Examining measurement properties of cognitive screening instruments used post-stroke

A Thesis by

Emma P J Elliott BSc (Hons)

Submitted for the degree of Doctor of Philosophy

To

The University of Glasgow

From

The Institute of Cardiovascular and Medical Sciences

College of Medical, Veterinary and Life Sciences

University of Glasgow

Glasgow Royal Infirmary

2021

Abstract

Background: Cognitive screening after a stroke is recommended by clinical guidelines, specialist societies and as part of national audit programs. However, due to vague recommendations, different cognitive syndromes, and differing opinions regarding cognitive screening instrument (CSI) choice and timing, a range of CSIs are being used in clinical practice and research. There are limited data related to the use of both brief CSIs (administered in ≤5 minutes) and stroke-specific CSIs. This means that some teams may be using CSIs without any supportive evidence that they are fit for purpose. I aimed to examine measurement properties of different brief generic CSIs and the Oxford Cognitive Screen (OCS).

Methods: I first conducted a study into the feasibility of various brief CSIs on a hyper acute stroke unit; I examined the completion rates, reasons for being untestable and examined associations with being untestable.

I conducted two systematic reviews of test accuracy; one to identify and evaluate shortened versions of the Montreal Cognitive Assessment (SF-MoCA) and the second to evaluate telephone-based CSIs.

Using the data from the Assessing Post-Stroke Psychology Longitudinal Evaluation study (APPLE), I examined completion rates and floor/ceiling effects of a range of brief CSIs and the OCS. I examined the accuracy of brief CSIs to detect pre-stroke cognitive impairment (against diagnosis in medical records) and to detect post-stroke single and multi-domain cognitive impairment, using the OCS as a reference standard. Finally, I investigated whether domain-specific results from the OCS completed at one-month post-stroke were associated with functional, mood and quality of life outcomes at six months.

Findings: A quarter of participants were untestable on at least one cognitive test item. Across the different CSIs examined, the clock drawing test (CDT) had the lowest completion rate, whereas there were no missing data using the 4 A's Test (4AT), due to scoring for untestable being incorporated.

In the first systematic review I identified thirteen SF-MoCAs. Across the published literature and in the external validation, the performance of the short forms varied but demonstrated a pattern of high sensitivity to detect multi-domain cognitive impairment, according to different reference standards.

In the second systematic review I identified 15 telephone-based CSIs to identify MCI or dementia. Four of these CSIs were used in participants post-stroke (Telephone Interview for Cognitive Status [TICS], TICS-modified, Telephone-Montreal Cognitive Assessment [T-MoCA], T-MoCA short). Of the limited data available in stroke, the telephone CSIs demonstrated high sensitivity to detect multi-domain cognitive impairment. Outside of stroke, the TICS and TICS-m had the greatest supportive evidence base to screen for dementia.

In the APPLE study, ceiling effects were highest for the Abbreviated Mental Test (AMT-4), Cog-4 and 4AT. Across eight brief CSIs, the pattern of accuracy for pre- and post-stroke cognitive syndromes was generally low sensitivity, high specificity, apart from the CDT and NINDS-CSN 5-min MoCA which exhibited the opposite pattern. The OCS had good completion rates, but fewer participants fully completed it in comparison to the brief CSIs. There were no issues of floor/ceiling effects. In unadjusted models, all OCS domains apart from memory were significantly associated with at least one six-month outcome. However, when controlling for confounding variables (such as age, education, pre-stroke disability and stroke severity), and adjusting for multiple testing, only one domain remained significant with one outcome: executive dysfunction had a modest association with reduced quality of life (measured using the EQ-5D).

Conclusions: To summarise, in the context of stroke, incomplete cognitive screening assessments should be expected. CSIs with fewer items or stroke specific CSIs do not necessarily have a higher completion rate. Clinicians and researchers should therefore make a-priori plans on how to address incomplete assessments.

Recommendations for CSI choice differ depending on the purpose of screening, including resources and plans for following up those with identified cognitive impairment. Most brief CSIs demonstrated low sensitivity, high specificity to detect post-stroke multi-domain cognitive impairment so would not be

recommended for clinical use. Telephone-based CSIs have some promising initial data in the stroke context, but further studies are needed before recommending for clinical use. There was insufficient evidence that results from the OCS at one month are associated with functional and mood outcomes at six months, but some evidence that executive dysfunction is independently associated with reduced quality of life. Further studies are necessary to understand the prognostic utility of the OCS.

Table of Contents

List of Tables

List of Figures

Preface: Context and thesis outline

All prospective data collection in this thesis were collected in a UK, NHS setting, therefore interpretations and recommendations are with this context in mind.

The overarching objective of the work presented in this thesis is to examine the following measurement properties of different cognitive screening instruments (CSIs): feasibility, accuracy, floor/ceiling effects, prognostic utility. The focus is primarily on the use of CSIs on a hyper acute stroke unit (HASU) in patients after ischaemic/haemorrhagic stroke or a transient ischaemic attack, however the two systematic review chapters also review data from other disease areas, such as Alzheimer's disease and multiple sclerosis, as comparators.

The outline of the thesis is provided in Figure 1. Chapter 1 provides the background and rationale for the work carried out. Chapter 2 addresses completion rates of brief CSIs in all patients who were admitted to a HASU. This real-world data set gives us a unique insight into reasons given when a CSI cannot be completed. Chapters 3 and 4 are systematic reviews, examining the accuracy of short forms of the Montreal Cognitive Assessment (MoCA) and telephone-based CSIs. Chapters 5-7 cover the methods and results from a prospective, observational, longitudinal study: Assessing Post-stroke Psychology Longitudinal Evaluation (APPLE). These chapters provide measurement properties of a range of brief CSIs and a stroke-specific multi-domain CSI: the Oxford Cognitive Screen. From these results a set of evidence-based recommendations specific to varying purposes of cognitive screening post-stroke are provided.

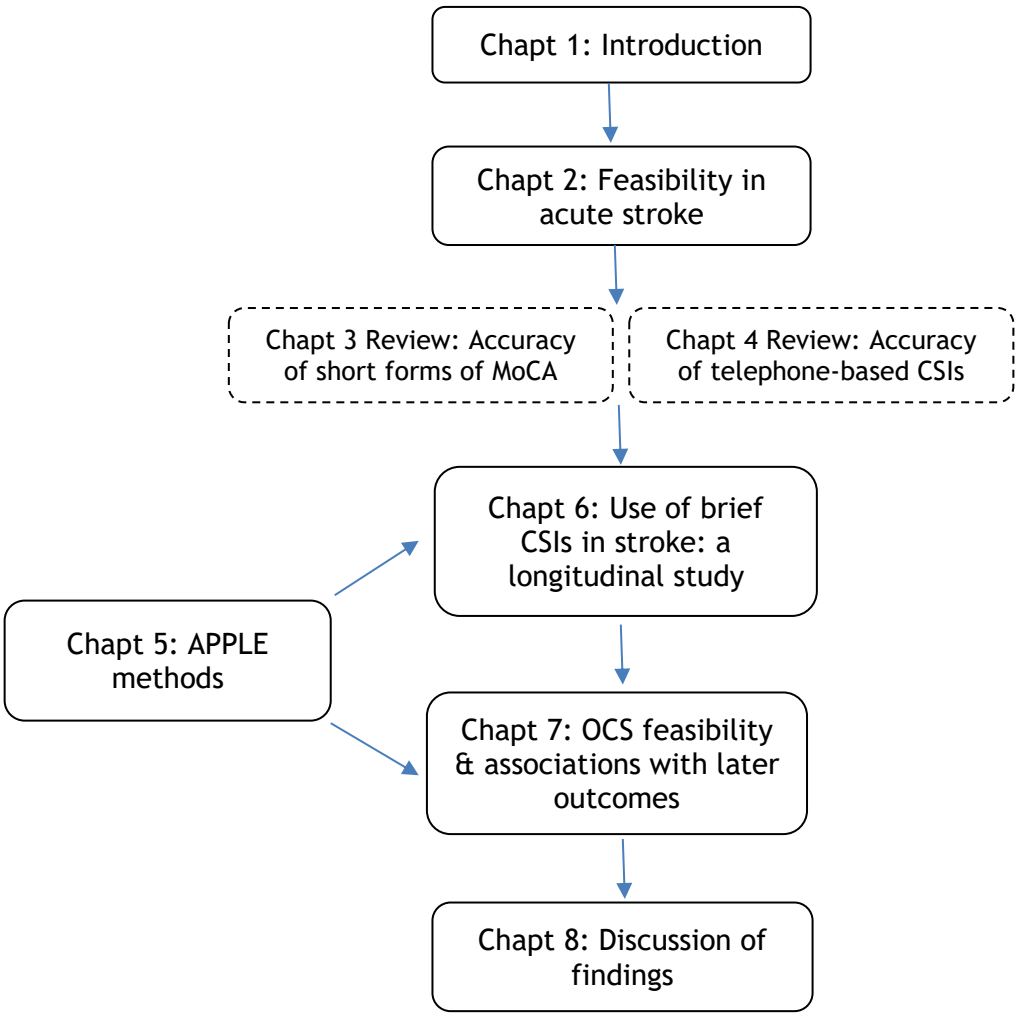


Figure 1 Thesis Structure

Abbreviations: APPLE, Assessing Post-stroke Psychology Longitudinal Evaluation; CSI, Cognitive screening instrument; MoCA, Montreal Cognitive Assessment; OCS, Oxford Cognitive Screen.

Acknowledgement

Firstly, I would like to thank my supervisors, Dr Terry Quinn and Professor Jesse Dawson, for their excellent supervision, advice and opportunities over the past four years. They are a great team to work with and I look forward to more collaborations in the future.

I would also like to thank my colleagues at the Glasgow Royal Infirmary for their continued support and making my time in Glasgow so enjoyable. I especially would like to thank my colleague and good friend, Bogna Drozdowska, who I saw almost every day; I am so happy that we got to undertake our PhDs together.

I am extremely grateful to all the participants that took part in the APPLE research study; it was a pleasure getting to know each of you over the 18-month period of the study. I would like to thank the dedicated research staff at all the APPLE hospital sites, especially Ruth Graham, and the Stroke Association for funding the research.

Finally, I would like to give a big thank you to my fantastic family and friends for your endless support.

Author's Declaration

My contribution to each chapter, along with significant contributions of others involved, is outlined below.

Dr Terence Quinn and Professor Jesse Dawson provided supervision, guidance, and feedback on each chapter of this thesis.

Chapter 1: I wrote the chapter.

Chapter 2: Dr Terry Quinn and I conceived the idea for this study. I wrote the chapter. I collected data for this chapter along with other researchers: Martin Taylor-Rowan, Robert Shaw, Bogna Drozdowska, Gillian Cuthbertson. I carried out the analysis.

Chapter 3: Dr Terry Quinn conceived the idea for this study. I wrote the chapter. I conducted screening, data extraction and quality assessment for the systematic review, along with Dr Jennifer McDicken and Dr Gareth Blayney. Dr Myzoon Ali performed the validation analyses.

Chapter 4: Dr Terry Quinn conceived the idea for this study. I wrote the chapter. I carried out the search, screening, data extraction and quality assessment, along with Claire Green. I conducted the meta-analysis.

Chapter 5: I wrote the chapter.

Chapter 6: Dr Terry Quinn and I conceived the idea for this study. I collected data for the APPLE study along with other PhD students and research nurses at each hospital site. I wrote the chapter and conducted analyses.

Chapter 7: I conceived the idea for this study, wrote the chapter and conducted analyses. The data came from the APPLE research study.

Chapter 8: I wrote the chapter.

Publications, Conferences and Awards

Publications related to thesis

- Chapter 1: Quinn, T. J., **Elliott, E**., and Langhorne, P. (2018) Cognitive and mood assessment tools for use in stroke. Stroke, 49(2), 483-490.

- Chapter 2: **Elliott, E**., Drozdowska, B.A., Taylor-Rowan, M.; Shaw, R.C.; Cuthbertson, G.; Quinn, T.J. (2019) Who Is Classified as Untestable on Brief Cognitive Screens in an Acute Stroke Setting? Diagnostics, 9(3), 95.

- Chapter 3: McDicken, J.A*, **Elliott, E**.*, Blayney, G., Makin, S., Myzoon, A., Larner, A. J., Quinn, T. J. (2019) Accuracy of the short form Montreal Cognitive Assessment – Systematic Review and Validation. Int J Geriatric Psychiatry, 34(10), 1515-1525. (*Joint first author)

- Chapter 4: **Elliott, E.,** Green, C., Llewellyn, D., Quinn, T. J. (2019) Diagnostic accuracy of Telephone Assessments of Cognition: Systematic Review and Meta-Analysis. Current Alzheimer research, 17(5), 460-471.

Publications not included in thesis

- **Elliott, E.**, Haldane, D., and Quinn, T. J. (2018) Pitfalls of neurocognitive testing in an occupational medical Setting. Occupational Medicine, 69(2), 83-85

- Shaw, R., Drozdowska, B., Taylor-Rowan, M., **Elliott, E.,** Cuthbertson, G., Stott, D., Quinn, T. J. (2019) Delirium in an acute stroke setting, occurrence and risk factors. Stroke, 50(11), 3265-3268.

- Shaw, R., Walker, G., **Elliott, E**., Quinn, T.J. (2019). Occurrence rate of delirium in acute stroke settings - systematic review and meta-analysis. Stroke, 50(11), 3028-3036.

- Taylor-Rowan, M., Keir, R., Cuthbertson, G., Shaw, R., Drozdowska, B., **Elliott, E.,** Evans, J., Stott, D. and Quinn, T. J. (2019) Pre-Stroke Frailty Is Independently Associated With Post-Stroke Cognition: A Cross-Sectional Study. Journal of the International Neuropsychological Society, 25(5), 501-506.

- Taylor-Rowan, M., Cuthbertson, G., Keir, R., Shaw, R., Drozdowska, B., **Elliott, E.**, Stott, D. J. and Quinn, T. J. (2019) The prevalence of frailty amongst acute stroke patients, and evaluation of method of assessment. Clinical Rehabilitation, 33(10), 1688-1696.

- Drozdowska, B. A., **Elliott, E**., Taylor-Rowan, M., Shaw, R. C, Cuthbertson, G., Langhorne, P., Quinn, T. J. (2020) Cardiovascular risk factors indirectly affect acute post-stroke cognition through stroke severity and prior cognitive impairment: a moderated mediation analysis. Alzheimers Res Ther, 16;12(1):85.

- Quinn, T. J., **Elliott, E.**, Hietamies, T. M., Martínez, G., Tieges, Z., Mc Ardle, R. (2012) Diagnostic test accuracy of remote, multidomain cognitive assessment (telephone and video call) for dementia. Cochrane Database of Systematic Reviews, Issue 9. Art. No.: CD013724.

- Calil, V., **Elliott, E.,** Borelli, W. V., Barbosa, B. J. A. P., Bram, J., Silva, F. de O., Cardoso, L. G. M., Mariano, L. I., Dias, N., Hornberger, M., & Caramelli, P. (2020). Challenges in the diagnosis of dementia: insights from the United Kingdom-Brazil Dementia Workshop. Dementia & Neuropsychologia, 14(3), 201-208.

Conference Oral presentations

- Prognostic ability of the Oxford Cognitive Screen (OCS) post-stroke. Presented at The British Neuropsychological Society (BNS) Autumn Meeting, 29th November 2019 (National Hospital for Neurology and Neurosurgery, London).

- Associations between results from the Oxford Cognitive Screen (OCS) and later functional and quality of life measures. Presented at the Organisation for Psychological research into Stroke (OPSYRIS) 2019 meeting, 4th October 2019 (University of Oxford).

- Who is classified untestable on brief cognitive screens in an acute stroke setting? Presented at the British Psychological Society Scottish brand Postgraduate research day, 27th August 2019.

- Cognitive screening tests: feasibility, accuracy and prognostic ability. Alzheimer's Research UK (ARUK) Scotland Network Centre Annual Meeting, 22nd August 2019

- Real-world feasibility of brief cognitive screens on a hyper-acute stroke unit: an exploration of untestable patients. Presented at the British Neuropsychological Society (BNS) Spring Meeting, 26th April 2019 (National Hospital for Neurology and Neurosurgery, London)

Poster presentations

- UK Stroke Forum (UKSF) 2017 (Liverpool), 2018 (Telford)

- European Stroke Conference (ESOC) 2019 (Milan)

- Scottish Dementia Research Consortium (SDRC) 2019 (Glasgow)

- British Neuropsychological Society (BNS) Spring meeting 2019 (London)

- ARUK Scotland Network Centre annual meeting 2019 (St Andrews)

- Glasgow Royal Infirmary (GRI) research day 2018, 2019 (Glasgow)

Awards

- I was awarded the Humphreys & Riddoch prize by the BNS at the Autumn 2019 meeting where I presented my work on Chapter 7.

- I won the University of Glasgow three-minute thesis (3MT) competition in 2019, where I summarised my thesis with a talk entitled 'Draw me a clock…'.

- I was runner up in the Universitas 21 network 3MT competition.

Definitions/Abbreviations

ACE: Addenbrooke's cognitive examination (-R denotes Revised version)

AD: Alzheimer's Disease

ADDTC: Alzheimer's Disease Diagnostic and Treatment Centers

ADL: Activities of Daily Living

AHA-ASA: American Heart Association-American Stroke Association

AMT: Abbreviated Mental Test

ANCOVA: Analysis of covariance

APPLE: Assessing post-stroke psychology longitudinal evaluation

AUROC (or AUC): Area under receiver operating characteristic curve

BCoS: Birmingham cognitive screen

BI: Barthel index

BMET: Brief memory and executive test

BNIS: Barrow Neurological Institute Screen for higher cerebral functions

BNS: Brief Neuropsychological Screening

CA: Conference abstract

CAMCOG: Cambridge Cognition examination

CAM-ICU: Confusion Assessment Method for the Intensive Care Unit

CASP: Cognitive Assessment scale for Stroke Patients

CAT: Computerised adaptive testing

CDT: Clock-drawing test

CG: Claire Green

CI: Confidence Interval

ClinRO: Clinician-reported outcome

COA: Clinical outcome assessment

Cognitive-FIM: Cognitive-Functional Independence Measure

CoMet: Cognitive screening Method for stroke patients

CRF: Case report form

CSI: Cognitive screening instrument

CTT: Classical Test Theory

DOC: Depression, Obstructive sleep apnoea and Cognitive impairment

DOR: Diagnostic odds ratio

DSM: Diagnostic and Statistical Manual (-V denotes version number)

DTA: Diagnostic test accuracy

EM-MoCA: Esclerose múltipla (Multiple sclerosis in Portuguese) Montreal Cognitive Assessment

ESO: European Stroke Organisation

FAI: Frenchay Activity Index

FIM: Functional Independence Measure

FN: False negative

FP: False positive

GCS: Glasgow Coma Scale

GP-cog: General Practitioner assessment of cognition

HASU: Hyper acute stroke unit

HCP: Healthcare professional

HF: Heart Failure

HSROC: Hierarchical summary receiver operating characteristic

IADL: Instrumental activities of daily living

ICH: Intracerebral haemorrhage

IMCT: Information memory concentration test

IQCODE: Informant questionnaire on Cognitive Decline in the Elderly

IQR: Inter-quartile range

IRT: Item Response Theory

IS: Ischaemic stroke

LACS: Lacunar syndrome

MA: Myzoon Ali

MCAS: Minnesota cognitive acuity screen

MCI: Mild Cognitive Impairment

MEPS: Mental performance in acute stroke

MMSE: Mini-Mental State Examination

MoCA: Montreal Cognitive Assessment

MoCA-Blind: Version of MoCA with visual items removed

mRS: Modified Rankin Scale

MS: Multiple sclerosis

MSQ: Mental status questionnaire

MuSCoW: Must Should Could Won't

MVCI: Mild Vascular Cognitive Impairment

NART: National Adult Reading Test

NHS: National Health Service (UK)

NIA-AA: National Institute on Aging and Alzheimer's Association

NICE: National Institute for Health and Care Excellence

NIHSS: National Institute of Health Stroke Scale

NINDS-AIREN: Neurological disorders and Stroke and Association Internationale pour la Recherché et l'Enseignement en Neurosciences

NINDS-CSN: National Institute of Neurological Disorders and Stroke and the Canadian Stroke Network

NPB: Neuropsychological battery

NPEC: Northwick Park Examination of Cognition

NPV: Negative predictive value

OCSP: Bamford/Oxfordshire community stroke project classification system

ObsRO: Observer-reported outcome

OCS: Oxford Cognitive Screen

OT: Occupational therapist

PACS: Partial Anterior Circulation Stroke

PCA: Principal Component Analysis

PD: Parkinson's disease

PerfO: Performance-rated outcome

PIS: Patient information sheet

POCS: Posterior Circulation Stroke

PPV:Ppositive predictive value

PRECiS: Patient-reported evaluation of cognitive state

PRISMA-DTA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies

PRO: Patient-reported outcome

QUADAS: Quality Assessment for Diagnostic Test Accuracy Studies

RBANS: Repeatable Battery for the Assessment of Neuropsychological Status

REC: Research ethics committee

RMT: Rasch measurement theory

ROC: Receiver operating characteristic

R-CAMCOG: Rotterdam Cambridge Cognition

SF-MoCA: Short forms of the MoCA

SPMSQ: Short Portable Mental Status Questionnaire

SPSS: Statistical Package for Social Sciences

SSNAP: Sentinel stroke national audit programme

STARDdem: Standards for Reporting Diagnostic Test accuracy studies in dementia

STIDA: Structured Telephone Interview for Dementia Assessment

STROND: Standards of Reporting of Neurological Disorders

SVD: Small vessel disease

TAB: Telephone assessment battery

TACS: Total Anterior Circulation Stroke

TCAB: telephone cognitive assessment battery

TIA: Transient ischaemic attack

TICS: Telephone Interview for Cognitive Status

TICS-m: Telephone Interview for Cognitive Status modified

T-MoCA: Telephone version of Montreal Cognitive Assessment

TN: True negative

TP: True positive

TRIACOG: Cognitive Screening Instrument for evaluating poststroke adults

T3MS: Telephone version of MMSE

VaD: Vascular dementia

VCI: Vascular Cognitive Impairment

VISTA: Virtual International Stroke Trial Archive

WAIS-III: Wechsler Adult Intelligence Scale (Version III)

4AT: 4-A's Test for rapid assessment of delirium

6-CIT: Six item cognitive impairment test

1 Introduction

1.1 What is cognition?

Throughout this thesis, cognition refers to the range of mental processes we use to acquire, process, understand, store, and retrieve information (1). These processes are used to make sense of the world around us. Cognition is therefore not a unitary concept; under this umbrella term are various domains or functions, yet even referring to these domains alone is reductionist. Domains are multifaceted, comprising numerous sub-domains, and many functions are not independent of one another. This makes studying cognition complex, and indeed many researchers spend their entire career researching just one of these domains.

There is no consensus regarding the classification of cognitive domains, although the domains of memory, language, attention, executive functioning, visuospatial processing, and processing speed are frequently differentiated. Cognitive abilities are also considered hierarchical in nature, for example executive functions are often referred to as higher order cognitive functions as they are more complex than the more basic perceptual abilities (2). Executive functioning abilities also exert control over more basic processes.

One of the main diagnostic guidelines used for cognitive disorders is the Diagnostic and statistical manual (DSM). The 5th version (DSM-5) (3) refers to six cognitive domains that can be affected in neurocognitive or neurodevelopmental disorders(4) (Figure 1-1).

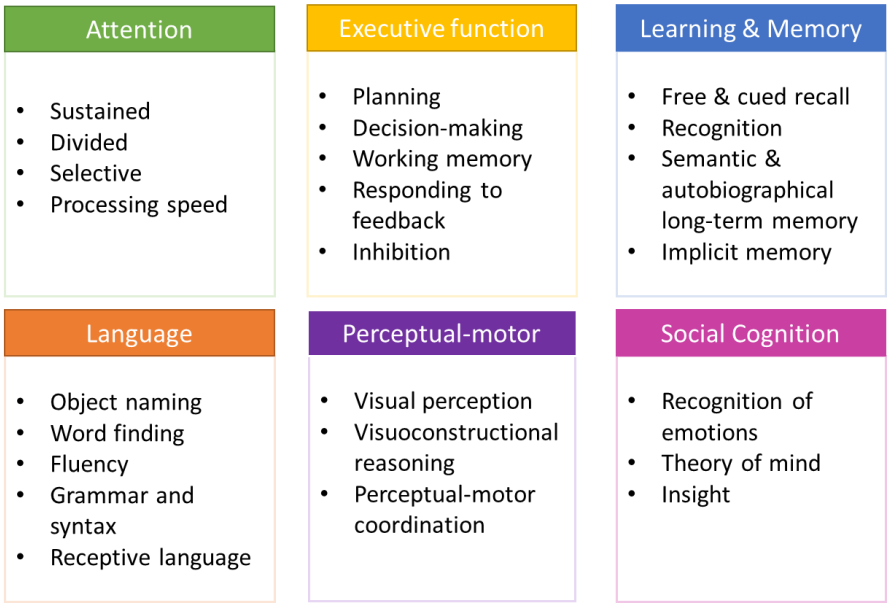


Figure 1-1 DSM-5 cognitive domains

1.2 Post-stroke cognitive impairment

Stroke is a risk factor for development of a cognitive disorder. Within the general population in the UK, over 100,000 people are diagnosed with a stroke (5) and 46,000 diagnosed with a transient ischaemic attack (TIA) each year, with an estimate of 1.2 million stroke survivors.

The prevalence of cognitive impairment post-stroke depends on the measurement and criteria used and the timing of assessment. In the acute period after a stroke, up to 70% of patients exhibit cognitive difficulties (6-8), whereas in the year following a stroke, the prevalence is roughly 40% (9). Patients with cognitive impairment after a stroke are also at risk of developing dementia (10), with approximately 10% of patients diagnosed after their first stroke, rising to 30% after multiple strokes (11). The negative consequences of cognitive impairment are well established, including increased length of hospital stay (12), functional impairment (13), lower quality of life (14) and increased risk of post-stroke depression (15).

There is a lack of consistent nomenclature describing post-stroke cognitive impairment in the literature, making it challenging and confusing for researchers to make comparisons across studies. This is due to varying severities in

impairment, different criteria being used and poor concordance between different diagnostic guidelines (16).

The nature of diagnosis is to think of conditions in binary terms; they are either present or absent, yet the distinction may not be clear cut in practice. If cognitive abilities and cognitive impairment exist along a spectrum, how we define impairment depends on normative data and understanding the process of normal cognitive aging (17), which we are continually learning more about. A binary outcome of impairment as present or absent allows for ease of analysis in research studies but lacks granularity and may mask group differences. Other approaches include creating hierarchical categories, assessment as a continuous scale, and assessing against population normative data.

Viewing cognitive impairment through a dementia paradigm alone is erroneous and one must differentiate between different syndromes, yet at the same time understand the relationship and potential overlap between them. These syndromes include dementia, mild cognitive impairment, delirium, and other non-degenerative types of post-stroke cognitive impairment, which will be discussed below. Within each of these syndromes, there are a variety of classifications. Throughout this thesis I will specify the syndrome of interest. Delirium is not the focus of any chapters, however since some chapters relate to cognitive screening on an acute medical unit, it should be taken into account when interpreting the results.

Vascular cognitive impairment (VCI) is a wide-reaching term encompassing all vascular contribution (not just stroke) to cognitive problems. VCI is also not specific to severity and covers a spectrum of syndromes. Despite terminology, it also includes mixed pathology dementia (18, 19).

1.2.1 Cognitive impairment, no dementia

The majority of post-stroke cognitive impairment is distinct from dementia in that it improves over time and is therefore not neurodegenerative. Since stroke itself is heterogenous in nature (different types, aetiologies, extent of damage), the cognitive profile of post-stroke cognitive impairment is unsurprisingly diverse. However, deficits in executive function and visual perception are more

frequent and characteristic in this group of patients in comparison to those with Alzheimer's Disease (AD), where memory impairments are more central (20-22).

Examining subgroups of stroke could provide some patterns, for example the profile of cognitive impairment associated with small vessel disease (SVD) includes executive dysfunction and reduced processing speed (23), whereas focal cognitive impairments such as aphasia and visuospatial neglect were found to be more common following a cardioembolic stroke, in comparison to large vessel or small vessel disease (24). Improvements or recovery from focal impairments can occur spontaneously, as seen in apraxia (25) and visuospatial neglect (26), or may improve as a result of rehabilitation/intervention (27), although evidence for effectiveness of interventions is lacking.

Mild cognitive impairment (MCI), referred to as minor neurocognitive disorder in the DSM-5, is a term used largely in non-stroke settings, yet has been adopted in some stroke studies, since it is a diagnostic term. In DSM-5 it is characterised by evidence of modest cognitive impairment that does not interfere with activities of daily living (ADL) (3, 28). Subtypes of MCI are sometimes described, including amnestic and non-amnestic (depending on whether memory is impaired) (19) and single or multi-domain MCI (depending on number of domains impaired). The syndrome of MCI is still considered a controversial and uninformative diagnostic label since the prognosis is unknown; some may return to their normal cognitive baseline (29), some may stay the same (30), whereas for others it is a transitional phase before progression to dementia (31).

1.2.2 Dementia

While dementia subtypes based on aetiology are still used, it is now well acknowledged that dementia is often not caused by a single pathology. Vascular pathology is found in more than half of diagnosed dementia cases, with many clinical diagnoses of AD confirmed as mixed or vascular on autopsy (32, 33).

The main distinction between MCI and dementia is based on the degree of functional deficit - in dementia, impairment affects a patient's independence to carry out daily tasks. Amongst the various diagnostic criteria for vascular dementia (VaD), the National Institute of Neurological Disorders and Stroke and

Association Internationale pour la Recherché et l'Enseignement en Neurosciences (NINDS-AIREN) requires memory to be impaired plus impairment in two other cognitive domains. However, in the American Heart Association-American Stroke Association (AHA-ASA) diagnostic criteria, a minimum of two domains must be impaired without the requirement of memory impairment. The DSM-5 includes a number of changes from its previous version; the term major neurocognitive disorder has replaced dementia, and memory impairment is no longer required for diagnosis. The four criteria used in DSM-5 are summarised in Figure 1-2. If these criteria are met, the subtype is then made based on aetiology. Other criteria also exist, for example those of the State of California Alzheimer's Disease Diagnostic and Treatment Centers (ADDTC) (34).

Major Neurocognitive disorder:

A. Evidence of significant cognitive decline from a previous level of performance in one or more cognitive domains (complex attention, executive function, learning and memory, language, perceptual-motor, or social cognition) based on:
   1. Concern of the individual, a knowledgeable informant, or the clinician that there has been a significant decline in cognitive function; and
   2. A substantial impairment in cognitive performance, preferably documented by standardized neuropsychological testing or, in its absence, another quantified clinical assessment.
B. The cognitive deficits interfere with independence in everyday activities (i.e., at a minimum, requiring assistance with complex instrumental activities of daily living such as paying bills or managing medications).
C. The cognitive deficits do not occur exclusively in the context of a delirium.
D. The cognitive deficits are not better explained by another mental disorder (e.g., major depressive disorder, schizophrenia).

Figure 1-2 DSM-V Major cognitive disorder criteria (2)

It is easy to mis-label patients with cognitive problems after stroke as having dementia, since stroke commonly causes physical impairment affecting independence in activities of daily living. It is therefore recommended that a diagnosis of dementia should not be given until at least 6 months after stroke onset (18).

1.2.3 Delirium

Delirium is an acute, transient, state of confusion and a common complication of acute illness, affecting around 25% of patients after a stroke within the first 6 weeks (35). Consideration and screening to detect delirium is therefore important in the acute stroke setting. Identifying delirium has implications for

both the short and longer term. Incident delirium can signal the emergence of a stroke-related complication, such as pneumonia, and in the longer term the presence of delirium is associated with poor outcomes.

1.3 Risk factors/mechanisms of post-stroke cognitive impairment

There are numerous potential mechanisms and risk factors involved in cognitive impairment post-stroke. Since this topic is not the focus of this thesis, only a brief overview will be provided. It is discussed in greater detail elsewhere (36-38).

Size and location of lesions impacts on the presence and type of cognitive impairment (36, 39, 40). Cerebral small vessel disease (SVD) is a common cause of cognitive impairment and dementia (41). SVD refers to lacunes, white matter hyperintensities, cerebral microbleeds, brain atrophy and enlarged perivascular spaces (42). SVD and AD pathology can also co-exist and overlap (38).

Risk factors for dementia include those that are not modifiable (age, sex) and those that are potentially modifiable: fewer years of education, hearing loss, traumatic brain injury, hypertension, alcohol, obesity, smoking, depression, social isolation, physical inactivity, air pollution and diabetes (43).

