

Personalised Cognitive Screening in Stroke Care: Development and Pilot Evaluation of a Modular, Adaptive, and Co-Produced Cognitive Screening Platform for Stroke

Sam S. Webb

sam.webb@ndcn.ox.ac.uk

University of Oxford

Paula Grammatikos

Maastricht University

Nele Demeyere

University of Oxford

Article

Keywords:

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11207437/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Post-stroke cognitive screening is complex, with cognitive impairment profiles varying in both severity and domains affected. Currently, there is no cognitive screening approach aligned with (inter)national clinical guidelines for stroke which adapts to individual patient abilities. We aimed to create a digital and adaptable domain-specific cognitive screening tool in the Oxford Cognitive Screening Platform (OCX) and complete pilot testing to investigate patient performances across the adaptive tests, provide initial acceptability evidence and validation of norm data within the new platform.

In study 1, the OCX was co-produced from an initial theory and expertise driven proposal for an adaptive testing framework based on existing standardised and validated cognitive screening tests. This initial skeleton assessment was refined through iterative co-production workshops with six (allied) healthcare professionals working in stroke and three cognitive neuropsychology experts, totalling 10 hours. Workshop recordings were transcribed and qualitatively analysed using Thematic template analysis. Finally, some included subtests were shortened following Item Response Theory. In study 2, the OCX was piloted with 22 stroke survivors and 86 neurologically healthy controls to assess initial performance ranges and differential adaptive pathway selection and descriptives in the patient group, acceptability and norm comparisons with the originally published normative data.

The OCX adaptive battery assesses six cognitive domains: language, attention, executive function, memory, praxis, and number. Each domain includes tiered test selection with test administration depending on impairment classification from prior tests. Pilot data suggests high acceptability, a range of performances, and differential adaptive pathway selection, with feedback indicating high acceptability.

Introduction

Stroke is a leading cause of disability and mortality (Feigin et al., 2021; Johnson et al., 2019), resulting in heterogenous profiles of cognitive impairment, with impairment rates in excess of 85% for stroke survivors acutely, and up to 65% in the long-term (Johnson et al., 2019; Kusec et al., 2023; Milosevich et al., 2023; Nys et al., 2007; Stroke Association, 2018). International guidelines advocate for domain-specific cognitive screening across the care continuum, from acute to community settings (Intercollegiate Stroke Working Party, 2023; Quinn et al., 2021; Rudd et al., 2017). However, these guidelines refrain from recommending specific tools, acknowledging the absence of a single instrument suitable for the heterogeneity of stroke presentations (McMahon et al., 2021).

Commonly used tests for stroke survivors across the stroke pathway (Quinn et al., 2018), like the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) or General Practitioner assessment of Cognition (GP-COG) (Brodaty et al., 2002) were designed primarily to detect dementia, with a focus on verbal memory assessment as the key marker for Alzheimer's disease. Performance on these cognitive screening tests is often confounded by common domain-specific impairments in stroke populations, particularly aphasia or motor impairments (e.g. see Demeyere et al., 2015, 2016; Tessaro et al., 2026). Moreover, such tools typically do not assess highly frequently impaired cognitive domain impairments including visuospatial neglect and apraxia and often provide only limited coverage of executive function.

In response, several stroke-specific screening tools have emerged (Bickerton et al., 2015; Chen et al., 2025; Demeyere et al., 2015, 2021; Gruia et al., 2023; Pagliarin et al., 2015; Robinson et al., 2021; Schmidt et al., 2023; S. S. Webb et al., 2021). Each tool has specific psychometric properties, and consequently different strengths and weaknesses for subpopulations. With stroke characterised by a highly heterogeneous, complex, and wide-ranging severity of cognitive impairment, no single screening tool fits the needs of every patient and clinical setting across the stroke pathway. The challenge in creating a comprehensive cognitive screening approach lies in balancing comprehensiveness with practicality. Assessments must be brief and easy to administer, limiting the burden on patients, as well as staff in time-

limited acute settings. At the same time, assessments must be sufficiently detailed to inform rehabilitation in chronic stages. Stroke-related impairments - especially in consciousness, communication, and attention - frequently preclude or confound the completion of standardised batteries (Pagliarin et al., 2015). In addition, expert clinical psychology support is often unavailable, necessitating tools that can be used by generalist healthcare staff without specialist neuropsychological expertise (McMahon et al., 2021).

There is a concern about the proliferation of screening tools without consensus on core components or psychometric comparability (Lees et al., 2014). The existing diversity of cognitive assessment practices risks leading to redundancy, inefficiency, and inconsistent patient care, with unqualified healthcare professionals choosing assessments based on experience rather than psychometric evidence (e.g. Quinn et al., 2018). Assessments should be selected based on a range of evidence-based factors, such as feasibility, validity, and purpose of the test (American Educational Research Association et al., 2014; Lees et al., 2014). A unifying, validated tool that adapts to patient performance could address these limitations by tailoring assessment depth to the individual's cognitive profile, allowing for brief and inclusive screening of severe cognitive impairment as well as highly sensitive screening of milder impairments, as clinically required.

Previous Computerised Adaptive Testing approaches (CAT; (Wainer et al., 2000) have used Item Response Theory (IRT), to estimate item parameters and determine subsequent item selection (Oladele & Ndlovu, 2021; Wainer et al., 2000). In the context of post-stroke cognitive screening, the Imperial Comprehensive Cognitive Assessment in Cerebrovascular Disease (IC3 - Gruia et al., 2025) was developed to provide an extensive self-administered remote adaptive cognitive assessment that can be completed once patients are home. The IC3 provides in-depth assessment of milder deficits, rather than our intended brief in-person cognitive screening of domain-specific impairments including severely impaired patients. A fundamental aspect of the CAT approach is the need for a large calibrated item response bank and extensive normative data to estimate item parameters before pathways are automated or chosen (Oladele & Ndlovu, 2021; Wainer et al., 2000). While this provides a robust psychometric framework, it typically requires redevelopment and recalibration of assessment content and precludes the use of validated cognitive tests with already established normative and validation data. An alternative approach to adaptive testing is to predefine standardised and validated assessment pathways combining established psychometric evidence, expert consensus, and co-production. As each test component has already undergone validation and standardisation, the methodological challenge here regards optimising the adaptive assessment pathway rather than developing test content.

