele-OCS was administered. Nine participants commented that there were limited differences in their experience between in-person and remote administration. Three participants had issues with their phone connection, which broke up the session slightly, but minimally impacted administration. One person had difficulty finding the correct pages and, due to a hemiplegia, required help from his partner to turn the pages. One participant required a repetition of one instruction once due to hearing difficulties. Otherwise, participant remote sessions were administered without major barriers.

Comparison of the Tele-OCS and standard OCS

We examined the proportion of subtasks impaired on the in-person versus Tele-OCS, and then the frequency of impairment classifications for each OCS subtask across modalities. Participants were impaired on an average of 5.71 (*SD*=7.79) subtasks on the in-person version and 5.71 (*SD*=8.59) subtasks in the remote version. In Figure 1 we present the Kendall rank correlation coefficient for proportion of subtasks impaired between remote and in-person OCS versions. We visually depict impairment classification distributions for all OCS scores in Figure 2 to illustrate the similarities / differences between modalities. Figure 2 shows that, in terms of count, there were a few individuals who were impaired only on the remote version of the task, but not in the in-person version. In Table 2, we present descriptive statistics for raw

Table 1. Demographics of the participants included in the final analysis.

Characteristic	N (missing%)	Value
Age (M (*SD*))	40 (0%)	69.30 (10.44)
Education (M (*SD*))	40 (0%)	13.45 (2.85)
Sex	40 (0%)	Female: 30%; Male: 70%;
Ethnicity	40 (0%)	Black-Caribbean: 2.5%; White-British: 97.5%;
Stroke type	40 (0%)	Subarachnoid haemorrhage: 2.5%; Intracerebral haemorrhage: 27.5%; Ischaemic: 70%;
Stroke side	40 (0%)	Bilateral: 10%; Left: 45%; Right: 45%;
NIHSS (median (IQR))	37 (8%)	7 (3–9)

Note. NIHSS refers to National Institute of Health Stroke Scale for stroke severity. IQR refers to interquartile range.

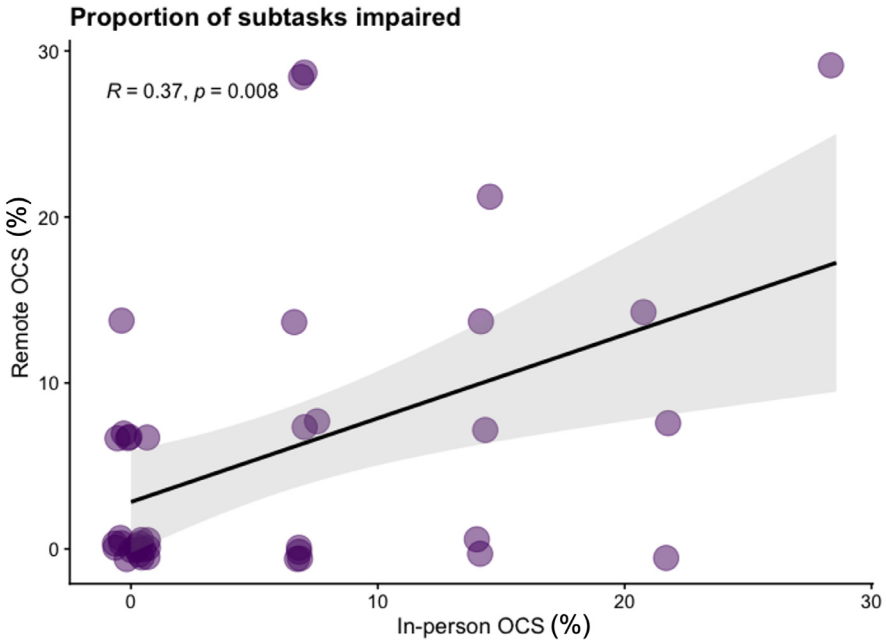

Figure 1. Scatter plot of the proportion of subtasks completed which are classified as impaired on the in-person Oxford Cognitive Screen (OCS) and Tele-OCS. R = Kendall rank correlation coefficient. Figure available under CC-BY 4.0 license (Figure 1, Webb *et al.*, 2023).

score performance on each shared OCS subtask as well as repeated-measures Cohen's *d* and significance of difference on *t*-tests between each subtask. Equivalence tests were additionally conducted, paralleling the t-test results presented here (see Supplementary Materials).

The results show that regardless of numerical differences in score, there were few statistically significant differences between raw performance scores on each subtask and overall proportion of tasks impaired. We present sensitivity/specificity analysis in the extended data as not all analysis could be completed due to low and differing numbers of impaired impairment classifications.

Discussion

We aimed to validate the use of the Tele-Oxford Cognitive Screen (Tele-OCS) by comparing performance within-subjects on the in-person OCS (OCS version B) and remote OCS (Tele-OCS version A), via telephone or video conferencing. Overall, the proportion of those impaired on the OCS subtasks was moderately related between in-person and remote administration. Within each of the subtasks, whilst we found small statistically significant deviations in raw scores in two tasks (picture naming and executive score), impairment classifications on each task were minimally affected by modality.

It is possible that differences in OCS versions (A vs B) led to increased variability in performance on OCS subtasks between modalities. This may be particularly the case for the picture naming task, where two participants were impaired on the Tele-OCS (based on version A), but not on the in-person OCS (based on version B). There are small differences between version A and B of the picture-naming task, irrespective of modality. For example, participants are more likely to make errors on version A, by calling the hippo a 'rhino' and calling the slices of melon, 'fruit' (see OCS control normative data, Demeyere *et al.*, 2015; in Webb *et al.*, 2023). For the executive score, although scores between modalities were statistically different, impairment ranges were nearly identical, with only one participant classed as impaired on the remote version and unimpaired on the in-person version.

Our results align with other validation studies of remote versions of commonly used cognitive screening tools. One longitudinal study comparing performance on telephone versus in-person administration of a cognitive test battery of multiple domains (attention, verbal learning and memory, verbal fluency, executive function, working memory and global cognitive functioning) (Rapp *et al.*, 2012). Statistical differences were found between in-person and remote assessment scores, with the most reliable test-retest data found using the same modality (Rapp *et al.*, 2012). Despite these small differences, remote and in-person administrations yielded equivalent interpretations in most cases, with the authors concluding that this remotely assessed battery was comparable to its in-person version. Katz *et al.* (2021) compared performance on the MoCA and Tele-MoCA, finding small but non-meaningful differences in performance between modalities (e.g., non-significant equivalence tests). Overall, prior research, along with the results of the present investigation, suggest that slight differences between remote and in-person versions of the same cognitive test are to be expected. It is possible that continuous performance scores, when

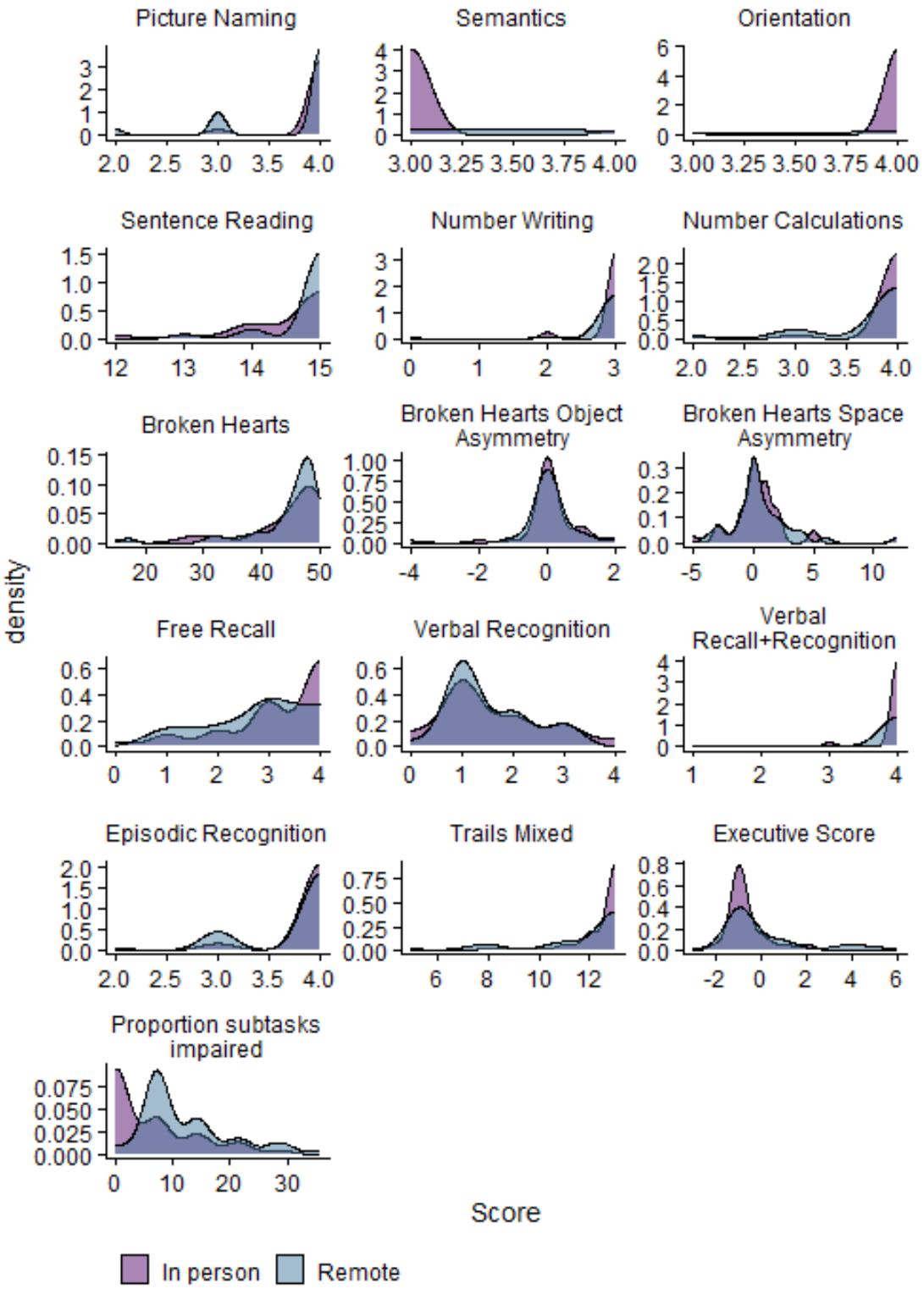

Figure 2. Illustrates the density of scores per subtask, and per OCS modality (coloured by in-person Oxford Cognitive Screen (OCS) or Tele-OCS). Figure available under CC-BY 4.0 license (Figure 2, Webb *et al.*, 2023).