The concept of cognitive reserve is used to illustrate how protective factors may moderate the relationship between brain pathology and clinical outcomes (44), for example having a greater number of years in education (45), social support (46) and regular exercise (47). Years in education is commonly used as a proxy of cognitive reserve. A recent study found that when controlling for age and relative lesion size, the number of years in education predicted performance in alertness, which is considered to be ‘education independent’ as well as other tasks considered ‘education dependent’ (48). Years in education however did not predict the presence of spatial neglect. It is recommended that proxies of cognitive reserve are considered even as early as the acute stroke phase.

1.4 Clinical outcome assessments (COAs)

There are four categories of clinical outcome assessments (COAs) used to measure patient outcomes and experience, based on how they are completed (Figure 1-3). Patient-reported outcomes (PROs) are measures completed by patients (self-report) to understand their experience of symptoms and the impact these have on their lives. These are commonly used to measure symptoms such as fatigue or pain. Clinician-reported outcomes (ClinROs) are measures carried out by a healthcare professional (HCP) or trained researcher. Most ClinROs involve some clinical judgement or interpretation of signs, symptoms, or behaviours. Performance outcomes (PerfOs) involve the patient carrying out a task, according to instructions that are administered by a healthcare professional/researcher. Observer-reported outcomes (ObsROs) are captured by someone who observes the patient in everyday life, for example a family member.

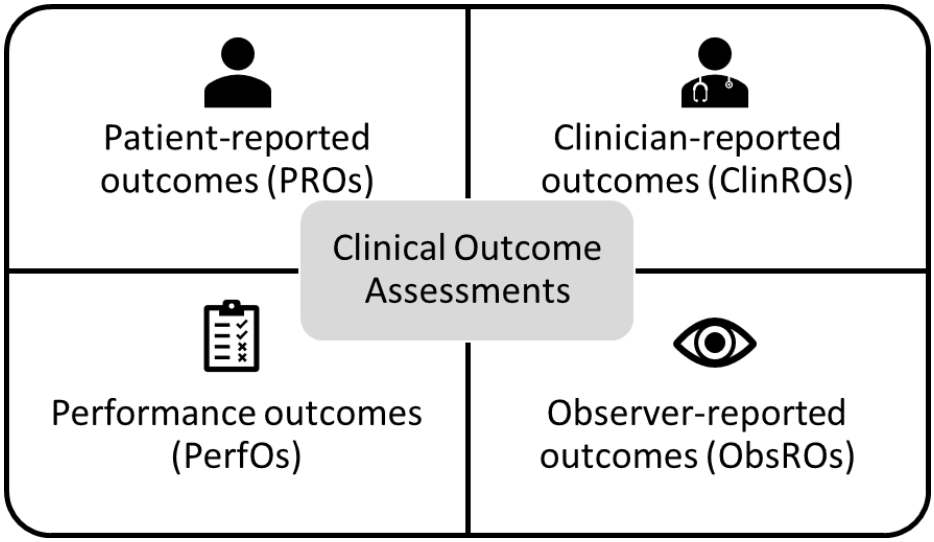


Figure 1-3 Types of Clinical Outcome Assessments

All four types of COAs will be used in this thesis. Cognition is the main focus and is measured using PerfOs, rather than any PRO scales that provide a subjective measure of cognitive decline, e.g. the cognitive change index (49) and the patient-reported evaluation of cognitive state (PRECiS) (50). PROs are used to measure quality of life, depression, and activities of daily living. ClinROs are used to measure disability and function and ObsROs are used for a family member's perspective of the patient's independence/function.

1.5 Cognitive screening instruments (CSIs)

Basic orientation questions are frequently used in clinical practice to determine a patient's level of consciousness (e.g. in the Glasgow Coma Scale (GCS) (51)). The GCS however is not considered a cognitive screening instrument (CSI). A CSI is a test comprising any number of tasks that aim to determine cognitive impairment in a patient. Occasionally items are taken out of existing instruments to create a CSI, for example the Cog-4 comprises the four cognitive items of the National Institute of Health Stroke Scale (NIHSS), which is routinely undertaken to measure stroke-related neurological deficits.

A distinction is made between CSIs and a full neuropsychological battery (NPB), which comprises numerous tests, requires expert interpretation and is considered the gold standard for detection of cognitive impairment. Most CSIs have one overall score and use a cut-off or threshold score to categorise patients into two groups (impaired or spared); they are referred to as global screening tests. Swartz et al. however suggest using an alternative method of using two cut-off scores to stratify patients into three groups (low, intermediate or high risk of cognitive impairment) (52). There are also a minority of CSIs that provide cut-off scores across different subtests/domains (domain-specific tests), with the attempt to create a halfway house between a CSI and NPB and to address the criticism of reducing cognition down to one single score.

It should be clear that there is no perfect CSI. Each CSI has its strengths and weaknesses and there cannot be one recommendation to fit all scenarios, nor even one recommendation for post-stroke. The purpose for administering the CSI will be a central theme running throughout this thesis, and it is this purpose that will drive the choice of CSI. As explained in an earlier section, there are a range of syndromes that may be screened for and different CSIs will be suited for each of these. Hospital resources and intended next steps for follow-up also play into decision-making. Some generic criteria can be set for CSI use on an acute stroke unit (Table 1-1).

Table 1-1 MuSCoW chart (Must Should Could Won't) detailing preferred properties for use of a CSI on an acute stroke unit

Must have	Should have
- Administration <20 mins - Appropriate test accuracy (for syndrome of interest)	- Aphasia friendly questions - Stroke normative values - Free training materials - Guidance on item non-completion
Could have	**Won't have**
- Translations - Adaptable to telephone or video calls - Parallel versions	- Diagnostic purpose - Require specialist training - Copyright issues

The current paradigm of cognitive assessment in stroke in the UK is a two-step system; patients scoring low on a CSI are triaged and often referred on to the neuropsychology department for more detailed, comprehensive testing (where staffing and resources allow). Therefore, for this system to work, good test accuracy of the CSI is vital so patients with impairment are not missed and those without any impairment are not subjected to unnecessary testing. There are two broad approaches to cognitive screening: universal (unselected) screening of all patients admitted to the stroke unit, or targeted assessment of a smaller group of people where there is concern about cognitive problems. The two approaches are not mutually exclusive, and a patient who passes a screening test but complains of cognitive issues should not be denied a more detailed assessment. Although we know that many institutions follow this 2-step paradigm, we also know that there is variability in terms of which patients are screened, the CSI used and the timing of administration.

In the UK, cognitive screening in acute stroke is generally carried out by occupational therapists (OTs). It is rare that psychologists are available in acute settings, and patients usually will have to wait for an outpatient appointment to access a psychologist. We also know that evidence does not always drive CSI choice. A qualitative study completed with OTs found variation in CSI choice in the community stroke setting and noted that the choice of tool was primarily

based on availability and familiarity rather than any psychometric properties, and the interpretation is often subjective (53).

Decision-making regarding the time-point most meaningful for screening (days/weeks/months after stroke event) is subject to debate and once again is dependent on the purpose of screening. Early identification of cognitive impairment is argued to have benefits, for example it should allow for targeted/personalised rehabilitation, appropriate goal setting and aid discharge planning (8). Informing and educating patients (and their families) about any cognitive issues is important so they understand how such problems could potentially affect them in their day to day lives. Finally, there is the argument that screening patients early on during their hospital admission is often the only opportunity to do so; patients can be discharged home quickly. Six-month reviews have been commissioned in England, Wales, and Northern Ireland but completion is far from universal. Those who argue against early cognitive screening cite the limited understanding of the natural history of post-stroke cognitive impairment, the variability of cognitive test performance during the acute period (18) (which could be due to the presence of delirium) and the lack of evidence regarding effectiveness of cognitive rehabilitation or treatments (54, 55).

There are many different CSIs to choose from, each varying in length, difficulty, and coverage of different cognitive domains. The Montreal Cognitive Assessment (MoCA) (56) and Mini-Mental State Examination (MMSE) (57) are often cited in stroke research (58, 59) as the two most commonly used CSIs, however the landscape is continually changing and since a charge was introduced with the MMSE, many clinicians have abandoned it, with the same expected to follow for the MoCA (60). Frequently, CSIs like the MoCA and MMSE that were developed for AD or MCI populations are used in the stroke setting. Any test used in a new context of use needs to be re-validated. The same is also true if any amendments are made to a CSI. CSIs that have not been developed with stroke patients in mind have a number of limitations that should be considered. They are often heavily reliant on language skills, which disadvantages aphasic patients, and there is often no assessment of neglect or apraxia. They were designed for outpatient settings and therefore may not be appropriate to

administer at the bedside on a busy acute medical unit, nor with patients with physical or speech impairments. Therefore, there is concern if a CSI used has not been validated in stroke patients.

CSIs are sometimes categorised according to administration time. The shortest CSIs, referred to as brief CSIs throughout this thesis, are most suited to acute medical settings since they can be administered in ≤5 minutes. However, a limitation is that they mainly focus on orientation and memory. Examples include the Clock-drawing test, Abbreviated mental test (AMT) (61), Six item cognitive impairment test (6-CIT) (62), Mini-cog (63) and GP-cog (64). CSIs that cover more cognitive domains include the MoCA(56), Addenbrooke's cognitive examination (ACE) (65), MMSE (57) and the Oxford Cognitive Screen (OCS) (66), but their administration time is longer. When choosing CSIs in both clinical practice and research, it is important to consider the trade-off between what is acceptable and feasible for both stroke patients and staff in these settings and what will provide the most meaningful data (for example, accurate, reliable, and prognostic).

Standardisation of administration/scoring of psychometric tests is essential to interpret them accurately. Therefore, any clinician or researcher administering a CSI should read the manual that details administration/scoring instructions for each test. Additional in-person training may also be necessary for more complex tests. The need for training should not be underestimated as errors are commonly made. A study carried out with 104 psychiatrists and geriatricians found inconsistencies with their use of the Abbreviated Mental Test (AMT) (67). Specifically, only 22% of doctors used the original 10 items, while the others omitted items (often orientation and memory). When asked to score the test for a fictional patient, only 16% arrived at the correct score. This is concerning since on a short test like this, a single point difference can result in a different categorisation of impaired/not impaired. It is impossible to know the past or present extent of incorrect use of a test, either clinically or in research, and the resulting impact.

Finally, it is important that caution is taken when interpreting scores from CSIs and scores should not be used in isolation. Despite CSIs not being diagnostic, low scores are often quoted to imply impairment (52) and are used as endpoints in

clinical trials. In comparison to detailed neuropsychological testing, CSIs have been found to underestimate cognitive impairment (68) and a single score provides a limited summary since it reflects a range of performance (52). Putting scores in context is essential as CSIs are affected by a variety of factors: age, education, socioeconomic status, culture and situational influences (emotional status, stress, fatigue, medication use) (28, 69). After receiving a stroke diagnosis, all these situational influences are likely to be present. CSIs also do not address the effect on function (52), nor do they take into account baseline cognitive status which is useful for comparison.

1.5.1 The Montreal Cognitive Assessment (MoCA)

The MoCA (56) was published in 2005 and developed with the purpose to identify MCI in community dwelling participants. Since then it has been utilised across a range of disease areas, such as multiple sclerosis (MS) (70), Parkinson's disease (PD) (71), cancer (72), Huntington's disease (73) and stroke (74).

The total score is out of 30, with <26 being the recommended threshold score. A lower threshold score is sometimes employed post-stroke, but there is not a consensus on which score should be used (52, 74, 75). The following items are included: trail making test, cube copy, clock-draw, animal naming, immediate recall (unscored), digit span forwards and backwards, tap each time letter A is heard, serial 7 subtraction from 100, repeating two sentences, verbal fluency, abstraction, delayed recall, and orientation. An extra point is added to the patients score when they have 12 or fewer years in education.

Strengths of the MoCA include evidence of content validity in Vascular Cognitive Impairment (VCI) (76), its wide validation with stroke patients, three alternative versions to allow for practice effects, and its translation into 65 languages (77). It is also preferred over other CSIs, such as the MMSE, because it is more sensitive to milder forms of cognitive impairment (76, 78, 79).

However, as the MoCA was not designed for an acute stroke setting, it has limitations. It is largely a language-dependent test requiring verbal output, so all sections apart from the first (visuospatial/executive) may be compromised in a patient with aphasia or if it is not conducted in the patient's native language.

The layout of the test is also not suited to patients with visual neglect as the items are not centralised. Finally, the lack of consensus regarding the cut-off score means there is uncertainty with interpretation. There is also some evidence that the MoCA has poorer sensitivity to right hemisphere deficits (80).

1.5.2 Mini-Mental State Examination (MMSE)

The MMSE (57) was published in 1975 and developed with the purpose to identify dementia. As previously discussed, it is one of the most commonly used CSIs and has been used across a wide range of disease areas including stroke (81), PD (81), and MS (82).

The total score is out of 30, with <24 being the recommended threshold score. The following tasks are included: current year, season, date, day, month, location (including state, county, town/city, hospital, floor), immediate recall of 3 unrelated words, serial 7 subtraction from 100 or spell 'World' backwards, delayed recall of 3 words, name 2 objects, repeat a phrase, follow a verbal instruction to fold a piece of paper in half with right hand and place it on the floor, follow a written instruction to close eyes, write a sentence and copy a figure of intersecting pentagons.

In the stroke context, the MMSE has several limitations including low sensitivity to milder forms of cognitive impairment, no optimum threshold score (83) and, like the MoCA, many tasks require verbal output.

1.5.3 Stroke specific CSIs

Some CSIs have been developed specifically for stroke patients (Table 1-2). Theoretically these CSIs should be superior to generic MCI/dementia screens in a number of ways, including inclusivity and sensitivity to stroke-related impairments. In terms of inclusivity, many of these tests include vertical, centralised layouts to account for neglect and options to answer non-verbally, to account for aphasia.

A systematic review published in 2019 found seven CSIs designed for stroke (59). However, one test (the Birmingham cognitive screen (BCoS)) takes roughly an hour to complete, which is too long for bedside screening and will not be

discussed further. From a search of the literature since this review was published, I found two additional stroke-specific CSIs. Many of these tests are unknown to those working in stroke, because the initial development paper is often the only available publication (e.g. Cognitive Screening Method for Stroke Patients (CoMet) (84), Brief Neuropsychological Screening (BNS) (85), Northwick Park Examination of Cognition (NPEC) (86)). While these screens have the potential to be useful measures, the lack of evidence means we know very little about their psychometric properties and therefore at present we cannot compare them and make recommendations.

Table 1-2 CSIs Designed for Stroke

Test Name	Designed or validated in acute settings (within roughly a week post-stroke)?	Administration time	Evidence level
BMET	No	13 mins	2 studies
BNS	Yes	5-10 mins	1 study
CASP	Yes	10 mins	2 studies
CoMet	No	20 mins	1 study
MEPS	Yes	Not specified	1 study
MVCI	Yes	5-13 mins	1 study
NPEC	Yes	30 mins	1 study
OCS	Yes	15 mins	>10 studies
TRIACOG	Yes	20 mins	1 study

Abbreviations: BCoS, Birmingham Cognitive Screen; BMET, Brief Memory and Executive Test; BNS, Brief Neuropsychological Screening; CoMet, Cognitive Screening Method for Stroke Patients; CASP, Cognitive Assessment for Stroke Patients; MEPS, Mental Performance in acute stroke; MVCI, Mild Vascular Cognitive Impairment; NPEC, Northwick Park examination of cognition; OCS, Oxford Cognitive Screen.
Content based on the systematic review mentioned and my own literature search.

1.5.3.1 Brief Memory and Executive Test (BMET) (23)

The BMET was published in 2012. It comprises 6 tasks, with no maximum score, since one item is scored based on length of time to complete. It was designed to target cognitive impairment arising in the context of SVD, as the authors argue this pattern of impairment is distinct from that seen in large vessel stroke and cortical dementia. The items include orientation, letter-number matching, five

item repetition, sequencing (motor, letter, letter-number), five-item recall and five item recognition.

The BMET was initially validated in patients with SVD and AD (23) and, in a second UK study, in 200 patients with lacunar stroke (87) (at an average of 20.5 months post-stroke) and 303 healthy controls. In this second study the BMET was administered by doctors and research nurses rather than neuropsychologists, which is important as this is how it would be used in practice.

1.5.3.2 Brief Neuropsychological Screening (BNS) (85)

The BNS was published in 2009. It comprises 14 tasks, with a maximum score of 68. The items include lexical decision and reading, shapes discrimination, incomplete letters, calculations, face recognition, auditory comprehension, written comprehension, naming of items and pantomime of use, matching objects, gesture imitation, verbal fluency, identifying coins from buttons, word recognition and proverb comprehension.

The BNS was initially developed and validated in 134 acute stroke inpatients and 247 healthy controls in Italy. Only those patients with severe aphasia were unable to complete the test.

1.5.3.3 Cognitive Assessment scale for Stroke Patients (CASP) (88)

The CASP was published in 2014, but there is a published conference abstract from 2012. It comprises 9 tasks, with a maximum score of 36. The items include naming, comprehension, cube copy, graphic series, inhibition/flexibility, line bisection, image recall, praxis, orientation.

The CASP was initially validated in 44 stroke patients (mean 42 days post-stroke) in a rehabilitation centre in France. Verbal expression and comprehension disorders resulted in 18% of the sample having incomplete data.

1.5.3.4 Cognitive Screening Method for Stroke Patients (CoMet) (84)

The CoMet was published in 2019. It comprises 14 scored tasks and an unscored self-evaluation, with a maximum score of 147. The tasks include orientation,

writing sentences, understanding instructions, word fluency, episodic memory, drawing, delayed episodic memory, object naming, object memory, sentence repetition, object replacement, object recognition, visual finding, number arranging.

The CoMet was initially validated in 77 stroke patients (mean of 51 days post-stroke) in Finland.

1.5.3.5 Mental Performance in acute stroke (MEPS) (89)

The MEPS was published in 2020. It comprises 14 tasks with a total score of 82. The tasks include temporal orientation, spatial orientation, orders comprehension, segments discrimination, reading & comprehension of sentences, immediate visual memory, digit span, visual exploration and attention, words repetition, clock drawing test, similarity judgements, ideomotor apraxia, picture naming, ideational apraxia. It provides data on both a domain and global functioning level.

The MEPS was initially validated in 129 acute stroke patients (mean of 5.5 days post-stroke) and 263 healthy control participants in Italy.

1.5.3.6 Mild Vascular Cognitive Impairment (MVCI) (90)

The MVCI was published in 2015. It comprises 13 tasks with a maximum score of 30. The tasks include orientation (time & place), immediate recall, recent memory, prospect memory, delayed recall, repetition, verbal fluency, comprehension, counting numbers, calculation, abstraction, visuospatial ability, problem solving (situation play). Tasks were taken from other CSIs: MMSE, Cambridge Cognition (CAMCOG) examination for mental disorders of the elderly, MoCA and Telephone Interview for Cognitive Status (TICS) and chosen to also be deliverable over the telephone.

The MVCI was initially validated in 60 stroke patients in South Korea, within 3 months of stroke onset, and with no previous cognitive disability.

1.5.3.7 Northwick Park examination of cognition (NPEC) (86)

The NPEC was published in 2016. It comprises 22 tasks and has a maximum score of 100. The following 5 domains are covered: orientation (5 points), reasoning/executive function (15 points), memory (20 points), language (25 points) and perception (16 points). It includes tests to assess praxis and neglect.

The NPEC was initially validated in 166 stroke patients admitted to a hyper acute stroke unit (HASU) (mean time to assessment was 5.6 days) and 100 healthy controls in the UK. The screen was used as part of routine clinical practice and as a result no data was available regarding the number of untestable patients.

1.5.3.8 The Oxford Cognitive Screen (OCS) (66)

The OCS was published in 2015 and comprises 10 tasks. There is no single maximum score, as each task has a different threshold score to indicate impairment. The tasks include picture naming, orientation, visual field test, sentence reading, number writing, calculations, broken hearts test, meaningless gesture imitation, delayed recall/recognition, and trails. The tasks are mapped onto the corresponding cognitive domain (language, attention, memory, praxis, and number processing). The administrator can use the visual snapshot of the patient's profile for documentation of the patient's strengths/weaknesses in the medical notes.

The OCS was initially validated in 207 acute stroke patients in the UK. To date (2021) it has been translated and validated in eight languages: Italian (91), Dutch (92), Spanish (93), Russian (94), Danish (95), Putonghua (96), Cantonese (97), and Brazilian-Portuguese (98), with other cultural and language adaptations underway. There is a second version available that can be used for multiple testing with the same patient, and there are plans to develop a third parallel version.

1.5.3.9 Cognitive screening (TRIACOG) instrument (99)

The TRIACOG was published in 2020. It comprises 22 tasks, with a total score of 144. The tasks include orientation to time, episodic semantic verbal memory (immediate and delayed), praxis (clock-draw, reproduction of a figure),

ideomotor apraxia (use a fork), visual memory (reproduction of figure), attention/working memory (digit span forward and backward), executive function (verbal fluency letter V), processing speed (rapid serial naming of shapes), language (naming objects and actions), oral and written comprehension, vocabulary, phrase reading, inference processing, spelling, repletion and numerical processing.

The TRIACOG was validated in 100 stroke patients (mean of 8.3 days post-stroke) and 100 healthy controls in Brazil.

1.5.4 Telephone-based cognitive screens

There are several CSIs which have been designed for telephone delivery, but evidence of their accuracy is lacking. The convenience of remote assessment means they are often used in studies with large samples, rather than clinical use. However, in recent times, remote CSIs are sought for clinical purposes too. Examples include the telephone interview for cognitive status (TICS)(100), its modified version (TICS-m)(101) and the telephone version of the MoCA (T-MoCA) (102).

1.5.5 CSIs for delirium

Screening tools for delirium are available, and many have good accuracy when compared with gold standard clinical assessment. The 4-A's test (4AT) (available online: www.the4AT.com) is a short screening tool for delirium that is available in several languages and is quick to administer with little training; it has some supportive data in stroke (103, 104). The Confusion Assessment Method also has proven accuracy for diagnosis of delirium in stroke. For patients with aphasia or other communication problems, the Confusion Assessment Method modified for use in Intensive Care Settings (CAM-ICU) can be used because it does not require any verbal response for completion.

1.5.6 Informant-based cognitive scales

Questionnaires completed by a relative, partner or caregiver (ObsROs) can provide useful additional information to CSIs, for example to capture change in

cognition over time. When completed immediately after stroke the method can be used to understand the patient's pre-morbid cognitive status.

One example is the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) (105). It consists of 26 tasks, which aim to measure change in cognition or functional performance over a 10-year period. There is limited literature on the use of informant assessment in stroke. A recent systematic review found no studies examining the accuracy of the IQCODE for assessing pre-stroke decline; the studies instead focused on its use in post-stroke assessment (106).

There are, however, limitations to informant questionnaires. The availability of an informant who is willing or able to comment on the patient's pre-stroke state is not guaranteed. If the tool is not used early after the stroke event, then recall bias can be an issue; informants may struggle to give an account of pre-stroke cognition and often describe the cognitive problems that they see after the stroke.

1.6 Neuropsychological batteries (NPBs)

Formal neuropsychological assessment is considered the gold standard for detection of cognitive impairment (58). Guidelines recommend that patients undergo this detailed testing if impairment is detected at the screening phase (107, 108). A NPB is not usually administered until much later in the stroke pathway (at least three months is recommended once the stroke has stabilised (6)), yet once again each institution varies with its approach.

Batteries are made up of individual tasks to examine each cognitive domain in depth. As a result, assessment can last several hours, and is therefore considered impractical for inpatients. There is not a preferred test battery to use in stroke and often neuropsychologists will choose tests tailored to each patient. Even within each domain of cognition, there is no consensus on which tasks to use. The National Adult Reading Test (NART) is often used as a proxy of premorbid intelligence, as vocabulary is said to be better preserved in neurodegenerative conditions compared to other cognitive abilities (109).

1.7 NINDS-CSN VCI protocols

In response to the lack of standardisation and to help detect VCI, a working group of the National Institute of Neurological Disorders and Stroke and the Canadian Stroke Network (NINDS-CSN) met in 2006 to produce clinical and research guidelines for vascular cognitive impairment. The neuropsychological working group produced three protocols of different lengths (5, 30 and 60 minute protocols) (110).

The publication states that the working group referred to the following test criteria to make recommendations: quality of the standardised sample, psychometric qualities, portability, cost, ease of use, domain specificity, availability of multiple forms, cross-cultural capability, lack of ceiling/floor effects and previous use in VCI samples.

The 5-minute protocol consists of items taken from the MoCA: immediate and delayed recall, orientation, and fluency. The 30-minute protocol consists of the animal naming test, verbal fluency (letters F, A, S), Wechsler Adult Intelligence Scale Version III (WAIS-III) digit-symbol coding, and Hopkin's verbal learning test, with the trail making test A and B listed as a supplemental test. The 60-minute protocol includes all the 30-minute protocol tests, with the addition of the Boston naming test and the Rey-Osterreith complex figure.

1.8 Clinical Guidelines and other recommendations

Clinical guidelines regarding cognitive assessment in stroke all recognise the importance of screening and detection but often provide quite general recommendations (Table 1-3). In the UK, the National Clinical Guidelines for Stroke (111) recommend routine cognitive screening to be undertaken using standardised measures but do not specify which particular screening tool to use, nor the time-point that screening will be most useful or informative. The UK NHS Improvement guidance is more specific with recommendations (timeframe for assessment within 6 weeks) and offers a cognitive assessment pathway where specific cognitive screening tests and timings are recommended (108). While this provides a useful guide for healthcare professionals, no empirical evidence is cited justifying the specific recommendations.

Details regarding cognitive screening are not mentioned in the latest European Stroke Organisation (ESO) guidelines (112). They state that there is insufficient evidence to support cognitive rehabilitation, therefore assessment is considered "desirable" since there are no proven treatments.

The AHA-ASA guideline (107) recommends routine assessment of cognition. The neurobehavioural cognitive status examination (now known as the 'cognistat') is the only screening tool specifically mentioned. The pathway detailed in the guideline mirrors the UK's approach; following detection of cognitive impairment with a screening tool, formal neuropsychological assessment is recommended. The guideline states that assessment should be sensitive to a wide range of abilities, acknowledging that screening tools inadequately assess higher-level cognitive functions.

The Stroke Foundation Australian guidelines (113) recommend for all patients to be screened after a stroke, but do not suggest which screening tests to use. Finally, the guideline from the Canadian stroke strategy (2019) (114) states that all patients with stroke or TIA should be considered at risk for VCI and considered for screening prior to discharge from acute care, using a validated CSI such as the MoCA. The recommendations provide a summary of the MoCA and other suggested screening tests, along with evidence of their psychometric properties. Other CSIs considered include the NINDS-CSN VCI protocols, the Cognitive-Functional Independence Measure (Cognitive-FIM), CAMCOG, the 'Depression, Obstructive sleep apnoea and Cognitive impairment' (DOC) screen, Frontal Assessment Battery, MMSE, and Repeatable Battery for the Assessment of Neuropsychological Status (RBANS).

Table 1-3 Clinical guidelines

Guideline	Statement regarding cognitive screening
UK National Clinical Guideline for Stroke 2016	"Routine screening should be undertaken… using standardised measures"
UK NHS Improvement 2011	Weeks 1-3: MoCA or ACE-R. Week 4: RBANS or Ravens coloured matrices if patient aphasic
European Stroke Organisation (ESO) 2008	"Assessment for cognitive deficits appears desirable"
American Heart Association-American Stroke Association (AHA-ASA) 2016	"Screening for cognitive deficits is recommended for all stroke patients before discharge home"
The Canadian Stroke Strategy 2019	"Screening for vascular cognitive impairment should be conducted using a validated screening tool, such as the Montreal Cognitive Assessment [Evidence level B]"
The Stroke Foundation (Australia) 2017	"All stroke survivors should be screened for cognitive and perceptual deficits by a trained person…using validated and reliable screening tools, ideally prior to discharge from hospital"

Abbreviations: ACE-R, Addenbrooke's cognitive evaluation revised; MoCA, Montreal Cognitive Assessment; NHS, National Health Service (UK); RBANS, Repeatable Battery for the Assessment of Neuropsychological Status.

1.9 Psychometrics and Clinimetrics

Psychometrics refers to the theory of psychological measurement. The term clinimetrics is sometimes used with regards to clinical measurements in medicine.

1.9.1 Classical test theory (CTT)

CTT is considered the first theory of measurement and sometimes referred to as the true score model. This approach relies on a small set of assumptions. Each person has an innate true score, summarised as X = T + E. X is an observed score, T is the true score, E is random error. The errors are assumed to be normally distributed with a mean of 0. CTT is widely used and considered an essential part of the development process of a scale.

CTT is a different approach to modern test theory, which includes item response theory and Rasch measurement theory (RMT). RMT conceptualises measurement scales like a ruler and, while this thesis does not employ RMT methodology, considering cognitive abilities and measurements across a ruler is helpful (Figure 1-4). Within cognitive tests, items will span a range of difficulty levels, and will therefore target different severities of impairment.

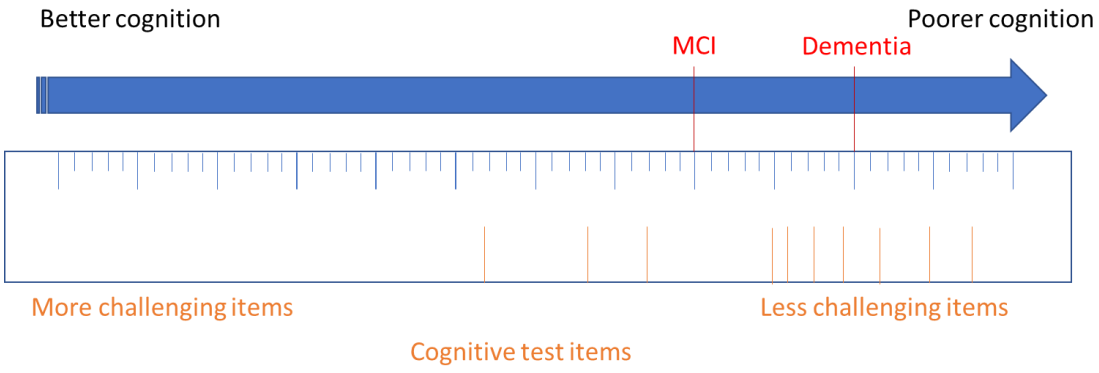


Figure 1-4 Spectrum of cognitive impairment and relation to measurement.

This schematic illustrates the spectrum of cognitive impairment (lines indicating MCI and dementia are arbitrary and for illustrative purposes only). This illustrates that different CSIs (and different CSI items) have varying levels of difficulty and will therefore target different ends of the spectrum.

1.9.2 Reliability

Reliability refers to the consistency of a measure. It is not a fixed property of a test; it is affected by aspects such as the rater, environment, administration method and sample (115). This is important to bear in mind with CSIs, where those administering the screen may have varying experience and the ward setting may compromise the reliability. There are different types of reliability (Table 1-4), however inter-rater and test-retest will not be covered in this thesis, since changes in cognitive scores are to be expected during the acute period, therefore it would be inappropriate.

Table 1-4 Reliability types

Type of reliability	Definition	Measure
Internal consistency	The degree of inter-relatedness amongst items and whether they measure the same construct	Cronbach's alpha
Inter-rater	Agreement of scoring between two or more raters	Kappa coefficient
Test-retest	Consistency at two different time-points (measurement repetition)	Intraclass correlation coefficient (ICC)

1.9.3 Validity

Validity refers to whether the scale measures what is intended, in this case cognitive abilities. An example where validity may be compromised in post-stroke cognitive testing is where limb weakness causes a patient to poorly complete a cognitive task. Some examples of the main types of validity are in Table 1-5.

Table 1-5 Validity measures

Type of validity	Definition	Measure(s)
Content validity	The degree in which a measure includes the necessary items to represent the concept	Mainly using qualitative methods
Criterion validity	How the score compares to a gold standard	Correlation, test accuracy measures
Construct validity (convergent, known-groups)	How well the items represent the construct	Correlations, mean scores across groups and associated p value

1.9.4 Floor and ceiling effects

Floor and ceiling effects are a measure of targeting of the CSI; the extent to which the range of cognitive abilities measured by the CSI matches the range of cognitive abilities in the sample. Floor/ceiling effects are defined as the proportion of participants scoring the highest (ceiling) or lowest (floor) possible score. Criteria is usually set at >15% (116). It is important to check whether these effects are present as they indicate the inability of the CSI to differentiate those at the low/high end of the spectrum, e.g., if a high proportion of participants score full marks, it suggests they all have the same cognitive ability. These effects are also important if the CSI is to be used longitudinally, as a CSI demonstrating ceiling effects leaves no room to identify any improvement in impairment over time.