Few published examples exist of co-produced (cognitive) assessment, though the digiDEM SCREEN (digital dementia screen) (Zeiler et al., 2025) was designed and evaluated with stakeholders.

We set out to co-produce a pre-defined expert driven rule-based adaptive cognitive assessment for use by health professionals working in stroke, with both clinical stroke professionals and expert cognitive neuropsychologists. Next, we piloted the prototype with a selection of stroke survivors and neurologically healthy adults to examine app functionality, performance, and acceptability. Planned future work will evaluate the psychometric properties of the finalised tool.

Methods

Two studies were conducted: Study 1) details the development of the OCX, an expert-derived adaptive cognitive assessment through workshops and qualitative analysis, and Study 2) reports pilot testing of the OCX prototype in stroke survivors and neurologically healthy adults.

Study 1: Development

Ethical approval

was granted via the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (REF 587077). All participants provided online written informed consent. All methods were performed in accordance with the relevant guidelines and regulations, including the Declaration of Helsinki.

Participants. Healthcare professionals (HPs) with experience in cognitive assessment in stroke were recruited. Participants were eligible if they were over the age of 18, fluent in English, located within Europe, were a (allied) healthcare professional, and had clinical experience of cognitive assessment in stroke for at least one year. We aimed to recruit ~ 7 HPs, including at least one clinical neuropsychologist with stroke experience to ensure independent expertise on post-stroke cognitive assessment, in addition to that of the authorship team. We advertised the study through the UK Stroke Forum 2024, via social media, and word of mouth via our clinical networks.

Initial skeleton platform concept. Before co-production workshops took place, authors SSW & ND designed the concept and skeleton of the new adaptive pathway of tests termed the Oxford Cognitive Screening Platform (OCX). We aimed for the OCX to assess six domains: language, attention, memory, executive function, praxis, and number processing, in line with existing cognitive assessments and (inter)national guidance on cognitive assessment in stroke and acquired brain injury (Intercollegiate Stroke Working Party, 2023; Quinn et al., 2021). The OCX was designed to not only be flexible in sub-test administration, but that it could be delivered in full or in part, without loss of participant progress (with exception of the memory domain, which needs to be delivered as a whole to ensure standardised retention intervals).

Test selection and structure were constrained by copyright: we selected established tools for which we either had intellectual property rights (the Oxford Cognitive Screen (OCS);(Demeyere et al., 2015)), OCS-Plus (Demeyere et al., 2021), the Oxford Digital Multiple Errands Test (OxMET) (Webb et al., 2021), Birmingham Cognitive Screen (BCoS) (Bickerton et al., 2015), the Mini-OCS (Webb et al., 2025), and the Oxford Competency Assessment (COMPASS) tool (Webb et al., 2025), or employed tests that were copyright-free. Tasks were selected to be maximally inclusive, with minimal language-dependency, special presentations optimised for hemi-spatial neglect, and allowing examiner-facilitated response options to mitigate motor impairments. In addition, tasks are read aloud to the participant to maximise participation in those with visual impairments. Each cognitive domain included up to 3 tiers of testing progression. Further details about the initial structure within each of the domains can be found in the supplementary materials.

Procedure. We conducted two sets of co-production workshops with healthcare professionals. The aim of both was to facilitate discussion and debate about the structure of the OCX, its content, and the three tiers of content within each domain. For the first workshop, the research team gave an introductory presentation to the lab history and our work with different cognitive screens and justification for creating a new platform. Then we introduced the OCX and our initial structure, content, and justification to the group. We described each section of the app and asked participants to jot down or remember any points they wished to raise for the discussion. The workshop discussions were audio-recorded via computer and smartphone audio recording devices simultaneously. The automatic transcription was manually edited to ensure accuracy.

Participants were first given 10 minutes to discuss the initial pre-designed order of administration of tasks that was in line with other cognitive screening tests (i.e., language – number – spatial attention – praxis – memory – executive function). The next task(s) included discussing each cognitive domain in isolation, ordered by our initial order of administration. Each domain was given 10 minutes to discuss. Where there was time in the maximum 2.5hr focus group, we revisited areas of discussion after each domain had been covered. We selected prior prompts for discussion, following a topic guide, with room to elaborate on the discussion. One individual at a time was permitted to raise their discussion point. The topic guide for the workshops is provided in Table S1.

Data analysis. Transcriptions of the workshop discussions were analysed using reflexive template analysis to identify consensus regarding (i) assessment inclusion, (ii) pathway structure, and (iii) progression criteria. Coding, via NVivo

(version 15,)(Lumivero, 2024) was primarily descriptive and focused on informing prototype development rather than theory generation. Further detail on our qualitative analysis philosophy and technique are detailed in supplementary materials. Whilst the research team's extensive knowledge of the app and participation in discussions was a key asset to the co-production process, we acknowledge that the research team who designed the OCX were the same team who analysed transcription of participant's responses, and this may have biased the data. We maintained to keep in mind our professional beliefs, experiences, and values when analysing data and not overinterpret.

Materials presented in the workshops are not openly available due to copyright limitations. We are open to share the presentations used within workshops with other researchers on a case-by-case basis. We adhered to the Standards for Reporting Qualitative Research (O'Brien et al., 2014).

Study 2: Prototype feasibility (dup: abstract ?)

Ethical NHS approval was granted via the South Central - Oxford C Research Ethics Committee (REF: 20/03/2025). All participants provided online written informed consent or were consented via consultee consent procedures. All methods were performed in accordance with the relevant guidelines and regulations, including the Declaration of Helsinki.

Participants. We recruited both stroke survivors and neurologically healthy older adults.