Table 2. Descriptive statistics for subtasks per domain on the Tele-OCS and in-person OCS.

Domain	Subtask	N	In-person M(SD)	Remote M(SD)	Cohen's *d*
Language	Picture Naming	40	3.92 (0.27)	3.70 (0.56)	0.36*
	Semantics	40	3.05 (0.22)	3 (0)	0.23
	Sentence Reading	40	14.53 (0.82)	14.80 (0.52)	-0.31
Number	Number Writing	39	2.92 (0.27)	2.87 (0.52)	0.11
	Number Calculations	40	3.90 (0.38)	3.75 (0.54)	0.23
Memory	Orientation	40	3.98 (0.16)	4 (0)	-0.16
	Verbal Recall Recognition	40	3.95 (0.22)	3.83 (0.59)	0.24
	Episodic Recognition	40	3.88 (0.4)	3.80 (0.41)	0.13
Attention	Broken Hearts	39	43.85 (7.96)	44.67 (6.41)	-0.20
	Broken Hearts Object Asymmetry	39	0.17 (0.68)	0 (0.83)	0.24
	Broken Hearts Space Asymmetry	39	0.50 (2.65)	0.67 (2.74)	-0.05
Executive	Trails Mixed	39	11.03 (2.88)	11.74 (2.11)	0.26
	Executive Score	39	-0.68 (0.86)	0.10 (2.02)	-0.46*
Overall	Subtasks Impaired (%)	40	5.71 (7.79)	5.71 (8.59)	0

Note. '*' signifies *p*-value <.05. Multiple comparison alpha corrected level is .005 (.05/11 = .005). '**' signifies *p*-values <.005).

delivered remotely, may be more likely to fluctuate, possibly due to technical problems, lack of examiner presence to guide correct instructions (e.g., for the trail-making task), and hearing or language difficulties that become exacerbated during remote cognitive assessment. However, overall differences are minor, and should not be a barrier to administering cognitive tests in remote format nor to overall interpretation of participant results. The Tele-OCS thus provides a highly comparable assessment of cognitive functioning.

There are several stroke-specific considerations concerning remote cognitive assessments. For instance, motor impairments may affect stroke survivors' ability to complete cognitive tests in remote format. One participant in the present study remarked that turning pages in the Tele-OCS booklet was difficult due to their hemiplegia. Motor difficulties such as these should be taken into consideration by clinicians when assessing stroke survivors remotely, so that best-practice may be established. Additionally, visuospatial neglect, a common post-stroke impairment, may impact remote cognitive test performance. Participants should be reminded to keep their remote testing booklet as close to their mid-line as possible so that cognitive performance in other domains is not impacted by neglect symptoms. The in-person and Tele-OCS are designed to be spatial neglect friendly, with stimuli presented centrally on all pages. Involvement of a family member or carer at home when conducting a remote assessment may be beneficial for participants that require additional support where in person assessment is not possible.

Remote cognitive assessment has multiple benefits. Delivering cognitive assessments in remote format can increase the accessibility of these assessments to a wider range of individuals (Webb *et al.*, 2022). Similarly, remote cognitive assessments can reduce participant burden for those participating in research studies, given that individuals do not have to be physically present to participate. Those who would not usually participate in research may be encouraged to do so if the modality of assessment is remote. It should be noted that the present investigation was the first to partially validate a *stroke-specific* cognitive screening tool for remote use. Given that commonly used remote assessments, such as the MoCA and Tele-MoCA, do not usually assess stroke-specific cognitive impairments (e.g., unilateral spatial neglect and aphasia; Laska *et al.*, 2001; Moore *et al.*, 2021), the Tele-OCS can improve the accessibility of adequate cognitive screening to stroke survivors during their recovery, increasing the likelihood of appropriate cognitive review if necessary.

Limitations

Although our study sample was statistically powered, the sample size was small and relatively homogeneous. For example, 97.5% of the sample was White-British, 70% were male, and around 70% were ischaemic stroke. Ideally, there would be around 82% White-British participants to reflect UK government 2021 census data for ethnicity (ONS, 2022), or >93% for older adults specifically, or 87% for general UK stroke survivors (Douiri *et al.*, 2021), approximate equality of males and female, and around 87–90% ischemic stroke (Douiri *et al.*, 2021;

King *et al.*, 2020). Strengths of our sample, however, include age and stroke severity, which represent the general UK stroke population (Douiri *et al.*, 2021).

Furthermore, given that many participants were in the chronic stage of stroke, they were often classified as unimpaired across subtasks in both modalities, limiting our interpretation for stroke survivors with more severe impairments. This may impact the generalisability of the present results to the wider stroke population. Although the majority of stroke survivors tend to be males, our weighting of males to females was biased as we had far more males. This may impact the generalisability of our results to females. We do not feel that the gender balance in our data impacts the results drastically, however, we acknowledge this as a limitation. We did not choose to purposively sample for sex specifically as there is limited and mixed evidence that sex alters global cognitive scoring after a stroke, however, sex differences do exist in pre-stroke risk factors, as well as type of stroke and non-cognitive recovery which we did not factor in this study (Duran *et al.*, 2024; Indraswari *et al.*, 2024). Future validation studies of the Tele-OCS may benefit from using larger, more heterogeneous samples, including those who have sustained more severe strokes, and more representative demographics. Moreover, it is necessary to consider that those who consent to take part in research may be different from those who do not, and as such more inclusive study protocols using those with in-capacity to consent to research may help contribute to more reflective demographics. There are reported differences in demographics between those who have capacity to consent to take part in a research study versus those who do not; typically those without capacity in this context are older, more neurologically impaired, and more frequently have aphasia (Lindley *et al.*, 2022; Thomalla *et al.*, 2017).

There is currently no normative data for performance on the Tele-OCS for comparison of scores, as our results rely on the in-person normative data to generate impairment classifications we cannot know if these impairment classifications are accurate. Further research will need to determine this. Furthermore, a small part of our study concerned different modalities of remote assessment (phone vs videocall). We did not aim to assess whether the modality of remote assessment affected performance. We relied on participant preference above all else to determine modality.

Additionally, we had too small a sample to do so. Given identical instructions and materials, there should be limited differences between modalities. Furthermore, it should be noted that stroke survivors, and indeed current older adults, do not always have the correct equipment or support to carry out assessment via videoconferencing. There are multiple barriers for older adults including patient acceptability, software used to make the call, hardware to do so, and health impairments related to independently hosting a videoconference and completing tasks simultaneously (Niyomyart *et al.*, 2024). Additionally, while internet access is common for stroke survivors, the ones with more access tend to be younger, white, and live longer, introducing further biases to research samples (Naqvi *et al.*, 2021; Zhu *et al.*, 2022). Indeed, other equivalence/videoconferencing studies have used younger, less impaired samples than the current study, potentially explaining their high use of videoconferencing and assumed practicality for older aged stroke survivors (Chapman *et al.*, 2021; Gnassounou *et al.*, 2022). However, it may be worthwhile for future studies with larger sample sizes to randomly allocate participants to either telephone or videoconferencing to investigate feasibility in older aged stroke survivors with cognitive impairment whether there are any differences between these types of remote assessment methods.

Finally, all participants were previously exposed to the OCS, whether as part of our research, or as part of their admission to hospital where OCS is given as standard where it can be. This means, there may be a chance for practice effects. We note, we did not observe many in the current study, so it is reasonable to assume less exist between first exposure and the current study. Alternatively, due to previous exposure, participants may be better at the OCS in general, which would explain our low rates of impairment. As other studies using the OCS have found impairment rates, even for those previously exposed, and at different time points post-stroke (Kusec *et al.*, 2023; Milosevich *et al.*, 2023), we do not feel this is a major issue in the current study.

Implications for research and clinical practice

The present investigation provides insight into how to best administer the Tele-OCS. Sending a paper pack in the post allows for clear presentation of instructions and stimuli while carrying out Tele-OCS assessment. However, there are also several drawbacks to this method; participants must be instructed not to look at the testing pack ahead of the testing session and it can be time-consuming (and potentially costly to NHS services) to send the pack via post. However, use of a paper testing pack allows stroke survivors who do not have access to videoconferencing to carry out the testing session over the telephone, without having to travel to a research or clinic site. In our sample, the majority of participants opted to complete the Tele-OCS over the phone, rather than using videoconferencing. This may be because they are less familiar with, or do not have access to, the required technology.

We encourage health professionals to take these factors into consideration when choosing to use the Tele-OCS. Clinicians who are new to remote assessment are encouraged to carry out practice administrations with the Tele-OCS before using it in a clinical context.

Conclusion

The OCS, when administered remotely, is a valid method of screening for cognitive impairment among stroke survivors.

Clinicians should use existing published normative data to interpret impairment on OCS sub-tests when delivered remotely.

Data availability

Underlying data

Open Science Framework: Introducing the Tele-OCS: A validated remotely administered version of The Oxford Cognitive Screen. https://doi.org/10.17605/OSF.IO/8RNY2 (Webb *et al.*, 2023)

This project contains the 'Reproducible manuscript' folder which you can download as a .zip file, and using the .Rproj file, can reproduce our base manuscript and analyses. This also contains 'data' subfolder with this research's underlying data.