1.9.5 Diagnostic test accuracy (DTA)

Test accuracy refers to the test's ability to discriminate between patients with the target condition/disease and those without it. The test of interest, in this case a CSI, is referred to as the 'index test'. Results of the index test are compared to a gold (or reference) standard. A reference standard should be the

best available method for detecting the target condition. The results of both the index test and reference standard are tabulated, as illustrated in Table 1-6.

Table 1-6 2x2 contingency table

	Reference standard positive (condition present)	**Reference standard negative (condition absent)**
Index test positive	True positive (a)	False positive (b)
Index test negative	False negative (c)	True negative (d)

From this table several different test accuracy metrics can be calculated. The formulas for each test metric are given in Table 1-7. Paired measures are most frequently used (e.g., sensitivity and specificity, positive predictive value (PPV) and negative predictive value (NPV)). Sensitivity refers to the true positive rate: the proportion of those with the condition that are correctly identified by the index test as cases, for example sensitivity of 80% means that 80% of individuals with the condition will test positive. Specificity refers to the true negative rate: the proportion of people without the condition that are correctly identified as non-cases, for example specificity of 80% means 80% of individuals without the condition will test negative. The two metrics are inversely proportional and depend on the cut-off value used; a higher cut-off will result in greater sensitivity but lower specificity. A highly sensitive test helps one rule out a condition when test negative, whereas a highly specific test helps one rule a condition in when test positive. Sensitivity and specificity are not affected by prevalence whereas PPV and NPV are. PPV and NPV can be thought of as the clinical relevance of a test. PPV is the probability that an individual with a positive test result, truly has the condition. NPV is the probability that an individual with a negative test result, truly does not have the condition.

Table 1-7 Test accuracy indicators

Test metric	Calculation
Paired measures	
Sensitivity	a/(a+c)
Specificity	d/(b+d)
PPV	a/(a+b)
NPV	d/(c+d)
Positive likelihood ratio (LR+)	Sensitivity/(1 – specificity)
Negative likelihood ratio (LR-)	(1 – sensitivity)/specificity
Unitary/single measures	
Accuracy	(a+d)/(a+b+c+d)
Youden index	(Sensitivity + Specificity) – 1
ROC curve	False positive rate (1 – Specificity) on the x axis plotted against sensitivity on the y axis
AUC or AUROC	Area underneath the ROC curve
DOR	ad/bc

Abbreviations: AUC: Area under curve; DOR, Diagnostic odds ratio; NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operating characteristic.

Occasionally a single measure of test accuracy is preferred and there are several global/unitary metrics. Each have limitations so one global accuracy measure cannot be considered superior (117). For example, test accuracy and the Youden index are affected by prevalence, whereas the diagnostic odds ratio (DOR) is not, but gives the most optimistic results, ignoring the test's weaknesses. Test sensitivity is often plotted against 1-specificity, on a graph called a receiver operating characteristic (ROC) plot (Example taken from Takwoingi & Quinn 2018 (118) in Figure 1-5). When the data for each cut-off point is plotted, it creates a ROC curve. The area underneath this curve (AUROC or AUC) is used as a global measure (higher values indicate greater accuracy).

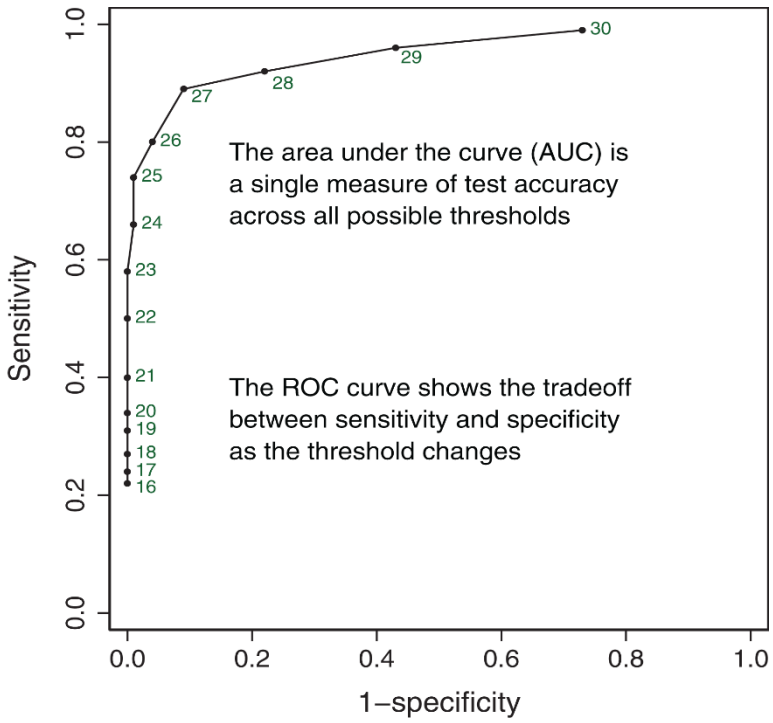


Figure 1-5 Example of a ROC curve using data for the MMSE (111)
The numbers in green illustrate the various threshold scores of the MMSE, and the corresponding sensitivity/specificity.

1.10 Accuracy of CSIs post-stroke (evidence-base)

The syndrome of interest (defined by the reference standard) impacts on the accuracy metrics, so it is important that this is clear and well defined. Most often, studies targeting more severe cognitive impairment, for example dementia or impairment across more than one domain, will have inflated

sensitivity, since many CSIs are better at detecting this end of the spectrum (119).

The reviews that have examined properties of CSIs post-stroke have variations in the choice of target syndrome and reference standard, which can explain any apparently different results. A review focusing on accuracy of CSIs for multi-domain cognitive impairment and clinical diagnosis of dementia after stroke found 35 studies covering 25 different CSIs (120). Pooled analysis used the following number of studies for each CSI: MMSE (12 studies), MoCA (6 studies) ACE-R (2 studies), Rotterdam-CAMCOG (R-CAMCOG) (2 studies). Sensitivity was highest for the ACE-R and MoCA, whereas specificity was highest using the MMSE or R-CAMCOG. Timing of CSI administration was compared; acute testing had higher sensitivity, but lower specificity. There was insufficient evidence regarding the use of brief CSIs in stroke, since only three studies were found and no studies using stroke-specific CSIs were found.

Other reviews have focused on comparing CSIs to NPB only and have included studies targeting single-domain cognitive impairment. Stolwyk et al. (58) reviewed 16 studies that used a CSI to detect either single or multi-domain cognitive impairment. Most of the studies noted to have adequate accuracy had a target condition of ≥2 domains rather than single domain. Use of the MMSE was discouraged, yet preliminary support was given to the MoCA (5 studies), the Repeatable battery for the Assessment of Neuropsychological status (RBANS) (1 study), Cognistat (1 study) and Barrow Neurological Institute Screen for Higher Cerebral Functions (BNIS) (1 study). Another systematic review with similar aims, included 21 papers examining 12 CSIs (121). Recommendations were made based on the syndrome of interest, although some scales recommended by the authors have limited evidence. Recommendations to detect any level of cognitive impairment are the Addenbrooke's cognitive evaluation revised (ACE-R), BNIS and Cognistat. For multi-domain impairments: ACE-R, T-MoCA, TICS-m. For dementia: TICS, CAMCOG, R-CAMCOG.

Finally, another review published in 2019 found 55 papers examining the psychometric properties of 26 CSIs in stroke (59). Most of these studies focused on validity rather than reliability and the data related to stroke-specific CSIs, (albeit a small amount) indicated encouraging findings.

Overall, the literature indicates that there are many studies examining properties of the MoCA and MMSE in stroke but few examining the many other CSIs that are available, including those designed for stroke. Studies addressing measurement properties of other CSIs should therefore be carried out, to meet this gap. As previously discussed, interpreting results across studies can be challenging, due to inconsistent definitions and differing reference standards. Limitations should also be acknowledged, such as case-control methodology and narrow inclusion criteria (limiting generalisability to those patients with more severe strokes, aphasia, and dementia).

1.11 Feasibility and acceptability

Feasibility and acceptability are terms often used in the context of completing questionnaires, for example PRO measures. Feasibility is typically defined as the time and cost needed to administer, score and interpret a measure, therefore capturing the burden and disruption for clinical staff, whereas acceptability is typically defined as the willingness or ability to complete an instrument from the patient perspective (122). Acceptability is often assessed through instrument completion, response rates and missing values.

In the context of administering CSIs post stroke, there is more overlap in feasibility and acceptability, which is why they are often used interchangeably. I therefore use the term feasibility in this thesis to cover both patient- and clinician- related barriers to completion. CSIs need to be acceptable to both patients and staff, but they also need to be feasible for the setting of an acute medical unit. In addition to these aspects, patient's physical stroke impairments, such as limb weakness, may affect both feasibility and acceptability.

Recent research has indicated that some patients after a stroke struggle or are unable to fully complete CSIs such as the MoCA (123, 124). However, there is evidence that they can be completed in mild to moderate strokes (125). Therefore, feasibility is a particular concern if you want to administer the same CSI to all stroke patients. It is thought that shorter tests taking less than 10 minutes to administer, or tests designed for stroke may perform better in terms

of feasibility and acceptability, although there is insufficient evidence available to confirm this.

1.12 Prognostic utility

The prognostic utility of a CSI relates to its ability to predict future outcomes. Outcomes of interest may be wide ranging and not limited to future cognitive issues, for example the ability to indicate future functional abilities, mood, and quality of life.

The PICOTS framework can be used as guidance for formulating a prognostic study. It involves defining the following components: Population, Index prognostic factor or model, Comparator prognostic factor or model, Outcomes, Timing and Setting. Different analysis methods may be used in prognostic studies; in this thesis I will use test accuracy and regression methods to understand the potential prognostic value of CSIs.

A review of the MoCA in stroke details that, when administered in the acute or subacute period, it provides a good prediction of later cognitive impairment at 3, 6 and 12 months (accuracy ≥90%) (74). Further research is needed to evaluate other CSIs and other outcome measures beyond cognition.

1.13 Summary

To summarise, there are limited data concerning psychometric properties of both brief CSIs (administered in ≤5 minutes) and stroke-specific CSIs in patients after stroke. This means that some teams may be using CSIs without supportive evidence that they are fit for purpose. The work that follows addresses these two research gaps.

2 Who are classified untestable on brief cognitive screening instruments in an acute stroke setting?

2.1 Introduction

As discussed in chapter 1, clinical guidelines recommend cognitive screening for all patients following a stroke, yet there are scenarios where a full cognitive screening instrument (CSI) or sections of a CSI cannot be completed. There can be various reasons for this, including individuals being too unwell, post-stroke fatigue or using tests which have not been designed for acute stroke. However regardless of test choice it should be acknowledged that completion of cognitive screening in a medically unwell person with recent neurological insult is challenging and, in some cases, inappropriate. If some people following a stroke are unable to be tested, this proves problematic for research studies using cognitive outcome measures and national audits such as the Sentinel Stroke National Audit Programme (SSNAP) (https://www.strokeaudit.org/) where each hospital's performance is monitored/benchmarked based on aspects such as whether cognitive screening has been carried out.

To examine completion rates in previous research, I carried out a scoping search in PubMed in 2019 to identify studies reporting on completion rates (Table 2-1). This was not a systematic search and should not be considered exhaustive. Previous research indicates that around 20% of people after a stroke cannot fully complete the CSIs that are routinely used, for example the MoCA and MMSE (125-127). However, since many studies do not report data on those who are untestable, it is difficult to ascertain the true incidence and characteristics of these individuals. Whilst non-completion can be a greater issue in acute settings, it is also reported later in the stroke pathway (Table 2-1). Results from previous studies can at times appear conflicting, with some studies concluding that it is feasible to administer a full battery of tests in acute stroke. It is important to emphasise that being able to complete such assessments with participants within a research study does not equate to feasibility for all stroke patients in all settings. Examining the participant characteristics of studies can help explain the apparently contradictory findings in the literature. Participants included in these studies are often not representative of a typical stroke unit population.

For example, studies may consciously or unconsciously favour the inclusion of those with minor strokes, little pre-morbid disability, and ability to provide written informed consent themselves, whilst patients with severe aphasia, or dementia are often excluded (128). The number of untestable patients is therefore reduced due to this selection bias.

Table 2-1 Previous Research on completion of Cognitive Tests in Stroke

Study	Test	Number of participants	Inclusion criteria relevant to feasibility	Timing after stroke	Completion rate
Acute setting					
Alderman 2013 (CA) (129)	8 tests	27	Mild strokes and TIA	≤24 hour	96%
Collas 2016 (130)	OCS	155	Unclear	5 days	89%
Horstmann 2014 (127)	MoCA	842	IS and ICH	2 days	81%
Pasi 2013 (125)	MoCA	137	IS and ICH	5-9 days	83%
Pendlebury 2015 (126)	AMT, MMSE	1097		4 days	76% partially testable, 69% fully testable
Van Zandvoort 2005 (131)	1.5-hour NPB	57	≤80, IS only. No previous stroke or psychiatric history, mRS 2-4	4-22 days	75%
Rehabilitation setting					
Barnay 2014 (132)	CASP, MMSE, MoCA	44	All aphasic patients	42 days	CASP 82% MMSE 64% MoCA 70%
Benaim 2015 (133)	CASP, MMSE, MoCA	50	Patients without aphasia	40 days	CASP 100% MMSE 100% MoCA 94%
Cumming 2011 (134)	MoCA	220	IS and ICH	3 months	Mild stroke: 87%, moderate stroke: 79%, severe stroke: 67%
Kwa 1996 (135)	CAMCOG	129	IS only	≥3 months	88%
Mancuso 2018 (136)	OCS, MMSE	325	No previous stroke or psychiatric/ neurological disease, able to consent themselves	33.9 days	Fully untestable: MMSE 2%, OCS 1%. Completion rate given for individual OCS tasks. Highest incompletion for trails (9%)

Lees 2017 (137)	ACE-III, MMSE, MoCA	51	IS, ICH	36 days	ACE-III 27% MMSE 43% MoCA 39%

Author's own table based on a scoping search of one database.
Abbreviations: ACE-III, Addenbrooke's cognitive evaluation; AMT, Abbreviated Mental test; CA, conference abstract; CAMCOG, Cambridge Cognition; CASP, Cognitive Evaluation for Stroke Patients; OCS, Oxford Cognitive Screen; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; NPB, Neuropsychological battery.

For data outside of a research study, SSNAP is a valuable, freely available resource. For patients admitted to and/or discharged from hospital between August 2017-November 2017, 15.9% were classified as 'not applicable' for cognitive screening for reasons of declining assessment or medically unwell for the entire admission and their total length of stay was ≥7 days (138). Of the patients where cognitive screening was deemed appropriate, 94.2% received an assessment. It is not documented why the other 5.8% were not screened.

Feasibility and acceptability of completing cognitive screens is multifactorial. Some aspects are specific to the environment (limited time and quiet space for testing on acute medical wards) and the stroke itself (extent of neurological damage, presence of aphasia or limb weakness). Other factors relate more generally to the nature of testing older adults with other comorbidities (e.g. hearing or visual impairment, arthritis and fatigue).

Incomplete assessments have clinical implications as, ultimately, they can risk a false positive or false negative categorisation of cognitive impairment. This is because the items which cannot be completed will either be assigned a score of zero or not scored at all. This can unintendedly cause harm, as patients without cognitive impairment may be asked to come back in for more detailed assessment and those with cognitive impairment with an unscored test may be missed. Test incompletion also complicates and biases research and audit; often patients with incomplete data are excluded from analyses and treated as missing (since a total score cannot be calculated and the recommended test threshold score cannot be used). Various approaches exist to incorporate incomplete data but there is no consensus on the best method (137).

There are different ways to address feasibility or acceptability issues, but the approach taken will depend on the aspect of greatest relevance. For example, one may decide to choose a test specifically designed for use in stroke. This

approach recognises that many traditional cognitive screens were designed for a memory clinic setting and are not suited to the specific challenges encountered in acute stroke settings. The stroke specific cognitive tests detailed in the introduction may be less biased by physical, communication and visuospatial impairments. A different approach would be to choose a shorter cognitive screen. This approach is attractive for time limited settings and where there is a shortage of trained staff. Shorter tests may also be attractive to patients as there is reduced test burden.

Stroke care is continuously evolving and differs internationally, yet there is a currently a paucity of feasibility evidence in an acute, National Health Service (NHS) context. No previous studies have explored feasibility of the shortest cognitive tests available within a stroke setting. This study aimed to meet these two gaps.

My primary aim was to describe the rate of full and partial test completion of various short cognitive screening tests in consecutive patients admitted to a UK hyper-acute stroke unit (HASU). My secondary aims were to explore the reasons for assessors assigning an untestable label and to describe factors associated with being untestable.

2.2 Methods

I conducted a cross-sectional study using routinely collected clinical data from a UK, teaching hospital. The database used for this study was designed in liaison with the Research Ethics Committee (REC); the data collected for this work did not go above routine, clinical care and data were fully anonymised before archiving, therefore written informed consent was not required for data usage. Collection of data was approved by the West of Scotland REC (ws/16/0001) on 04/02/16 (Appendix 1). I followed best practice guidelines (Standards of Reporting of Neurological Disorders (STROND) for the design, conduct and reporting of the study. I collected data alongside four other University of Glasgow students who were enrolled on one of the following courses: Psychology PhD/Neuroscience MSc/undergraduate MBChB. We were all trained prior to collecting data. One researcher went daily onto the HASU to check for new

admissions on weekdays and checked for new admissions over the weekend at the start of the week.

2.2.1 Setting and population

Participants were consecutive admissions to the Glasgow Royal Infirmary HASU. The unit admits all individuals with suspected stroke or TIA, without any age, disability, or comorbidity exclusions. Data from patients admitted during four timepoints were included: May 2016-Febraury 2017, April-June 2017, October-December 2017, and July-August 2018. These timepoints were based on when a researcher was available to go daily to the HASU to collect data. The only inclusion criteria set was that participants had to be treated (but not necessarily subsequently confirmed) as a suspected stroke at the time of assessment in order to be included.

2.2.2 Clinical and demographic data

Anonymised demographic data (age at time of admission, sex) and a range of clinical data were collected from the medical case notes or through prospective assessment. Data were collected onto a proforma which was approved by REC (Appendix 2). Only relevant variables from the proforma were used for this study. Medical history included previous stroke/TIA and previous diagnosis of dementia. The Bamford/Oxfordshire Community Stroke Project classification system (OCSP) (139) was completed for participants who had an ischaemic or haemorrhagic stroke. There are four syndromes: total anterior circulation stroke (TACS), partial anterior circulation stroke (PACS), posterior circulation stroke (POCS) and lacunar stroke (LACS). Classification can be made on clinical symptoms, without imaging data.

Stroke severity was measured using the National Institutes for Health Stroke Scale (NIHSS) score completed on hospital admission. In cases where the NIHSS was not completed and documented by the clinical team, it was retrospectively scored using the details of their symptoms on admission. Although not as accurate as direct assessment, this method of obtaining scores has high reliability and validity (140). The NIHSS is a ClinRO scored from 0-42, with higher scores indicating greater stroke severity. The areas covered are level of

consciousness, questions, commands, gaze, visual fields, facial palsy, motor arm (left & right), motor leg (left & right), ataxia, sensory, language, dysarthria, and extinction/inattention. This measure is considered the gold standard for rating stroke severity (141) and is widely used, both in clinical practice and research. Limitations, however, are that the items covered are biased towards left hemispheric strokes (these score higher than right hemispheric strokes of a larger volume) (142) and the symptoms associated with posterior circulation strokes are largely neglected (143).

Pre-stroke functioning was measured using the modified Rankin Scale (mRS). The mRS is a ClinRO and was designed to measure post-stroke global disability on a scale of 0 (no disability) to 6 (death). It is the most widely used measure of function in stroke clinical research, and many clinical trials have used it to measure pre-stroke function for inclusion criteria, as well as using is an outcome measure for post-stroke function. The scale performs well in measuring post-stroke disability in terms of high test-retest reliability and construct and convergent validity (144) but inter-rater reliability ranges from poor to near perfect across different studies (145). Its performance to measure pre-stroke function is limited in terms of inter-rater reliability and construct validity (146). In cases where the mRS was not documented by the clinical team, scoring was attempted using details recorded in the medical notes (e.g. living independently) or directly asking the patient.

2.2.3 Cognitive assessment

The cognitive assessment was attempted during the first week of stroke unit admission. The assessment consisted of a set of 13 questions which are included within nine different CSIs. These questions were asked in the same order for all patients (order presented in Figure 2-4). After completion of these 13 questions, the nine CSIs could be scored.

The CSIs compared were: the Clock-drawing test (CDT), Mini-Cog (63), Abbreviated MoCA (147), National Institute of Neurological Disorders and Stroke-Canadian Stroke Network (NINDS-CSN) 5-min MoCA (110), 10-point Abbreviated Mental Test (AMT-10) (61) and its shorter version AMT-4, General Practitioner Assessment of Cognition (GPCOG) (patient section), Six-item Cognitive

Impairment test (6-CIT), and the 4 ‘A’s Test (4AT) (Available online: www.the4AT.com). The justification for choosing these CSIs is that the administration time to deliver these tests individually is under 5 minutes, making them all suitable for acute clinical practice. They are generic cognitive screens (not specifically designed for stroke) but have some supporting validation work in primary and geriatric care (148) and they share a number of tasks in common. Tasks within each test are detailed in Table 2-2.

Table 2-2 Items within each cognitive test

Test	Tasks
Clock-drawing test	1. Clock draw
Mini-Cog (63)	2. 3-word delayed recall 3. Clock draw
Abbreviated MoCA (147)	1. 5-word delayed recall 2. Clock draw
AMT-4 (149)	1. Age 2. Year 3. Place 4. Date of Birth
4AT (Available online: www.the4AT.com)	1. Alertness 2. Age 3. Date of Birth 4. Place 5. Year 6. Months of the year backwards
6-CIT (62)	1. Time 2. Month 3. Year 4. Count backwards from 20 5. 5-part delayed recall 6. Months of the year backwards
GPCOG (64)	1. Date 2. Month 3. Year 4. Date of Birth 5. 5-part delayed recall 6. Clock draw (numbers, hands) 7. News item
NINDS-CSN 5 min MoCA (110)	1. Date 2. Month 3. Year 4. Day 5. Place 6. City 7. 5-word delayed recall 8. Fluency (letter F)

AMT-10 (61)	1. Age 2. Time 3. Year 4. Place 5. 2-person recognition 6. Date of Birth 7. Year of WW1 8. Current prime minister 9. Count backwards from 20 10. 2-part delayed recall

Abbreviations: AMT, Abbreviated mental test; GPCOG, General Practitioner assessment of cognition; NINDS-CSN, National Institute of Neurological Disorders and Stroke-Canadian Stroke Network; 4AT, the 4 'A's Test; 6-CIT, Six-item Cognitive Impairment test.

Participants were only approached once for assessment, apart from the following scenarios: the assessment was interrupted for another clinical investigation (e.g., scan), the participant requested for the assessment to be done at a later time-point or over two sessions. Participants were not approached at all (and categorised as fully untestable) if the parent clinical team reported that the participant was too unwell to undergo a cognitive screen or if the assessor felt that any form of direct testing would not be possible. If participants could not be directly assessed, we examined the medical notes to check if a cognitive screen had already been completed and documented since admission.

2.2.4 Defining Untestable Outcomes

In discussion with my supervisor and based on literature, I created descriptions to categorise completion of CSIs. These were created before I began analysis of the data. The operationalised definition for 'fully untestable' was where no part of the cognitive assessment was attempted with the participant or where there was no response to questions when testing was attempted. Partially untestable was defined where at least one item within the cognitive assessment could not be completed or was not attempted (a decision made by either by the participant, parent clinical team or researcher) but at least one question was completed.

If a participant could not be screened, all researchers were instructed to document on the data collection proforma; and to provide a reason. These free text responses detailing reasons why a participant could not be assessed (untestable) were collated and assessed. A list of untestable categories was created based on the combination of previous literature, clinical experience, and initial scoping of free text responses. I went through the free text reasons documented for each untestable participant and collated them into categories (e.g., aphasia and dysarthria both captured under speech problems). These decisions were made in discussion with the stroke consultant (Dr Quinn). Where an assessor had listed more than one reason for being untestable, one primary factor was chosen that was deemed to have the greatest impact (e.g., a participant documented as both acutely confused and dysarthric, was categorised under confusion). In cases where a test item was attempted but poorly completed, for example, a participant with limb weakness who attempted the clock-drawing task with their weak or non-dominant hand, were classified as testable.

2.2.5 Statistical Analysis

I determined the completion rates for each question, and these were used to determine the completion rates for each of the various tests. I included the full sample of participants in all analyses. The decision to retain participants whose diagnosis was later determined to be non-stroke was to understand feasibility of the measures within all participants admitted with a suspected stroke in a real-world setting. We also retained the admission NIHSS for these participants where it was completed.

I carried out univariate and multivariate logistic regression. I chose variables for inclusion in the model based on previous literature (126, 132) and plausible associations with feasibility. I included the following 12 covariates in both univariate and multivariate analyses: age, sex, NIHSS, Bamford stroke classification, (TIA, PACS, TACS, POCS, LACS and non-stroke (used as reference group)), pre-morbid mRS, presence of intracerebral haemorrhage (ICH), previous diagnosis of dementia and previous stroke or TIA. I did not include delirium in the model since the only consistently applied measure of delirium available was the 4AT and this was one of the tests under investigation. I described

associations as odds ratios (OR) with corresponding 95% confidence intervals. I used the rule of 10 outcome events per predictor variable to determine the number of covariates we could include in the model and so required 120 ‘cases’. I ran analyses twice to account for how partially untestable participants are treated differently in the literature; in the first analysis they were treated as testable and in the second treated as untestable (grouped with fully untestable). I performed all data analyses using the statistical software package SPSS (version 25 IBM, Armonk, NY, USA).

2.3 Results

The total sample included 703 participants. Participants had a mean age of 69.4 years (SD 13.7); 392 (54%) were male; 429 (61%) had an ischaemic stroke, 22 (3%) had an ICH; median NIHSS 2 (interquartile range (IQR) 1-5). There were 119 (17%) participants classified as fully untestable (they did not attempt or complete any of the questions within the assessment) and 58 (8%) participants classified as partially untestable (≥1 question was not attempted). Full characteristics of the total sample, testable and untestable participants are available in Table 2-3.

Table 2-3 Characteristics of the Sample

Variables	Total sample (N=703)	Fully testable (N=526)	Partially untestable (N=58)	Fully untestable (N=119)
Sex (Male)	382 (54%)	296 (56%)	27 (47%)	59 (50%)
Age, years (Mean, SD) *Missing data*	69.4 (13.7) *N=2*	66.9 (13.5)	76.6 (9.7)	76.8 (12.5) *N=2*
IS ICH TIA Non-stroke *Missing data*	429 22 137 109 *N=6*	302 10 121 89 *N=4*	42 4 5 7	85 8 11 13 *N=2*
Bamford classification (completed for IS and ICH) *Missing data*	66 TACS 174 PACS 100 POCS 111 LACS *N=6*	13 TACS 118 PACS 80 POCS 101 LACS *N=4*	3 TACS 25 PACS 12 POCS 6 LACS	50 TACS 31 PACS 8 POCS 4 LACS *N=2*
NIHSS (Median, IQR)	2 (1-5)	2 (1-3)	4 (3-7)	8 (4-16)
Pre-morbid mRS (Median, IQR)	1 (0-3)	1 (0-2)	2 (0-3)	3 (0-3)
Previous stroke (IS/ICH) or TIA (N, %)	218 (31%)	162 (30%)	20 (34%)	36 (30%)
Previous diagnosis of dementia (N, %)	61 (9%)	23 (4%)	8 (14%)	30 (25%)

Abbreviations: ICH, intracerebral haemorrhage; IS, ischaemic stroke; IQR, interquartile range; LACS, lacunar stroke; mRS, modified Rankin scale; N, Number; NIHSS, National Institutes for Health Stroke Scale; PACS, partial anterior circulation stroke; POCS, posterior circulation stroke; TACS, total anterior circulation stroke; TIA, transient ischaemic attack; SD, standard deviation.

2.3.1 Reasons for untestable status

The reasons documented by the assessors for categorisation of fully untestable fell under 8 categories, with neurological deterioration (e.g., participants who were unresponsive, medically unwell, palliative) being the main reason (over half of the sample: 54%). The other reasons are detailed in Figure 2-1.

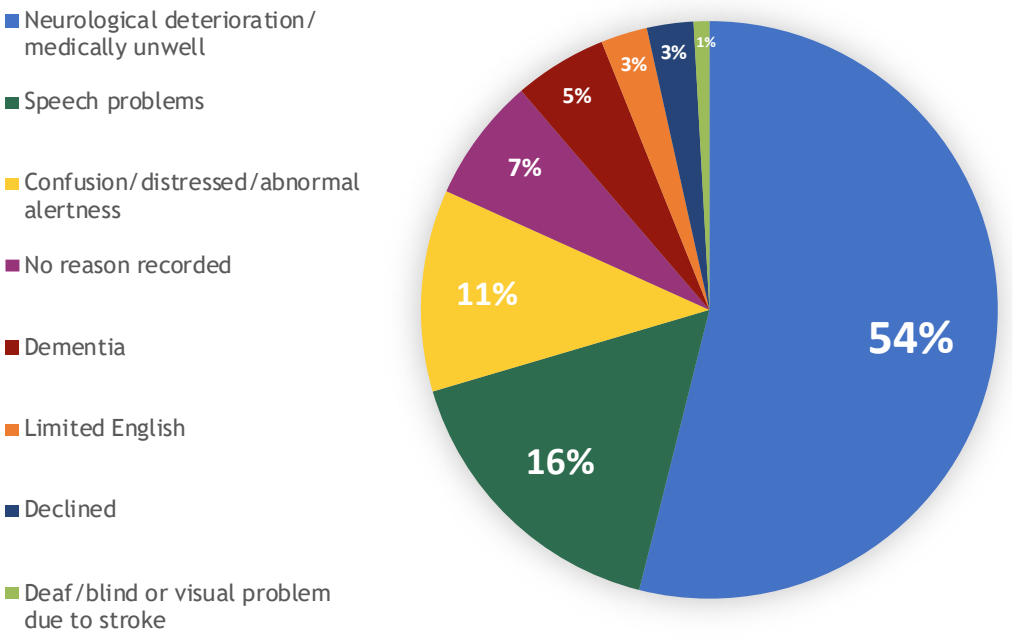


Figure 2-1 Reasons for Participants classified as fully untestable (n=119)

The reasons documented by the assessors for categorisation of partially untestable fell under 9 categories (8 of these the same as fully untestable) but the main two reasons were limb weakness (26%) and speech problems (22%) (Figure 2-2). Only 12 participants (2%) within the full sample declined the cognitive assessment; three declined the full assessment and nine declined certain questions. Characteristics of those that declined: 7 (58%) male, mean age of 74.3 years (SD 13.9), median NIHSS of 3 (IQR 2-5), and diagnoses consisted of 1 non-stroke, 3 TIAs, 3 PACS, 3 POCS and 2 LACS. Participants whose final diagnosis was non-stroke were a diverse group; diagnoses included migraine, subarachnoid haemorrhage, and vasovagal events. Of these 20 (18%) were untestable in some way.

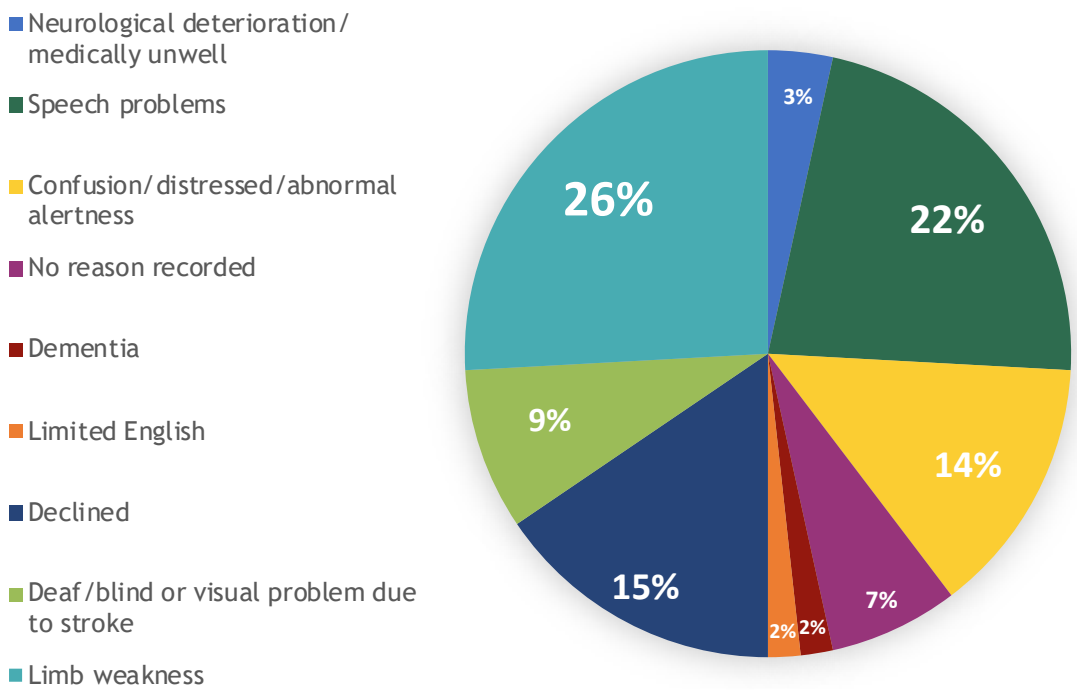


Figure 2-2 Reasons for Participants classified as Partially Untestable (n=58)

For the participants in the partially untestable group, I examined the completion of each question within the assessment. Age was completed by the greatest number of participants (57/58 (98%)) and clock-drawing by the fewest (7/58 (12%)) (Figure 2-4). In 25/58 (43%) of participants, the clock-draw was the only task that they did not attempt for the following reasons: limb weakness (N=14), visual impairment (N=4), declined (N=5) and no reason documented (N=2). There was a general downward trend of completion, with items at the beginning of the test having the highest completion.