Stroke survivors. For stroke survivors, eligibility criteria included being over 17 years of age, able to concentrate for at least 5 minutes overall, sufficient baseline understanding of English to understand the orientation questions, and having had, or suspected to have had, a stroke. Exclusion criteria were: having severe current psychiatric complaints or having severe sensory or perceptual impairments which made attempting the OCX impossible even with reasonable adjustments (e.g. blindness). Participants were recruited at the acute stroke unit of the John Radcliffe Hospital in Oxford and the Oxfordshire Stroke Rehabilitation unit in the Abingdon community hospital.

Neurologically healthy adults. Age- and education-matched control participants were recruited through an existing research database. Exclusion criteria were self-reported severe psychiatric illness and neurocognitive comorbidities (including e.g. stroke, transient ischaemic attack, head injury, dementias, Parkinson's disease).

Procedure. We worked with an in-house programmer to develop the cross-platform pilot OCX application in Android and iOS formats. Stroke survivors and neurologically healthy ageing adults completed the research visits as inpatients on the stroke wards, or at the University of Oxford, or were seen at their home, depending on time post stroke, access requirements, and participant choice. We administered the OCX and a custom survey using the Task Load Index (Hart & Staveland, 1988) and additional specific – to OCX – items (see Figure S1 in supplementary materials). For acceptability, we sought to indicate ease of test, understanding, enjoyment or dislike, length/burden, cognitive demand, and willingness to repeat the test. We benchmarked values based on quartile ranges, low and high scores are considered bottom and top quartile, respectively.

During early piloting we noted any functional issues with any test or the app itself and iteratively implemented changes. Data collection for neurologically healthy controls started when the final OCX version was available and compared the data from the new controls to the existing normative data for each individual test within the OCX. The first collection of normative data only completed the memory domain to generate new norms for the newly designed visual memory task, before the latter half completed all of the adaptive pathway. We visually describe the methodology in Fig. 1.

Note

(A)HP = (Allied) healthcare professionals.

Data analyses. For the pilot data testing, we briefly summarised participant performance, including which tiered tests were selected and if any clear patterns emerged (e.g., all participants went down the same routes, indicating a test may not be able to differentiate). We examined overall scores on the novel survey for load and acceptability and compared stroke to control participants for feedback. Items regarding repetition, understanding, and performance were reverse coded. The first six items from the Task Load Index (Hart & Staveland, 1988) were summed. The remaining items were summed, with higher scores reflecting higher acceptability. To interpret our novel questionnaire, we used lower quartile and upper quartile as categories for low burden and high acceptability, respectively.

Finally, we compared performance of the new controls to the original normative data of each test within the OCX by using appropriate statistical tests depending on statistical assumptions being met (e.g., normality and skew). This is presented in the supplementary materials.

Our manuscript analyses were written in Rmarkdown using R studio version 2026.09.0 + 174 (2026.09.0 + 174) (Posit team, 2024) and R version 4.4.2 (2024-10-31) (R Core Team, 2022). Analysis scripts and underlying data are available from our Github repository (Webb et al., 2026).

Results

Study 1: Development

Expert participants. A total of six healthcare professionals were recruited for the workshops in addition to authors (SSW and ND). The average age of participants was 35.66 ($SD = 13.69$), with most having an undergraduate level of education ($n = 4$; postgrad $n = 1$, doctoral $n = 1$), most being White: English, Welsh, Scottish, Northern Irish or British ($n = 5$; White: Other White $n = 1$), most being Occupational Therapists ($n = 4$; Speech and Language therapist $n = 1$, neuropsychologist $n = 1$), most working in stroke rehabilitation settings ($n = 3$; early supportive discharge $n = 1$, acute stroke $n = 1$, general neurorehabilitation $n = 1$). The average years of experience with stroke care was 10.50 years ($SD = 9.03$) and average years of experience with cognitive screening was 9.83 years ($SD = 7.70$).

Workshop 1. To ensure manuscript brevity, we summarise major changes made to OCX following the first workshop in text and present key quotes to support all changes in Table S3 in the supplementary materials.

Key changes included (i) the order of administration for domain modules, (ii) corrections to ensure consistency, (iii) improvements to clarify instructions, including improving practice trials, and (iv) changes to maximise inclusion and minimise burden.

With regards to the order of administration, the changes reflected perceived clinical priorities (e.g. the Number domain moved later in the ordering as seen to be lower priority as well as more challenging). Changes to ensure consistency included ensuring that both sides of the possible pathways from tier 0 had similar modalities and areas (e.g. if semantics is assessed in tier 1, it should be done for both pass and fail of tier 0) and not repeating similar tasks, as well as ensuring each test is sufficiently different and necessary to maximise breadth whilst trying to minimise assessment burden. Examples here included extending the praxis tests to include both ideomotor and ideational elements. Improving clarity and instructions involved additional practice trials and a re-design of the rule finding task. Finally, with regards to inclusion-based changes, a clear focus was on ensuring most domains would include tests using different modalities, not just verbal or written (e.g., visual as well as verbal memory tasks). In addition, cultural appropriateness was considered with some items removed or changed where they were deemed too dated or culturally specific.

Workshop 2. Detailed information on all discussions and codes per domain is provided in Table S4 for feedback per app section. To summarise the discussions and changes, these included finalising the order of the “in-step tasks” prior to the domain testing and determining a finalised suggested order of administration for the domains. The general flow and

structure of each domain was considered acceptable, and whilst further discussion included considerations of additional elements missing in OCX, including assessments of risk or sequencing of actions, maintaining a balance between brevity and completeness meant no further additions were made.

The final OCX prototype. The OCX is structured as in Fig. 2 and Table 2.

Test inclusion. We detail the final full list of tests included in OCX by cognitive domain. Where test items were removed based on co-production changes, we detail analyses of these subtests using Item Response Theory in the supplementary materials. In brief, we shortened the test items and correlated the short form with the original full form subtest to evidence the construct validity of the short form (average correlation was $r = .85$, range .78 to .922). We then adjusted normative data cut offs with the original control data isolating only the final short form items. Modified tasks are noted explicitly in Table 2.