Extended data

Open Science Framework: Introducing the Tele-OCS: A validated remotely administered version of The Oxford Cognitive Screen. https://doi.org/10.17605/OSF.IO/8RNY2 (Webb *et al.*, 2023)

This project contains the 'Supplementary materials' folder.

Acknowledgements

We want to thank the whole OX-CHRONIC team for their help in remote assessment of participants and for provision of materials. With special thanks to Evangeline Grace Chiu for their help in further developing the Tele-OCS booklet and instructions.

References

Amy YX, Kapral MK: **More people are surviving after acute stroke.** *BMJ.* 2019; **365**: l2150.
Publisher Full Text

Benge JF, Kiselica AM: **Rapid communication: preliminary validation of a telephone adapted Montreal Cognitive Assessment for the identification of mild cognitive impairment in Parkinson's disease.** *Clin Neuropsychol.* 2021; **35**(1): 133–147.
PubMed Abstract | **Publisher Full Text**

Brandt J, Spencer M, Folstein M: **The telephone interview for cognitive status.** *Neuropsychiatry Neuropsychol Behav Neurol.* 1988; **1**(2): 111–117.
Reference Source

Brown AD, Kelso W, Eratne D, *et al.*: **Investigating equivalence of in-person and telehealth-based neuropsychological assessment performance for individuals being investigated for Younger Onset Dementia.** *Arch Clin Neuropsychol.* 2024; **39**(5): 594–607.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Chapman JE, Gardner B, Ponsford J, *et al.*: **Comparing performance across in-person and videoconference-based administrations of common neuropsychological measures in community-based survivors of stroke.** *J Int Neuropsychol Soc.* 2021; **27**(7): 697–710.
PubMed Abstract | **Publisher Full Text**

Chapman JE, Ponsford J, Bagot KL, *et al.*: **The use of videoconferencing in clinical neuropsychology practice: a mixed methods evaluation of neuropsychologists' experiences and views.** *Aust Psychol.* 2020; **55**(6): 618–633.
Publisher Full Text

Demeyere N, Riddoch MJ, Slavkova ED, *et al.*: **The Oxford Cognitive Screen (OCS): validation of a stroke-specific short cognitive screening tool.** *Psychol Assess.* 2015; **27**(3): 883–894.
PubMed Abstract | **Publisher Full Text**

Demeyere N, Riddoch MJ, Slavkova ED, *et al.*: **Domain-specific versus generalized cognitive screening in acute stroke.** *J Neurol.* 2016; **263**(2): 306–315.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Douiri A, Muruet W, Bhalla A, *et al.*: **Stroke care in the United Kingdom during the COVID-19 pandemic.** *Stroke.* 2021; **52**(6): 2125–2133.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Duran T, Gaussoin SA, Latham LA, *et al.*: **Examining a Preclinical Alzheimer's Cognitive Composite for telehealth administration for reliability between in-person and remote cognitive testing with neuroimaging biomarkers.** *J Alzheimers Dis.* 2024; **99**(2): 679–691.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Esmael A, Elsherief M, Eltoukhy K: **Prevalence of cognitive impairment in acute ischaemic stroke and use of Alberta Stroke Programme Early CT Score (ASPECTS) for early prediction of post-stroke cognitive impairment.** *Neurol Neurochir Pol.* 2021; **55**(2): 179–185.
PubMed Abstract | **Publisher Full Text**

Fox-Fuller JT, Rizer S, Andersen SL, *et al.*: **Survey findings about the experiences, challenges, and practical advice/solutions regarding teleneuropsychological assessment in adults.** *Arch Clin Neuropsychol.* 2022; **37**(2): 274–291.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Gnassounou R, Defontaines B, Denolle S, *et al.*: **Comparison of neuropsychological assessment by videoconference and face to face.** *J Int Neuropsychol Soc.* 2022; **28**(5): 483–493.
PubMed Abstract | **Publisher Full Text**

Hammers DB, Stolwyk R, Harder L, *et al.*: **A survey of international clinical teleneuropsychology service provision prior to and in the context of COVID-19.** *Clin Neuropsychol.* 2020; **34**(7–8): 1267–1283.
PubMed Abstract | **Publisher Full Text**

Hong WJ, Tao J, Wong AWK, *et al.*: **Psychometric Properties of the Chinese (Putonghua) Version of the Oxford Cognitive Screen (OCS-P) in Subacute Poststroke Patients without Neglect.** *Biomed Res Int.* 2018; **2018**: 6827854.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Huygelier H, Schraepen B, Demeyere N, *et al.*: **The Dutch version of the Oxford Cognitive Screen (OCS-NL): normative data and their association with age and socio-economic status.** *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn.* 2020; **27**(5): 765–786.
PubMed Abstract | **Publisher Full Text**

Indraswari F, Yaghi S, Khan F: **Sex specific outcomes after ischemic stroke.** *J Stroke Cerebrovasc Dis.* 2024; **33**(8): 107754.
PubMed Abstract | **Publisher Full Text**

Intercollegiate Stroke Working Party: **National clinical guideline for stroke for the UK and Ireland**. Intercollegiate Stroke Working Party, 2023.
Reference Source

Intercollegiate Stroke Working Party: **Stroke guidelines.** RCP London, 2016.
Reference Source

Katz MJ, Wang C, Nester CO, *et al.*: **T-MoCA: a valid phone screen for cognitive impairment in diverse community samples.** *Alzheimers Dement (Amst).* 2021; **13**(1): e12144.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

King D, Wittenberg R, Patel A, *et al.*: **The future incidence, prevalence and costs of stroke in the UK.** *Age Ageing.* 2020; **49**(2): 277–282.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Kong AP, Lam PH, Ho DW, *et al.*: **The Hong Kong version of the Oxford Cognitive Screen (HK-OCS): validation study for Cantonese-speaking chronic stroke survivors.** *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn.* 2016; **23**(5): 530–548.
PubMed Abstract | **Publisher Full Text**

Kusec A, Milosevich E, Williams OA, *et al.*: **Long-term psychological outcomes following stroke: the OX-CHRONIC study.** *medRxiv.* 2023; 2023–03.
Publisher Full Text

Laska AC, Hellblom A, Murray V, *et al.*: **Aphasia in acute stroke and relation to outcome.** *J Intern Med.* 2001; **249**(5): 413–422.
PubMed Abstract | **Publisher Full Text**

Lindley RI, Kane I, Cohen G, *et al.*: **Factors influencing the use of different methods of consent in a randomized acute stroke trial: the third International Stroke Trial (IST-3).** *Int J Stroke.* 2022; **17**(5): 553–558.
PubMed Abstract | Publisher Full Text | Free Full Text

Mancuso M, Demeyere N, Abbruzzese L, *et al.*: **Using the Oxford Cognitive Screen to detect cognitive impairment in stroke patients: a comparison with the Mini-Mental State Examination.** *Front Neurol.* 2018; **9**: 101.
PubMed Abstract | Publisher Full Text | Free Full Text

Mancuso M, Varalta V, Sardella L, *et al.*: **Italian normative data for a stroke specific cognitive screening tool: the Oxford Cognitive Screen (OCS).** *Neurol Sci.* 2016; **37**(10): 1713–1721.
PubMed Abstract | Publisher Full Text

Merriman NA, Sexton E, McCabe G, *et al.*: **Addressing cognitive impairment following stroke: systematic review and meta-analysis of non-randomised controlled studies of psychological interventions.** *BMJ Open.* 2019; **9**(2): e024429.
PubMed Abstract | Publisher Full Text | Free Full Text

Milosevich ET, Moore M, Pendlebury ST, *et al.*: **Domain-specific cognitive impairment 6 months after stroke: the value of early cognitive screening.** *medRxiv.* 2023; 2023–06.
Publisher Full Text

Mole JA, Demeyere N: **The relationship between early post-stroke cognition and longer term activities and participation: a systematic review.** *Neuropsychol Rehabil.* 2020; **30**(2): 346–370.
PubMed Abstract | Publisher Full Text

Moore MJ, Vancleef K, Riddoch MJ, *et al.*: **Recovery of visuospatial neglect subtypes and relationship to functional outcome six months after stroke.** *Neurorehabil Neural Repair.* 2021; **35**(9): 823–835.
PubMed Abstract | Publisher Full Text | Free Full Text

Morrissey S, Gillings R, Hornberger M: **Feasibility and reliability of online vs in-person cognitive testing in healthy older people.** *PLoS One.* 2024; **19**(8): e0309006.
PubMed Abstract | Publisher Full Text | Free Full Text

Munro Cullum C, Hynan LS, Grosch M, *et al.*: **Teleneuropsychology: evidence for video teleconference-based neuropsychological assessment.** *J Int Neuropsychol Soc.* 2014; **20**(10): 1028–1033.
PubMed Abstract | Publisher Full Text | Free Full Text

Naqvi IA, Montiel TC, Bittar Y, *et al.*: **Internet access and usage among stroke survivors and their informal caregivers: cross-sectional study.** *JMIR Form Res.* 2021; **5**(3): e25123.
PubMed Abstract | Publisher Full Text | Free Full Text

Newkirk LA, Kim JM, Thompson JM, *et al.*: **Validation of a 26-point telephone version of the Mini-Mental State Examination.** *J Geriatr Psychiatry Neurol.* 2004; **17**(2): 81–87.
PubMed Abstract | Publisher Full Text

Niyomyart A, Ruksakulpiwat S, Benjasirisan C, *et al.*: **Current status of barriers to mHealth access among patients with stroke and steps toward the digital health era: systematic review.** *JMIR mHealth uHealth.* 2024; **12**: e54511.
PubMed Abstract | Publisher Full Text | Free Full Text