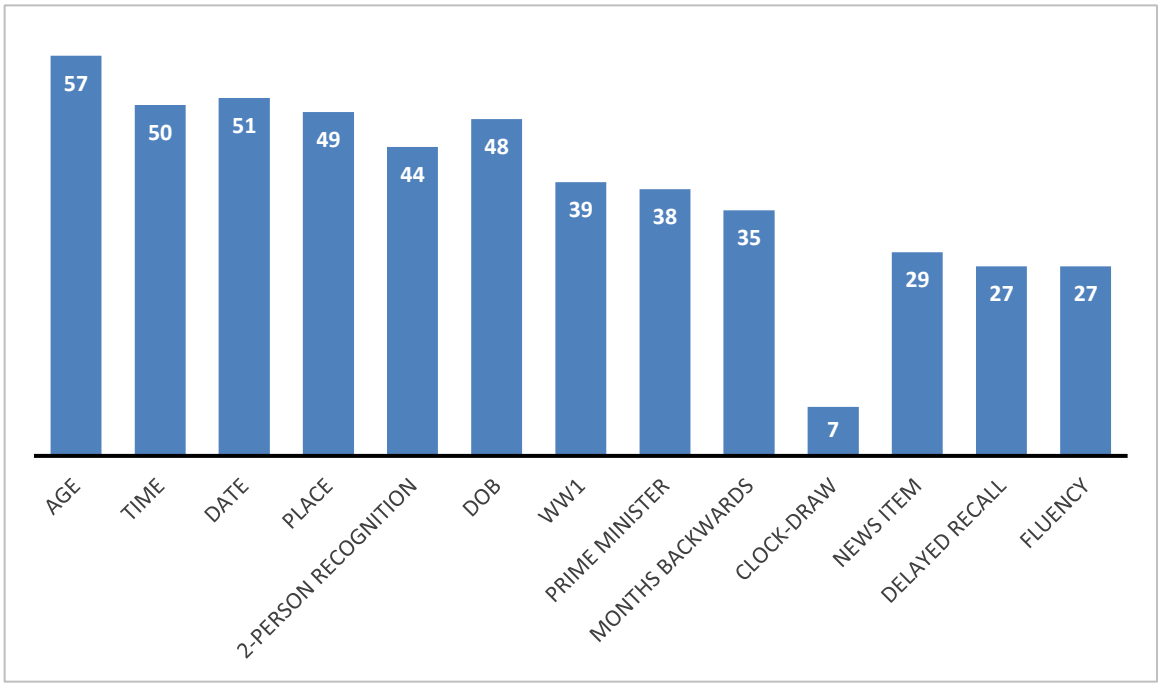


Figure 2-4 Items completed by those in the partially untestable group (listed in order of administration) (N=58)

Using this data, I could calculate the completion rate of each test across the full sample. Completion rates ranged 75-100%. The only test which could be scored in full for all participants was the 4AT. The tests with the lowest completion rate were the clock-drawing test and other tests which also include this item: Abbreviated MoCA, Mini-Cog and GP-Cog.

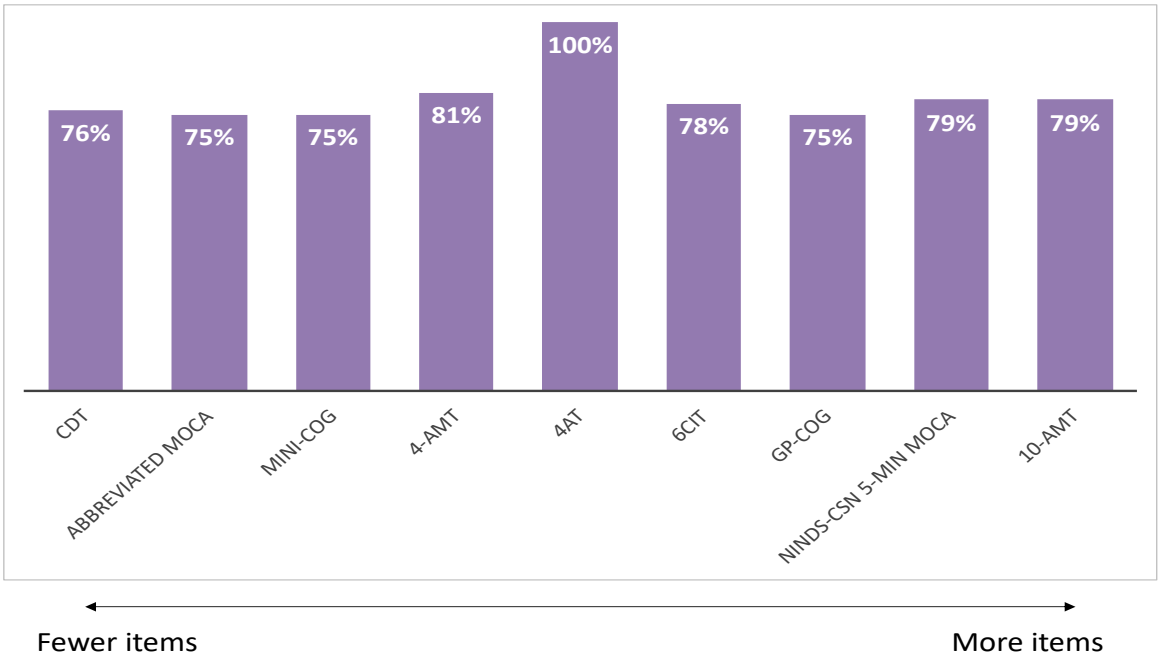


Figure 2-3 Percentage of CSIs scored in full (tests ordered by number of items)

2.3.2 Associations with untestable status

In the univariate analyses: higher age, TACS, ICH, higher NIHSS, higher pre-morbid mRS and a pre-stroke diagnosis of dementia were associated with being untestable whilst a lacunar stroke was associated with being testable. In the first multivariate regression analysis (n=680), independent associations with fully untestable status were: higher NIHSS score (OR 1.18, 95% CI 1.11-1.26), higher pre-morbid mRS (OR 1.28, 95% CI 1.02-1.60) and pre-stroke dementia (OR 3.35, 95% CI 1.53-7.32). A lacunar stroke diagnosis was associated with being testable (OR 0.19, 95% CI 0.06-0.65) (Figure 2-5). In the second analysis (where partially untestable and fully untestable groups were combined), the above variables remained significant. In addition, the following associations were found for being untestable: older age (OR 1.04, 95% CI 1.02-1.06) and presence of ICH (OR 3.44, 95% CI 1.13-10.44), whilst a TIA classification was associated with being testable (OR 0.45, 95% CI 0.20-0.997) (Table 2-4).

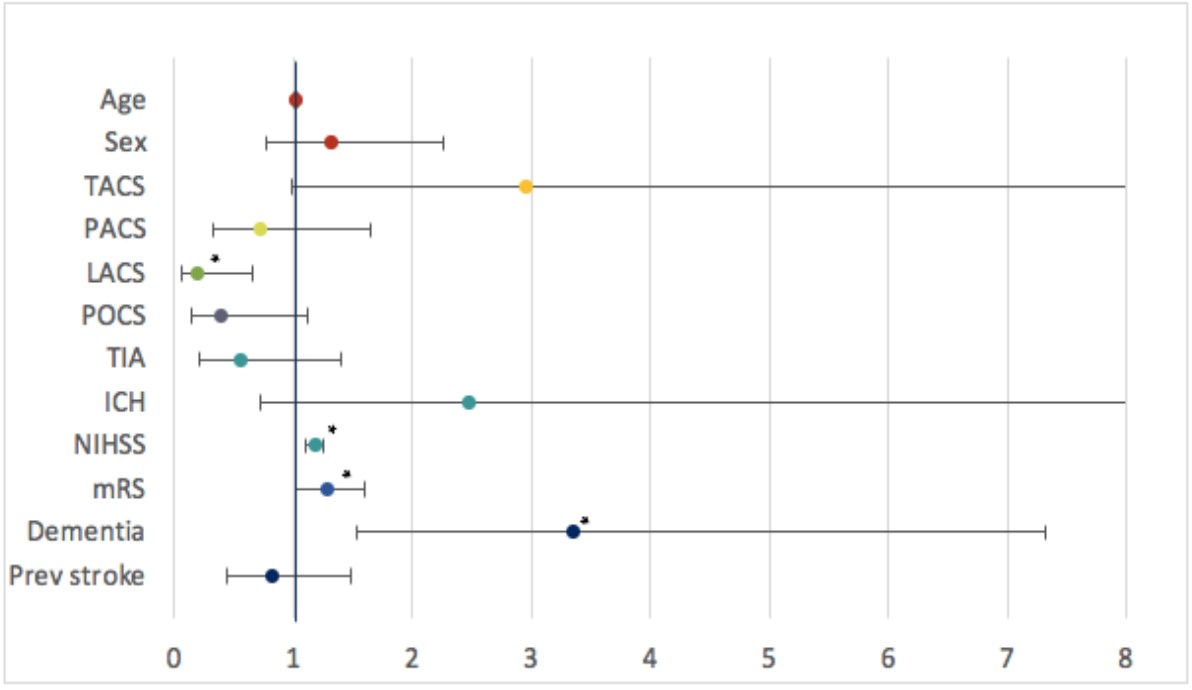


Figure 2-5 Independent associations for fully untestable
Odds ratios (OR) for each variable in the multivariate analysis for status of fully untestable (*p<0.05)

Table 2-4 Associations with untestable status

Variables	Univariate for fully untestable	Multivariate (partially treated as testable)	Multivariate (partially treated as untestable)
	OR (95% CI)		
Age (years)	1.06 (1.04-1.08)*	1.02 (1.00-1.04)	1.04 (1.02-1.06)*
Sex (Male)	0.80 (0.54-1.18)	1.32 (0.77-2.26)	0.97 (0.62-1.51)
Stroke classification (Non-stroke used as reference category):			
TACS PACS LACS POCS TIA	23.08 (10.29-51.76)* 1.60 (0.80-3.22) 0.28 (0.08-0.88)* 0.64 (0.25-1.62) 0.65 (0.28-1.50)	2.96 (0.98-8.93) 0.73 (0.32-1.65) 0.19 (0.06-0.65)* 0.39 (0.14-1.12) 0.55 (0.21-1.40)	1.47 (0.50-4.34) 0.92 (0.46-1.83) 0.26 (0.10-0.64)* 0.73 (0.33-1.61) 0.45 (0.20-1.00)*
ICH	2.96 (1.21-7.23)*	2.48 (0.72-8.59)	3.44 (1.13-10.44)*
NIHSS	1.30 (1.23-1.36)*	1.18 (1.11-1.26)*	1.23 (1.14-1.31)*
Pre-morbid mRS	1.64 (1.41-1.91)*	1.28 (1.02-1.60)*	1.24 (1.03-1.50)*
Pre-stroke diagnosis of dementia	6.01 (3.47-10.42)*	3.35 (1.53-7.32)*	2.74 (1.32-5.70)*
Previous stroke (IS, ICH) or TIA	0.96 (0.62-1.47)	0.82 (0.45-1.48)	0.91 (0.56-1.49)

*significant at $p<0.05$

Abbreviations: CI, confidence interval; ICH, intracerebral haemorrhage; IS, ischaemic stroke; LACS, lacunar stroke; mRS, modified Rankin scale; NIHSS, National Institutes for Health Stroke Scale; OR, odds ratio; PACS, partial anterior circulation stroke; POCS, posterior circulation stroke; TACS, total anterior circulation stroke; TIA, transient ischaemic attack.

2.4 Discussion

In an unselected sample of 703 participants admitted to our HASU, a quarter were classified as partially or fully untestable on a range of brief cognitive screening tests. The completion rate was lowest for the clock-drawing task so tests which include this (Abbreviated MoCA, Mini-Cog, GP-COG) had the lowest completion rate. The 4AT was the only test which could be scored in full for all participants as it includes scoring for untestable. Factors associated with being fully untestable were previous diagnosis of dementia, higher pre-morbid mRS and higher NIHSS on admission, whilst a diagnosis of lacunar stroke was associated with being testable.

2.4.1 Research in context

The results are generally in keeping with the limited literature on test completion. The associations of non-completion with stroke severity and dementia have face validity and the reasons given by the assessors for a label of untestable were similar to those described in previous studies (limb weakness (125, 126, 137), aphasia (125, 126), pre-morbid functional status (125) and reduced consciousness (126)). These results reinforce that non-completion is driven by both stroke specific and non-stroke related factors.

This work focused on some of the shortest cognitive screens available, in the anticipation that they would be more practical for both the participant and clinician. The completion rates were found to be similar to that quoted in previous studies which used longer, multi-domain screens (e.g., MoCA). As this work did not directly compare with these tests in our sample, one cannot assume that the rate of incompletion is equivalent. I can however comment on varying test length since the number of items ranged 1-10 across the nine CSIs. In one sense length is an important factor, since tasks placed later in the assessment generally had lower completion. On the other hand, the screens with the fewest items in our study (Mini-cog and Abbreviated MoCA) had the lowest completion rate. Therefore, within short screens, a focus on length alone is too simplistic and test content should be considered equally important.

The poor completion of the clock-drawing task mirrors previous findings; this task was more frequently refused by older medical inpatients (150) and tasks requiring copying or drawing were found to have the lowest rates of completion in a stroke population (137). While the decision to immediately remove drawing tasks may be tempting, it is one that should be considered carefully; the decision should be informed by other psychometric properties and always depend on the specific research question, since there is evidence that such tasks (when completed) can predict later outcomes (151) and can be highly sensitive as explored in Chapter 6.

With regards to test choice in an acute stroke setting, one may wonder whether a stroke specific test would have higher completion rates. Considering the reasons for participants having a fully untestable label, it is likely that the rates would be similar regardless of which screening tool was used. In the partially untestable group however, limb weakness and speech impairment were the two main reasons for non-completion, therefore a stroke-specific test, designed with these impairments in mind, may be superior.

2.4.2 Strengths and limitations

A major strength of this study is that I had access to an unbiased, real-world sample; something which is essential for work addressing feasibility. Once the process of written informed consent is involved, you inevitably have a biased sample to a certain extent. Therefore, I had valuable data on participants who are often excluded from research (e.g., those with dementia or severe aphasia). Whilst using clinical data have these benefits, I also acknowledge that due to the messy reality of acute clinical practice, data are often missing. The approach I took allowed me to retrospectively derive some of this missing data from various sources, including inpatient medical records, primary care data and consultation with the parent clinical team. As discussed in the methods, with the example of the NIHSS, retrospective scoring is considered to be a valid and reliable method. However, this assumes that a comprehensive examination was done and that findings were documented in detail. With this in mind, NIHSS scores may have been underestimated, since there is uncertainty whether all areas covered on the NIHSS were assessed in the initial clinical examination (e.g., neglect and ataxia).

There were some interesting aspects of testing where data were not recorded. It would be useful to know the frequency of participants with limb weakness that attempted a drawing task (classified as testable) and whether they used their weak or non-dominant hand. Many of these participants lose points or score zero for poor completion. We also did not record if an assessment had to be completed over two sessions or if any part of the assessment was interrupted. In the context of a research study, scenarios such as these should be anticipated, with guidance provided at the start of a study, to ensure consistency.

Although the concept of partially and fully untestable was operationalised, there is subjectivity in the interpretation. In psychometrics, completion rates are referred to as data completeness. Cognitive tests are different to other clinical outcome assessments or patient reported outcomes, due to this subjectivity of deciding whether a participant can complete a question. Therefore, the same participant could have missing data or a full data set depending on who is administering it. It is essentially a judgement call by the clinician whether participants with aphasia, limb weakness and visual problems can complete a task (if the participant does not decline themselves). The same individual could therefore be classified differently purely based on who assessed them. This is particularly relevant in this study, where like in clinical practice and large research studies, differing assessors performed the cognitive testing. It should be acknowledged that while all assessors received training, we each have varying levels of experience and therefore could make different judgements.

A final important limitation is the methodology of administering the 13 questions in the same order for each participant (without any randomisation). I acknowledge that this approach inherently introduces bias. However, test items included in each of the eight tests were generally spread out (e.g., the two tasks used in the shortest tests were not the first two questions). The aim of the database used for this study was to utilise clinical data collected as part of standard of care. Introducing randomisation of test items would therefore not be possible, without the work being undertaken as part of a research study requiring written informed consent, which as discussed earlier has limitations. Data collection for the database also commenced prior to my work on this study, so I could not change aspects of its design.

2.4.3 Recommendations for future research and practice

The strict administration and scoring criteria required for cognitive tests can be problematic for use in the stroke setting. Clinicians and researchers can therefore expect to encounter a number of people that will be untestable on certain tasks, or those who are testable, but their stroke-related impairments result in a misleading test score. This will also be the case in other neurological conditions which affect motor function or language (e.g., multiple sclerosis, Parkinson's disease, primary progressive aphasia). Whilst in clinical practice, an assessment is put into context of these other impairments to create a diagnostic formulation, in research test data are absolute and so it is important that a-priori rules are set for dealing with incomplete tests. The importance of doing this is highlighted by the fact that our analyses showed different results depending on how partially untestable participants were classified. Numerous approaches exist to deal with missing data (137), but to maximise the utility of the data collected, we recommend that researchers make full use of incomplete participant data where possible, rather than applying a complete case-analysis approach. This is also arguably the most ethical approach to avoid data wastage when patients have spent time providing it.

The results of this work have implications for cognitive test design. Firstly, with regards to how test items are ordered. It may be beneficial to have items placed in a priority order, since items placed earlier have higher completion. Assessment may end prematurely due to fatigue (a frequent complaint following stroke) or poor concentration (ward settings are often noisy). The assessment could also be interrupted for another clinical investigation. Another area regards acknowledging that, even with a very modified test that is designed for stroke impairments, some people will still be classified untestable. The 4AT is unique to other cognitive screens in that it incorporates scoring for untestable and refusal of tasks. The Mini-Cog (63) also includes scoring for "inability or refusal to draw a clock" but not for the delayed recall component. This approach is helpful and pragmatic and should be adopted when designing new screens. For existing screening tools, some consensus guidance and resources should be made available for challenging cases to improve reliability, since different assessors can score the same patient differently. These types of resources exist for scoring

other stroke scales, such as the NIHSS, in patients who are comatose, intubated, or aphasic.

Test completion rates are just one measure of feasibility/acceptability. There are numerous factors, relating to the person being assessed, the assessor and the ward setting which affect cognitive screening (examples in Figure 2-6) and all of these perspectives should be considered in decision-making. To date, there has been little data published about the patient and clinician's experience of cognitive assessment and how environmental factors can affect assessment on the ward (noise, space, interruptions). Qualitative or mixed methods research is vital to explore and understand these factors in a stroke setting. Some recent work, completed with patients with brain tumours, compared the feasibility and acceptance of the MoCA administered pre- and post-operatively through use of a questionnaire (152). Patients indicated feeling distracted at the pre-operative time-point, highlighting an increased burden prior to a procedure. This scenario could be viewed similarly to acute stroke when patients are often awaiting multiple investigations and results and is therefore a distressing time.

With the increased use of computerised versions of cognitive screens in the future, many factors from the clinician/assessor's perspective are likely to improve, both in terms of efficiency (automatic scoring saves time) and in terms of accuracy (reducing scoring errors and subjectivity). Future work should also make use of routinely collected clinical data, such as that collected in SSNAP. One could also argue that any study using a researcher to administer a scale (or collect data), rather than a clinical member of staff, does not truly address broader feasibility and implementation issues. Finally, in terms of CSI choice, other psychometric properties should be considered which will be discussed in the following chapters.

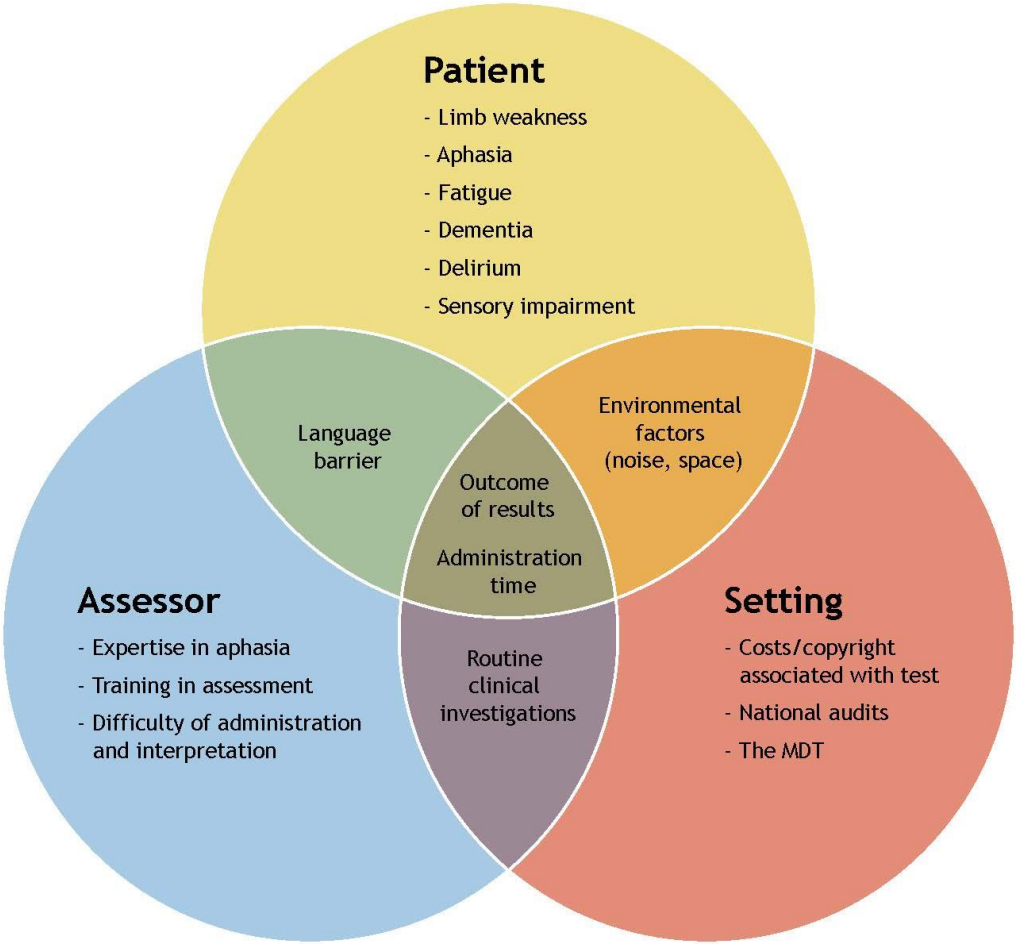


Figure 2-6 Examples of factors affecting delivery and interpretation of cognitive screening

To conclude, in a real-world sample of people admitted to our HASU, a quarter were classified as fully or partially untestable on cognitive screening (comprised of 9 brief CSIs). Cognitive screens with fewer items do not necessarily have a higher completion rate and the 4AT was the only test which could be scored in full for all participants. Clinicians and researchers should make a-priori plans on how to address incomplete assessments.

3 Accuracy of short forms of the Montreal Cognitive Assessment: Systematic review and validation

3.1 Introduction

In the previous chapter I focused on the feasibility of completing brief cognitive screening instruments (CSIs). Accuracy of a CSI is another important measurement property to assess which will be the focus in this chapter. The time required to administer a CSI is considered a key determinant as to whether it will be used in clinical practice (153) and it has been suggested that short tests achieve better cooperation from patients (61). Certain clinical factors, such as if an individual has physical, speech or cognitive impairments, make the duration of assessment longer than what is documented in the literature (21) or even cause them to be untestable on certain items, as illustrated in Chapter 2.

As discussed in the introduction chapter, the MoCA has been recommended for use after stroke due to sensitivity to milder forms of cognitive impairment (79) and as a result is a popular choice for both clinical and research use. The administration time of the MoCA is quoted as 10 minutes (154), however this can be an unrepresentative estimate in practice, since no two patients are the same and completion can take up to 30 minutes for some individuals after a stroke (21). Therefore, this can make completion of CSIs a challenge in acute medical settings, where time is limited, caseloads are high, and other investigations are prioritised.

If a suitable short test does not already exist, one possibility is to derive one from an established, validated test. The goal of any shortened version is to increase feasibility and acceptability of testing, while maintaining classification accuracy. For tests comprising many items, one may choose to retain only those items that have a high discriminative value, and discard those which do not. Discriminative ability will depend on the purpose and the clinical population in which the test will be used. There may be other reasons to derive a short form of course; one may want to discard items which are not suitable for individuals who have hearing or visual impairment.

Shortened versions of popular scales can be found across different areas, for example there are short forms of the Beck Depression Index (155), Stroke Impact scale (156) and Barthel Index (157). The same is true for popular cognitive tests and shortened versions of the MoCA (SF-MoCA) have been described in the literature (158). As discussed in Chapter 1, researchers in the National Institute of Neurological Disorders and Stroke and the Canadian Stroke Network (NINDS-CSN) met in 2006 to produce test protocols to assess vascular cognitive impairment (110). The 5-minute protocol that they recommended consists of subtests taken from the MoCA and is therefore a shortened version. The number of shortened variations of the MoCA is currently unknown, and no paper has compared all SF-MoCAs against each other using the same cohort. Despite much interest in a SF-MoCA, validity should not be assumed. Even if the process of developing the shortened version is robust and accuracy metrics favourable, it is still necessary to externally validate the test in an independent sample and across different settings.

I first aimed to carry out a systematic review to determine the number of different SF-MoCA versions used across the literature and to collate their published evidence on test accuracy for detecting cognitive impairment. The second aim was to carry out an external validation of the SF-MoCA versions using two independent data sets.

3.2 Methods

3.2.1 Systematic review

I carried out a systematic review of the literature, following best practice guidelines in all aspects of design, conduct and reporting: Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies (PRISMA-DTA) (see Appendix 3 for PRISMA checklist) (159). All aspects of screening, data extraction and quality assessment, were performed independently by at least two researchers (Dr Jennifer McDicken, Dr Gareth Blayney or myself), and using a third arbitrator (Dr Terry Quinn) to resolve conflicts as required. The protocol is registered with the research registry database www.researchregistry.com (Unique Identification Number: reviewregistry298).

A Cochrane information specialist assisted with the search strategy. The search terms were developed using a concept-based approach. The first concept of interest was the MoCA and its synonyms, including names of existing short forms known. The second concept was around short-form tests and item reduction. The terms used for the second concept were taken from a previous systematic review addressing shortened versions of the Barthel index (157). The full search strategy is available in Appendix 4. I searched the following multidisciplinary electronic databases from 2005 (year of publication of the original MoCA paper) to April 2017: MEDLINE (Ovid), Embase (Ovid), Health and Psychosocial Instruments (Ovid), PsycINFO (EBSCO), and CINAHL (EBSCO) and applied no language restrictions. In addition to the database search, I screened papers already known to me. I included published conference abstracts in the initial data synthesis in the attempt to identify all shortened versions of MoCA being used across the literature, however I did not assess the quality of reporting or risk of bias in the abstracts, due to insufficient details.

I screened the titles and abstracts generated by the initial searches for relevance and proceeded to full-text review for potentially eligible studies that I checked against the inclusion criteria. I also screened reference lists of included studies and relevant reviews, repeating the process until no new titles were found.

The index test of interest was any SF-MoCA. I defined the SF-MoCA as any test including >1 question from the original MoCA and designed to detect any level of cognitive impairment. I did not include the MoCA-Basic(160) under this definition, since it has different test items designed specifically for individuals who are illiterate or have less than 5 years of education and therefore not classed as a shortened version. I included studies which used a test accuracy design, comparing the SF-MoCA to either clinical diagnosis or another longer cognitive screen (of which I included the full MoCA). Studies in any setting (primary, secondary, community) and for any intended use were also included.

I extracted data to a study specific proforma. I created tables describing the characteristics of included studies, with details of the index tests and the method used to derive the short form. Where accuracy data were not presented in the paper, I created 2 x 2 contingency tables to derive metrics of sensitivity

and specificity (161). I contacted authors to obtain data or clarify methods, where needed. For data synthesis, I grouped studies by disease area and examined the target cognitive syndrome, index test, threshold score and reference standard, to determine whether a meta-analysis could be undertaken.

I assessed methodological quality and risk of bias using the Quality Assessment for Diagnostic Test Accuracy Studies (QUADAS-2) tool (www.bris.ac.uk/quadas/quadas-2) (162) as recommended by the Cochrane Collaboration (https://methods.cochrane.org/sdt/handbook-dta-reviews). QUADAS-2 assesses four key domains: patient selection, application of index test, application of reference standard, and patient flow/timing. I assessed quality of reporting using the dementia-specific extension to the Standards for Reporting of Diagnostic Accuracy (STARDdem) tool (163).

3.2.2 External Validation

Working with a research collaborator (Dr Myzoon Ali [MA]), I examined the psychometric properties of the MoCA and the SF-MoCAs identified using two independent data sets. The first data set included people with a diagnosis of ischaemic stroke (IS), intracerebral haemorrhage (ICH) or transient ischaemic attack (TIA) and was obtained from the Virtual International Stroke Trials Archive (VISTA), a not-for-profit repository of anonymized data from stroke trials or observational cohorts (164). I included data from any VISTA studies containing MoCA assessment and an appropriate reference standard. Recognizing the difficulty of applying a dementia label in an acute stroke setting, and to align with the systematic review, I included data sets where the comparator was another multidomain cognitive assessment (other than the MoCA). All participants in this data set had MMSE data, so this was used as the reference standard.

The second data set was provided from a memory clinic (the Walton Centre, Liverpool, UK). This data set was included to examine whether test properties differed in another population. It provided a representative cohort as the data was routinely collected as part of clinical practice. The data set included two patient cohorts. The first cohort covered new patient referrals consecutively recruited between September 2009-March 2011 and the second recruited

between June 2015-May 2016. The MoCA was administered in clinic and then a clinical diagnosis (according to DSM-IV criteria) was made by a clinician, blinded to the MoCA score (165, 166).

The data sets contained individual patient level data on each scored item of the MoCA and a reference standard comparator. This data was used to score each SF-MoCA identified through the systematic review. Therefore, there were numerous index tests (each differing SF-MoCA version). To align with the systematic review, the reference standard was clinical diagnosis of dementia or scores from an alternative multi-domain cognitive assessment.

The MoCA has individual test items that are grouped into 8 categories/domains (visuospatial/executive, naming, memory, attention, language, abstraction, delayed recall, orientation). In this work where I refer to 'each MoCA item' this does not necessarily refer to each individual subdomain question. For some domains (e.g., visuospatial/executive) each individual question was examined separately, whereas other domains (e.g., animal naming) were grouped together. This resulted in 12 items: trails, cube copy, clock-draw, naming (3 sub-items), digit list (2 sub-items), list of letters, serial 7 subtraction (5 sub-items), sentence repetition (2 sub-items), fluency, abstraction, delayed recall (5 sub-items), orientation (6 sub-items).

I examined the floor (number and percentage of participants scoring the minimum score (0)) and ceiling effects (number and percentage of participants with the maximum score) of each of the 12 MoCA items. I used Cronbach's alpha as a measure of internal consistency (reliability) of the MoCA. To identify potentially redundant items in the MoCA, I used Spearman coefficient to describe the correlation between individual test items and total MoCA score (item-total correlation) and then described the effect on internal consistency if that item was removed. If internal consistency of the complete scale is unchanged when an item is removed, it suggests that the item is not contributing independent of other items and could potentially be removed without compromising test performance. I described rank correlation of each MoCA item with another (item-item correlations).

MA formatted the data and ran the validation analyses due to data access. My role was to assist with the design of the analysis plan and interpretation of the statistical read-out. Validation analyses were undertaken using SAS version 9.4 (SAS Institute, Cary) software.