Study 2: Prototype feasibility

Participants. We collected pilot data from 22 stroke survivors with a mean age of 72.40yrs (SD = 12.4), a mean education of 14.60yrs (SD = 3.72), median NIHSS of 5 (range = 0 to 18), and median days since stroke 18.50 (range = 1 to 15600), and 86 neurologically healthy controls (mean age = 69.50, SD = 15.40, range 19–92; mean education years = 16.20, SD = 2.73). There were no statistical differences between controls and stroke participants in age ($p = .75$) or education ($p = .07$). Note, the first 6 controls only completed the memory and executive function subtests of the OCX initially. Additionally, for stroke survivors, the average MoCA score was 16 (SD = 9), with 90.91% scoring below a cut off 26, and 72.73% scoring below 23.

Pilot performance data. We first descriptively explored the different adaptive pathways that the pilot sample of 22 stroke participants took through the OCX, to determine whether any clear patterns of pre-set pathways were common. We provide the percentages of participants who pass at each tier level across all domains and mark the minimum and maximum observed pass rates overall. Table S5 shows some tier levels had higher pass rates than others, but that the lower the pathway progresses, the fewer pass rates there are (increasing difficulty or further investigation). In Fig. 3, we visualise pass rates of each domain separately.

Pilot acceptability data. For controls, the average feedback score for Task Load was 21.65 (SD = 13.61, range = 2–66) and their average acceptability was 82.05 (SD = 14.24, range = 34–100). For stroke survivors the average Task Load score was 16.73 (SD = 7.50, range = 2–27) and average acceptability was 85.45 (SD = 12.83, range = 62–100). Across samples, the average task load was 20.90 (SD = 12.96, range = 2–66) and acceptability was 82.58 (SD = 14, range = 34–100). There were no statistically significant differences between controls and stroke participants for task burden nor feedback ($W = 406.50$, $p = .32$, 95% CI = -3.99 to 11.00, and $W = 284$, $p = .47$, 95% CI = -10.99 to 5.00, respectively).

Comparison of new pilot control sample to the original normative data.

We found no statistical differences (Welch Two Sample t-test) between the original control sample and the current normative samples across age and education (age: mean difference = -2.61, $t(123.90) = -1.42$, $p = .16$, 95% CI -6.24 to 1.03; education: mean difference = -1.79, $t(133.16) = -4.72$, $p = .06$, 95% CI 14.38 to 16.17). We examined differences in performance across groups on shared variables from the OCX and the original paper-based tests.

Differences between controls groups, corrected with Bonferroni-Holm multiple comparison correction, were found for the following tasks out of 32 subtest comparisons (r effect size for Wilcoxon comparisons, average $r = .14$): orientation ($N = 82$, $W = 36248$, $p = .003$, $r = -0.01$), picture naming ($N = 76$, $W = 32668$, $p < .001$, $r = -0.02$), word/word matching ($N = 3$, $p = .032$, $r = -0.19$), picture/picture matching ($N = 73$, $W = 2275$, $p < .001$, $r = -0.44$), single word reading ($N = 73$, $W = 3689$, $p < .001$, $r = -0.20$), mini-Broken hearts ($N = 76$, $W = 6348$, $p = .013$, $r = .18$), free story recall ($N = 47$, $W = 1968$, $p < .001$, $r =$

-0.35), number/price/time comparison ($N = 76$, $W = 3790$, $p = .001$, $r = -0.20$), number writing ($N = 68$, $W = 1800$, $p < .001$, $r = .45$), written number decision ($N = 8$, $W = 304$, $p = .001$, $r = -0.17$), OxMET ($N = 26$, $W = 798$, $p < .001$, $r = -0.32$). Note that some comparisons had small sample sizes due to protocol differences between the first 57 controls and the rest to 86 whom did not get specific tests. Calculated percentile cut offs per task however did not meaningfully differ between the original and new pilot control groups, all original and new cut offs as well as statistical comparisons are presented in Table S6.

Discussion

We aimed to co-produce and provide initial pilot psychometric and acceptability data for a novel adaptive cognitive screen application (OCX), suitable for in-person assessments of stroke survivors presenting with a wide range of cognitive impairments and severity of impairment. Iterative co-production of the app with (allied) healthcare professionals through workshops finalised the application tested in study 2. The OCX showed that, in a pilot stroke sample, there was no single pathway through the adaptive design, with differential performances depending on domain-specific cognitive abilities (i.e., not everyone passed each tier test). The OCX showed promising acceptability and was perceived to have a low burden by both neurologically healthy adults and an initial small sample of 22 stroke survivors.

Co-production of (cognitive) assessments

We present a well-documented and inclusive co-production test development process, including expert stakeholders who provided both clinical end-user views as well as neuropsychological assessment expertise.

Limitations and Future research

We have transparently reported that we did not base our decision trees and pathways upon a pre-calibrated (Item Response Theory [IRT]) bank of items, and instead used existing normative data, psychometrics and expert co-production to construct the adaptive testing. With the current data we are unable to calculate ability levels (theta in IRT) for participants on our test. In line with the Standards of Education and Psychological Testing (American Educational Research Association et al., 2014), we maintain psychometric validity by designing a test that (i) has a specific intended sample group, (ii) assesses particular cognitive impairments, and (iii) enables interpretation of results in an intended way (clinical practice by those who would use it, e.g., occupational therapists, neuropsychologists). Additional planned research will aim to provide further evidence of validity to evidence the psychometric justification for our choices and the validity of interpretations. We

aim to establish the concurrent, criterion, predictive, convergent, and discriminant validity of the OCX and examine the equivalence of the OCX subtests to their original pen-and-paper counterparts for stroke survivors across the care continuum and chronicity post-stroke. Further evidence will be gathered regarding the diagnostic accuracy of this screen in comparison to the OCS and OCS-Plus.