Nys GM, van Zandvoort MJ, van der Worp HB, *et al.*: **Early cognitive impairment predicts long-term depressive symptoms and quality of life after stroke.** *J Neurol Sci.* 2006; **247**(2): 149–156.
PubMed Abstract | Publisher Full Text

Office for National Statistics: **Ethnic Group, England and Wales: Census 2021.** ONS website, statistical bulletin, 2022.
Reference Source

Pendlebury ST, Welch SJV, Cuthbertson FC, *et al.*: **Telephone assessment of cognition after transient ischemic attack and stroke: modified telephone interview of cognitive status and telephone Montreal Cognitive Assessment versus face-to-face Montreal Cognitive Assessment and neuropsychological battery.** *Stroke.* 2013; **44**(1): 227–9.
PubMed Abstract | Publisher Full Text | Free Full Text

Posit team: **RStudio: integrated development environment for R**. Posit Software, PBC, 2024.
Reference Source

Quinn TJ, Richard E, Teuschl Y, *et al.*: **European Stroke Organisation and European Academy of Neurology joint guidelines on post-stroke cognitive impairment.** *Eur Stroke J.* 2021; **6**(3): I–XXXVIII.
PubMed Abstract | Publisher Full Text | Free Full Text

R Core Team: **R: A language and environment for statistical computing.** R Foundation for Statistical Computing, 2021.
Reference Source

Ramos CCF, Amado DK, Borges CR, *et al.*: **Oxford Cognitive Screen - Brazilian Portuguese version (OCS-Br) a pilot study.** *Dement Neuropsychol.* 2018; **12**(4): 427–431.
PubMed Abstract | Publisher Full Text | Free Full Text

Rapp SR, Legault C, Espeland MA, *et al.*: **Validation of a cognitive assessment battery administered over the telephone.** *J Am Geriatr Soc.* 2012; **60**(9): 1616–1623.
PubMed Abstract | Publisher Full Text | Free Full Text

Raymer J: **Oxford Cognigive Screen remote administration version now available!** 2020.
Reference Source

Robotham RJ, Riis JO, Demeyere N: **A Danish version of the Oxford cognitive screen: a stroke-specific screening test as an alternative to the MoCA.** *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn.* 2020; **27**(1): 52–65.
PubMed Abstract | Publisher Full Text

Rudd AG, Bowen A, Young G, *et al.*: **National clinical guideline for stroke: 2016.** Clinical Medicine, 2017.
Reference Source

Sexton E, McLoughlin A, Williams DJ, *et al.*: **Systematic review and meta-analysis of the prevalence of cognitive impairment no dementia in the first year post-stroke.** *Eur Stroke J.* 2019; **4**(2): 160–171.
PubMed Abstract | Publisher Full Text | Free Full Text

Simmons JP, Nelson D, Simonsohn U: **A 21 Word Solution.** 2012.
Publisher Full Text

Thomalla G, Boutitie F, Fiebach JB, *et al.*: **Effect of informed consent on patient characteristics in a stroke thrombolysis trial.** *Neurology.* 2017; **89**(13): 1400–1407.
PubMed Abstract | Publisher Full Text | Free Full Text

Valera-Gran D, López-Roig S, Hurtado-Pomares M, *et al.*: **Validation of the Spanish version of the Oxford Cognitive Screen (S-OCS): Psychometric properties of a short cognitive stroke-specific screening tool.** *Clin Rehabil.* 2019; **33**(4): 724–736.
PubMed Abstract | Publisher Full Text

Valério D, Almeida J, Demeyere N, *et al.*: **The European Portuguese version of the Oxford Cognitive Screening (OCS-Pt): a screening test for acute stroke patients.** *Neurol Sci.* 2022; **43**(6): 3717–3728.
PubMed Abstract | Publisher Full Text

von Elm E, Altman DG, Egger M, *et al.*: **The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.** *Lancet.* 2007; **370**(9596): 1453–1457.
PubMed Abstract | Publisher Full Text

Webb SS, Carrick C, Kusec A, *et al.*: **Introducing the Tele-OCS: a validated remotely administered version of The Oxford Cognitive Screen [Registration].** *OSF.* 2023.
http://www.doi.org/10.17605/OSF.IO/6MFJX

Webb SS, Kontou E, Demeyere N: **The COVID-19 pandemic altered the modality, but not the frequency, of formal cognitive assessment.** *Disabil Rehabil.* 2022; **44**(21): 6365–6373.
PubMed Abstract | Publisher Full Text

Wickham H: **Ggplot2.** (2nd ed.). Springer International Publishing, 2016.
Publisher Full Text

Wickham H: **tidyr: Tidy messy data.** [Manual], 2021.
Reference Source

Wickham H: **httr: Tools for working with URLs and HTTP.** (R package version 1.4.2), 2022.
Reference Source

Wickham H, Bryan J: **readxl: Read Excel Files.** (R package version 1.3.1), 2019.
Reference Source

Wilke CO: **cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'.** (R package version 1.1.1), 2020.
Reference Source

Xie Y: **bookdown: Authoring books and technical documents with r markdown.** [Manual], 2021.
Reference Source

Zhu H: **kableExtra: Construct complex table with 'kable' and pipe syntax.** [Manual], 2021.
Reference Source

Zhu C, Tran PM, Dreyer RP, *et al.*: **Disparities in internet use among US stroke survivors: implications for telerehabilitation during COVID-19 and beyond.** *Stroke.* 2022; **53**(3): e90–e91.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ? ✔ ?

Version 2

Reviewer Report 07 July 2025

https://doi.org/10.21956/healthopenres.14761.r29448

?

Hayley Wright
[1] Coventry University, Coventry, England, UK
[2] Coventry University, Coventry, England, UK

Thank you for the opportunity to review this paper. This is a novel and strong contribution to an under-researched and under-resourced area of stroke care. Overall, the paper is excellent and gives a thorough account of the validity-testing of remote vs in-person testing of a stroke-specific and domain-specific cognitive screen.

Minor comment for the introduction (first paragraph) is to cite more recently updated stroke guidelines, such as the National Clinical Guideline for Stroke (2023).

Methods: "In the current study we attempted to reflect the UK stroke population in key demographics of age, education, sex, stroke type, stroke side, and stroke severity." Can the authors please clarify how this was done? Perhaps stratified sampling or similar? Can they also clarify how this reconciles with the following statement further down "Recruitment was conducted irrespective of participant demographics..."

Can the authors add some justification for why the Tele-OCS was based on Version A of the OCS (rather than version B, or creating a remote version for both)?

By recruiting participants from the OCS programme and OXCHRONIC, are you introducing any test-retest bias to the results? That is, do practice effects improve performance, such that any differences between modalities are minimised by exposure, rather than the method of administration per se? This is mentioned in the limitations section in the discussion but does not satisfy my doubts as a reader. Can further evidence be presented to support the limited practice effects from other studies?

The limitations section of the discussion is thoughtful but needs supporting with references. In paces, the suggestions feel anecdotal or vague. For example:

- "We did not choose to purposively sample for sex specifically as it is not thought to alter the

results of global cognitive scoring after a stroke" - reference needed.
- "Moreover, it is necessary to consider that those who consent to take part in research may be different from those who do not, and as such more inclusive study protocols using those with in-capacity to consent to research may help contribute to more reflective demographics." - reference to support that ways in which those who lack capacity to consent may hold different demographic characteristics to those who have capacity.
- "Furthermore, it should be noted that stroke survivors do not always have the correct equipment to carry out assessment via videoconferencing" - reference? This feels like a sweeping statement - could potentially be perceived as stereotyped / prejudiced in its tone. Can you be clear about whether this was borne out in your data, or where the evidence comes from to support this, and how it can be remedied in future research? E.g. providing / loaning smartphones or tablets for participants to complete remote assessments with.

Minor note about grammar - some sentences are very long and hard to follow, e.g. "We did not aim to assess whether the modality of remote assessment affected performance, as we relied on participant preference above all else, had too small a sample to do so, given identical instructions and materials, there should be limited differences". Breaking down into shorter sentences will help readers (particularly lay, or less-strong, readers) to understand the importance and impact of the work.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Cognitive psychology; digital health interventions

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 11 April 2024

https://doi.org/10.21956/healthopenres.14761.r28061

Dimosthenis Tsapekos
[1] King's College London, London, England, UK
[2] King's College London, London, England, UK

The article has been much improved.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Neuropsychology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 17 August 2023

https://doi.org/10.21956/healthopenres.14374.r27335

?

Dimosthenis Tsapekos
[1] King's College London, London, England, UK
[2] King's College London, London, England, UK
[3] King's College London, London, England, UK

- In the abstract a value of d=0 is reported, is this a typo? It's quite unusual for an effect size metric to be exactly 0.
- The introduction sufficiently motivates the need to validate a remotely administrated version of the OCS and the conduct of the study.
- The issue of previous participation in research needs to be clarified. Does this involve previous administration of the OCS? This could affect the outcomes through learning effects.
- The two versions of the measure are sufficiently described.
- Data analysis description is complete.
- Visual presentation of data facilitates the understanding of the findings.
- The authors present impairment rates for subtasks. It's not clear whether there is impairment classification based on the overall score of the two OCS versions. This would be very useful to clarify and report.
- Limitations/Implications are sufficiently described.
- I think it's important for the authors to delineate the previous experience participants had with the two OCS measures. This was the only confusing part.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Neuropsychology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 15 Mar 2024
Sam Webb

Title: 'Introducing the Tele-OCS: Preliminary evidence for a remotely administered version of The Oxford Cognitive Screen'
Response to reviewers We thank the reviewers and the editors for providing feedback on this manuscript draft. Please find below a detailed response to each reviewer comment. Reviewer comments are in italics, our reply is in bullet point format in plain text, and all changes made in the manuscript text are in red.