Principal component analysis and exploratory factor analysis were used to assess the underlying structure of the MoCA, i.e., how many differing constructs were being assessed by the scale. Factor loadings are standardized regression coefficients; high loadings were defined as >0.7. MA also ran the following analyses: correlation of each SF-MoCA with the original MoCA and sensitivity, specificity, negative/positive predictive values (NPV/PPV) and classification accuracy for each SF-MoCA (using recommended threshold scores from the literature) against clinical reference standard.

3.3 Results

3.3.1 Systematic review

The search retrieved 710 titles. Once duplicates were removed, 578 titles were screened, and 140 full text articles reviewed. Three additional papers were identified separately, resulting in 21 studies being eligible and included (18 full papers and three conference abstracts (102, 147, 158, 166-182)) (Figure 3-1). The number of participants included in studies ranged N=59 to N=1850 (N=6477 in total).

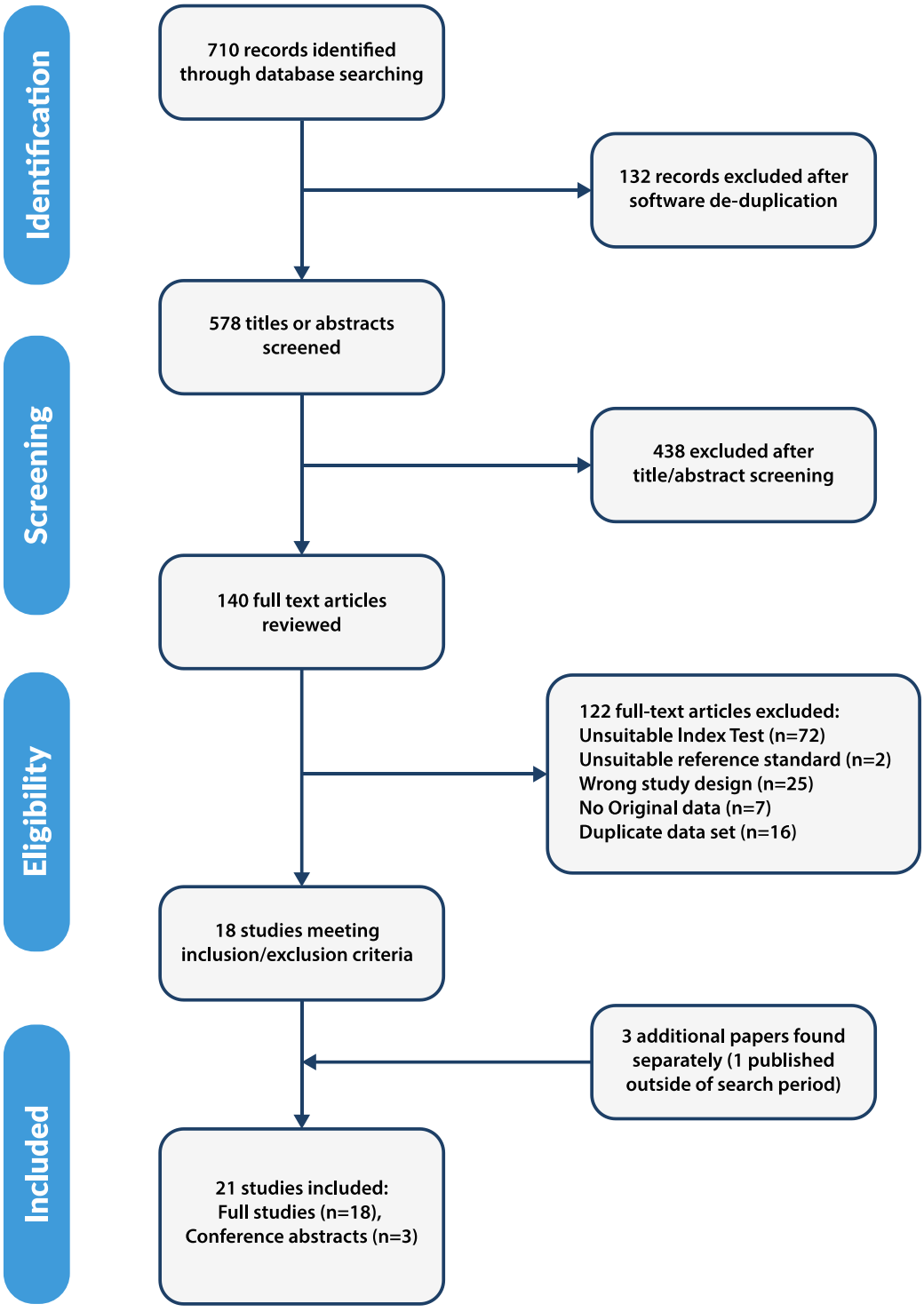


Figure 3-1 PRISMA flow diagram

There were 13 different published SF-MoCAs, each with differing content or scoring. The number of test items included ranged 2-8 and total scores ranged 8-

30. All SF-MoCAs retained the delayed recall item, whilst the 'lion' and 'camel' items within the animal naming section were omitted by all versions (Figure 3-2). The most frequently described SF-MoCA was the 5-minute protocol recommended by the National Institute of Neurological Disorders and Stroke and the Canadian Stroke Network (NINDS-CSN) (n=7 papers). Three SF-MoCA versions, with the same content and scoring, were given different names by different authors, for example the 'New short MoCA' and 'Mini-MoCA' both comprised trails, cube copy, delayed recall, fluency, and abstraction. There were also SF-MoCA versions with different items but sharing the same title, for example two 'Mini-MoCAs', both with a total score of 10, but comprising different items (Table 3-1).

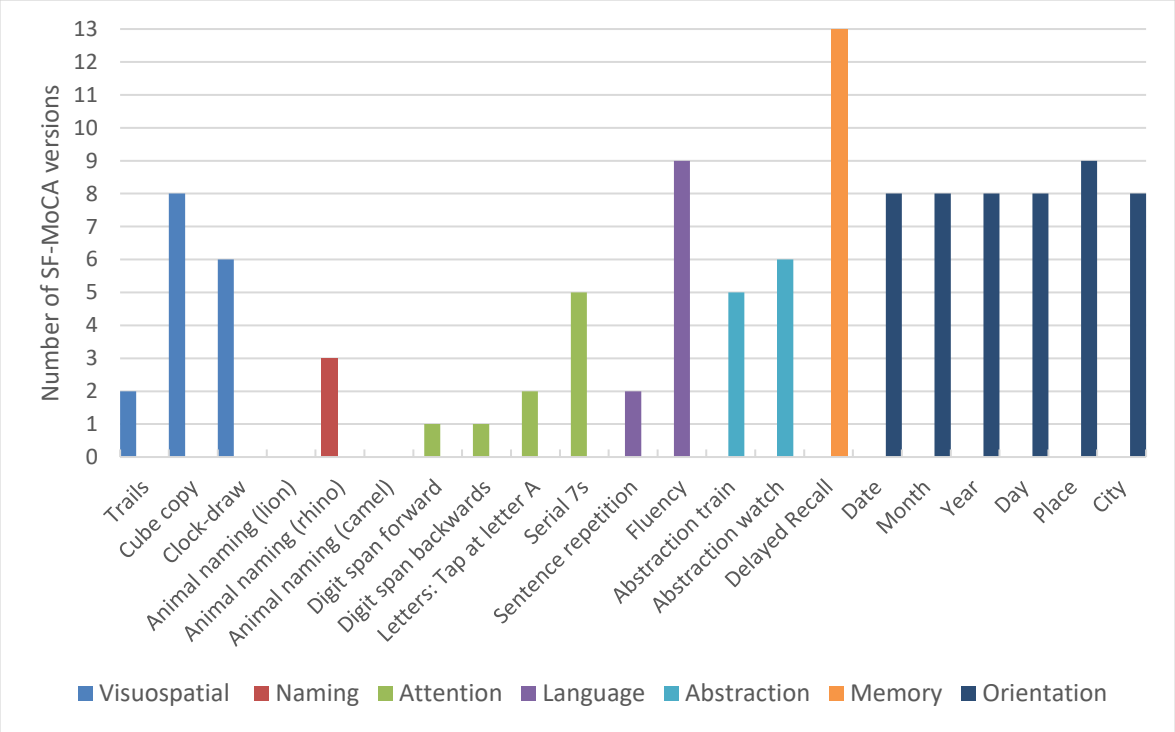


Figure 3-2 Items retained across different SF-MoCA versions

Different methods were used to derive the items retained in each SF-MoCA: regression (170, 173, 182), item response theory (IRT) and computerised adaptive testing (CAT) (158), Cramer's V (168), receiver operating characteristic (ROC) analysis (147, 181), random forest analysis (183), z scores (171), analysis of covariance (ANCOVA) (175) and expert opinion (110). Two papers removed all visual MoCA items, but were designed for different purposes; One for use with visually impaired individuals (176) and the other designed with the purpose of

being deliverable over the telephone (102). The other included studies chose to validate existing SF-MoCAs.

Scoring and content of the MoCA 5-min protocol(167) was not consistent with the original MoCA; an extra scoring component was added for immediate recall and the verbal fluency component was altered to animal fluency rather than letter and used a scoring system of 0.5 points for each word. Some shortened versions split up the MoCA domains, for example retaining one out of three animals (rhino) for the naming section in the MoCA reduced (173) and S-MoCA (158). Threshold scores used to categorise patients with cognitive impairment varied across the papers, even where the same short form was used (Table 3-1) and some studies did not state a threshold score. I did not attempt to pool data to create summary estimates of SF-MoCA test accuracy, due to the significant heterogeneity in test content, thresholds, and populations.

Table 3-1 Content and scoring of published SF-MoCAs

Name of short-form MoCA	Number of items	Test items and scoring	Threshold scores suggested by each paper, with cognitive syndrome
Abbreviated MoCA (147)	2	Clock draw (3) Delayed recall (5) *Total score: 8*	Panenkova 2016 (MCI): <4
NINDS-CSN 5-min protocol (110) Telephone-MoCA short (102)	3	Delayed recall (5) Fluency (1) Orientation (6) *Total score: 12*	Lim 2017 (Dementia): <7 Bocti 2013a (CI), Cameron 2016 (CI), Pendlebury 2013 (MCI): <10 Kaur 2013 (CI): <11 Dong 2015 (CI): <13 Lin 2016 (CI)
Mini-MoCA (170, 182)	3	Clock-draw (3) Delayed recall (5) Abstraction (2) *Total score: 10*	Mai 2013 (CI): <7
SF-MoCA (168)	3	Delayed recall (5) Serial 7s (3) Orientation (6) *Total score: 14*	Horton 2015: <9 (AD) or <12 (MCI)
MoCA 5-min protocol (167)	4	Immediate recall (5) Delated recall (10)* Fluency (9)** Orientation (6) *Total score: 30* **2 points for each free recall, 1 point for cued recall* ***Animal Fluency, 0.5 point for each word (max 9 points)*	Wong 2015 (CI): <15
MoCA reduced (173)	4	Clock-draw (3) Animal naming (rhino) (1) Delayed recall (5) Orientation (6) *Total score: 15*	Cecato 2016 (Dementia): <9
Four-item mini-MoCA (171)	4	Cube copy (1) Delayed recall (5) Serial 7s (3) Fluency (1) *Total score: 10*	Bocti 2012 (MCI): <9

New short MoCA (171) Mini-MoCA (174)	5	Trails (1) Cube copy (1) Delayed recall (5) Fluency (1) Abstraction (2) *Total score: 10*	Bocti 2013, Campbell 2016 (CI): <7
5-min MoCA (172)	5	Clock-draw (3) Delayed recall (5) Serial 7s (3) Fluency (1) Orientation (6) *Total score: 18*	Dong 2015 (CI): <13
EM-MoCA (175)	7	Trails (1) Cube copy (1) Clock-draw (3) Delayed recall (5) Fluency (1) Abstraction (2) Orientation (6) *Total score: 19*	Freitas 2018 (CI): <17
MoCA reduced (173)	7	Animal naming (rhino) (1) Delayed recall (5) Tap at letter A (1) Sentence repetition (2) Fluency (1) Abstraction (2) Orientation (6) *Total score: 18*	Cecato 2016 (MCI): <14
S-MoCA (158)	8	Trails (1) Clock-draw (3) Animal naming (rhino) (1) Delayed recall (5) Serial 7s (3) Fluency (1) Abstraction (watch) (1) Orientation (place) (1) *Total score: 16*	Roalf 2017, Larner 2017 (Dementia): <12
MoCA-Blind Telephone-MoCA (102)	8	Delayed recall (5) Digit span (2) Tap at letter A (1) Serial 7s (3) Sentence repetition (2) Fluency (1) Abstraction (2) Orientation (6) *Total score: 22*	Wittich 2010 (Dementia): <19 Pendlebury 2013 (multi-domain MCI): <18

Abbreviations: AD, Alzheimer's Disease; CI, Cognitive impairment; EM-MoCA, Esclerose múltipla (Multiple sclerosis in Portuguese) Montreal Cognitive Assessment; MCI, Mild Cognitive Impairment; S-MoCA, Short form MoCA; T-MoCA, Telephone Montreal Cognitive Assessment; NINDS-CSN; SF-MoCA, Short form MoCA.

3.3.1.1 Cognitive impairment post-stroke

Nine studies (N=2,514 participants) used six different SF-MoCAs to evaluate cognition post-stroke (Table 3-2). Settings and timing of assessment using the index test (SF-MoCA) post-stroke varied: the majority of papers administered the tests more than 3 months post stroke, one paper used the SF-MoCA in the acute period following stroke (less than or equal to 2 weeks) (169), and two papers in the setting of a stroke prevention clinic (therefore, some non-stroke patients were also included) (170, 182).

Different studies used a SF-MoCA to target varying severities of cognitive impairment, with accuracy data available for single and multi-domain mild cognitive impairment combined through to dementia. The reported accuracy varied across the included studies: median sensitivity was 0.88 (range: 0.70-1.00) and median specificity was 0.70 (range: 0.39-0.92).

3.3.1.2 Cognitive impairment in older adults

Eight studies (N=4,367 participants) used seven SF-MoCAs to evaluate cognition in older adults (Table **3-3**). Once again, different severities of cognitive impairment were targeted. In studies where a SF-MoCA was used to identify dementia in older adults (N=7), median sensitivity was 0.87 (range: 0.62-0.98) and specificity was 0.87 (range: 0.07-0.98). In studies where a SF-MoCA was used to identify MCI (N=3), median sensitivity was 0.84 (range 0.82-0.89) and specificity was 0.72 (range 0.64-0.85).

3.3.1.3 Cognitive impairment in other conditions

Four studies (N=461 participants) used two SF-MoCAs to evaluate cognition in the context of other health conditions: two in multiple sclerosis (MS), one in heart failure (HF) and one in Parkinson's disease (PD) (Table 3-4). Both sensitivity and specificity were higher in the two MS studies (175, 177), compared to the HF and PD studies.

Heterogeneity across the conditions, index tests, thresholds and reference standards precluded a meta-analysis being undertaken.

3.3.1.4 Quality Assessment and study reporting

Using QUADAS-2, one study was considered to have a low risk of bias in all four areas(166). Potential for bias in the other studies was generally around patient selection (n=17) (inappropriate exclusions and non-consecutive samples), use of index test (n=11) (no pre-specified cut-off), and the timing between the index test and reference standard not reported or ambiguous (n=9) (Figure 3-3). Eight papers were of particular concern (rated high or unclear risk of bias across three areas). Study reporting was variable, and no study reported all items recommended in STARDdem guidance (Table 3-5).

Table 3-2 Stroke Papers (N=9)

Study	Participants (n)	Target condition	Index test	Setting/Timing	Reference standard	Threshold score	Sensitivity	Specificity
Bocti 2013	386	CI	New Short MoCA	3 months	MoCA	<7/10	91%	83%
			NINDS-CSN 5 min protocol			<10/12	87%	74%
Campbell 2016	72	CI	Mini-MoCA	Rehab unit	Cognistat	<7/10	93%	92%
Davies 2011	102	CI	Mini-MoCA	Stroke prevention clinic	MoCA	not reported	not reported	not reported
Dong 2015 (CA)	327	CI	5 min MoCA	3-6 months	NPB	<13/20	70%	87%
Lim 2017	308	Dementia	NINDS-CSN 5 min protocol	≤2 weeks	NPB	<7/12	82%	67%
Lin 2016	83	CI	NINDS-CSN 5 min protocol	3-18 months	MDT Consensus	<15/30	81%	55%
Mai 2013	102	CI	Mini-MoCA	Stroke prevention clinic	MoCA	<7/10	99%	78%
Pendlebury 2013		Single & multi-domain MCI	T-MoCA	>1-year post stroke	NPB	<19/22	89%	46%
			T-MoCA Short (NINDS-CSN)			<10/12	96%	39%
	68	Multi-domain MCI	T-MoCA			<18/22	100%	52%
			T-MoCA Short (NINDS-CSN)			<10/12	83%	48%
Wong 2015	104	CI	MoCA 5 min protocol	39 days	CDR	<15/30	84%[a]	73%[a]

[a]Data obtained through contacting the author.
Abbreviations: CA, conference abstract; CDR, Clinical dementia rating; CI, Cognitive impairment; MCI, Mild cognitive impairment; MDT, Multidisciplinary team; MoCA, Montreal Cognitive Assessment; NINDS-CSN, National Institute of Neurological Disorders and Stroke and the Canadian Stroke Network; NPB, Neuropsychological battery; T-MoCA, Telephone Montreal Cognitive Assessment

Table 3-3 **Older Adults Papers (N=8)**

Study	Participants (N)	Target condition	Index test	Reference standard	Threshold score	Sensitivity	Specificity
Bocti 2012 (CA)	341	MCI vs HC	Mini-MoCA	MDT workup including MoCA	<9/11	84%	85%
Horton 2015	Derivation Group =317	AD vs MCI+HC	SF-MoCA	MDT Consensus	Unknown	95%[a]	87%[a]
	Validation Group = 91				unknown	80%[a]	95%[a]
Cecato 2016	Total = 136 AD = 53 MCI = 44 HC = 39	AD vs MCI	Reduced MoCA	DSM IV	<8.5/18	85%	87%
		MCI vs HC			<13.5/18	82%	72%
Larner 2017	Cohort 1: 150	Dementia vs MCI	S-MoCA	DSM IV	<12/16	94%	25%
	Cohort 2: 260					98%	7%
Panenkova 2016	540	MCI	Abbreviated MoCA	MoCA*	<4/8	89%	64%
Roalf 2017	1850	All cause dementia vs HC	s-MoCA	DSM IV	<12/16	62%	86%
Wittich 2010	277	AD	MoCA-Blind	NPB	<18/22	87%	98%
Xu 2016	405	CIND, dementia	NINDS-CSN 5 min protocol	MDT Consensus	Not reported	Not reported	Not reported

*Defined as one standard deviation (SD) below the norm.
[a]Data obtained from ROC curve.

Abbreviations: AD, Alzheimer's Disease; CI, Cognitive impairment; CIND, Cognitive impairment no dementia; DSM, Diagnostic and Statistical Manual; HC, Healthy Control; MCI, Mild Cognitive Impairment; MDT, Multidisciplinary team; MoCA, Montreal Cognitive Assessment; NINDS-CSN, National Institute of Neurological Disorders and Stroke and the Canadian Stroke Network; NPB, Neuropsychological battery; s-MoCA, short form MoCA.

Table 3-4 Cognitive Impairment in other disease areas' Papers (N=4)

Study	Participants (n)	Disease Area	Index test	Reference standard	Threshold score	Sensitivity	Specificity
Cameron 2016	221	HF	NINDS-CSN 5-min protocol	MoCA	<10/12	89%	71%
Dong 2015	101	PD	NINDS-CSN 5-min protocol	CDR	<9/12	77%	78%
Freitas 2018	59	MS	EM-MoCA	Formal neuropsychological evaluation	<17/19	94%	87%
Kaur 2013	80	MS	NINDS-CSN 5-min protocol	Not reported	<10.5/12	97%	90%

Abbreviations: CDR, Clinical Dementia Rating; EM-MoCA, Multiple Sclerosis (Portuguese)-MoCA; HF, Heart Failure; MoCA, Montreal Cognitive Assessment; NINDS-CSN, National Institute of Neurological Disorders and Stroke and the Canadian Stroke Network; MS, Multiple Sclerosis; PD, Parkinson's Disease.

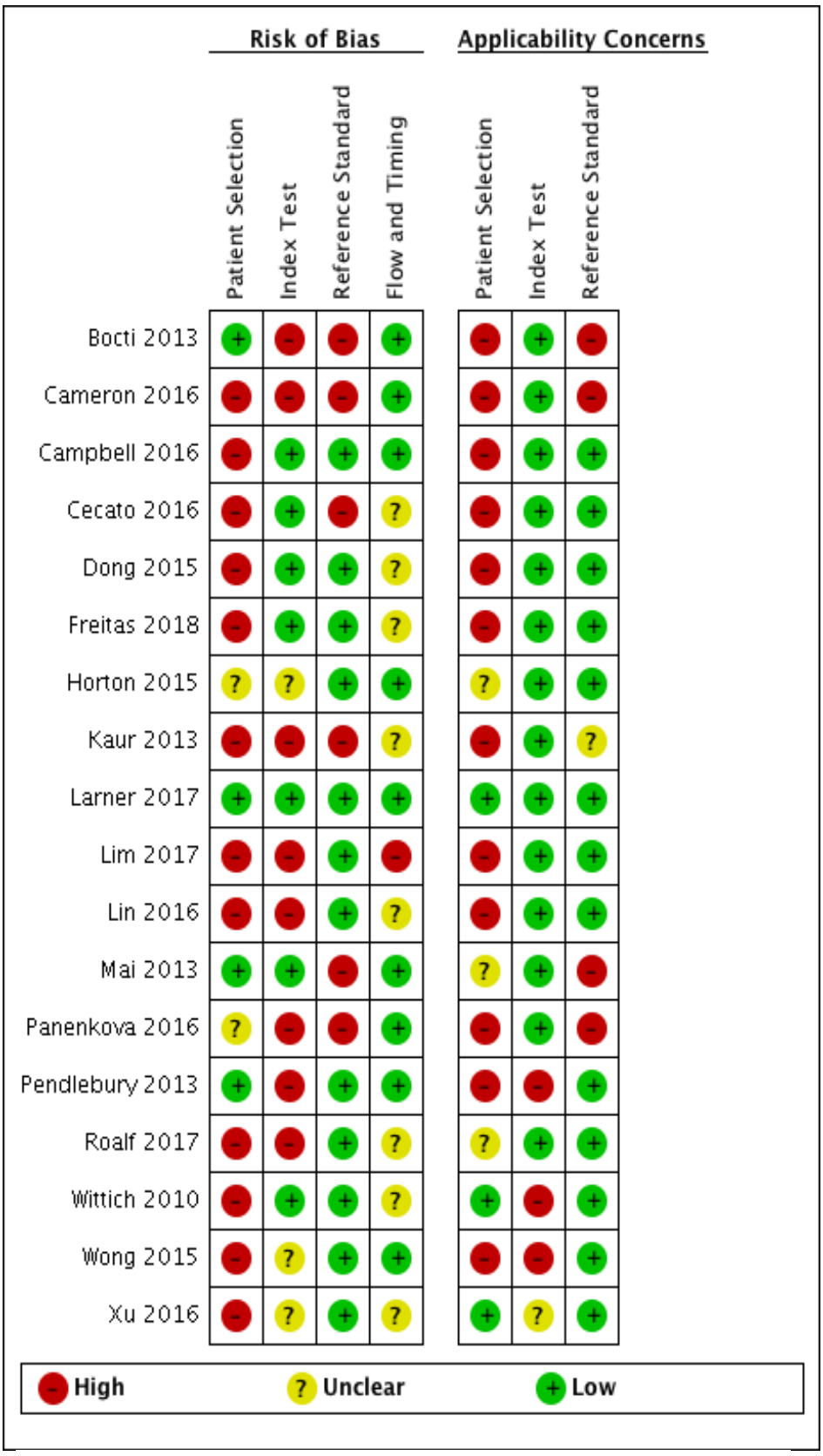


Figure 3-3 Risk of bias judgements using QUADAS-2

QUADAS-2 criteria are detailed in Appendix 6.

Table 3-5 Quality of reporting using STARDdem

	STARDdem	Bocti 2013	Cameron 2016	Campbell 2015	Cecato 2016	Dong 2015 (PD)	Freitas 2018	Horton 2015	Kaur 2013	Larner 2017	Lim 2017	Lin 2016	Mai 2013	Panenkova 2016	Pendlebury 2013	Roalf 2016	Wittich 2010	Wong 2015	Xu 2016
Title/Abstract	1	y	y	y	y	y	y	y	y	y	y	y	y	n	y	y	y	y	y
Introduction	2	y	y	y	y	y	y	y	y	y	y	y	y	n	y	y	y	y	y
Methods	3	y	y	y	y	n	y	y	n	y	y	y	y	y	y	y	y	y	y
	4	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y
	5	y	n	y	n	y	y	n	n	y	y	n	y	y	y	n	n	y	y
	6	y	y	y	y	y	y	y	n	y	y	y	y	y	y	y	y	y	y
	7	y	y	y	y	y	y	y	n	y	y	y	n	y	y	y	y	y	y
	8	y	n	y	y	y	y	y	n	y	y	y	y	y	y	y	y	y	y
	9	y	y	y	y	y	y	y	n	y	n	y	y	n	n	y	y	y	y
	10	n	y	y	y	y	y	n	n	n	n	y	n	n	n	n	n	n	y
	11	n	y	n	y	y	y	n	n	n	n	y	y	n	n	y	n	n	n
Stats	12	n	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	
	13	n	n	n	n	n	n	y	n	n	n	n	n	n	n	y	n	y	n
Results	14	y	n	y	y	n	y	y	n	y	n	n	n	n	y	y	n	y	n
	15	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y
	16	y	n	n	y	y	n	n	n	y	y	n	y	y	n	n	n	y	n
	17	y	y	y	n	n	n	n	n	y	n	n	y	y	y	n	n	y	n
	18	y	y	y	y	y	y	y	n	y	y	y	y	y	y	y	y	y	y
	19	y	y	y	y	y	y	y	n	y	n	n	y	n	y	y	y	y	n
	20	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
Estimates	21	n	n	n	y	y	y	y	n	y	n	y	n	y	y	y	n	y	y
	22	y	y	n	n	n	n	n	n	y	y	y	y	n	n	n	n	y	n
	23	n	y	n	y	n	y	y	n	n	n	y	y	n	y	y	n	y	y
	24	n	y	n	n	n	n	y	n	n	n	n	n	n	n	y	n	y	n
Discussion	25	n	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y

Reporting of each STARDdem item: Y = yes; N = no. STARDdem items available in Appendix 5.

3.3.2 External Validation

The stroke data set included 787 patients with a median age of 70, median NIHSS of 4, median MoCA of 21, and 289 (37%) had dementia or post stroke cognitive impairment (Table 3-6). Assessments were performed in the acute period (first weeks) following stroke. The memory clinic data set included 410 patients, with median age of 60 (IQR 51, 58), median MoCA of 23 (IQR 18, 26) and 79 (19%) had dementia.

Table 3-6 Characteristics of the stroke sample

Variable	**IS (N=728)**	**ICH (N=59)**	**Total (N=787)**
Age (Median, IQR)	70 (62, 78)	67 (56, 77)	70 (62, 78)
Sex, male N (%)	421 (58%)	45 (%)	466 (59%)
Baseline NIHSS *Missing data (N)*	4 (2,6) *19*	5 (3,10) *4*	4 (2,6) *23*
MoCA (Median, IQR)	21 (16, 25)	21 (15, 25)	21 (16, 25)
MMSE (Median, IQR)	26 (22, 28)	26 (21, 28)	26 (22, 28)
TIA (N, %)	15 (2%)	-	15 (2%)
Hypertension (N, %)	502 (69%)	46 (78%)	548 (70%)
Diabetes (N, %)	275 (38%)	18 (31%)	293 (37%)

Abbreviations: ICH, Intracerebral haemorrhage; IQR, Interquartile range; IS, Ischaemic stroke; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment, NIHSS, National Institute of Health Stroke Scale; TIA, Transient ischaemic attack.

The validation work focused on 12 of the 13 differing versions of SF-MoCA found in the systematic review. The 13th version includes an additional scoring component which is not recorded as part of the original MoCA (immediate recall)(184) and I could not retrospectively score it. The 12 index tests resulted in varying numbers of patients being assigned as impaired (test positive). In stroke participants the percentage of those screening positive ranged 37-92%, and in the memory clinic participants it ranged 26-89% (Figure 3-4).

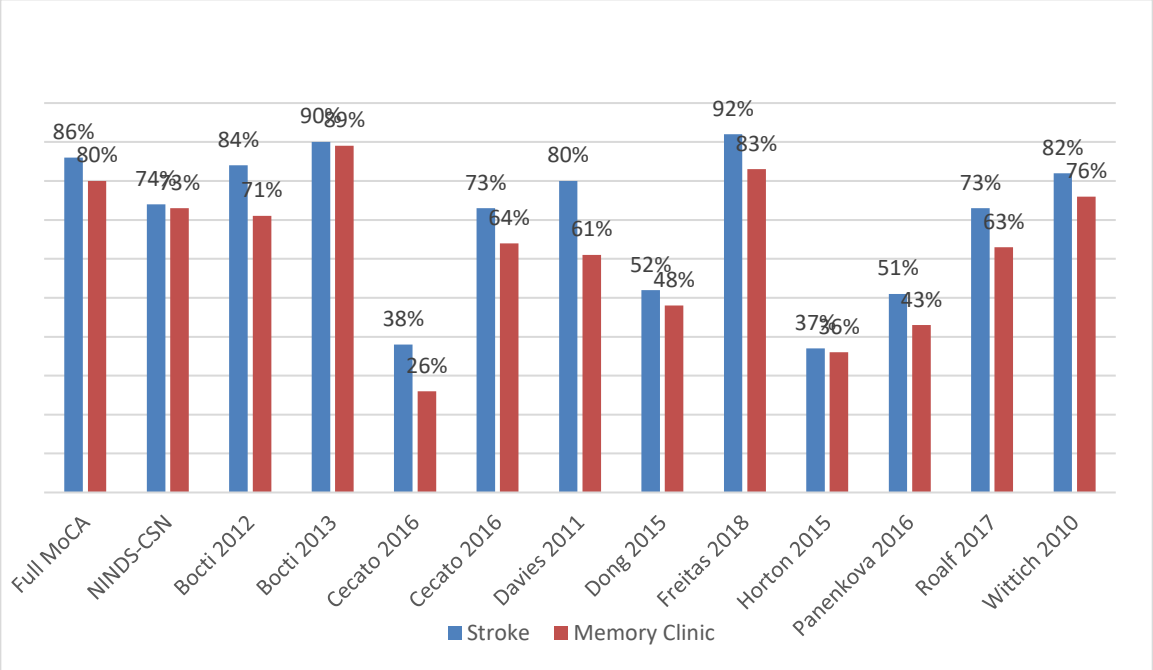


Figure 3-4 Comparison of percentage of patients classified impaired on the MoCA and the 12 different SF-MoCA versions across the two data sets.

SF-MoCA versions are listed under the author that originally published the scale

In the memory clinic cohort, the task with the highest floor effect was verbal fluency (56%), whereas the task with the highest ceiling effect was animal naming (78%). In the stroke cohort, the task with the highest floor effect was cube copy (71%), whereas sentence repetition had the highest ceiling effect (86%) (Table 3-7).

Table 3-7 Floor & Ceiling effects of each item

Item (score range)	Memory clinic (N=407)			Stroke (N=787)		
	Missing (N)	Floor (score of 0)	Ceiling (scored full marks)	Missing (N)	Floor (score of 0)	Ceiling (scored full marks)
Trails (0-1)	3	202 (50%)	205 (50%)	3	423 (54%)	361 (46%)
Cube copy (0-1)	3	164 (40%)	243 (60%)	3	553 (71%)	231 (29%)
Clock-draw (0-3)	3	8 (2%)	277 (68%)	3	93 (12%)	326 (29%)
Naming (0-3)	3	2 (<1%)	316 (78%)	0	51 (6%)	430 (55%)
Digit span (0-2)	3	30 (7%)	273 (67%)	0	41 (5%)	514 (65%)
Letters (0-1)	3	109 (27%)	298 (73%)	0	248 (32%)	539 (68%)
Serial 7s (0-3)	3	28 (7%)	255 (63%)	0	63 (8%)	442 (56%)
Repeat sentences (0-2)	3	41 (10%)	253 (62%)	0	17 (2%)	680 (86%)
Fluency (0-1)	3	226 (56%)	181 (44%)	0	293 (37%)	494 (63%)
Abstraction (0-2)	3	83 (20%)	156 (38%)	0	489 (62%)	85 (11%)
Delayed recall (0-5)	3	154 (38%)	34 (8%)	1	312 (40%)	46 (6%)
Orientation (0-6)	3	2 (<1%)	218 (54%)	0	7 (1%)	468 (59%)

No reasons were available for missing data. Floor effects refer to the number of participants scoring the lowest value (0). Ceiling effects refer to the number of participants scoring the highest value for the test item.