OCS was designed as a first-line cognitive screen and may have more limited sensitivity in memory and executive functioning domain tasks (Murphy et al, 2023), due to the test prioritising specificity and inclusive testing of aphasia. OCS-Plus on the other hand focusses on domain-general impairments and was shown to have high sensitivity for detecting mild impairments in a stroke sample, showing excellent agreement in detecting impairments through a more extensive neuropsychological battery (S. S. Webb et al., 2022). Whilst by design we envisage the OCX to have both high sensitivity and specificity, this remains an empirical question for future research.

We further acknowledge limitations in the new control data collected in this pilot study. Due to the adaptive nature of the test, controls did not complete all of the possible subtests within the OCX as the app was not designed to choose specific subtests out of context. Whilst the results show that the new and original control data did not meaningfully differ in the

tests that were completed, future research will need to gather new control data for all included tests to provide a standardised and fully normed OCX.

Conclusion

The OCX is a novel multistage computerised cognitive screening tool designed to be suitable for stroke survivors. OCX provides adaptive testing across 6 cognitive domains, aiming to provide both a more in-depth cognitive screen for mildly impaired patients, as well as being suitable for more severely impaired patients, within a brief, in-person screening app. Initial pilot performance variation and high acceptability provide strong justification for future research to determine validity and diagnostic accuracy.

Declarations

Nele Demeyere is a developer of the “Oxford Cognitive Screen”. Nele Demeyere and Sam S Webb are developers of the “Oxford Cognitive Screen-Plus”. The OCS and OCS-Plus are licenced through Oxford University Innovations which holds the copyright. Nele Demeyere and Sam S. Webb have previously received funds from Brain Stimulation AB, but this company had no input in the concepts, ideas, design, methods, results, or publication of the present study.

Author Contribution

According to CRediT taxonomy: S.S.W was responsible for conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualisation, writing - original draft, and writing - review & editing; P.G. was responsible for data curation, investigation, validation, writing - original draft, and writing - review & editing. N.D. was responsible for supervision, resources, funding acquisition, and conceptualisation. All authors reviewed the manuscript.

Acknowledgement

We would like to thank all participants who gave their time to aid in this study, and Nick Goodall who programmed and implemented the pilot OCX on the app stores. This research was funded by the Stroke Association through a postdoctoral fellowship award to S.S.W. (PF25_100003), N.D. (Advanced Fellowship NIHR302224) is funded by the National Institute for Health and Care Research (NIHR). The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR, NHS or the UK Department of Health and Social Care.

Data Availability

Data and analysis scripts/qualitative data/transcripts are available in our Github repository:
<https://github.com/SamSWebb/Development-and-Pilot-Evaluation-of-OCX#>

References

1. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds). *Standards for educational and psychological testing* (American Educational Research Association, 2014).
2. Bickerton, W. L. et al. The BCoS cognitive profile screen: Utility and predictive value for stroke. *Neuropsychology* **29** (4), 638–648. <https://doi.org/10.1037/neu0000160> (2015).

3. Brodaty, H. et al. The GPCOG: A New Screening Test for Dementia Designed for General Practice. *J. Am. Geriatr. Soc.* **50** (3), 530–534. <https://doi.org/10.1046/j.1532-5415.2002.50122.x> (2002).
4. Chen, N. Y. C., Tan, M. Y. L., Xu, J., Zuo, L. & Dong, Y. VasCog Screen test: Sensitive in detecting cognitive impairment in patients who had a stroke or with heart failure. *Stroke Vascular Neurol.* **10** (1). <https://doi.org/10.1136/svn-2023-002701> (2025).
5. Demeyere, N. et al. Introducing the tablet-based Oxford Cognitive Screen-Plus (OCS-Plus) as an assessment tool for subtle cognitive impairments. *Sci. Rep.* **11** (8000). <https://doi.org/https://doi.org/10.1038/s41598-021-87287-8> (2021).
6. Demeyere, N., Riddoch, M. J., Slavkova, E. D., Bickerton, W. L. & Humphreys, G. W. The Oxford Cognitive Screen (OCS): Validation of a stroke-specific short cognitive screening tool. *Psychol. Assess.* **27** (3), 883. <https://doi.org/http://dx.doi.org/10.1037/pas0000082> (2015).
7. Demeyere, N. et al. Domain-specific versus generalized cognitive screening in acute stroke. *J. Neurol.* **263** (2), 306–315. <https://doi.org/10.1007/s00415-015-7964-4> (2016).
8. Feigin, V. L. et al. Global, regional, and national burden of stroke and its risk factors, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet Neurol.* **20** (10), 795–820. [https://doi.org/10.1016/S1474-4422\(21\)00252-0](https://doi.org/10.1016/S1474-4422(21)00252-0) (2021).
9. Gruia, D. C. et al. Mitigating the impact of motor impairment on self-administered digital tests: A longitudinal cohort study in stroke. *eClinicalMedicine*, **88**. (2025). <https://doi.org/10.1016/j.eclinm.2025.103469>
10. Gruia, D. C. et al. IC3 protocol: A longitudinal observational study of cognition after stroke using novel digital health technology. *BMJ Open.* **13** (11), e076653. <https://doi.org/10.1136/bmjopen-2023-076653> (2023).
11. Hart, S. G. & Staveland, L. E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139–183). Elsevier. (1988). <https://www.sciencedirect.com/science/article/pii/S0166411508623869>
12. Intercollegiate Stroke Working Party. *National Clinical Guideline for Stroke for the UK and Ireland*. Intercollegiate Stroke Working Party. www.strokeguideline.org (2023).
13. Johnson, C. O. et al. Global, regional, and national burden of stroke, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* **18** (5), 439–458. [https://doi.org/10.1016/S1474-4422\(19\)30034-1](https://doi.org/10.1016/S1474-4422(19)30034-1) (2019).
14. Kusec, A. et al. Long-term psychological outcomes following stroke: The OX-CHRONIC study. *BMC Neurol.* **23**, 426 (2023).
15. Lees, R. et al. Test Accuracy of Cognitive Screening Tests for Diagnosis of Dementia and Multidomain Cognitive Impairment in Stroke. *Stroke* **45** (10), 3008–3018. <https://doi.org/10.1161/STROKEAHA.114.005842> (2014).
16. Lumivero. *NVivo* (Version 15) [Computer software]. (2024). <https://lumivero.com/products/nvivo/>
17. McMahon, D., Micallef, C. & Quinn, T. J. Review of clinical practice guidelines relating to cognitive assessment in stroke. *Disabil. Rehabil.* **0** (0), 1–9. <https://doi.org/10.1080/09638288.2021.1980122> (2021).
18. Milosevich, E. T., Moore, M. J., Pendlebury, S. T. & Demeyere, N. Domain-specific cognitive impairment 6 months after stroke: The value of early cognitive screening. *medRxiv*, 2023–2006. (2023).
19. Nasreddine, Z. S. et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **53** (4), 695–699. <https://doi.org/https://doi.org/10.1111/j.1532-5415.2005.53221.x> (2005).
20. Nys, G. M. S. et al. Cognitive Disorders in Acute Stroke: Prevalence and Clinical Determinants. *Cerebrovasc. Dis.* **23** (5–6), 408–416. <https://doi.org/10.1159/000101464> (2007).