1. ***In the abstract a value of d=0 is reported, is this a typo? It's quite unusual for an effect size metric to be exactly 0.***

- *The d* = 0 effect size was due to rounding the value to 2 decimal places. We have now adjusted this in the abstract to be more precise, as below:

--- We found that the proportion of OCS subtasks impaired did not differ across modalities (*d* = <.001). ---

2. The introduction sufficiently motivates the need to validate a remotely administrated version of the OCS and the conduct of the study.

- We thank the reviewer for their kind words.

3. *The issue of previous participation in research needs to be clarified. Does this involve previous administration of the OCS? This could affect the outcomes through learning effects. I think it's important for the authors to delineate the previous experience participants had with* **the two OCS measures. This was the only confusing part.**

- Reviewer 1 has also addressed this point (see comment 1), We have now clarified that all participants, and most stroke survivors in Oxfordshire, have attempted the OCS at least once while in acute care therefore practice effects are to some degree unavoidable. We have also addressed this in our limitations section:

--- Finally, all participants were previously exposed to the OCS, whether as part of our research, or as part of their admission to hospital where OCS is given as standard where it can be. This means, there may be a chance for practice effects. We note, we did not observe many in the current study, so it is reasonable to assume less exist between first exposure and the current study. Alternatively, due to previous exposure, participants may be better at

the OCS in general, which would explain our low rates of impairment. As other studies using the OCS have found impairment rates, even for those previously exposed, and at different time points post-stroke (Kusec *et al*., 2023; Milosevich *et al*., 2023), we do not feel this is a major issue in the current study. ---

4. *The two versions of the measure are sufficiently described.* ***Data analysis description is complete. Visual presentation of data facilitates the understanding of the findings. Limitations/Implications are sufficiently described.***

- We thank the reviewer for their kind comments on our methods, analyses, figures, and limitations section.

5. ***The authors present impairment rates for subtasks. It's not clear whether there is impairment classification based on the overall score of the two OCS versions. This would be very useful to clarify and report.***

- We thank the reviewer for highlighting this for clarification, we hope we have now clarified which tasks we are referring to; see below:

--- We examined the proportion of subtasks impaired on the in-person versus Tele-OCS, and then the frequency of impairment classifications for each OCS subtask across modalities. Participants were impaired on an average of 5.71 (*SD*=7.79) subtasks on the in-person OCS and 5.71 (*SD*=8.59) subtasks in the Tele-OCS. ---

Competing Interests: Nele Demeyere is one of the original authors of the Oxford Cognitive Screen but does not have a financial conflict of interest.

Reviewer Report 04 August 2023

https://doi.org/10.21956/healthopenres.14374.r27286

?

Léon Franzen
[1] Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Germany
[2] Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Germany
[3] Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Germany
Frederike Schröpfer
[1] Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Schleswig-Holstein, Germany
[2] Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Schleswig-Holstein, Germany
[3] Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Schleswig-Holstein, Germany

Webb et al. attempt to provide an important development of the OCS by attempting to validate the remote version of the Oxford Screening Tool administered via telephone or videocall. The Covid-19 pandemic has shown us that many studies and assessments can also be delivered online providing somewhat comparable results. The benefits of remote assessments and studies are

undisputed. However, remote delivery comes with considerable challenges including much reduced control over the test administration and testing environment. This potential pitfall is exacerbated when delivering an assessment via telephone to potentially cognitively impaired patients. In the present manuscript, by postulating similar impairment classification of the same participants even under changed conditions, Webb et. al. test their remote OCS version and conclude it to be a valid version across telephone and videocall delivery channels. While I appreciate their open science practices, particularly their sampling and testing practices come with a number of considerable shortfalls that do not allow for drawing the all-encompassing conclusion of the remote OCS being a widely validated test after this rather small sample study. I will elaborate on and discuss these points in more detail below.

Major Issues

Recruitment and sample characteristics

With their study to validate the remotely administered oxford cognitive screen (OCS), Webb et al. attempted a previously analogue assessment of cognitive performance. One crucial issue of the study lies in the recruitment of participants. The authors used the strategy of community dwelling to recruit their participants but these individuals were sourced from completed studies of the Oxford Cognitive Screening Programme or other studies which tested or trained the cognitive performance of patients after a stroke. In my opinion, this creates a recruitment bias. Due to the prior participation in other studies, whereby the selection of participants is not at random and thus the sample is not representative of the underlying population. As well, since patients were recruited as part of a 6 month follow up and also after other studies, it would be important to clarify whether these participants had already done the OCS at least once within the first months after their stroke, as I presume the potential for practice effects exists. However, this piece of information is not provided. This could be tested by comparing reliability between participants who had taken the test before starting this study and those who have not.

Similarly, the existing study design results in an increased risk of selection bias, as only participants who completed the previous studies were included. If one recruits only from the group of test persons who have already completed a study, one may suspect that these are particularly motivated and performance-oriented individuals. Hence, generalisability to the general population may not be given.

As previously mentioned, the Information regarding the participants is rather limited. Webb et al. reported most participants being white males as a weakness of the study. They did also not address the presence of ischemic strokes either, which account for 70% of their sample. This is much lower than the prevalence reported in previous studies of around 87% (Busch & Kuhnert, 2017; Saini et al., 2021), which may affect generalisability. It may also come with different symptoms compared to the presence of subarachnoid and intracerebral haemorrhages. A more precise, detailed presentation of the included participants would be desirable for the reader.

Following on from this point, generalisability to the entire British population or beyond and to various remote delivery channels is also questionable, given the small sample (N=40), the heavy reliance on white British males and most of these participants taking the test on the telephone. To warrant the authors claims of presenting a successfully validated test, the sample size would need to be increased and sample's characteristics more balanced.

Further, the authors do not address the interval between the stroke and the time of the assessment. Here, too, a more detailed description and analysis would be desirable due to the following reasons.

The time post-stroke in which participants took part varied largely. The IQR suggests that some had a twice as long (or an even longer) recovery period before they participated in the study. Some participants were included after more than a 6.5 times longer recovery time (1500/228 days) compared to the median. Since the study design was within-subjects this aspect may not matter too much by itself but contributes to much heterogeneity in recovery possibilities within an already small sample.

Following on from this, the within-subject test-retest interval between the in-person and remote tests shows some considerable range as well. Here, I'm doubtful that a gap of 6 days is comparable with one of 72 days. Hence, I'd suggest to the authors to ensure that participants with a shorter interval did not show more consistency in their results compared to those with an interval longer than the mean. This could also be tested using a correlative approach correlating test-retest interval and the difference in % of the proportion of subtasks impaired.

The inclusion/exclusion criteria of the study were described by Webb et al. but appear to miss some important information. On the one hand, participants have a history of a confirmed stroke, but the severity or location of the stroke is not described. May this have an effect? I would also like to see the authors provide a more detailed description of the included participants and how their characteristics relate to the results, as no information on relevant symptoms such as neglect or hemiparesis among the participants other than may be indirectly obtained from the broken hearts subtest is provided.

Also, bad hearing is an exclusion criterion but one of the participants had trouble performing the test due to hearing difficulties. Why was this patient included regardless?

In sum, I would like to encourage Webb et al. to elaborate on and show that these issues relating to the sample had no effect and/or add these points on the risk of sampling bias and sample characteristics to their limitation section.

Furthermore, I consider the simultaneous validation of two different forms of presentation (telephone or video conference) within one study to be unfavourable. As described in the study, the two forms of presentation place different demands on the patients and their environment. It is very patient-friendly to adapt the implementation to the preferences of the individual, but in my view, it also distorts the result. In the case of the present study, not only was the basic decision for one of the methods discussed with the patients, but also throughout the course of the test patients switched between video conferencing and telephone administration. Thereby, changing the way of administering a rather visually based test from a visual form of administration (video conferencing) to a purely auditory delivery form does not really evaluate the same test material due to differences in the required instructions and materials in my eyes. Webb et al., for example, report different instructions that the participants received during the implementation.

In order to give a clear picture of the strengths of the remote-OCS, a randomised allocation of the patients to a fixed form of administration would benefit the results. Particularly, since the two

different forms of implementation also place different demands on the participant as well as the person administering the test. Here, too, standardisation would be interesting for the sake of comparability. In this way, not only a differentiated statement could be made about the difficulties of implementing the test via telephone vs videocall, but also its strengths could be recognised and weighed against each other.

Analyses

I was surprised to see the main preregistered hypothesis and analyses revolving around the intraclass correlation coefficient (ICC) as the measure of choice, given that the original OCS validation paper has already shown issues between the parallel versions. If I understand this argument correctly, this ICC problem had already been known to the authors before registering the present study's analysis plan. Then switching their analysis approach entirely may be acceptable, if they also report the originally planned analyses in their supplementary materials at least. However, the mentioned supplementary materials folder on the OSF appears empty. Please double-check this folder.
Further the manuscript does not specify which form of the ICC was calculated before discarding this approach. considering a lower ICC bound of .5 to be enough for validating test-retest consistency appears rather low for such an established test, as it only indicated moderate reliability. Most importantly, changing their analysis approach also invalidates the power analysis performed for an ICC analysis. Taken together, I would encourage the authors to elaborate on their rationale and provide actual results to the reader for transparency.

One of the main motivations for the OCS was to avoid overall binary classifications as in impaired vs unimpaired. However, the new analysis that aims to convince the reader of the tele-OCS's validity relies purely on these binary classifications as the key metric, albeit for subtests. This approach discards much information on actual test-retest scores, even if it is subtle information, which the actual raw scores could provide. Please consider reporting these scores.

Even more concerning is the author's reliance on binary impairment classification and norms from the in-person test due to a lack of tele-OCS data. This point is being acknowledge as a major limitation by the authors themselves. This raises the question of the validity of their analysis used for making their claims of having successfully validated the tele-OCS. Once again, it could be addressed by collecting a larger sample and adapting the analysis.