Cronbach's alpha for the full MoCA was 0.88 in the stroke data set. This value decreased if any of the items were deleted. Cronbach's alpha for the full MoCA was 0.82 in the memory clinic and similarly this decreased for all items when deleted, apart from 'letters – tapping for letter A' attentional task, where alpha stayed the same (Table 3-8). In the stroke data set, clock drawing was the single item most correlated with total score, while sentence repetition was least correlated. In the memory clinic data set, orientation questions were most correlated and 'letters – tapping for letter A' was least correlated (Table 3-8).

Table 3-8 Table of correlation with full MoCA and Cronbach's alpha if item deleted

Item	Stroke		Memory Clinic	
	Correlation with Total MoCA	Alpha if item deleted	Correlation with Total MoCA	Alpha if item deleted
Trails	0.60	0.85	0.54	0.80
Cube	0.51	0.86	0.47	0.81
Clock	0.71	0.85	0.53	0.80
Naming	0.58	0.85	0.39	0.81
Digits	0.55	0.87	0.45	0.81
Letters	0.55	0.86	0.36	0.82
Subtraction	0.64	0.85	0.54	0.80
Repetition	0.31	0.87	0.45	0.81
Fluency	0.55	0.86	0.40	0.81
Abstraction	0.42	0.86	0.49	0.81
Recall	0.57	0.86	0.46	0.81
Orientation	0.60	0.85	0.55	0.80

Table 3-9 Rank correlation between items of the MoCA

	TRAILS	CUBE	CLOCK	NAMING	DIGITS	LETTERS	SUBTRACTION	REPEAT	FLUENCY	ABSTRACTION	RECALL	ORIENTATION
TRAILS	1	0.48	0.50	0.39	0.38	0.38	0.46	0.17	0.38	0.30	0.38	0.40
CUBE	0.40	1	0.46	0.31	0.30	0.28	0.35	0.14	0.29	0.35	0.36	0.31
CLOCK	0.38	0.38	1	0.53	0.46	0.44	0.54	0.27	0.43	0.31	0.45	0.51
NAMING	0.23	0.24	0.23	1	0.39	0.40	0.41	0.20	0.39	0.28	0.37	0.42
DIGITS	0.27	0.17	0.19	0.23	1	0.33	0.49	0.27	0.31	0.26	0.36	0.35
LETTERS	0.24	0.16	0.23	**0.15**	0.30	1	0.42	0.21	0.36	0.28	0.34	0.42
SUBTRACTION	0.37	0.33	0.40	0.38	0.31	0.27	1	0.22	0.44	0.30	0.38	0.51
REPEAT	0.27	0.15	0.34	**0.22**	0.36	0.23	0.26	1	0.18	0.11	0.21	0.20
FLUENCY	0.30	0.24	0.21	**0.14**	0.33	0.23	0.26	0.26	1	0.25	0.44	0.20
ABSTRACTION	0.29	0.32	0.28	0.29	0.23	0.21	0.32	0.33	0.19	1	0.31	0.27
RECALL	0.32	0.30	0.28	0.19	0.24	**0.16**	0.26	0.23	0.20	0.36	1	0.41
ORIENTATION	0.38	0.31	0.46	0.23	0.26	0.19	0.39	0.29	0.26	0.31	0.45	1

Grey cells = stroke data, white cells = memory clinic data.

No items were highly correlated (>0.8). Those items where correlation was not significant at <0.0001 are in bold type

In both data sets, correlation of one item with another did not suggest a redundant item (no correlation greater than 0.6); (Table 3-9). Exploratory factor analyses and principal components analysis suggested a unidimensional scale, with only clock drawing highly loaded (0.76) in the stroke data set and no items highly loaded in the memory clinic data (Table 3-10).

Table 3-10 Factor loadings for MoCA items

	Stroke	Memory Clinic
TRAILS	0.649	0.597
CUBE	0.556	0.527
CLOCK	0.762	0.600
NAMING	0.628	0.434
DIGITS	0.590	0.488
LETTERS	0.587	0.400
SUBTRACTION	0.697	0.609
REPEAT	0.331	0.501
FLUENCY	0.594	0.441
ABSTRACTION	0.446	0.535
DELAYED RECALL	0.607	0.525
ORIENTATION	0.653	0.623

The test accuracy of the published SF-MoCAs varied when assessed in the independent data sets (Table 3-11). Test accuracy of the full MoCA was included for comparison at the conventional threshold of less than 26. Accuracy of MoCA was similar in the two data sets: sensitivity: 1.00 in both, specificity: 0.22 in stroke, and 0.26 in the memory clinic data set. In both data sets, the SF-MoCA versions were highly correlated with the full MoCA (all were greater than 0.80).

In the stroke trial data set, median sensitivity was: 1.00 (range: 0.80-1.00); median specificity: 0.39 (range: 0.14-0.87), PPV: 0.49 (range: 0.40-0.77), NPV: 1.00 (range: 0.88-1.00). In the memory clinic data set, median sensitivity: 0.96 (range: 0.72-1.00); median specificity: 0.40 (range: 0.14-0.86), PPV: 0.28 (range: 0.24-0.55), and NPV: 0.98 (range: 0.93-1.00). In both data sets Cecato's MoCA reduced (AD) had the lowest sensitivity: 0.80 in stroke and 0.72 in memory clinic. Classification accuracy was highest using Cecato's AD version in the memory clinic cohort and using Horton's version in the stroke cohort.

Table 3-11 Accuracy of each SF-MoCA in the external validation

Test	Threshold score	Correlation with MoCA	Sensitivity	Specificity	Positive predictive value	Negative positive value	Accuracy
Stroke patients							
Full MoCA	<26/30	-	1.00	0.22	0.43	1.00	0.51
NINDS-CSN	<10/12	0.88	0.98	0.40	0.49	0.97	0.61
BOCTI 2012	<9/10	0.92	1.00	0.16	0.41	1.00	0.47
BOCTI 2013	<7/10	0.92	1.00	0.25	0.43	1.00	0.53
CECATO 2016 (AD)	<9/15	0.98	0.80	0.87	0.77	0.88	0.84
CECATO 2016 (MCI)	<14/18	0.95	0.98	0.42	0.49	0.97	0.62
DAVIES 2011	<7/10	0.92	0.99	0.32	0.46	0.99	0.57
DONG 2015	<13/20	0.96	0.92	0.72	0.65	0.94	0.79
FREITAS 2018	<17/19	0.97	1.00	0.14	0.40	1.00	0.45
HORTON 2015	<9/14	0.91	0.82	0.89	0.82	0.90	0.87
PANENKOVA 2016	<4/8	0.89	0.87	0.70	0.62	0.90	0.76
ROALF 2017	<12/16	0.97	1.00	0.42	0.50	1.00	0.64
WITTICH 2010	<19/22	0.97	1.00	0.28	0.45	1.00	0.55
Memory clinic patients							
Full MoCA	<26/30	-	1.00	0.26	0.24	1.00	0.40
NINDS-CSN	<10/12	0.87	0.95	0.33	0.25	0.97	0.45
BOCTI 2012	<9/10	0.91	1.00	0.14	0.27	1.00	0.30
BOCTI 2013	<7/10	0.92	0.96	0.36	0.26	0.97	0.48
CECATO 2016 (AD)	<9/15	0.89	0.72	0.86	0.55	0.93	0.83
CECATO 2016 (MCI)	<14/18	0.95	0.96	0.44	0.29	0.98	0.54
DAVIES 2011	<7/10	0.89	0.94	0.47	0.30	0.97	0.56
DONG 2015	<13/20	0.94	0.91	0.62	0.36	0.97	0.68
FREITAS 2018	<17/19	0.96	1.00	0.22	0.24	1.00	0.37
HORTON 2015	<9/14	0.90	0.78	0.74	0.42	0.93	0.75
PANENKOVA 2016	<4/8	0.84	0.80	0.66	0.36	0.93	0.69
ROALF 2017	<12/16	0.96	0.96	0.46	0.30	0.98	0.56
WITTICH 2010	<19/22	0.97	0.98	0.30	0.25	0.99	0.43

3.4 Discussion

Thirteen cognitive screens purport to be a shortened form of the MoCA. The available SF-MoCAs are not interchangeable as they have differing test items, application, and test properties. The external validation of the SF-MoCA confirmed differences in test properties and across different settings (stroke setting vs memory clinic). In general, the SF-MoCAs had a pattern of high sensitivity and lower specificity, with corresponding high NPV and lower PPV.

In terms of psychometrics, it is debatable whether the MoCA content should be reduced at all. The analyses suggested no obviously redundant item in the original MoCA and do not necessarily favour the creation of a shorter form. Aside from test properties, it is important to acknowledge there are different motivations for a shortened form. In spite of whether one should shorten tests, certain scenarios, such as test administration by telephone or assessing a blind person, necessitate that certain items from the original scale are discarded, effectively creating a short-form assessment.

Various approaches to developing short versions of longer tests are described (185) and the processes used to develop the various published SF-MoCA varied. The process of developing the SF-MoCA should also be evaluated, as this aspect is not captured through QUADAS-2. Modern psychometric approaches (i.e., item response theory (IRT)) are more robust and stringent to classical test theory (CTT). For example, a test may perform well according to CTT, while IRT may reveal problematic items. Only one SF-MoCA was developed through this method (158). The SF-MoCA with the greatest validation work (NINDS-CSN 5-min protocol) was derived through expert opinion. While the authors state that aspects such as psychometric qualities, cost, ease of use, availability of multiple forms, deliverable over the telephone, cross-cultural capability, and previous use in VCI were considered in making the recommendations, there is no supportive data of these aspects provided in the paper.

3.4.1 Research in context

In addition to the literature found in the systematic review, three relevant studies have been published more recently. Seven of the SF-MoCAs found in this

review were compared in a large sample of 4,606 participants recruited from Alzheimer disease (AD) centres across the US(186). The study authors concluded that the only short version comparable to the full MoCA was the s-MoCA, created by Roalf et al (158).

Another recently published study used IRT methods to create another SF-MoCA for use in mild cognitive impairment in PD (187). The 8 items included in this SF-MoCA were unique to SF-MoCA versions found in this review, comprising trail making, clock-draw, digit span backwards, serial 7s, repeat sentence (cat), verbal fluency, abstraction (watch) and delayed recall. The team who developed the original MoCA have also recently added a shortened version online, along with accuracy data comparing the test to the full MoCA, however this work is yet to be published. This 5-minute version has the same items as the NINDS-CSN version, however scoring for the fluency item is altered (scored out of four rather than one), so the total score is 15. Including these additional versions brings the total number of SF-MoCA versions to 15. Work is also currently underway to develop alternative versions of the MoCA for those with hearing impairment (188).

There have been limited studies using modern psychometric methods to address properties of the MoCA, but those that have employed these methods provide useful insight. Using the Rasch model, Freitas et al. in one study found an overall good fit of both the items and the person's values (189) and in another study found that 'delayed recall' was the most difficult item, whereas 'orientation' was the easiest (190). IRT methods have also identified tasks which are more or less influenced by education (191). The tasks 'cube copy' and 'clock-draw' (numbers and hands) were found to be less influenced by education.

3.4.2 Strengths and limitations

I acknowledge some limitations in this work. For the systematic review, I was constrained by the methodology and reporting of the original research. Many of the original papers had substantial risk of bias. I adopted an inclusive approach and accepted papers where SF-MoCA was compared to the original MoCA. This is problematic, since the MoCA itself has limitations and is a poor choice of reference standard. In the external validation, I derived SF-MoCA data from the

original MoCA test data. I recognize that the properties of a SF-MoCA may differ if used directly rather than if retrospectively derived. I also used a mixed reference standard of clinical diagnosis or multi-domain assessment. While this could be criticised, this approach is representative of real-world practice where a diagnosis of post stroke dementia is rarely made in the acute period. I also acknowledge that using a neurology-led memory clinic population for validation has some limitations. These patients are likely to be selected and may have already been triaged using a cognitive screen, so brief screens are less useful in this specialist setting. These factors all potentially limit the generalisability of these findings.

The different rates of accuracy found in the validation analyses in comparison with the studies identified in the systematic review are likely due to methodological limitations, differing case-mix, and differing comparator groups. Finally, due to the validation being a secondary analysis, I was unable to adjust the results for education, which would usually be done in practice. Strengths of this work include the use of a comprehensive search strategy allied with comprehensive assessments of reporting and bias. The SF-MoCA tests were validated in a large sample across two settings, one being a real-world sample.

3.4.3 Implications for research and clinical practice

The terminology used to describe the short versions of the MoCA is potentially confusing. Some of the short-form tests were presented under the same name yet contained different items, for example, there were two versions of the "mini-MoCA" (170, 182) and two of the "MoCA reduced" (173). Conversely, some SF-MoCA had identical content and scoring but had a different title, for example, the "new short MoCA" and "mini-MoCA" were the same test (171, 174). Abbreviations also potentially add to the confusion with "MoCA-B" being used to describe both the "MoCA-Basic" and "MoCA-Blind" tests (160, 176). I encourage researchers and clinicians to be explicit about the test content and scoring when using a SF-MoCA.

This work has a number of clinical implications. The SF-MoCAs were used in a variety of patient populations. Many of the populations assessed represented neurodegenerative diseases where patients are likely to have mixed physical and

cognitive impairments (MS, PD and stroke). In these settings, a short cognitive test may have particular utility as patients may struggle to complete a longer assessment (137). More specifically, the choice of SF-MoCA could be tailored, for example, removing the first three questions which require drawing for patients with limb weakness. The MoCA test items that were most discriminating differed between stroke and memory clinic patients. This finding has biological plausibility as the predominant dementia pathologies will also differ in these patient groups, with greater impairment of executive function in the stroke group (120). This suggests that the optimal short form will depend on the population to be tested. The NINDS-CSN 5-minute protocol was recommended specifically for vascular cognitive impairment (VCI); however, this was also the choice of test in papers studying non-vascular groups, e.g., MS and PD populations.

Across the different SF-MoCA versions, a general pattern of high sensitivity and lower specificity was demonstrated, with corresponding high NPV and lower PPV. These results are not surprising since the MoCA was designed to detect mild cognitive impairment and this work had a focus on dementia. The preferred trade-off of test accuracy metrics depends on the purpose of testing and the context of use (161). The test accuracy findings for the SF-MoCAs suggest that some of these CSIs would be useful for ruling out dementia. This means that testing negative for dementia on a SF-MoCA makes it unlikely that a person has dementia. However, with specificity being low across many of the SF-MoCAs, a test positive (or abnormal) result is less helpful and will need to be followed by further assessment, since there will be many false positives. In this regard, the properties of SF-MoCA are similar to the original MoCA, where sensitivity for a dementia diagnosis is around 94% and specificity less than 60% (192).

All SF-MoCAs retained the delayed recall memory item, however the number of other tasks in between where the words are first introduced (immediate recall) and where they are tested (delayed recall) varied. Therefore, a practical, yet important consideration facing clinicians and researchers should be to consider timing between immediate and delayed recall parts of the test. For example, a test only including clock-draw and delayed recall, like the ‘Abbreviated MoCA’ (147), will result in a shorter duration of recall, compared to the full MoCA. This

would make it easier to remember the five words and likely not comparable to the full version.

While short tests have a theoretical advantage of increased feasibility and acceptability, the results from Chapter 2 illustrated that test content was more important than length of test in terms of completion rates. It is also important to acknowledge the limitations of shortening tests. As discussed in Chapter 1, cognitive ability and cognitive impairment exist on a spectrum. Tests with fewer items will inevitably have less conceptual coverage, which could lead to conceptual gaps or the 'ruler being too short'. Shortening a test may therefore risk content validity. Previous work has also demonstrated that there is a trade-off between speed and accuracy in cognitive screening (193). The context of use (targeted concept, targeted population, decision to be informed) should guide the trade-off.

Computerised adaptive testing (CAT) could provide the answer to these issues. A recent study administered the Cambridge Cognition examination (CAMCOG) via CAT and then administered the full CAMCOG supplemented with extra neuropsychological tests (194). Testing time of the CAMCOG through CAT was reduced by 37% or more, in comparison to using the full test and there was excellent agreement between the estimated cognitive ability levels of both approaches. The CAT method provides a promising area for future research.

3.4.4 Conclusion

The cognitive screens named 'mini-MoCA', 'short-form MoCA', '5 minute-MoCA' etc. describe a variety of differing CSIs with differing content and test properties. The psychometric properties of the MoCA do not suggest a preferred content of a shorter version and so choice of SF-MoCA should be based on the context of use. Test accuracy of the various published SF-MoCAs suggest that they may be useful as initial CSIs if the purpose of testing is to rule out dementia. However, such an approach should be prospectively validated in an independent sample before being used in a clinical setting.

4 Accuracy of telephone-based cognitive screening instruments: Systematic review and meta-analysis

4.1 Introduction

The previous chapters have investigated feasibility and accuracy of CSIs which are delivered in person. Chapter 2 illustrated that not all patients can be screened in full, or at all, at an early time-point following a stroke. In addition, hyper-acute stroke units have a fast turnaround, with some patients discharged home within a day. It is therefore likely that these individuals will not have been screened for cognitive impairment. Cognitive screening is also not routinely undertaken in TIA clinics, although a third of patients with a TIA diagnosis have cognitive impairment (195). These scenarios are concerning as without formal screening, problems may go undetected, leaving patients unsupported once they are home. One method to improve rates of cognitive screening would be to remotely screen patients following discharge using a telephone-based cognitive screen but whether this a valid mode of screening is yet to be determined.

Telephone-based assessments of cognition were first published over thirty years ago, offering a practical, time and cost-effective alternative to in-person assessment. Telephone assessments are particularly suited to assessing high volume or geographically dispersed populations. They have been used in seminal trials and observational cohorts, for example the Health and Retirement study (196).

Certain groups are particularly disadvantaged when it comes to accessing healthcare. For instance, people living in remote areas, people with limited mobility or people with chronic health issues may all struggle to attend an assessment centre or clinic (197). With regard to these aspects, feasibility and acceptability may therefore be improved through telephone delivery. However, in a stroke setting, since aphasia rates are reported as one in three in the acute phase (198), a telephone-based test may not be suitable for all.

Despite the convenience of telephone assessments, they are not routinely used in clinical practice (stroke, dementia, or other settings). This may be due to a

general lack of awareness and familiarity of telephonic cognitive screens amongst healthcare professionals or concerns that accuracy is inferior to screening in person. Their use is therefore generally confined to research. Improved feasibility should not be at the cost of poor accuracy. As discussed in Chapter 1, the focus of cognitive screening post-stroke is not necessarily to identify dementia. This is also the case in other neurological conditions with heightened risk of cognitive impairment (e.g., PD and MS). With this in mind, accuracy of screening tools must be demonstrated in identifying milder forms of cognitive impairment, in order for them to be recommended in such settings.

In light of this complexity, a review of telephone-based cognitive tests, and their accuracy, is needed. Previous systematic reviews on this topic (197, 199-201) have described several different tools, yet no quantitative synthesis of their accuracy has been presented. Greater clarity on the accuracy of telephone-based tests would assist clinicians and researchers in determining the optimal test strategy. My objective was to determine the test accuracy of telephone-based cognitive screening tools for the identification of dementia or MCI.

4.2 Methods

I carried out a systematic review of the literature addressing telephone-based cognitive screens. I followed best practice guidelines in all aspects of design, conduct and reporting (see Appendix 8 for PRISMA checklist) (159, 202). All aspects of title searching, data extraction and quality assessment, were performed independently by myself and another researcher trained in systematic review (Claire Green [CG]), using a third arbitrator (Dr Quinn)) for any conflicts. The protocol is registered with PROSPERO: CRD42017055967.

4.2.1 Search strategy

The search terms were developed using a concept-based approach. Concepts were dementia, cognition, telephone assessment. In addition, names of specific telephone administered cognitive screening tests were also used as search terms (see Appendix 7 for search strategy). I searched the following multidisciplinary, international, electronic databases from inception to 29th July 2018: ALOIS (Cochrane Dementia and Cognitive Improvement Group), CINAHL (EBSCOhost),

EMBASE (OvidSP), MEDLINE (OvidSP) and PsycINFO (EBSCOhost). I applied no language or date restrictions but only included full-text papers published in peer reviewed scientific journals. I checked reference lists of relevant studies and reviews for potentially eligible studies, repeating the process until no new titles were found.

The search strategy was checked by an information specialist. Despite using a sensitive search strategy, there is always a risk that some relevant studies are not detected. Therefore, any studies identified through other sources were also screened.

4.2.2 Study Selection

I carried out screening using the Covidence systematic review software (203). The target condition was all-cause dementia or MCI (resulting from any neurological event or disorder). The index tests of interest were any telephone-based cognitive, screening test assessing more than one cognitive domain. Studies using CSIs incorporating an informant section were included only where this was an additional component (not replacement) to the participant section. If data for participant/informant sections were presented separately, we extracted the participant only section. The reference standard was formal, face to face diagnostic assessment, using neuropsychological testing and/or clinical diagnosis. Within this diagnostic rubric, I accepted any validated, multi-domain neuropsychological battery that provided quantitative data and any clinical diagnosis made according to accepted international criteria. These could be disease specific (e.g. National Institute of Neurological Disorders and Stroke-the Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN) (204)) or general (e.g. Diagnostic and Statistical Manual (DSM) (3)).

I excluded studies that set out to assess cognition or intelligence in a cognitively healthy population; studies using single-domain cognitive tests; studies using the same telephone screen or another screening test as the in-person reference standard; studies where only index ‘test positive’ participants received the reference standard testing and studies that used other information technology or

telehealth as a means of assessing cognition such as smart phone applications or videoconferencing.

4.2.3 Data extraction

Data were extracted to a study specific proforma. I created tables describing the characteristics of included studies, with details of the index test and threshold, reference standard, sensitivity, specificity, positive and negative predictive values. I contacted authors to obtain data or clarify methods, where needed.

4.2.4 Risk of bias and applicability

I assessed methodological quality using the revised Quality Assessment of Diagnostic Accuracy Studies tool (QUADAS-2) (www.bris.ac.uk/quadas/quadas-2) (162). QUADAS-2 assesses studies in terms of internal and external validity across four domains (patient selection, index test, reference standard, flow, and timing). I used the standard QUADAS-2 anchoring questions and developed review specific criteria (202).

4.2.5 Synthesis and analysis

Study data were grouped based on whether the target condition was dementia, MCI or any level of cognitive impairment (dementia and MCI groups combined). I created forest plots of sensitivity and specificity for all the possible combinations of tests, thresholds, and diagnoses. I calculated positive and negative predictive values for all studies apart from those using a case-control design, since this type of sampling only provides indirect estimates (205). This is because the investigator chooses the ratio of cases to controls, and this determines the 'prevalence'.

There are two hierarchical methods that can be used for meta-analysis when studies report sensitivity and specificity: the bivariate and the hierarchical summary receiver operating characteristic (HSROC) models. The two approaches share statistical properties and are mathematically equivalent but have different aims and parameters (206). The focus of the bivariate approach is the summary sensitivity and specificity point at a common threshold, whereas the focus of the HSROC model is the estimation of a summary curve from studies using different

thresholds. Where >1 study reported accuracy data using the same telephone screen (or where minor changes had been made to the screen) and a common threshold score, I created summary estimates of pooled sensitivity and specificity using a random effects bivariate model. I plotted summary estimates in ROC space and described 95% confidence intervals of the summary estimate. Where studies included accuracy data at various cut-off points, I carried out multiple bivariate analyses to explore different thresholds common to more than one study. As a separate subgroup analysis, I examined studies with an exclusive stroke population, to evaluate whether accuracy is compromised within this group.

As a post-hoc analysis I explored the effect of varying TICS threshold, plotting diagnostic odds ratios (DORs), then sensitivity and specificity, against threshold score for TICS dementia studies in a meta-regression. DORs provide a single, overall accuracy metric as described in Chapter 1. Where sensitivity or specificity was reported as 100% in any of the studies, I added a correction factor of 0.5 to cells in the 2 by 2 table (true positive, false positive, true negative, false negative) as recommended (207). I used RevMan (version 5.3) (208), a bespoke test accuracy software (MetaDTA: Diagnostic Test Accuracy meta-analysis (version 1.2) (209)) and Comprehensive Meta-Analysis (version 3) to carry out analyses.

4.3 Results

Of 11,731 titles screened, 34 studies were eligible and included, 17 of which contained data suitable for meta-analyses (Figure 4-1 for PRISMA flow diagram).

There were 26 studies providing data on dementia, 5 studies providing data on any level of cognitive impairment and 14 studies providing data on MCI. Some studies provided data for more than one of these three groups. Thirteen studies (n=1437) were included in meta-analyses for dementia, eight studies (n=791) in meta-analyses for MCI.

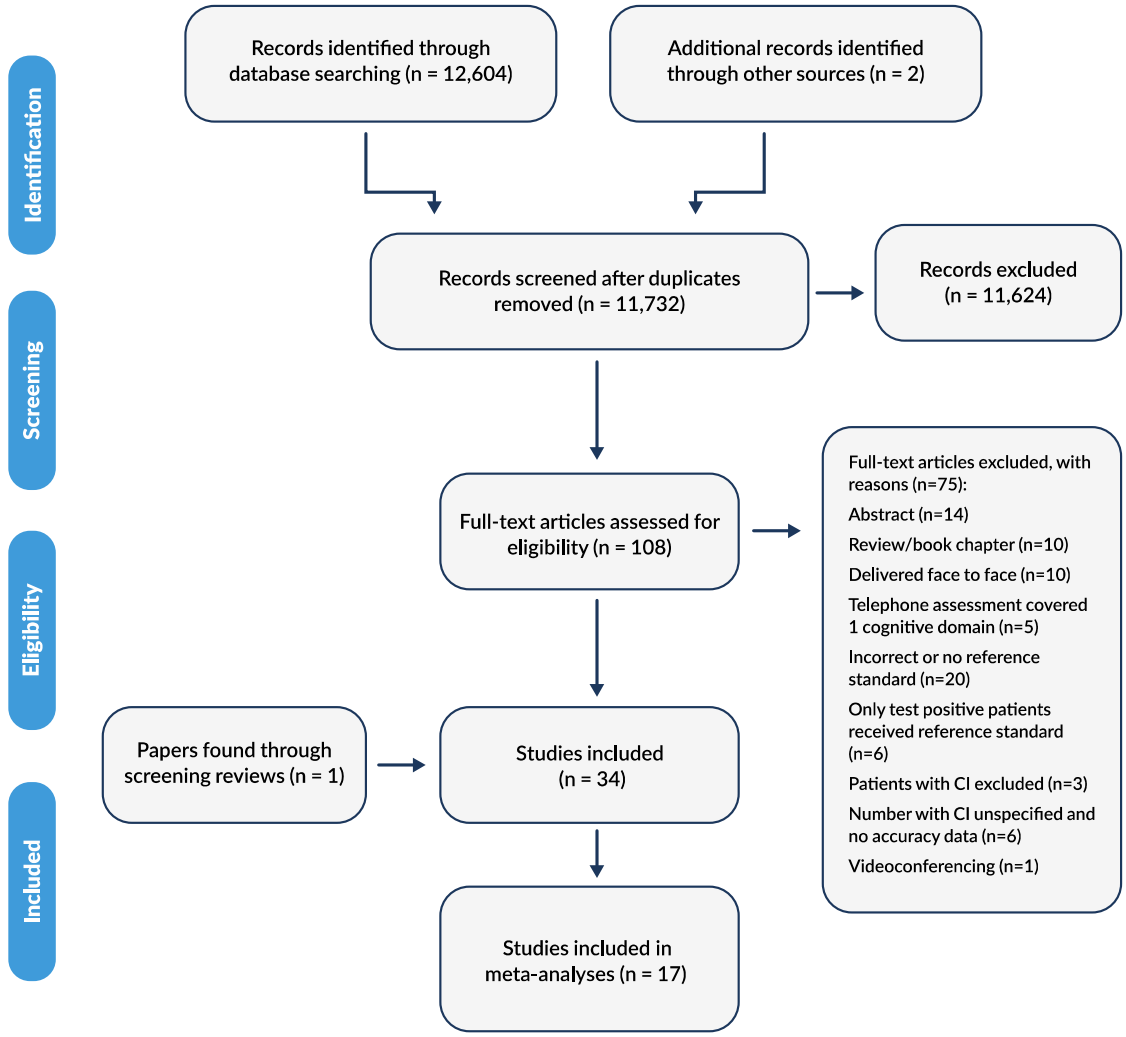


Figure 4-1 PRISMA flow diagram

Research was undertaken in the context of older adults (n=30 studies) and stroke (n=4 studies). Aetiology of dementia was largely Alzheimer's disease (AD) in the samples. Different diagnostic criteria were used for dementia diagnosis, the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorder Association criteria (NINCDS-ADRDA) being most common (11 studies). For MCI diagnosis, Petersen or its modified criteria were used across most studies (11 studies). Assessments were carried out in seven different languages (English, Dutch, Finnish, Korean, Italian, Japanese, Portuguese).

Fifteen different telephone assessments of cognition were used across the studies. The length, component tasks and cognitive domains tested varied across the included tests. The most prevalent test was based on the MMSE: The Telephone Interview for Cognitive Status (TICS) or one of its two modified versions (18 studies (101, 102, 210-226)). There was heterogeneity in TICS interpretation, with optimal threshold scores in included studies ranging from 20 to 33 (out of 41) to differentiate dementia from MCI or cognitively intact, and 30 to 37 (out of 41) to differentiate MCI from cognitively intact individuals. Threshold scores for the TICS-m also varied across included studies, ranging from 24 to 31 (out of 50) to differentiate dementia and 25 to 34 (out of 50) to differentiate MCI. Of the other telephone-based screens, two were also derived from the MMSE (184, 227) two from the MoCA (102), two from the Mental Status Questionnaire (228, 229), one designed as a self-test (230) and the remainder were bespoke telephone assessments (Table 4-1).

Table 4-1 Telephone-based CSIs

Measure	Description	Studies
Mini-Mental State Examination (MMSE) based (n=2)		
26-point Telephone-MMSE	26 items, score/26	Wong 2009
Telephone modified MMSE (T3MS)	34 items, score/100	Alexopoulos 2006
Montreal Cognitive Assessment (MoCA) based (n=2)		
Telephone MoCA (T-MoCA)	8 items, score/22	Pendlebury 2013 Zietemann 2017
T-MoCA short	3 items, score/12	Pendlebury 2013 Wong 2015
Telephone Interview for Cognitive Status (TICS) based (n=3)		
TICS	11 items, score/41	Brandt 1988 Dal Forno 2006 (Italian) Desmond 1994 Go 1997 Kempen 2007 (Dutch) Konagaya 2007 (Japanese) Lipton 2003 Manly 2011 Seo 2011 (Korean)
TICS	11 items, score/38 (Modifications: Name scored out of 1 instead of 2. House number and vice-president removed)	Jarvenpaa 2002
TICS modified (TICS-m)	12 items, score/50	Cook 2009 Crooks 2005 Graff-Radford 2006 Knopman 2010 Meng 2005 Pendlebury 2013 Plassman 1994 Salazar 2014 Seo 2011 (Korean) Welsh 1993 Vercambre 2010
Mental Status Questionnaire (MSQ) based (n=2)		
Short Portable Mental Status Questionnaire (SPMSQ)	10 items, score/10	Roccaforte 1994
TELE interview	17 items, score/20	Gatz 1995 Gatz 2002 Jarvenpaa 2002
Other (n=6)		
Information memory concentration test (IMCT)	27 items, score/37	Zhou 2004

Minnesota Cognitive Acuity Screen (MCAS)	9 items, score/no upper limit as dependent on words generated in fluency task	Tremont 2011 Pillimer 2018
MCAS modified	As above but changes to instructions, delayed recall and recognition added	Pillimer 2018
Structured Telephone Interview for Dementia Assessment (STIDA)	2 sections: Subject and Informant, score/81	Go 1997
Telephone cognitive assessment battery (TCAB)	6 Neuropsychological tests combined	Debanne 1997
Telephone cognitive self-test	8 items, score/45	Van Mierlo 2017

4.3.1 Quality Assessment

No studies were considered to have low risk of bias across all four QUADAS-2 areas (231). Common issues of concern were present across both dementia and MCI studies: case control methodology (n=16 studies), no pre-specified threshold score for the index test (n=19 studies) and time between index test and reference standard unspecified (n=17 studies) (Figure 4-2).