21. O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A. & Cook, D. A. Standards for Reporting Qualitative Research: A Synthesis of Recommendations. *Acad. Med.* **89** (9), 1245. <https://doi.org/10.1097/ACM.0000000000000388> (2014).
22. Oladele, J. I. & Ndlovu, M. A Review of Standardised Assessment Development Procedure and Algorithms for Computer Adaptive Testing: Applications and Relevance for Fourth Industrial Revolution. *Int. J. Learn. Teach. Educational Res.* **20** (5), 1–17. <https://doi.org/10.26803/ijlter.20.5.1> (2021).
23. Pagliarin, K. C. et al. Montreal–Toulouse Language Assessment Battery: Evidence of criterion validity from patients with aphasia. *J. Neurol. Sci.* **357** (1–2), 246–251 (2015).
24. Posit team. *RStudio: Integrated Development Environment for R* (Posit Software, 2024). <http://www.posit.co/>
25. Quinn, T. J., Elliott, E. & Langhorne, P. Cognitive and mood assessment tools for use in stroke. *Stroke* **49** (2), 483–490 (2018).
26. Quinn, T. J. et al. European Stroke Organisation and European Academy of Neurology joint guidelines on post-stroke cognitive impairment. *Eur. Stroke J.* 23969873211042192. <https://doi.org/10.1177/23969873211042192> (2021).
27. R Core Team. *R: A language and environment for statistical computing* [Manual]. R Foundation for Statistical Computing. (2022). <https://www.R-project.org/>
28. Robinson, G. A., Shi, L., Nott, Z. & Ceslis, A. A Brief Executive Language Screen for Frontal Aphasia. *Brain Sci.* **11** (3). Article 3. <https://doi.org/10.3390/brainsci11030353> (2021).
29. Rudd, A. G., Bowen, A., Young, G. & James, M. A. National clinical guideline for stroke: 2016. *Clinical Medicine.* (2017).
30. Schmidt, L. F. S., Rodrigues, J. C., Bandeira, D. R. & de Salles, J. F. Validity evidence for the Cognitive Screening Test in Stroke Patients. *Psico-USF* **27**, 735–749. <https://doi.org/https://doi.org/10.1590/1413-82712022270411> (2023).
31. Stroke Association. *Lived Experience of Stroke Report.* (2018). http://www.stroke.org.uk/sites/default/files/conferences/nisc/documents/lived_experience_of_stroke_chapter_1.pdf
32. Tessaro, B. et al. Assessment of cognition in aphasia: Perspectives from clinicians and researchers. *Neuropsychological Rehabilitation.* **36** (2), 278–317. <https://doi.org/10.1080/09602011.2025.2506598> (2026).
33. Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F. & Mislevy, R. J. *Computerized adaptive testing: A primer* (Routledge, 2000). <https://api.taylorfrancis.com/content/books/mono/download?identifierName=doi&identifierValue=10.4324/9781410605931&type=googlepdf>
34. Webb, S. et al. *The Oxford COMPASS: A Brief Cognitive Screen to help guide Competency Assessments.* PsyArXiv. (2025). https://doi.org/10.31234/osf.io/vygbf_v1
35. Webb, S. S. et al. The Oxford Digital Multiple Errands Test (OxMET): Validation of a Simplified Computer Tablet Based Multiple Errands Test. *Neuropsychological Rehabilitation.* <https://doi.org/10.1080/09602011.2020.1862679> (2021).
36. Webb, S. S., Grammatikos, P. & Demeyere, N. *OCX pilot study* [Data set]. Github. (2026).
37. Webb, S. S. et al. Validation of the UK English Oxford cognitive screen-plus in sub-acute and chronic stroke survivors. *Eur. Stroke J.* 23969873221119940. <https://doi.org/10.1177/23969873221119940> (2022).
38. Webb, S. S., Sun, L., Tang, E. & Demeyere, N. *The Mini-Oxford Cognitive Screen (Mini-OCS): A Very Brief Cognitive Screen for Use in Chronic Stroke* (European Stroke Journal, 2025).
39. Zeiler, M. et al. A User-Centered Design Approach for a Screening App for People With Cognitive Impairment (digiDEM-SCREEN): Development and Usability Study. *JMIR Hum. Factors.* **12** (1), e65022. <https://doi.org/10.2196/65022> (2025).

Tables

Table 1 is not available with this version.

Table 2 is available in the Supplementary Files section.