Lastly, the authors switched in their analysis towards using several t-tests and effect sizes for their claims on (non-)significant differences. However, these results may suffer from two issues. First, multiple t-tests come with the problem of an increased false positive rate, as the authors tested data from the same participants. Please report whether these tests were corrected for multiple comparisons. Second, in the context of this investigation no difference (i.e., evidence for the null hypothesis) is the desired test outcome. T-tests cannot provide this information. Therefore, I'd suggest switching from t-tests to BayesFactors for being able to report the evidence for the null hypothesis directly.

Figures

The figures are missing elaborate figure captions that explain all aspects of the figure. For example, the impaired and unimpaired labels and the fact that some bars appear to be missing

for some tests, which is simply due to the value being 0. Relating to my previous point on plotting all the raw scores, here, I'd suggest plotting the continuous data as raincloud (https://doi.org/10.12688/wellcomeopenres.15191.2), violin or similar plots and avoid barplots in order to provide the reader with additional, valuable information and thereby facilitate trust in their interesting results.

Minor Issues

Information on the test's delivery as part of the procedure section could be enhanced by providing the links to the online videos hosted on Youtube (https://www.youtube.com/watch?v=CC_QVJHqFvk) or the videos could be uploaded to the OSF repository.

References

1. Busch MA, Kuhnert R: 12-Monats-Prävalenz einer koronaren Herzkrankheit in Deutschland. *Journal of Health Monitoring*. 2017. Reference Source
2. Saini V, Guada L, Yavagal DR: Global Epidemiology of Stroke and Access to Acute Ischemic Stroke Interventions.*Neurology*. 2021; **97** (20 Suppl 2): S6-S16 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Cognitive neuroscience and psychology, psychiatry, remote assessments

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Author Response 15 Mar 2024

Sam Webb

Title: 'Introducing the Tele-OCS: Preliminary evidence for a remotely administered version of The Oxford Cognitive Screen'

Response to reviewers We thank the reviewers and the editors for providing feedback on this manuscript draft. Please find below a detailed response to each reviewer comment. Reviewer comments are in italics, our reply is in bullet point format in plain text, and all changes made in the manuscript text are in red.

Reviewer 1: ***Webb et al. attempt to provide an important development of the OCS by attempting to validate the remote version of the Oxford Screening Tool administered via telephone or videocall. The Covid-19 pandemic has shown us that many studies and assessments can also be delivered online providing somewhat comparable results. The benefits of remote assessments and studies are undisputed. However, remote delivery comes with considerable challenges including much reduced control over the test administration and testing environment. This potential pitfall is exacerbated when delivering an assessment via telephone to potentially cognitively impaired patients. In the present manuscript, by postulating similar impairment classification of the same participants even under changed conditions, Webb et. al. test their remote OCS version and conclude it to be a valid version across telephone and videocall delivery channels. While I appreciate their open science practices, particularly their sampling and testing practices come with a number of considerable shortfalls that do not allow for drawing the all-encompassing conclusion of the remote OCS being a widely validated test after this rather small sample study. I will elaborate on and discuss these points in more detail below.***

Major Issues Recruitment and sample characteristics 1) ***With their study to validate the remotely administered oxford cognitive screen (OCS), Webb et al. attempted a previously analogue assessment of cognitive performance. One crucial issue of the study lies in the recruitment of participants. The authors used the strategy of community dwelling to recruit their participants but these individuals were sourced from completed studies of the Oxford Cognitive Screening Programme or other studies which tested or trained the cognitive performance of patients after a stroke. In my opinion, this creates a recruitment bias. Due to the prior participation in other studies, whereby the selection of participants is not at random and thus the sample is not representative of the underlying population. As well, since patients were recruited as part of a 6 month follow up and also after other studies, it would be important to clarify whether these participants had already done the OCS at least once within the first months after their stroke, as I presume the potential for practice effects exists. However, this piece of information is not provided. This could be tested by comparing reliability between participants who had taken the test before starting this study and those who have not. Similarly, the existing study design results in an increased risk of selection bias, as only participants who completed the previous studies were included. If one recruits only from the group of test persons who have already completed a study, one may suspect that these are particularly motivated and performance-oriented individuals. Hence, generalisability to the general population may not be given.***

- We thank the reviewer for their comments. With regards to your comment on potential for practice effects due to repeated exposure to the OCS, within Oxfordshire all stroke survivors would have at least attempted the OCS as part of their routine

care, as this is recommended standard practice in the UK (Royal College of Physicians, 2016). Because of this, it would be difficult to recruit stroke survivors within Oxfordshire who had not completed the OCS beforehand and therefore comparing reliability between participants who had completed the OCS prior to the current study would not be meaningful. We have made this point more clearly as below

--- We used a prospective cross-sectional within-subjects design. Due to the standard practice of in stroke care within Oxfordshire, most stroke survivors admitted to a hospital for stroke will complete, or partially complete, the OCS. We recruited chronic stroke survivors in the UK between November 2021 to August 2022. All participants completed the OCS when originally admitted to hospital, separate to this study. ---

- We also note this as a limitation in our discussion.

--- Finally, all participants were previously exposed to the OCS, whether as part of our research, or as part of their admission to hospital where OCS is given as standard where it can be. This means, there may be a chance for practice effects. We note, we did not observe many in the current study, so it is reasonable to assume less exist between first exposure and the current study. Alternatively, due to previous exposure, participants may be better at the OCS in general, which would explain our low rates of impairment. As other studies using the OCS have found impairment rates, even for those previously exposed, and at different time points post-stroke (Kusec *et al*., 2023; Milosevich *et al*., 2023), we do not feel this is a major issue in the current study. ---

- With regards to your comment on potential risk of recruitment and selection bias, both OCS-Recovery or OX-CHRONIC studies had broad inclusion criteria (i.e., diagnosis of a stroke, age ≥18, ability to provide informed consent, ability to concentrate for at least 20 minutes) to maximise those who are recruited. Due to this, there are few barriers to participation and enhanced our sample representativeness. As reflected in our demographics, we have recruited a representative pool of participants from the Oxfordshire area, reflecting the typical age, education, and ethnicity of stroke survivors, as well as stroke hemisphere, and stroke severity. We note we also well reflected the ethnicity and age statistics of older stroke survivors, whereby most (>95%) of stroke survivors are white and white-British in the UK if they are over 65 (Census data 2021, older persons report, see https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/articles/profileoftheolderpopulationl 04-03),

- We acknowledge that, as in any research study, potential for selection bias exists. We have expanded on this in the discussion as below:

- We used a prospective cross-sectional within-subjects design. Due to the standard practice of inpatient stroke care within Oxfordshire, most stroke survivors admitted to a hospital for stroke will complete, or partially complete, the OCS, prior to commencing any research study. We recruited chronic stroke survivors in the UK between November 2021 to August 2022. All participants completed the OCS when originally admitted to hospital, separate to this study. We contacted long-term stroke survivors from the OX-CHRONIC (Kusec et al., 2023) study who had recently (within 30 days at time of contact) completed the Tele-OCS as part of the OX-CHRONIC protocol. Additionally, we asked all participants from the Oxford Cognitive Screening

programme, who had just (at time of being contacted) completed an in person 6-month follow up OCS (Demeyere *et al.*, 2015; Milosevich, Moore, Pendlebury, & Demeyere, 2023), to take part. We aimed to get both the OCS and Tele-OCS completed within 30 days of each other, but this was not always possible. Inclusion criteria for both the Oxford Cognitive Screening programme and OX-CHRONIC are very inclusive and aligned with the current inclusion criteria. This means the sampling approaches originally used have minimal bias (except bias of those who are willing to participate), and the underlying samples are representative of the UK stroke population. In the current study we attempted to reflect the UK stroke population in key demographics of age, education, sex, stroke type, stroke side, and stroke severity.

--- from discussion Moreover, it is necessary to consider that those who consent to take part in research may be different from those who do not, and as such more inclusive study protocols using those with in-capacity to consent to research may help contribute to more reflective demographics. ---

2) As previously mentioned, the Information regarding the participants is rather limited. Webb et al. reported most participants being white males as a weakness of the study. They did also not address the presence of ischemic strokes either, which account for 70% of their sample. This is much lower than the prevalence reported in previous studies of around 87% (Busch & Kuhnert, 2017; Saini et al., 2021), which may affect generalisability. Following on from this point, generalisability to the entire British population or beyond and to various remote delivery channels is also questionable, given the small sample (N=40), the heavy reliance on white British males and most of these participants taking the test on the telephone. To warrant the authors claims of presenting a successfully validated test, the sample size would need to be increased and sample's characteristics more balanced. In sum, I would like to encourage Webb et al. to elaborate on and show that these issues relating to the sample had no effect and/or add these points on the risk of sampling bias and sample characteristics to their limitation section.

- We thank the reviewer for their comments. We have now restructured our limitations section to elaborate on the homogeneity of our sample. We note that we have already called for larger studies, but have specifically edited this now (see next with changes in red:

--- Although our study sample was statistically powered, the sample size was small and relatively homogeneous. For example, 97.5% of the sample was White-British, 70% were male, and around 70% were ischaemic stroke. Ideally, there would be around 82% White-British participants to reflect UK government 2021 census data for ethnicity (ONS, 2022), or >93% for older adults specifically, or 87% for general UK stroke survivors (Douiri et al., 2021), approximate equality of males and female and around 87-90% ischemic stroke (Douiri et al., 2021; King *et al.*, 2020). Strengths of our sample, however, include age and stroke severity, which represent the general UK stroke population (Douiri et al., 2021). Furthermore, given that many participants were in the chronic stage of stroke, they were often classified as unimpaired across subtasks in both modalities, limiting our interpretation for stroke survivors with more severe impairments. This may impact the generalisability of the present results to the wider stroke population. Although the majority of stroke survivors tend to be males, our weighting of males to females was biased as we had far more males. This may impact the generalisability of our results to females. We do not feel that the

gender balance in our data impacts the results drastically, however, we acknowledge this as a limitation. We did not choose to purposively sample for sex specifically as it is not thought to alter the results of global cognitive scoring after a stroke. Future validation studies of the Tele-OCS may benefit from using larger, more heterogeneous samples, including those who have sustained more severe strokes, and more representative demographics. Finally, all participants were previously exposed to the OCS, whether as part of our research, or as part of their admission to hospital where OCS is given as standard where it can be. This means, there may be a chance for practice effects. We note, we did not observe many in the current study, so it is reasonable to assume less exist between first exposure and the current study. Alternatively, due to previous exposure, participants may be better at the OCS in general, which would explain our low rates of impairment. As other studies using the OCS have found impairment rates, even for those previously exposed, and at different time points post-stroke (Kusec *et al*., 2023; Milosevich *et al*., 2023), we do not feel this is a major issue in the current study. ---

3) It may also come with different symptoms compared to the presence of subarachnoid and intracerebral haemorrhages. A more precise, detailed presentation of the included participants would be desirable for the reader.