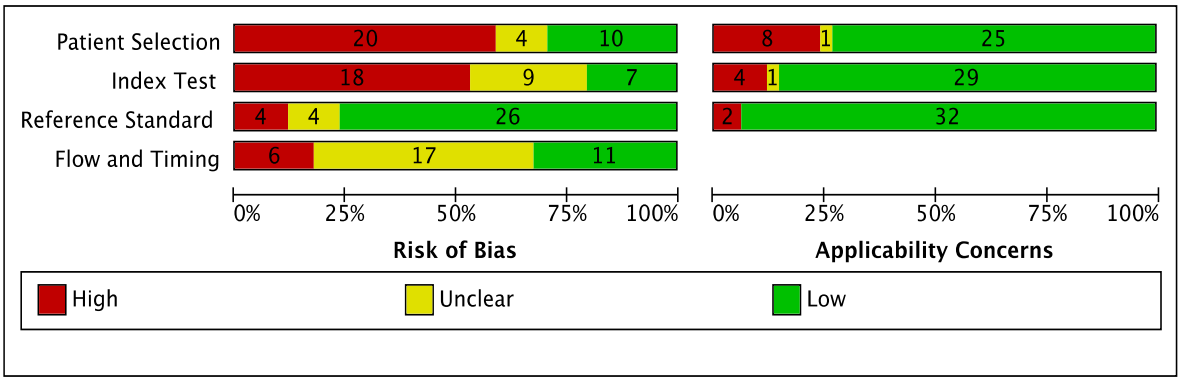


Figure 4-2 Summary of Quality assessment across all studies

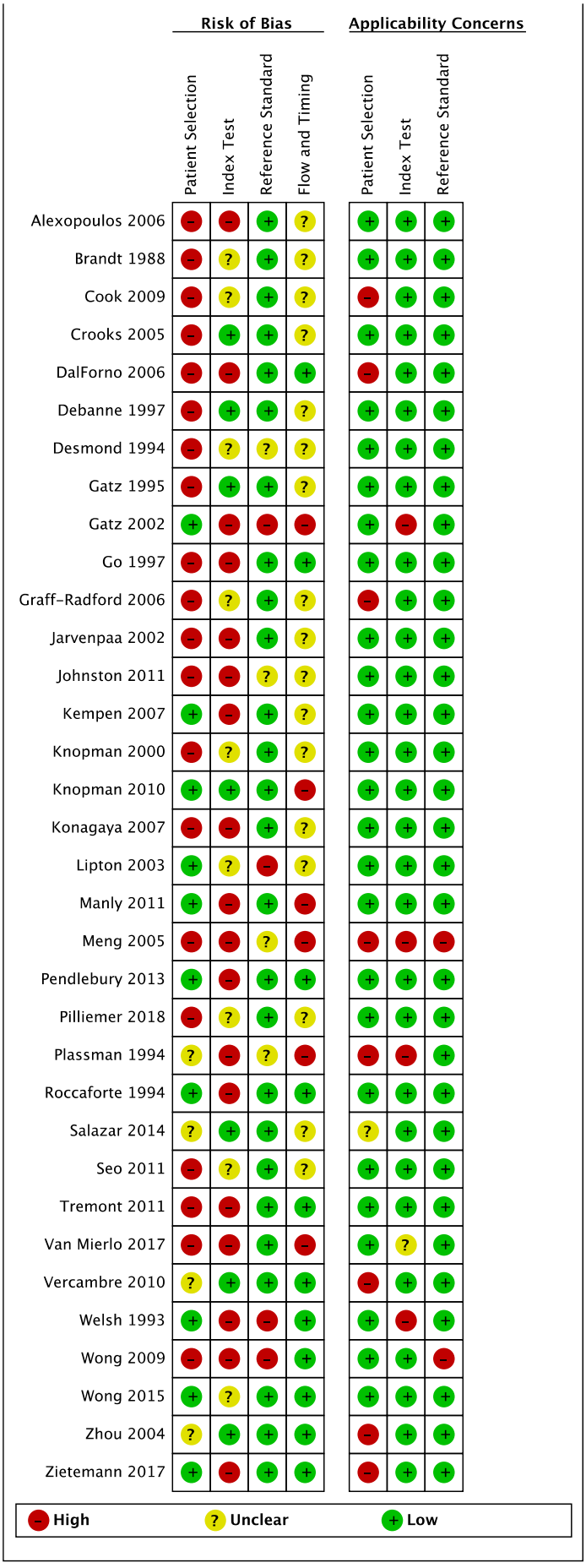


Figure 4-3 Risk of Bias in each individual study

4.3.2 Accuracy in dementia

In total 26 studies (n=3129 participants) assessed the accuracy of telephone screens in identifying dementia (Table 4-2). Twelve different CSIs were examined: IMCT (1 study), TCAB (1 study), TELE (3 studies) TICS (11 studies), TICS-m (6 studies), T3MS (1 study), T-CMMSE (1 study), T-MoCA short (1 study), MCAS (2 studies), telephone self-test (1 study). STIDA (1 study). SPMSQ (1 study) (Full names of the CSIs are available in Table 4-1). Often studies reported accuracy across numerous thresholds, for example 6 of the 11 TICS studies (55%) had data for more than one threshold score.

Table 4-2 Test Accuracy in Dementia

Study	Subjects (N)	N (%) with Dementia	Dementia type	Reference standard (Diagnostic criteria)	Index test (threshold <)	Sensitivity	Specificity	PPV	NPV
Alexopoulos 2006*	30	16 (53%)	AD	ICD-10, NINCDS-ADRDA	T3M3 (85)	1.00	1.00	N/A	N/A
Brandt 1988*	49	16 (48%)	AD	NINCDS-ADRDA	TICS (31)	0.94	1.00	N/A	N/A
Crooks 2005*	38	6 (16%)	AD, mixed	NINCDS-ADRDA	TICS-m (28)	0.83	1.00	1.00	0.97
Dal Forno 2006*	109	45 (41%)	AD	NINCDS-ADRDA	TICS (28)	0.84	0.86	N/A	N/A
Debanne 1997	80	40 (50%)	AD	NINCDS-ADRDA	TCAB (not stated)	0.98	0.85	0.87	0.98
Desmond 1994*^	72	6 (8%)	Post-stroke dementia	Unclear	TICS (25)	1.00	0.83	N/A	N/A
Gatz 1995*	34	12 (35%)	AD	NINCDS-ADRDA	TELE (algorithm)	1.00	0.91	N/A	N/A
Gatz 2002	269	22 (8%)	AD, VaD	DSM-III-R, NINCDS-ADRDA, NINDS-AIREN	TELE (16)	0.86	0.90	0.43	0.99
Go 1997	28	15 (54%)	AD	CDR ≥ 0.5	TICS (29)	0.80	0.77	0.80	0.77
					STIDA (10)	0.80	0.85	0.86	0.79
Jarvenpaa 2002*	56	30 (54%)	AD	NINCDS-ADRDA	TICS (26)	0.87	0.89	N/A	N/A
					TELE (17)	0.90	0.89		
Johnston 2011	27	13 (48%)	Unspecified	DSM-IV-TR	TICS (31)	0.92	0.50	0.63	0.87
Kempen 2007	51	14 (28%)	Unspecified	DSM-IV	TICS (28)	0.87	0.78	0.60	0.94
Knopman 2010	167	42 (25%)	AD, other cause	DSM-IV	TICS-m (29)	0.83	0.82	0.61	0.93

Knopman 2000*	210	99 (47%)	AD, VaD, other cause	Clinical diagnosis, criteria unspecified	MCAS (not stated)	1.00	0.87	N/A	N/A
Konagaya 2007*	135	49 (36%)	AD	DSM-IV & NINCDS-ADRDA	TICS (33)	0.98	0.91	N/A	N/A
Lipton 2003	300	27 (9%)	AD, VaD, LBD, frontotemporal	DSM-III-R	TICS (28)	0.74	0.86	0.74	0.86
Manly 2011	377	53 (14%)	AD, VaD, Parkinsons, DLB	DSM-III	TICS (23)	0.88	0.87	0.53	0.98
Meng 2005*	116	64 (55%)	Unclear	Unclear	TICS-m (28)	0.99	0.90	N/A	N/A
Plassman 1994	67	11 (16%)	AD, dementia unknown aetiology	NPB	TICS-m (not stated)	Correlational data between NPB tests & TICS-m, range 0.27-0.8			
Roccaforte 1994	100	66 (66%)	Unspecified	CDR ≥1	SPMSQ (8)	0.74	0.79	0.88	0.61
Seo 2011	155	85 (55%)	AD, non-AD dementia	DSM-IV	TICS (25)	0.87	0.90	0.87	0.90
					TICS-m (24)	0.88	0.90	0.91	0.86
Tremont 2011*	150	50 (33%)	AD	NINCDS-ADRDA	MCAS (43)	0.86	0.77	N/A	N/A
Welsh 1993	208	20 (10%)	AD	NINCDS-ADRDA	TICS-m (31)	0.85	0.83	0.35	0.98
Wong 2009*	65	34 (52%)	Unspecified	DSM-IV	T-CMMSE (16)	1.00	0.97	N/A	N/A
Wong 2015^	104	51 (49%)	Post-stroke cognitive impairment	CDR (0.5-1)	T-MoCA short (15)	0.84	0.73	0.75	0.83
Zhou 2004	132	65 (49%)	AD, VaD, mixed, other	DSM-IV	IMCT (stratified by education)	0.80	0.81	0.80	0.81

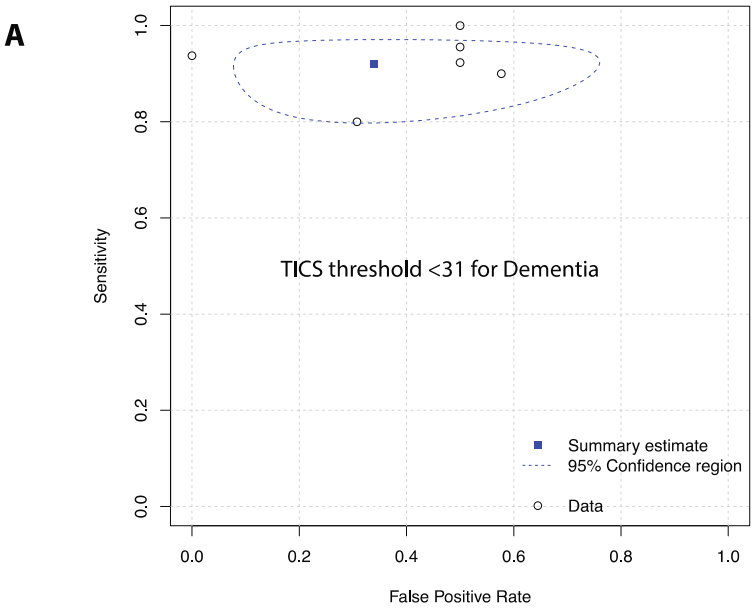
*Case-Control methodology. ^Stroke studies. **Abbreviations**: CDR, Clinical Dementia Rating; NPB, neuropsychological battery; TICS, Telephone interview for cognitive status; MCAS, Minnesota cognitive acuity screen; NIA-AA, National Institute on Aging-Alzheimer's Association; NINCDS-ADRDA, National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorder Association criteria; NPB, neuropsychological battery; N/A, Not Applicable; SPMSQ, Short Portable Mental Status Questionnaire; STIDA, Structured Interview for dementia Assessment; TAB, Telephone Assessment Battery; TCAB, Telephone cognitive assessment battery; T3MS, Modified Mini-mental State Examination; TICS, Telephone interview for cognitive status; T-MoCA, Telephone Montreal Cognitive Assessment

Across the range of telephone screening tests, sensitivity ranged 67-100%, specificity ranged 50-100%. I included thirteen studies (n=1437) in the dementia meta-analyses and pooled test accuracy data for three tests: TICS (4 thresholds: <31, <29, <28, <25), TICS-m (2 thresholds: <31 and <28) and TELE (threshold <16) (Table 4-3 and Figure 4-4). One study was included in all TICS meta-analyses(218). The balance of sensitivity, specificity differed according to cut-off score, however sensitivity was greatest using the original cut-off score of 31. Only one meta-analysis did not include case-control studies: TICS at threshold of 29.

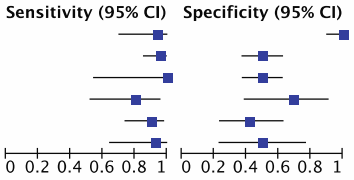
Table 4-3 Meta-Analyses of dementia studies

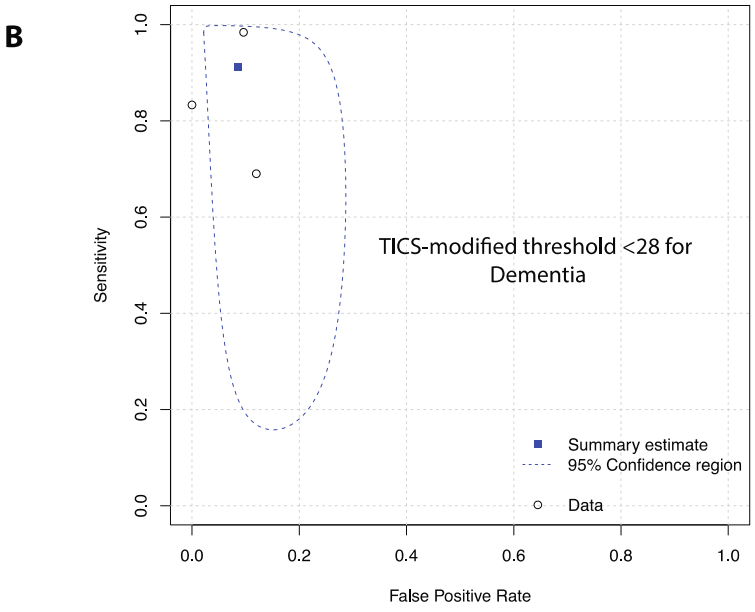
Test (Threshold)	Studies (Participants)	Cognitive impairment, n (%)	Sensitivity (95% CI)	Specificity (95% CI)	Positive Likelihood Ratio (95% CI)	Negative Likelihood Ratio (95% CI)
TICS (<31/41)	6 (341)	125 (37%)	0.92 (0.86-0.96)	0.66 (0.39-0.86)	2.70 (0.73-4.68)	0.12 (0.04-0.21)
TICS (<29/41)	4 (435)	86 (20%)	0.85 (0.74-0.92)	0.73 (0.61-0.83)	3.17 (1.94-4.40)	0.21 (0.10-0.32)
TICS (<28/41)	5 (588)	122 (21%)	0.87 (0.76-0.94)	0.77 (0.65-0.85)	3.73 (2.28-5.18)	0.17 (0.06-0.27)
TICS (<25/41)	5 (362)	150 (41%)	0.79 (0.64-0.89)	0.90 (0.83-0.94)	7.97 (3.84-12.10)	0.23 (0.09-0.37)
TICS-m (<28/50)	3 (321)	112 (35%)	0.91 (0.63-0.98)	0.91 (0.85-0.95)	10.52 (3.42-17.63)	0.10 (-0.06-0.26)
TELE (<17/20)*	2 (303)	52 (17%)	0.89 (0.77-0.95)	0.90 (0.86-0.93)	8.63 (5.48-11.77)	0.13 (0.03-0.23)

*One study used a threshold of <17 and the other <16.



Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Brandt 1988	15	0	1	33	0.94 [0.70, 1.00]	1.00 [0.89, 1.00]
DalForno 2006	43	32	2	32	0.96 [0.85, 0.99]	0.50 [0.37, 0.63]
Desmond 1994	6	33	0	33	1.00 [0.54, 1.00]	0.50 [0.37, 0.63]
Go 1997	12	4	3	9	0.80 [0.52, 0.96]	0.69 [0.39, 0.91]
Jarvenpaa 2002	27	15	3	11	0.90 [0.73, 0.98]	0.42 [0.23, 0.63]
Johnston 2011	12	7	1	7	0.92 [0.64, 1.00]	0.50 [0.23, 0.77]





Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Crooks 2005	5	0	1	32	0.83 [0.36, 1.00]	1.00 [0.89, 1.00]
Knopman 2010	29	15	13	110	0.69 [0.53, 0.82]	0.88 [0.81, 0.93]
Meng 2005	63	5	1	47	0.98 [0.92, 1.00]	0.90 [0.79, 0.97]

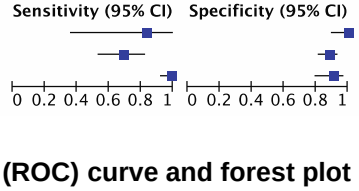


Figure 4-4 Summary receiver operating characteristic (ROC) curve and forest plot describing test accuracy studies of *(A)* Telephone Interview for Cognitive Status (TICS) at threshold of <31/40; and *(B)* TICS-m at threshold of <28/50 to identify dementia
Circles = individual studies; square = summary estimate; dotted line = confidence interval

4.3.3 Accuracy in Dementia/MCI combined

In total 5 studies (n=957 participants) assessed the accuracy of telephone screens in identifying any level of cognitive impairment (MCI and dementia groups combined) (Table 4-4). The rest of the sample in the study by Van Mierlo et al. were described as participants with subjective cognitive decline.

Table 4-4 Test Accuracy for MCI/Dementia Combined

Study	Partici-pants	No. with CI	Test	Reference standard	Threshold	Sens	Spec	PPV	NPV
Manly 2011	377	121	TICS	Clinical diagnosis	<27	0.73	0.77	0.73	0.77
Knopman 2010	167	84	TICS-m	Clinical diagnosis	<32	0.83	0.78	0.79	0.82
Vercambre 2010	120	10	TICS-m	NPB	<30	0.89	0.68	0.20	0.99
Tremont 2011	200	150	MCAS	NPB	<52.5	0.91	0.78	0.93	0.74
Van Mierlo 2017	93	63	Telephone self-test	Clinical diagnosis (NIA-AA)	Z score cut off only	0.73	0.73	0.85	0.56

4.3.4 Accuracy in Mild Cognitive Impairment

In total 14 studies (n=1,684 participants) assessed the accuracy of telephone screens in identifying MCI (Table 4-5). Eight different tests were examined: TICS (3 studies), TICS-m (8 studies), T3MS (1 study), T-MoCA (2 studies), T-MoCA short (1 study), MCAS (2 studies), MCAS modified (1 study), telephone self-test (1 study). Eight studies (n=791) were included in meta-analyses. I pooled test accuracy data for one test: TICS-m (3 thresholds: <33, <29, <28) (Table 4-6, Figure 4-5).

Table 4-5 Test Accuracy in MCI

Study	Subjects (N)	N (%) with Dementia	MCI type	Reference standard (Diagnostic criteria)	Index test (threshold <)	Sensitivity	Specificity	PPV	NPV
Alexopoulos 2006*	32	18 (56%)	MCI	Petersen	T3MS (89)	0.83	1.00	N/A	N/A
Cook 2009	71	17 (24%)	aMCI	Petersen	TICS-m (34)	0.82	0.87	0.67	0.94
Crooks 2005*	38	4 (11%)	MCI	Petersen	TICS-m (28)	0.50	0.93	N/A	N/A
Graff-Radford 2006	128	8 (6%)	MCI	Petersen	TICS-m (29)	0.63	0.86	0.23	0.97
Knopman 2010	125	42 (34%)	MCI	Petersen	TICS-m (32)	0.71	0.78	0.62	0.84
Manly 2011	324	68 (21%)	MCI	Petersen	TICS (30)	0.79	0.58	0.33	0.91
Meng 2005*	116	18 (16%)	MCI	Unclear	TICS-m (33)	0.89	0.92	N/A	N/A
Pendlebury 2013^	68	27 (40%)	Any MCI (single/multi-domain)	Modified Petersen criteria	TICS-m (25)	0.85	0.56	0.56	0.85
					T-MoCA (19)	0.89	0.46	0.52	0.86
					T-MoCA short (11)	0.96	0.39	0.51	0.94
		12 (18%)	Multi-domain MCI		T-MoCA (18)	1.00	0.52	0.31	1.00
					T-MoCA short (10)	0.83	0.48	0.26	0.93
Pillemer 2018	60	30 (50%)	aMCI	Petersen	MCAS (not stated)	0.97	0.87	0.88	0.97
					MCAS modified (not stated)	0.97	0.97	0.97	0.97
Salazar 2014	184	60 (33%)	MCI	Clinical diagnosis, criteria unclear	TICS-m (28)	Not reported		N/A	N/A
Seo 2011	145	75 (52%)	MCI	Petersen	TICS (29)	0.69	0.69	0.70	0.68
					TICS-m (29)	0.73	0.67	0.70	0.70
Tremont 2011*	200	100 (50%)	aMCI	Petersen	MCAS (53)	0.86	0.78	N/A	N/A
Van Mierlo 2017	93	22 (24%)	MCI	NIA-AA	Telephone self-test (Z value cut-off)	0.59	0.73	0.40	0.85
Zietemann 2017^	100	22 (22%)	Any MCI (single/multi-domain)	Modified Petersen criteria	TICS (37)	0.82	0.44	0.29	0.90
					T-MoCA (19)	0.81	0.73	0.45	0.94
		8 (8%)	Multi-domain		TICS (36)	0.87	0.61	0.18	0.98
					T-MoCA (18)	0.87	0.82	0.33	0.98

*Case-control. ^Stroke studies. **Abbreviations**: MCAS, Minnesota cognitive acuity screen; NIA-AA, National Institute on Aging-Alzheimer's Association; NPB, neuropsychological battery; N/A, Not Applicable; SPMSQ, Short Portable Mental Status Questionnaire; TAB, Telephone Assessment Battery; T3MS, Modified Mini-mental State Examination; TICS, Telephone interview for cognitive status; T-MoCA, Telephone Montreal Cognitive Assessment

Table 4-6 Meta-analyses of MCI studies

Test (Threshold)	Studies (Participants)	Cognitive impairment, n (%)	Sensitivity (95% CI)	Specificity (95% CI)	Positive Likelihood Ratio (95% CI)	Negative Likelihood Ratio (95% CI)
TICS-m (<33/50)	3 (196)	77 (39%)	0.82 (0.70-0.90)	0.87 (0.72-0.95)	6.35 (1.27-11.43)	0.21 (0.10-0.31)
TICS-m (<29/50)	4 (469)	142 (30%)	0.56 (0.33-0.77)	0.89 (0.61-0.98)	5.17 (-0.66-11.01)	0.49 (0.30-0.68)
TICS-m (<28/50)	4 (362)	71 (20%)	0.34 (0.18-0.56)	0.96 (0.87-0.99)	7.85 (0.55-15.14)	0.69 (0.50-0.87)

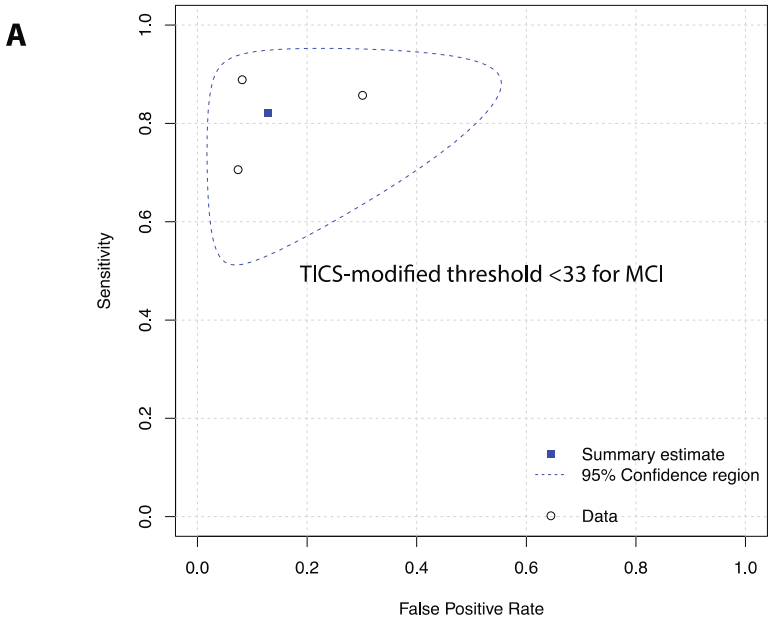
4.3.5 Subgroup analysis: Accuracy in stroke population

In total 4 studies (n=344 participants) addressed post-stroke cognitive impairment. In terms of cognitive impairment severity, two studies described post-stroke dementia, and the other two provided data for multi-domain MCI and single and multi-domain MCI combined. No studies addressed the impact of aphasia on assessment. I pooled data for T-MoCA (threshold <19) for any type of MCI (single and multi-domain combined) and T-MoCA (2 thresholds <17, <18) for multi-domain MCI (Table 4-7, Figure 4-5).

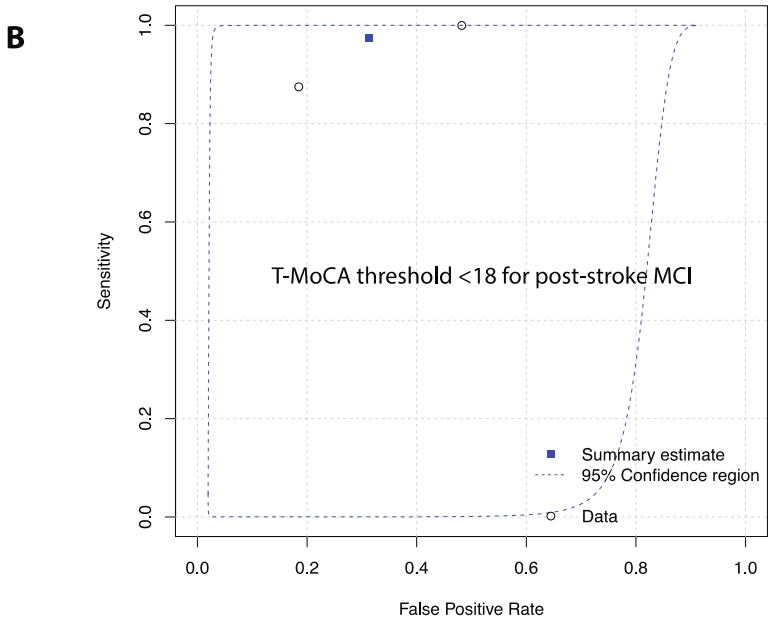
Table 4-7 Meta-analyses of cognitive impairment post-stroke studies

Test (Threshold)	Studies (Participants)	Cognitive impairment, n (%)	Sensitivity (95% CI)	Specificity (95% CI)	Positive Likelihood Ratio (95% CI)	Negative Likelihood Ratio (95% CI)
T-MoCA (<17/22)*	2 (168)	20 (12%)	0.80 (0.55-0.93)	0.83 (0.63-0.93)	4.70 (0.67-8.73)	0.25 (0.03-0.46)
T-MoCA (<18/22)*	2 (168)	20 (12%)	0.98 (0.30-1.00)	0.69 (0.45-0.86)	3.12 (1.12-5.11)	0.04 (-0.12-0.19)
T-MoCA (<19/22)**	2 (168)	49 (29%)	0.86 (0.71-0.94)	0.61 (0.41-0.78)	2.21 (1.19-3.24)	0.23 (0.07-0.40)

*Multi-domain MCI. **Single and multi-domain MCI combined.



Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Cook 2009	12	4	5	50	0.71 [0.44, 0.90]	0.93 [0.82, 0.98]		
Knopman 2010	36	25	6	58	0.86 [0.71, 0.95]	0.70 [0.59, 0.79]		
Meng 2005	16	8	2	90	0.89 [0.65, 0.99]	0.92 [0.85, 0.96]		



Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Pendlebury 2013	12	27	0	29	1.00 [0.74, 1.00]	0.52 [0.38, 0.65]		
Zietemann 2017	7	17	1	75	0.88 [0.47, 1.00]	0.82 [0.72, 0.89]		

Figure 4-5 Summary receiver operating characteristic (ROC) curve and forest plot describing test accuracy studies of (A) Telephone Interview for Cognitive Status modified (TICS-m) at threshold of <33/50 to detect MCI; and (B)T-MoCA at threshold of <18/22 to identify multi-domain MCI post-stroke

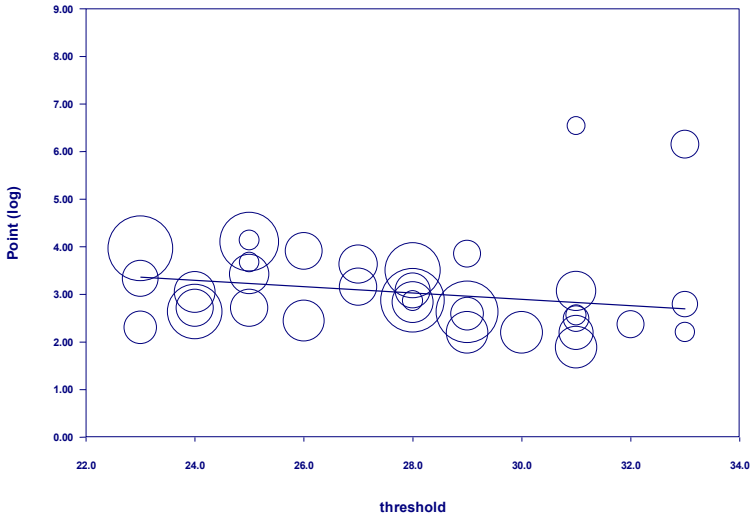
4.3.6 Meta-regression

The diagnostic odds ratios used for the meta-regression are listed in Table 4-8. Meta-regression suggested no relationship between TICS threshold score and overall accuracy in identifying dementia (slope -0.07, [SE, 0.05]; p=0.1525). Subsequent meta-regression found a trade-off between sensitivity and specificity across TICS thresholds. A significant relationship was found between threshold score and sensitivity (slope 0.02, [SE, 0.01]; p<0.0001) and a significant inverse relationship for specificity (slope -0.04, [SE, 0.01]; p<0.0001). Thus, as TICS threshold score increased, sensitivity for identification of dementia increased and specificity decreased (Figure 4-6).

Table 4-8 Diagnostic odds ratios used in meta-regression

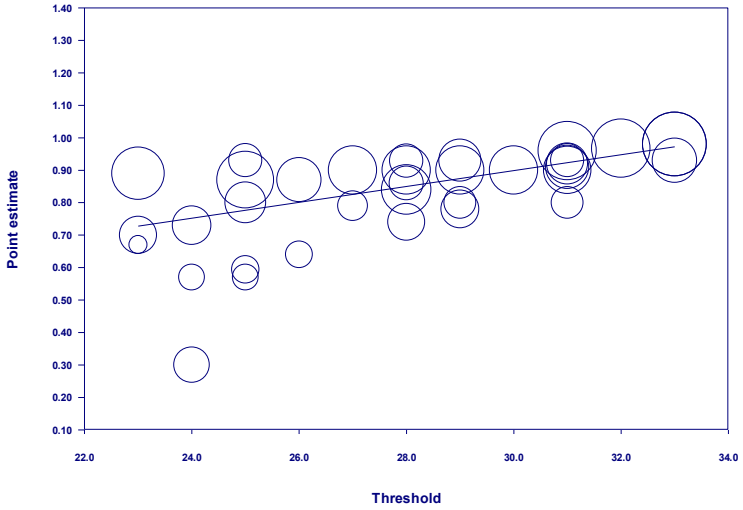
Study	Threshold	DOR (95% CI)
Konagaya 2007	33	468.0 (56.8 – 3859.1)
Jarvenpaa 2002	33	9.1 (0.4-184.6)
	32	10.7 (1.2-93.9)
	31	6.6 (1.6-27.4)
	30	9.0 (2.2-37.2)
	29	9.0 (2.2-37.2)
	28	17.0 (4.0-71.8)
	27	37.8 (8.1-176.5)
	26	49.8 (10.1-246.5)
	25	30.7 (6.8-137.3)
	24	21.1 (4.9-89.9)
	23	28.0 (5.4-144.4)
Brandt 1988	31	692.3 (26.7 – 17976.4)
Dal Forno 2006	31	21.5 (4.8 – 96.4)
	28	33.2 (11.4 – 96.8)
Desmond 1994	31	13.0 (0.7 – 240.1)
	28	17.6 (1.0 – 324.6)
	25	62.7 (3.3 – 1193.6)
	23	10.0 (1.6 – 61.5)
Go 1997	33	16.3 (1.6 – 163.4)
	31	9.0 (1.6 – 50.7)
	29	13.3 (2.2 – 81.2)
	25	39.5 (2.0 – 787.7)
Johnston 2011	31	12.0 (1.2 – 118.9)
Kempen 2007	29	47.1 (5.3 – 416.6)
	28	21.8 (4.0 – 117.8)
	27	23.5 (4.8 – 114.7)
	26	11.5 (2.7 – 48.8)
	25	15.1 (3.1 -73.8)
	24	15.1 (3.1 – 73.8)
Lipton 2003	29	13.9 (5.3 – 36.0)
	28	17.7 (7.0 – 44.6)
	24	13.9 (4.7 – 41.3)
Seo 2011	25	60.5 (22.2 – 165.5)
Manly 2011	23	52.6 (21.2 – 130.6)

A



Regression of Point estimate on Threshold

B



Regression of Point estimate on Threshold

C

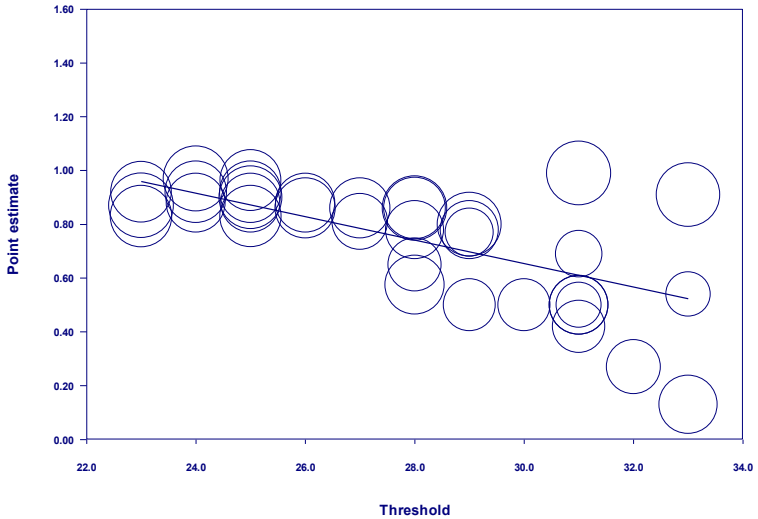


Figure 4-6 Meta-regression of TICS threshold score against (A) overall accuracy (diagnostic odds ratios); (B) Sensitivity; (C) Specificity

4.4 Discussion

This systematic review identified 34 studies describing 15 differing telephone-based CSIs. The best available test accuracy evidence was for the TICS and TICS-m. The pattern of test properties, with high sensitivity and lower specificity at conventional thresholds, suggest that these tools could be used as an initial screen for potential dementia. In the identification of MCI however, TICS-m was more specific, than sensitive. Subgroup analyses suggested that telephone-based cognitive assessments (T-MoCA) were useful in stroke patients, but the small number of studies limits recommendations.