Figures

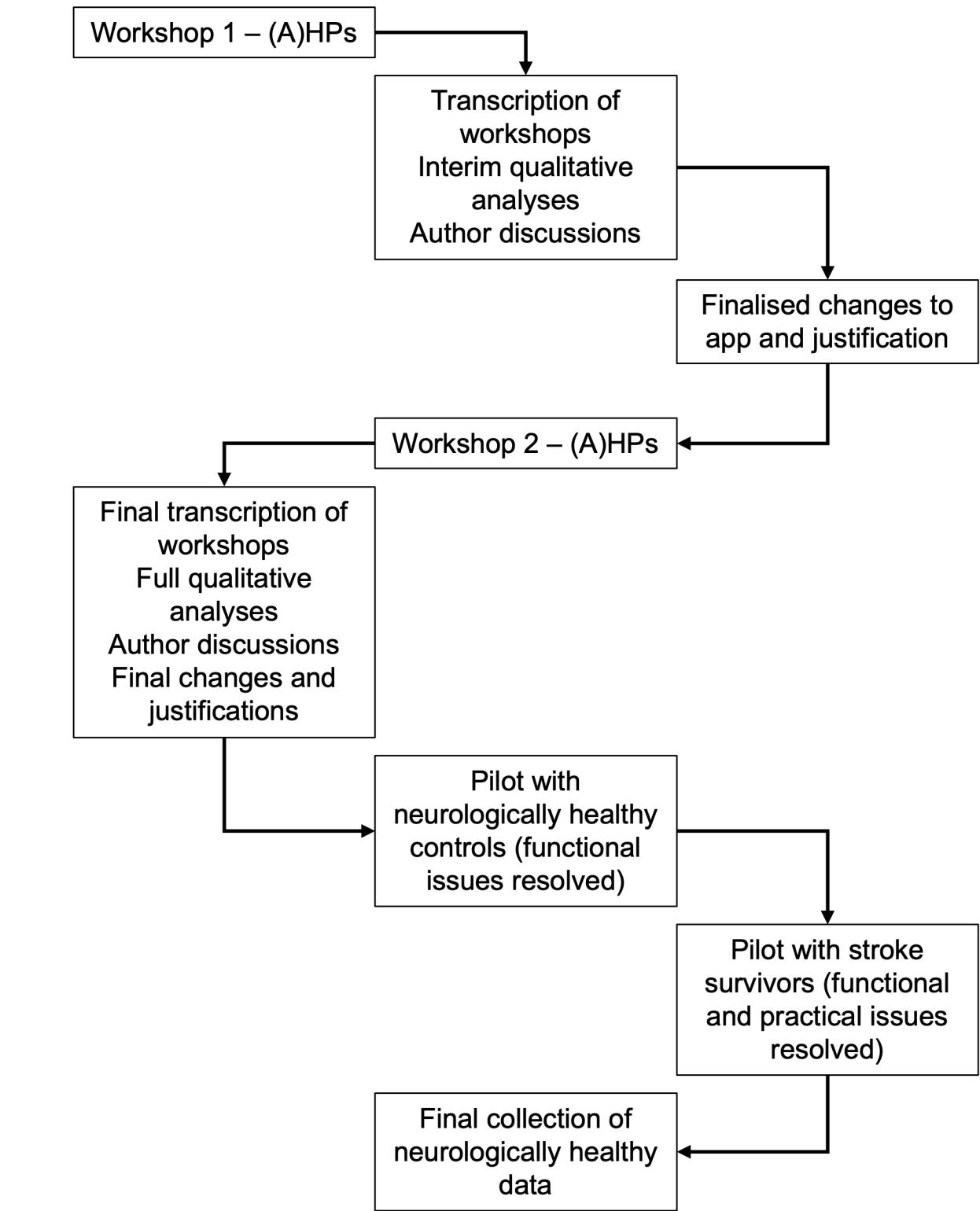


Figure 1

Flow diagram for the design of the whole study, including workshops, piloting, and re-standardisation of the app.