- Due to (sub) sample size, evaluating performance differences by stroke subtype or other demographics was not considered meaningful. We have noted this potential sample bias in the limitations section. We regret to not have further details available for the sample. We have added acute OCS data which was collected prior to the study from participant medical files. We have now included performance on the OCS subtasks in the form of mean, SD, and range to show the range of performance, to reveal potential biases. See supplemental table S1 and extract from the data analysis section below:

--- Data analysis Prior to formal analyses we examined sample characteristics, including demographics and prior OCS performance taken within 20 days of stroke. Acute OCS performance is presented in Table S2 in the supplementary materials. ---

4) Further, the authors do not address the interval between the stroke and the time of the assessment. Here, too, a more detailed description and analysis would be desirable due to the following reasons. The time post-stroke in which participants took part varied largely. The IQR suggests that some had a twice as long (or an even longer) recovery period before they participated in the study. Some participants were included after more than a 6.5 times longer recovery time (1500/228 days) compared to the median. Since the study design was within-subjects this aspect may not matter too much by itself but contributes to much heterogeneity in recovery possibilities within an already small sample.

- Given the Tele-OCS is intended to be used in community-based samples, we did not aim to recruit participants from a specific time post-stroke, and instead capitalised on existing data (to reduce participant burden) from projects comprising stroke survivors at both 6-months to $\geq$ 2 years post-stroke. To clarify this, we now note in the participant recruitment section that we recruited chronic stroke survivors, see below from participants section

--- We used a prospective cross-sectional within-subjects design. Due to the standard practice of in stroke care within Oxfordshire, most stroke survivors admitted to a hospital for stroke will complete, or partially complete, the OCS. We recruited chronic stroke survivors in the UK between November 2021 to August 2022. All participants completed the

OCS when originally admitted to hospital, separate to this study. We contacted long-term stroke survivors from the OXCHRONIC (Kusec et al., 2023) study who had recently (within 30 days at time of contact) completed the tele-OCS as part of the OXCHRONIC protocol. The OXCHRONIC population were at minimum 2 years post-stroke with an average of 4.5 years. Additionally, we asked all participants from the Oxford Cognitive Screening programme, who had just (at time of being contacted) completed an in person 6-month follow up OCS (Demeyere *et al.*, 2015; Milosevich, Moore, Pendlebury, & Demeyere, 2023), to take part. We aimed to get both the OCS and Tele-OCS completed within 30 days of each other, but this was not always possible. ---

- To investigate whether time post-stroke had any detectable effect on performance, we conducted a correlation analysis between time post-stroke and proportion of OCS subtasks impaired, in which no significant relationship was found (copied below):

--- The time in days between in-person and Tele-OCS administration ranged from 6 to 72, with an average of *M*=27.70 days (*SD*=12.29). 23 participants were administered the Tele-OCS first (57.5%). Time interval was not correlated with any differences in subtasks impaired on either modality ($r(38)= -0.03$, $p=.85$). Administration of the Tele-OCS took an average time of 17.17 minutes (*SD*=2.73, range = 14 - 22). The Tele-OCS was administered using phone to the majority of participants (*n*=36, 90%). ---

- As presented in our results, there were limited cases of trends in improvement in across modality, meaning it is unlikely that there would have been recovery in between assessment time points.

5) Following on from this, the within-subject test-retest interval between the in-person and remote tests shows some considerable range as well. Here, I'm doubtful that a gap of 6 days is comparable with one of 72 days. Hence, I'd suggest to the authors to ensure that participants with a shorter interval did not show more consistency in their results compared to those with an interval longer than the mean. This could also be tested using a correlative approach correlating test-retest interval and the difference in % of the proportion of subtasks impaired.

- see figure 1 linked here: https://healthopenresearch.s3.eu-west-1.amazonaws.com/linked/182865.13291-_comments_image.pdf
- We thank the reviewer for this comment highlighting that there may be differences in performance for those with shorter vs longer assessment intervals. In investigating this relationship, we found no correlation between the difference in proportion of tasks impaired and the time between assessments ($r(38)= -0.03$, $p=.85$). We have additionally plotted this correlation, with proportion of tasks impaired on the y-axis, and time between sessions on the x axis, indicating no clear pattern of performance across time intervals.

6) The inclusion/exclusion criteria of the study were described by Webb et al. but appear to miss some important information. On the one hand, participants have a history of a confirmed stroke, but the severity or location of the stroke is not described. May this have an effect? Stroke location (hemisphere) and stroke severity are detailed in Table 1. We do not have immediate access to scan data as this was not part of our protocol, so we cannot report on exact location of stroke. We separately examined the effect of sex, stroke type, and stroke hemisphere on subtasks impaired on the OCS across modality and found no significant main effect or interaction of demographics on performance. Results are detailed in the supplementary and references now in text.

7) I would also like to see the authors provide a more detailed description of the included participants and how their characteristics relate to the results, as no information on relevant symptoms such as neglect or hemiparesis among the participants other than may be indirectly obtained from the broken hearts subtest is provided.

- We unfortunately did not collect information for characteristics such as hemiparesis. Neglect among stroke survivors in the Oxfordshire area is tested with the Broken hearts task, and can thus be used to determine the rate of impairment in our sample. However, we feel our sample is too small for further subgroup analyses of different sub-categories of participants.
- We have now better described the sample and provided prior OCS performance scores for reference. It can be seen via the range of scores that impairments in scores were seen throughout the OCS subtasks.

8) Also, bad hearing is an exclusion criterion but one of the participants had trouble performing the test due to hearing difficulties. Why was this patient included regardless?

- This exclusion criteria refer to hearing difficulties that could not be reasonably adjusted for. Any hearing difficulties that became evident whilst testing were overcome by repeating a single instruction twice. To clarify this point, we have stated in the manuscript:

--- Participant identification Eligibility criteria were: 1) history of confirmed stroke; 2) 18 years of age or older at time of stroke; 3) capacity to consent to research; 4) able to remain alert for at least 20 minutes; and 5) spoke and understood sufficient English. Participants were excluded if they had hearing, language, or visual impairments (not including visual neglect) that would not allow for remote or in-person assessment outside of reasonable adjustments. Any hearing difficulties that became evident whilst testing were overcome by repeating a single instruction twice. ---

9) Furthermore, I consider the simultaneous validation of two different forms of presentation (telephone or video conference) within one study to be unfavourable. As described in the study, the two forms of presentation place different demands on the patients and their environment. It is very patient-friendly to adapt the implementation to the preferences of the individual, but in my view, it also distorts the result.

- We agree with the reviewer that including different remote modalities (i.e., telephone vs in person) may be a weakness of the paper because we do not directly test their equivalence (nor did we aim to) due to low numbers in any comparison. However, the administration protocol and materials are identical for each remote modality, as detailed in the Tele-OCS description within the manuscript. When the Tele-OCS is conducted via videoconferencing, there are no additional demonstrations or instructions, therefore any potential differences are not currently testable. Participant preference of remote modality in our study is reflective of using the Tele-OCS in clinical settings where choice of modality would be offered.
- We acknowledge that differences in remote modality should be investigated in future studies within our limitations, as below):

--- There is currently no normative data for performance on the Tele-OCS for comparison of scores, as our results rely on the in-person normative data to generate impairment classifications we cannot know if these impairment classifications are accurate. Further research will need to determine this. Furthermore, a small part of our study concerned

different modalities of assessment (telephone vs videocall). We did not aim to assess whether the modality of remote assessment affected performance, as we relied on participant preference above all else, had too small a sample to do so, and we feel that as the instructions and materials are identical, there should be limited differences. ---

10) In the case of the present study, not only was the basic decision for one of the methods discussed with the patients, but also throughout the course of the test patients switched between video conferencing and telephone administration. Thereby, changing the way of administering a rather visually based test from a visual form of administration (video conferencing) to a purely auditory delivery form does not really evaluate the same test material due to differences in the required instructions and materials in my eyes. Webb et al., for example, report different instructions that the participants received during the implementation. Apologies this was not transparent, to be clear:

- There is no case reported of any participants switching modalities within a session, all were consistent.
- As noted in our response above, administration of the Tele-OCS both via telephone and videoconferencing have identical instructions and testing materials.
- We have now added some clarification to the manuscript, see below:

--- Procedure Once participants from the Oxford Cognitive Screening programme were identified as having just (within 30 days) completed an in-person OCS in their homes, we asked if they would complete the Tele-OCS. For those who agreed, we posted the Tele-OCS pack at least 1 week before the appointment and reminded them not to open the pack until instructed. In the session, informed consent was taken and the Tele-OCS was administered either by phone or videoconferencing depending on the preference of the participant. Shifting modality during a session was not permitted, nor did this come up during data collection. We provided a pre-paid envelope for participants to post their packs back. If participants were recruited from the OXCHRONIC study, we contacted them within 30 days of having completed the Tele-OCS for the OXCHRONIC protocol, and administered the OCS in person at their homes. ---

11) In order to give a clear picture of the strengths of the remote-OCS, a randomised allocation of the patients to a fixed form of administration would benefit the results. Particularly, since the two different forms of implementation also place different demands on the participant as well as the person administering the test. Here, too, standardisation would be interesting for the sake of comparability. In this way, not only a differentiated statement could be made about the difficulties of implementing the test via telephone vs videocall, but also its strengths could be recognised and weighed against each other.