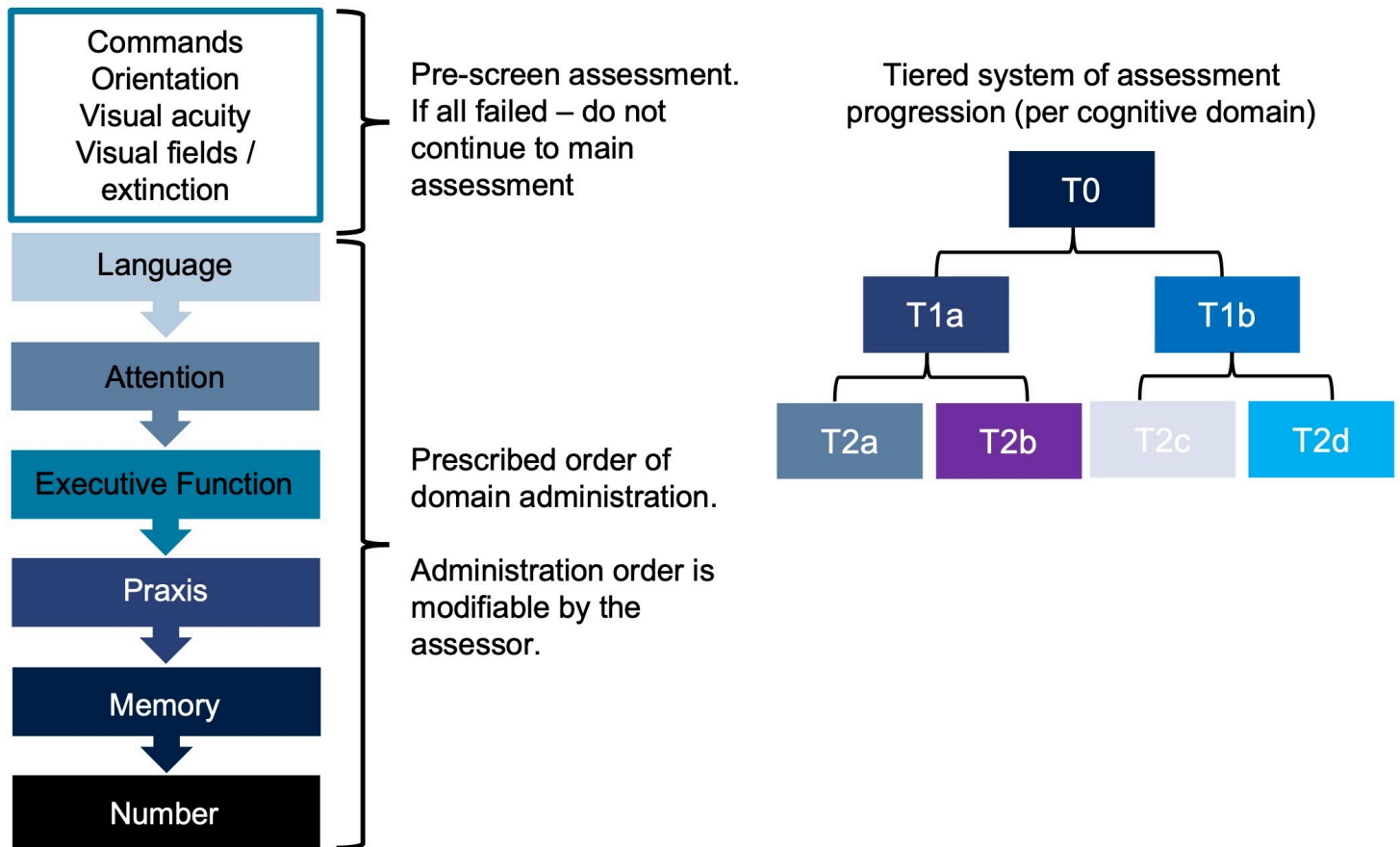


Figure 2

Overall structure and flow of assessment for the Oxford Cognitive Screening Platform (OCX). Four initial “in-step” tasks are given to establish basic cognitive and visual abilities to aid interpretation of performance before progression to either a pre-determined order of assessed domains or flexible ordering. On the right of Figure 2, we demonstrate how each domain is designed, where ‘T’ stands for ‘tier’. Some domains have two tests that are run together on a single tier and may be labelled ‘t1a1’ and ‘T1a2’, for example.

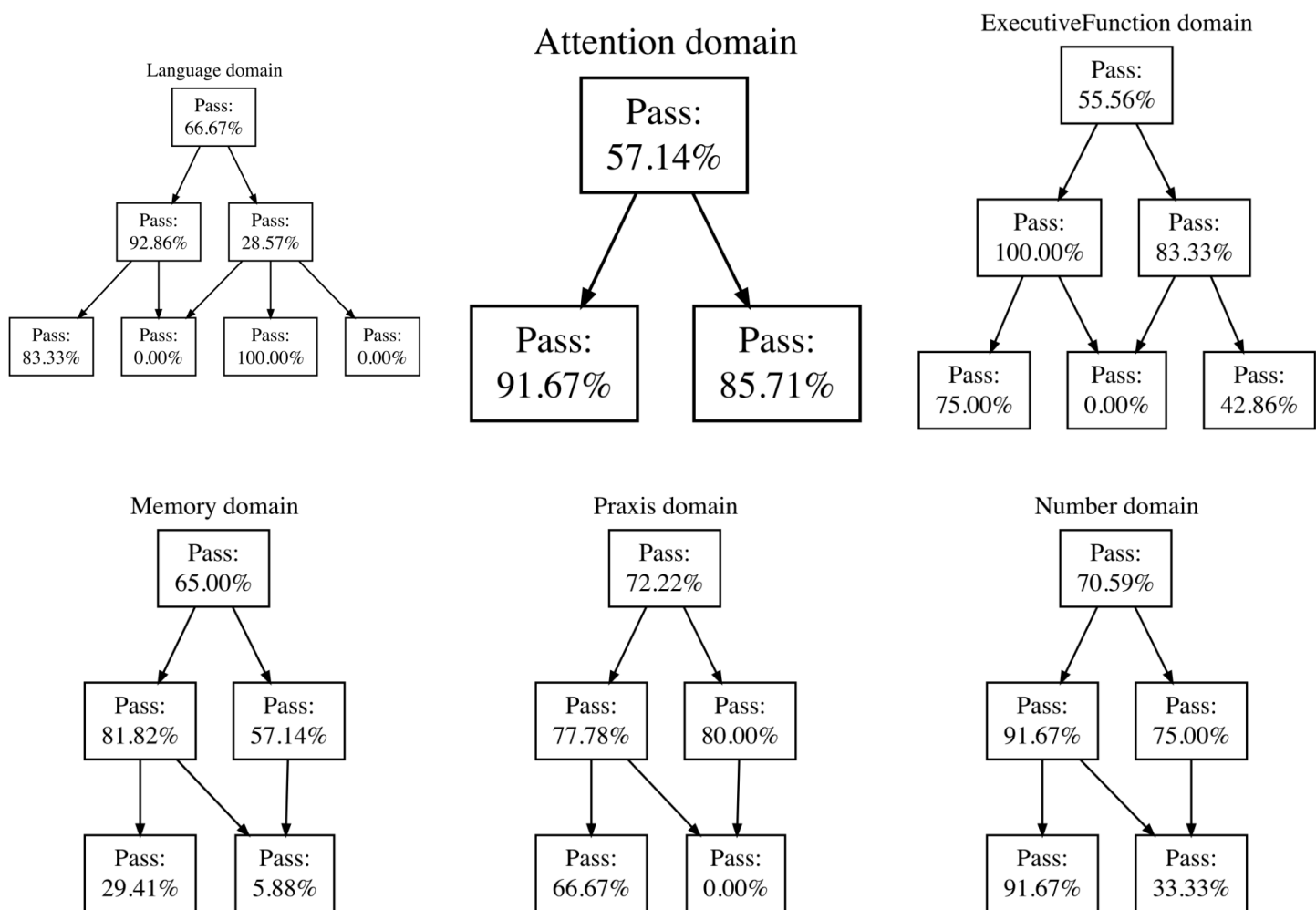


Figure 3

Pass (not outside of cut off) rates at each tier for stroke participants across each cognitive domain in the OCX.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SuppscoproductionV3.docx](#)
- [Table2.docx](#)