- We thank the reviewers for this insight and agree that it may be worthwhile to randomise the allocation of participants to telephone vs videocall modalities in future studies. This could elucidate potential differences between these types of assessment method. This has now been added to the limitations section of the discussion. However, we want to re-iterate that it was not possible to randomly allocate participants to modalities in the current study for a number of reasons (which are noted in the limitations section).
- Moreover, and more generally from the current study, randomisation may not be feasible/possible with stroke participants as some individuals do not have any

familiarity with videoconferencing/do not have access to internet/restricted internet in their geographical area. Additionally participants may have smart phones but may not a tablet or computer to complete the assessment in the same way - unless you recruited people who can do both at the outset (but this would likely result in greater bias in the sample).

- Changes to the manuscript discussion are in red:

--- There is currently no normative data for performance on the Tele-OCS for comparison of scores, as our results rely on the in-person normative data to generate impairment classifications we cannot know if these impairment classifications are accurate. Further research will need to determine this. Furthermore, a small part of our study concerned different modalities of remote assessment (phone vs videocall). We did not aim to assess whether the modality of remote assessment affected performance, as we relied on participant preference above all else, had too small a sample to do so, given identical instructions and materials, there should be limited differences. Furthermore, it should be noted that stroke survivors do not always have the correct equipment to carry out assessment via videoconferencing. However, it may be worthwhile for future studies with larger sample sizes to randomly allocate participants to either telephone or videoconferencing to investigate whether there are any differences between these types of remote assessment methods. **---**

Analyses 12) I was surprised to see the main preregistered hypothesis and analyses revolving around the intraclass correlation coefficient (ICC) as the measure of choice, given that the original OCS validation paper has already shown issues between the parallel versions. If I understand this argument correctly, this ICC problem had already been known to the authors before registering the present study's analysis plan.

- Our original ICC analysis were pre-registered as they were an appropriate test to use for continuous subtask scores in theory, however, the small ranges of variation observed within our sample precluded the use of ICCs in a sensible manner. Therefore, it was the high concordance between time points in our observed data, rather than the OCS subtests itself, that meant we had to switch analysis plan. This has now been clarified in the paper as below:

--- This project was originally pre-registered at the start of data collection (3 participants had taken part). The preregistered analysis plan was inspired by ICCs between OCS parallel versions A and B in the original OCS normative article (Demeyere *et al*., 2015), though once data was collected, it became clear that this was not the best approach for the current data that was observed. The small variation of scores observed in the subtasks meant the raw continuous scores ICCs between in-person and remote subtest comparisons were not appropriate. We instead opted to shift towards test of difference between modalities, to retain the consistent statistical power. The sample size calculations remained the same as we were similarly powered for equivalent comparisons via difference based *t*-tests. ---

13) Then switching their analysis approach entirely may be acceptable, if they also report the originally planned analyses in their supplementary materials at least. However, the mentioned supplementary materials folder on the OSF appears empty. Please double-check this folder.

- Our apologies for the oversight in not uploading the original planned analyses within

our supplementary materials. Supplementary materials are now uploaded to OSF, which include the results of the ICCs mentioned in the main manuscript (and the underlying code to reproduce them).

14) Further the manuscript does not specify which form of the ICC was calculated before discarding this approach.

- We thank the reviewer for pointing out this oversight which has now been corrected both in the manuscript and in supplemental materials.

--- Sample size was pre-determined in our pre-registration at a minimum of 30 participants to detect a two-way random effects single unit agreement based intraclass correlation coefficient (ICC) ---

15) Considering a lower ICC bound of .5 to be enough for validating test-retest consistency appears rather low for such an established test, as it only indicated moderate reliability. Most importantly, changing their analysis approach also invalidates the power analysis performed for an ICC analysis. Taken together, I would encourage the authors to elaborate on their rationale and provide actual results to the reader for transparency.

- We agree that the manuscript would benefit from providing greater information power for the adjusted analyses and on our decisions for a sufficient ICC for the Tele-OCS. We firstly note that there was little difference in the power analysis for sample size for a t-test or ICC for this study, but this has now been clarified in our manuscript:

--- Changes from pre-registration This project was originally pre-registered at the start of data collection (3 participants had taken part). The preregistered analysis plan was inspired by ICCs between OCS parallel versions A and B in the original OCS normative article (Demeyere *et al*., 2015), though once data was collected, it became clear that this was not the best approach for the current data that was observed. The small variation of scores observed in the subtasks meant the raw continuous scores ICCs between in-person and remote subtest comparisons were not appropriate. We instead opted to shift towards test of difference between modalities, to retain consistent statistical power. The sample size calculations remained the same as we were similarly powered for equivalent comparisons via difference based *t*-tests. We present ICC analyses in extended data. ---

- Regarding our decision to have a lower bound of 0.50 ICC, this was considered to be the minimum acceptable ICC for the Tele-OCS and not the expected ICC estimate. As you have noted, this lower bound would reflect moderate reliability and our study was powered to detect this.

16) One of the main motivations for the OCS was to avoid overall binary classifications as in impaired vs unimpaired. However, the new analysis that aims to convince the reader of the tele-OCS's validity relies purely on these binary classifications as the key metric, albeit for subtests. This approach discards much information on actual test-retest scores, even if it is subtle information, which the actual raw scores could provide. Please consider reporting these scores.

- We agree that a strength of the OCS is that it provides insight into domain-specific strengths and weaknesses of the stroke survivors being assessed. The OCS was developed to have impairment classifications per subtask, as this is how it would be used in clinical settings to determine domain-specific impairments. The original OCS aims to classify impairment based on specific cognitive domains, rather than

providing a single domain-general impairment classification that provides limited information. In this sense, the OCS and the Tele-OCS provide more granular information about cognition that is used pragmatically in clinical settings. In Table 2, we have also presented raw data on continuous scores in the manuscript. As stated in the paper, the use of impairment classifications is more in line with both the aims of the OCS, and with how the OCS is used in practice.

17) Even more concerning is the author's reliance on binary impairment classification and norms from the in-person test due to a lack of tele-OCS data. This point is being acknowledge as a major limitation by the authors themselves. This raises the question of the validity of their analysis used for making their claims of having successfully validated the tele-OCS. Once again, it could be addressed by collecting a larger sample and adapting the analysis. We agree with these comments and await further research about the generation of new normative scores for the Tele-OCS. We have updated the text throughout as well as the title of the manuscript to refer to "Preliminary evidence" and also discuss this explicitly in the limitations.

18) Lastly, the authors switched in their analysis towards using several t-tests and effect sizes for their claims on (non-)significant differences. However, these results may suffer from two issues. First, multiple t-tests come with the problem of an increased false positive rate, as the authors tested data from the same participants. Please report whether these tests were corrected for multiple comparisons.

- We thank the reviewer for identifying this oversight. In Table 2 where we report Cohen's d between OCS and Tele-OCS subtasks, we have now added a note to our table legend indicating the corrected alpha, as below:

[i] Note. '*' signifies p-value <.05. Multiple comparison alpha corrected level is .005 (.05/11 = .005). '**' signifies p-values <.005).

19) Second, in the context of this investigation no difference (i.e., evidence for the null hypothesis) is the desired test outcome. T-tests cannot provide this information. Therefore, I'd suggest switching from t-tests to BayesFactors for being able to report the evidence for the null hypothesis directly.

- Thank you for your comments, we have now included equivalence tests in addition to t-tests to account for different interpretations. Initially we wished to evaluate any differences in performance, hence our decision to use t-tests. We have now included whether the differences are smaller than the smallest effect size we can detect. The results of our equivalence tests (presented in the *Supplementary Materials)* are in line with the t-test results. For consistency with the remainder of the manuscript and literature on the OCS, we have opted for frequentist statistics which are more understandable by the general scientific community, and equivalence tests here (and t-test with our original aim) are appropriate to answer our questions.

In Table 2, we present descriptive statistics for raw score performance on each shared OCS subtask as well as repeated-measures Cohen's *d* and significance of difference on *t*-tests between each subtask. Equivalence tests were additionally conducted, paralleling the t-test results presented here (see Supplementary Materials).

Figures 20) The figures are missing elaborate figure captions that explain all aspects of the figure. For example, the impaired and unimpaired labels and the fact that some bars

appear to be missing for some tests, which is simply due to the value being 0. Relating to my previous point on plotting all the raw scores, here, I'd suggest plotting the continuous data as raincloud (https://doi.org/10.12688/wellcomeopenres.15191.2), violin or similar plots and avoid barplots in order to provide the reader with additional, valuable information and thereby facilitate trust in their interesting results.

- We agree that more detailed captions were needed and have now elaborated on them to improve ease of understanding.
- Prior to initial submission, we attempted raincloud plots but these became difficult to read when clustered together as a single figure, and to visually compare performance of the OCS to the Tele-OCS. We therefore chose to use density distributions to appropriately characterise distributions of data and overlap between in-person and remote administrations of the OCS. We hope this is acceptable. See plot: see figure 2 linked here: https://healthopenresearch.s3.eu-west-1.amazonaws.com/linked/182866.13291-_comments_image_2.pdf

Minor Issues 21) Information on the test's delivery as part of the procedure section could be enhanced by providing the links to the online videos hosted on Youtube (https://www.youtube.com/watch?v=CC_QVJHqFvk) or the videos could be uploaded to the OSF repository.

- There are currently no videos which reflect the administration of the Tele-OCS, videos that do exist were created in 2020 as part of our interim remote OCS advice. These are being developed.

Competing Interests: Nele Demeyere is one of the original authors of the Oxford Cognitive Screen but has no financial conflict of interest.