The literature on telephone assessments was characterised by substantial heterogeneity. Heterogeneity was evident even within the publications describing a single test. For example, TICS was the test with the greatest supporting test accuracy literature. Within the TICS label, there was the original TICS (most often used to identify dementia), and its modified version (with the additional delayed recall component), often described as preferable for MCI identification. However, other modified TICS versions were found in the literature. This illustrates a situation seen in other areas of dementia research where we should not assume the content of a test based on the name given by researchers (232). Tests are often altered to be country or culture specific, but there were also examples of modification to scoring or content where the rationale for the change was unclear. In future studies, researchers should specify the specific cognitive test components and scoring systems that they have used and avoid altering these aspects of published tests unless fully justified.

The threshold scores used to define a person as ‘test positive’ varied even where the same screening test was used. Taking TICS as an example, most of the included studies did not use the original recommended threshold score of 31. Indeed, most did not report a pre-specified threshold score at all, but rather reported accuracy across a range of different potential cut-off scores. This can over-estimate the accuracy of the test, as researchers may preferentially report the threshold that performs best in their data. Meta-regression suggested that by altering the threshold of the test, one test property (sensitivity or specificity) can be favoured over the other.

Considering both the target condition and the choice of comparator group are very important when interpreting accuracy data, and especially in the decision-making for data synthesis. For example, two studies may both be targeting dementia, but one may want the test to discriminate from MCI, and the other to discriminate from healthy cognition. Studies poorly defining the target condition (e.g., cognitive impairment), are unhelpful as we are unable to ascertain what severity of impairment the authors are referring to and results can be misinterpreted. Agreement on MCI diagnosis can also be challenging due to the numerous definitions and diagnostic criteria (233). The MCI studies that we included targeted different types of MCI (amnestic only, single, multi-domain), therefore there is argument that even these should not be combined. Our review highlights the importance of standardisation in definitions, and on which diagnostic criteria are used for diagnosis.

4.4.1 Research in context

Faced with a choice of methods for administering cognitive tests, clinicians may wonder whether the accuracy of telephone-based assessment is comparable to that of traditional face-to-face assessments. Comparing the summary estimates for accuracy of TICS, TICS-m, T-MoCA against recent reviews of MMSE and MoCA (192, 234, 235) would suggest that there is no substantial decrement in accuracy when using the telephone. To definitively assess comparative accuracy would require comparison of telephone and in-person assessment in the same population against the same gold standard. I found no studies using this approach and so our indirect comparisons are the best available evidence at present.

4.4.2 Strengths and limitations

This is the first systematic review to include meta-analyses to investigate the accuracy of telephone-based cognitive assessments. Strengths of this work include a comprehensive search strategy informed by an information specialist and following best practice guidance for the conduct of test accuracy reviews (including multiple data sets; ensuring two researchers carried out screening and data extraction independently and offering systematic consideration of bias for each of the included studies).

Limitations of the included studies should be considered when interpreting our findings. Problematic design issues were frequent, such as the use of case-control methodology. Like other test accuracy reviews, we highlight issues to be considered in the design and reporting of future cognitive test accuracy studies, for example, the need for standardised reporting of the content and application of cognitive tests.

The pooled analyses were limited (few studies for each outcome) due to variation in the telephone-based tests used, different threshold scores employed, and different diagnostic criteria for dementia and MCI. The limited number of studies also precluded the incorporation of quality assessment into the summary data.

4.4.3 Implications for clinical practice and research

There are no set values of sensitivity or specificity in which recommendations can be made for all settings. As these metrics are inversely proportional, comparison of different tests is dependent on the context of how the test will be used. TICS and TICS-m appear to be sensitive tests for the identification of dementia when using the original test thresholds (31 and 28 respectively). This means they could be useful for first-line screening, eliminating those unlikely to have dementia and selecting a group who require further testing. This however comes at the cost of a high number of false positives. Unqualified accuracy metrics can often seem abstract and illustrating test properties using a theoretical example can aid understanding. Based on the data for TICS: in a theoretical population of 1000 community dwelling older adults, including 80 people living with dementia; 74 of these would be correctly classified using TICS with the conventional threshold. However, 267 people without dementia would also screen positive and may, as a result, receive additional unnecessary investigations.

A different pattern of accuracy was demonstrated in the identification of MCI so this context of use should be considered separately. The STIDA is unique to other telephone-based screens, including sections on medical history, patient-reported cognitive and functional abilities, and informant-based questions, in addition to the formal cognitive screening questions. This is worth highlighting since

independence in activities of daily living (ADLs) is a key differentiation between MCI and dementia and it is rare that questions evaluating the impact of cognitive impairment on ADLs are included within cognitive screens (in-person or otherwise). Data regarding functional ability are often obtained using an informant-based questionnaire, for example, the IQCODE (236). While the STIDA (222) was the only scale to incorporate a proxy section, published informant-based questionnaires were found amongst other included studies. These were administered over the telephone with a family member, following completion of the cognitive screen with the patient. We only extracted accuracy data from the patient section of the STIDA, to be consistent with the other included screening tests, however the study reported that accuracy was higher when used in combination with the informant section.

Although these results are encouraging, telephone-based screens have inherent limitations and face-to-face assessment should not be abandoned or replaced. A main limitation of this mode of screening is that it prevents the evaluation of visuospatial functions. The purpose of using the screening test is therefore central to this discussion. If used in a clinical setting, as part of the pathway to reach a dementia diagnosis, then just like face-to-face screens, they should be followed up with a comprehensive physical and neuropsychological examination before a diagnosis is reached. However, the purpose of cognitive screening in some settings is not necessarily to detect dementia. Many neurological conditions have visuospatial sequela e.g. spatial neglect following a stroke (237), which would not be detected by telephone-based screens. This means that if the clinical or research purpose is to identify any level of cognitive deficit, they will be less useful in conditions where these abilities are frequently impaired (e.g. posterior cortical atrophy (PCA) and PD (238)). The vast majority of studies included in this review were targeting AD type dementia and although examination of visuospatial abilities provides valuable information (239), memory impairment is the predominant early and central feature of AD (240). Therefore, as telephone-based cognitive screening tools major on memory, they would be suitable for first-line screening.

As telephone evaluation does not allow for lip reading or non-verbal cues, people living with hearing impairment may be further disadvantaged. This is

especially a concern since there is evidence that hearing loss negatively impacts test scores in healthy, older adults when assessed in-person (241). Since the included studies did not address the impact of hearing impairment on test accuracy, I cannot provide conclusions regarding this. However, experience of using telephone-based cognitive screens within my research group has confirmed that screening participants with hearing impairment or even where participants have different accents is challenging. Tasks most at risk are those which have strict instructions regarding repetition, for example reading out words for memory tasks as participants may be unable to hear test stimuli, whereas other items which can be repeated (e.g., asking someone their age) are less problematic. This emphasises the importance of reminding patients to wear hearing aids and the importance of practical aspects such as having a good phone connection. Previous work completed in-person has shown that a headset with an amplifier has helped those with hearing impairment to complete a cognitive assessment (103). Telephone technologies allow for amplification and headsets could be provided to patients. The studies conducted in a stroke setting also did not provide information on whether participants had aphasia. Future research should provide data on whether samples include participants with hearing impairment or aphasia; and should explore whether accuracy is impacted within these groups.

With these limitations in mind, the data suggest that telephone-based screening may have a particular role when in-person assessments are not feasible (e.g., for those who cannot attend clinic appointments) and in large trials, cohorts or registries that require a cognitive outcome measure at scale. Greater use of screening via videoconferencing (VC) is anticipated, which offers the convenience of telephone with some of the advantages of in-person testing. Through this format, minimal modifications to the original measures would need to be made and visuospatial abilities can be assessed. However, this represents a new approach to testing and so should be subject to the same scrutiny of test properties as any other novel assessment. A recent systematic review comparing neuropsychological assessment delivered face-to-face or via VC found that scores dropped for some tasks, whilst others were unchanged (242). Some preliminary supporting evidence also exists in community-based stroke survivors from a recent study comparing face-to-face and VC administrations of the MoCA using a

randomised crossover design (243). A review considering the diagnostic accuracy of VC is needed, where clinical diagnosis is used as the comparator. Although there was no clearly superior telephone screening test, this does not imply that new tests should be developed. There were many telephone assessments included in the review, yet the number of telephone cognitive assessments available are greater still, since a number did not meet our inclusion criteria. Further research should therefore be done evaluating the psychometric properties of available tests so summary estimates can be more reliable in the future. There is also an argument for selecting one or two tests as the preferred measures and ensuring that researchers and clinicians are trained in application and scoring. This is aligned with and moves towards core outcome sets in other aspects of neurology and dementia.

To conclude, this review found the TICS and TICS-m to have high sensitivity in the identification of dementia in non-specialist settings. Telephone-based cognitive screens should be considered as an alternative screening method when face-to-face assessment is not viable.

5 Assessing Post-Stroke Psychology Longitudinal Evaluation (APPLE) study methods

Chapter 1 detailed our current understanding and evidence gaps in research concerning post-stroke cognitive assessment. Few acute stroke studies report data using brief cognitive screens or stroke-specific cognitive screens. Chapter 2 demonstrated that around a quarter of participants do not fully complete brief cognitive screens, due to a range of reasons, and very few CSIs provide scoring where participants are untestable. Chapter 3 demonstrated that some shortened forms of the MoCA have similar sensitivity and specificity to that of the full scale. Building on this work, the study detailed in this Chapter, aims to provide data on some of these research gaps. This Chapter describes the study methods and Chapters 6 and 7 provide the results.

5.1 Overview of the study

Assessing Post-Stroke Psychology Longitudinal Evaluation (APPLE) is a multicentre, observational cohort study. The broad aims of the study are to understand the neuropsychological consequences of stroke, at both the acute stage and longer-term. The study is funded by the Stroke Association and Chief Scientist Office of Scotland; funding reference: PPA 2015/01_CSO. The protocol is registered on research registry (ID: 1018) and available in Appendix 9.

This longitudinal study follows participants over a period of 18 months but for the purpose of my thesis I am using the data collected up to the 6-month follow-up. In this chapter I describe the COAs and CSIs used for the two chapters that follow; I do not describe any of the other measures used in the APPLE study which are not relevant to my work; these are described in two other PhD theses.

5.1.1 Ethical Approval

Ethics committee and local Research and Development approval was obtained for all sites (REC number 16/SS/0105) (Appendix 11).

5.1.2 Study Aims

The overall aim is to examine psychometric properties of eight generic, brief CSIs and a stroke-specific CSI: The Oxford Cognitive Screen (OCS).

My specific research aims using the APPLE data set are detailed below.

Chapter 6 aims (using generic, brief CSIs):

- Completion rates.
- Longitudinal comparison of participants classified as cognitively impaired (screen positive) using different brief CSIs.
- Floor/ceiling effects across the CSIs.
- Accuracy of brief CSIs in detecting pre-stroke cognitive impairment when compared to clinical diagnosis.
- Accuracy of brief CSIs in detecting cognitive impairment on the OCS.

Chapter 7 aims (using the OCS):

- Completion rates and floor/ceiling effects.
- To examine whether impairments in individual OCS cognitive domains or a global OCS score are associated with later functional, mood and quality of life outcomes at six months.

Chapter 6 aims are detailed in the APPLE protocol in Appendix 9. Chapter 7 was designed by myself, after the APPLE study had commenced.

5.2 Patient and public involvement

People living with stroke have been involved in all stages of the study. Over-burdening the patient is a concern in studies involving cognitive testing and multiple scales/questionnaires. At the planning and design stages, my supervisor

and the APPLE team gained feedback on the proposed tests from stroke survivors. This helped refine the final scales which would be considered acceptable and subsequently included. A patient advisory group was also set up as part of the study to comment on both the design and progress of the study. The group consists of stroke survivors and healthcare professionals.

5.3 Inclusion & Exclusion criteria

Participants were generally recruited through acute stroke units, with a minority recruited through a TIA clinic. In the attempt to include a representative cohort, all participants admitted with suspected stroke due to any aetiology (ischaemic/haemorrhagic) or TIA were considered eligible. We sought to include participants that are often excluded from research, for example those with known cognitive impairment/dementia or aphasia. The inclusion and exclusion criteria which were set are detailed in Table 5-1.

Table 5-1 Inclusion & Exclusion criteria

Inclusion Criteria	Exclusion Criteria
1. Clinical diagnosis of stroke or TIA at time of assessment. 2. Age greater than 18 years. 3. Clinical team happy that patient is suitable for some form of psychological testing.	1. No spoken English prior to stroke.

5.4 Informed consent

Participants were required to have the patient information sheet (PIS) for at least 24 hours prior to providing written informed consent. For participants deemed to lack mental capacity, consent was sought from their nearest relative/guardian. Mental capacity was determined by the researcher, with input

from the clinical multidisciplinary team, where required. According to the mental capacity act, a person needs to be able to understand the information relevant to the decision, retain that information and use or weigh up that information. Therefore, if there was a concern that any of these criteria were not met, a relative signed the consent form. Where available, relatives were separately recruited into the study in order to complete the informant-based assessments. The PIS can be found in Appendix 10.

5.5 Schedule of events

Longitudinal data were collected across five timepoints in the study, in order to capture any potential changes in cognition. In this thesis I have used data from the first three timepoints: baseline, one month and six months.

5.5.1 Baseline

The baseline visit was completed as soon as possible following the participant's admission to the stroke unit. After eligibility criteria were checked, written informed consent was obtained. Clinical and demographic information for each participant were collected from their medical notes. This included documenting any history of MCI or dementia recorded in the medical record and capturing years of education, hearing/visual impairment from the participant directly. Visual impairment captured those who were partially sighted or blind but did not include those who wear glasses/contact lenses. Stroke severity was measured using the National Institutes for Health Stroke Scale (NIHSS) (244) score completed on hospital admission or shortly afterwards. A full description of this scale can be found in Chapter 2. The Confusion Assessment Method for the intensive care unit (CAM-ICU) (245) was used to screen for delirium.

A range of COAs were completed to capture the patient's pre-morbid functional abilities:

- Global disability was assessed using the modified Rankin Scale (mRS) (246): scores range from 0 (no disability) to 6 (death). Full description of the measure is covered in Chapter 2.

- Independence in basic activities of daily living was measured by the Barthel index (BI) (247). The BI covers 10 areas (feeding, bathing, grooming, dressing, bowels, bladder, toilet use, transfers, mobility, stairs) and is scored from 0-100 (with higher scores indicating greater independence) and covers feeding, bathing, grooming, dressing, bowels, bladder, toilet use, transfer, mobility, and stairs.

- Independence in instrumental activities of daily living scale (IADL) was measured using a variation of Lawton's IADL scale(248), scored 0-14 (with higher scores indicating greater independence). The scale covers telephone use, getting to places outside of walking distance, shopping for groceries/clothes, preparing meals, housework, taking medicine and handling money. Each of the seven areas is scored either 0 (completely unable), 1 (with some help) or 2 (without help).

5.5.1.1 Baseline Cognitive Assessment

Our baseline assessment, referred to throughout as 'AMT-plus', has additional questions to the 10-point Abbreviated Mental Test. As many brief CSIs share common questions, I scored eight different CSIs from one set of questions: CDT, Abbreviated MoCA, 4AT, Cog-4, 6-CIT, NINDS-CSN 5-min MoCA, AMT-4 and AMT-10. The tasks in these tests largely assess learning and memory, with only a few other tasks covering other DSM-5 cognitive domains (Table 5-3).

It is important to note that through taking this approach some changes to the original scales have been made. Delayed recall in our assessment includes the five words used in the MoCA (face, velvet, church, daisy, red). This is different to the delayed recall component in the 6-CIT which includes a 5-part name and address, and the AMT-10 includes a 3-part address. Most of the cognitive assessment data were collected prospectively with the exception of a few patients who were untestable, but the AMT-4 had already been completed by the clinical team and was documented in the medical notes and the Cog-4 which was scored from the admission NIHSS score.

Table 5-2 Differences between published CSIs and use in this study

Test	Differing component	Instructions in original test	Content or scoring in this study
6-CIT	Delayed recall	5 components: John, Smith, High St, Bedford	5 words In MoCA
NINDS-CSN	Total score	12	11 (day of the week not asked)
AMT-10	Delayed recall	2 components: 42 West Street	5 words In MoCA

Table 5-3 Cognitive tests within our assessment and the DSM-5 cognitive domains covered

Domain (areas covered) / **CSI**	**Executive Function** (planning, decision-making, working memory, inhibition, flexibility)	**Learning and Memory** (free recall, recognition, semantic & autobiographical long-term memory)	**Language** (fluency, object naming, receptive language, grammar & syntax)	**Complex attention** (sustained, divided, selective, processing speed)	**Perceptual-motor** (visual perception, visuoconstructional reasoning, perceptual-motor coordination)
Clock-drawing test	Clock-draw				Clock-draw
Abbreviated MoCA	Clock-draw	5-word delayed recall			Clock-draw
4-AMT		Age, Year, Place, DOB			
4AT		Age, Year, Place, DOB		Months backwards	
Cog-4 (NIHSS items)		Age, Month	Aphasia assessment	Inattention (visual, tactile, auditory, or personal)	
6-CIT		Time, Month, Year, 5-part delayed recall		Count backwards from 20, Months backwards	
NINDS-CSN 5 min MoCA	Fluency (letter F)	Date, Month, Year, Day, Place, City, 5-word delayed recall	Fluency (letter F)		
10-AMT		Age, Time, Year, Place, 2-person recognition, DOB, Year of WW1, Current Prime Minister, 3-part delayed recall		Count backwards from 20	

It is acknowledged that it is debatable as to which domains some tasks fall under. In some cases, tasks are listed under more than one domain.

Clock drawing test (CDT)

The CDT is traditionally used to assess visuo-spatial and constructive abilities, but errors can also be due to memory or attentional dysfunction. It is included within numerous cognitive screening tests and occasionally used as a standalone test. It has proven to be useful in screening for dementia (249, 250), detecting cognitive impairment such as spatial neglect and has been validated in various settings. It is considered less useful for detecting milder forms of cognitive impairment (251).

There are different methods to administer the task, for example some CSIs include a pre-drawn circle and different times can be used. Scoring methods can vary substantially in terms of complexity and can be quantitative or qualitative. In the APPLE study this task is administered and scored according to the MoCA guidelines which score the task out of three (Figure 5-1). Some other tests which include the CDT state that the hand length is not scored e.g., the Mini-Cog (63), therefore I was unable to derive this test.

One point is allocated for each of the following three criteria:

- Contour (1 point): the clock face must be a circle with only minor distortion acceptable (e.g., slight imperfection on closing the circle);
- Numbers (1 point): all clock numbers must be present with no additional numbers; numbers must be in the correct order and placed in the approximate quadrants on the clock face; Roman numerals are acceptable; numbers can be placed outside the circle contour;
- Hands (1 point): there must be two hands jointly indicating the correct time; the hour hand must be clearly shorter than the minute hand; hands must be centred within the clock face with their junction close to the clock centre. A point is not assigned for a given element if any of the above criteria are not met.

Figure 5-1 MoCA scoring guidelines for the clock-drawing task(154)

Abbreviated MoCA (147)

This is the shortest of the SF-MoCA versions described in Chapter 3. The tasks retained from the Czech version of the original MoCA are clock-draw and delayed recall and the total score is eight. The validation paper included individuals aged 60 or over, without known cognitive impairment.

Cog-4 (252)

The Cog-4 uses the following four items taken from the NIHSS: level of consciousness which measures orientation (month & age), ability to follow commands, language, and extinction & inattention. It is a cognitive measure that is easy to obtain since the NIHSS is routinely completed. The Cog-4's limitations however have been documented by previous studies: poor accuracy when compared to the MoCA (253), significant floor effects and scores are dependent on side of stroke (254).

Six-Item Cognitive Impairment Test (6-CIT) (255)

The 6-CIT, previously known as the 'short Blessed test' or the 'six-item orientation-memory-concentration test', was created as a shortened version of the Mental Status Test (MST)(256). It is scored out of 28 but unlike most tests, higher scores indicate worse cognition. The items include current year and month, immediate recall of five components (unscored), current time, count backwards from 20, months of the year backwards, delayed recall of five components. The following score categories can be used: 0-7 normal, 8-9 mild cognitive impairment, 10-28 significant cognitive impairment. The 6-CIT has been used in both primary and secondary care, yet there are few validation studies available(257).

10-point Abbreviated Mental Test Score (10-AMT) (61) and 4-point AMT (4-AMT) (149)

The 10-AMT was designed from a list of 26 questions, and initially validated in hospital inpatients aged 65 or above (61). The items include age, current time, immediate recall of two components (unscored), current year and location,

recognising two people, DOB, year of WW1, name current president/prime minister, and count backwards from 20. A shortened 4-point AMT (4-AMT) was later introduced retaining four items: age, date of birth, place and year. It was initially validated against the 10-AMT in an outpatient population (149). Both tests have been widely used within hospitals in the UK and validated in acute medical settings (258).

The 4 'A's Test (4AT) (www.the4AT.com)

The 4AT is a four-item screening test mainly designed and used in delirium detection. Score thresholds however are also provided for possible cognitive impairment. The test consists of an observational assessment of alertness, the AMT-4, months of the year backwards and the final item asks whether there is evidence of acute change or fluctuation in mental status arising over the last 2 weeks and still evident over the past day. A score of 0 indicates that delirium or cognitive impairment is unlikely, 1-3 indicates possible cognitive impairment, ≥4 indicates possible delirium and cognitive impairment.

NINDS-CSN 5-min protocol (110)

As discussed in the introduction, the National Institute of Neurological Disorders and Stroke-Canadian Stroke Network (NINDS-CSN) recommended a 5-minute assessment which could be administered both in person and over the telephone. It is also one of the shortened forms of the MoCA detailed in Chapter 3, consisting of a five-part delayed recall, orientation (date, month, year, day of the week, place, city) and verbal fluency (total score 12). In the APPLE study we have data for all items apart from day of the week, therefore the total score in this study is 11.

5.5.2 One Month follow-up

The one-month follow-up was completed either at the hospital's clinical research facility, the patient's home or in hospital if the participant was still an inpatient. We repeated the same cognitive assessments as described at baseline (AMT-plus). In addition, the Oxford Cognitive Screen (OCS) was attempted with the participant. The OCS has been designed for the acute stroke setting and

creates a cognitive profile of strengths and weaknesses, rather than an overall score or global categorisation. A full description of the test can be found in Chapter 1. A copy of the OCS is provided in Appendix 13.

5.5.3 Six Month follow-up

At the six-month follow-up, we used one of three versions: a full-length version incorporating all assessments and scales, a shortened version or a telephone-based version. In-person assessments were prioritised and carried out at a clinical research facility or at the participant's home. The full-length version was also prioritised and used with the majority of participants. The shortened version was used in scenarios such as where participants had severe aphasia, hearing impairment, dementia or struggled at the last follow-up. Each case was considered separately and no a priori rules were set. We attended follow-ups with both versions available to use if required. Participants were reassured in all versions of assessment that they could stop at any time.

The telephone-based assessment was mainly used for participants who could not attend due to mobility issues or geographical barriers but was also convenient to obtain follow-up data for those participants with time constraints (e.g., those who work full-time).

5.5.3.1 Full length assessment

The full-length assessment took approximately 45 minutes – 1 hour to administer, consisting of a neuropsychological battery of tests followed by a range of questionnaires. Cognition was first assessed using the 'AMT-plus' again, followed by the NINDS-CSN 30-minute neuropsychological protocol.

Functional, mood and quality of life outcomes

For global disability and basic activities of daily living, the mRS and the Barthel were completed again. The 3-level version of the EQ-5D (259, 260) was used to assess health-related quality of life (HRQoL). There are two components to the scale: the EQ-5D-3L descriptive system and the EQ visual analogue scale (VAS). The first part asks the participant to rate whether they have no problems, some problems, or extreme problems across five areas (mobility, self-care, usual

activities, pain/discomfort, and depression/anxiety). This generates a 5-digit number reflecting responses for each question, for example 11111 would reflect no problems across the five areas. This 5-digit number was then transformed into a single summary utility index score, using the published UK time trade-off (TTO) validation (261). The VAS instructs the participant to rate how good their health is at the present moment, in their own opinion, on a scale from 0 (worst imaginable health state) to 100 (best imaginable health state).

The Center for Epidemiologic studies Depression Scale Revised (CESD-R) (262) contains 20 items covering nine symptom areas: sadness, loss of interest, appetite, sleep, concentration, guilt, fatigue, agitation, and suicidal ideation. The participant is asked how often they have experienced each symptom using five response options. The score range for the measure is 0 to 60, with higher scores indicating more frequent depressive symptoms. The symptoms reflect DSM-V criteria for a major depressive episode.

5.5.3.2 Short version

The shortened version took approximately 20 minutes to administer. Cognition was assessed using the 'AMT-plus', followed by the Montreal Cognitive Assessment (MoCA). The MoCA is scored out of 30, with higher scores reflecting better performance. Full description of the MoCA can be found in Chapter 1. The CESD-R, mRS, EQ-5D are also completed.

5.5.3.3 Telephone version

The telephone version took approximately 20 minutes to administer. For the cognitive assessment we used the TICS-m (101). A full description of the TICS-m and other telephonic CSIs are available in Chapter 4.

We also include the items from the 'AMT-plus' that can be delivered over the telephone: Time, Place, Date of Birth, Year of World War 1, News item, Months of the year backwards, Fluency (letter F). Only one of the brief CSIs (AMT-4) could be scored from the included questions.

The following scales were also completed over the telephone: mRS, CESD-R and the EQ-5D (descriptive section, not VAS).

5.6 Recruitment

Patients were recruited into the study across 11 UK hospital sites (8 NHS trusts): the Glasgow Royal Infirmary (GRI), Queen Elizabeth University Hospital (QEUH), Royal Alexandra Hospital, University Hospital Monklands, University Hospital Hairmyres, Aberdeen Royal Infirmary, Victoria Hospital (Fife), Forth Valley Royal Hospital, Perth Royal Infirmary, Charing Cross Hospital and Swansea (Figure 5-2).

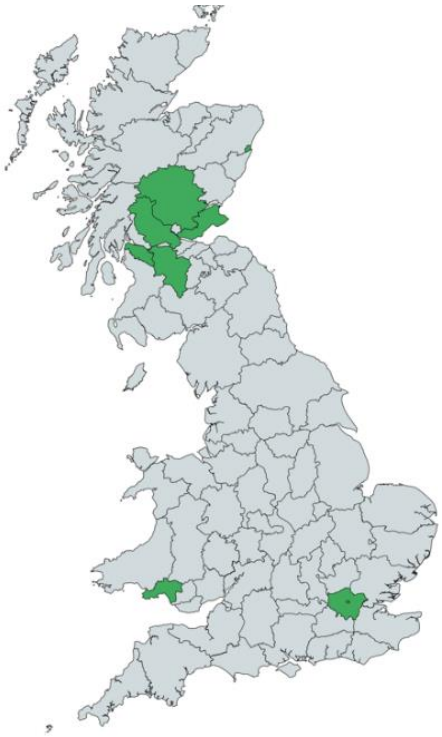


Figure 5-2 Map of APPLE recruitment sites

Recruitment commenced straight after ethical approval was obtained. The first participant was recruited into the study in November 2016 and the final in February 2019. We recruited 354 stroke participants and 151 informants in total. Full breakdown of recruitment per site is given in Table 5-4.

5.7 Case report forms

I collected and recorded anonymised participant data on paper case report forms along with two other PhD students and the research nurses involved at each hospital site. All files are securely filed in a locked room at each hospital, with another copy held securely at the Robertson centre for Biostatistics (RCB), University of Glasgow. The RCB managed the APPLE database for all sites,

including generating queries and dealing with database locks. All data was collected and handled according to International Conference on Harmonisation Good Clinical Practice (ICH-GCP) guidelines.

Table 5-4 Recruitment (stroke participants) across each NHS trust

NHS Health board and hospital	Number of participants recruited (N)
NHS Greater Glasgow & Clyde *GRI* *QEUH* *RAH*	Total: 196 *150* *22* *24*
NHS Fife	56
NHS Lanarkshire *Monklands* *Hairmyres*	47 *45* *2*
NHS Grampian	30
Imperial College Healthcare NHS trust	12
NHS Forth Valley	7
NHS Tayside	3
Swansea Bay University Health board	3
Total	354

6 Use of Brief Cognitive Screening Instruments post stroke: a longitudinal study

6.1 Introduction

For cognitive screening instruments (CSIs) to be implemented in acute medical settings, they need to be simple to deliver, since administration time is considered a key determinant of usage (153, 263). As discussed in Chapter 2, feasibility, and acceptability of administering a CSI in acute stroke is multifaceted; a range of factors relating to the participant, the HCP delivering the screen and the hospital setting itself should be acknowledged.

In addition to administration time, a CSI that can be used by any HCP, with minimal training, is also preferred. The SSNAP 2019 acute organisational audit report highlighted that only 7% of stroke units have access to a qualified clinical psychologist working full time (https://www.strokeaudit.org/results/Organisational/National-Organisational.aspx). The assessment paradigm recommended is the use of brief CSIs during the hyperacute period for initial triage, followed up with more detailed testing at later time-points (264). For this reason, sensitivity of tests is often favoured over specificity.

As described in Chapter 1, brief CSIs in this thesis refer to those CSIs that can be administered in under 5 minutes. In Chapter 3, various SF-MoCAs demonstrated high sensitivity for detecting cognitive impairment in the stroke validation sample. In addition to brief tests derived from the MoCA, there are many other brief CSIs available, for example those designed for primary care, yet many of these have not been validated in a stroke population. A previous systematic review examining test accuracy of CSIs for detection of multi-domain cognitive impairment or dementia in stroke concluded that there was no evidence that CSIs with longer administrations perform better in terms of accuracy (120). However only three studies with data on brief CSIs were found and included, providing data on the three cities test (265), CDT (266), Cog-4, AMT (10- and 4-point versions) and 4AT in stroke (267). Two of these studies used either the MoCA or MMSE as the reference standard, both of which have limitations, as discussed in Chapter 1. Although this systematic review did not address the

accuracy of CSIs to detect milder levels of cognitive impairment, there is evidence that being impaired in a single cognitive domain can have significant impact on functioning (268, 269).

The primary aims of this study were to calculate the rates of participants classified as 'test positive' according to eight different brief CSIs across three timepoints and to determine the accuracy of them using different reference standards: pre-stroke cognitive impairment (against clinical diagnosis) and post-stroke cognitive impairment (against the OCS). Secondary aims were to report on completion rates and floor/ceiling effects of each CSI across three timepoints.

6.2 Methods

6.2.1 Study design

This is a prospective, observational, longitudinal study using data from the APPLE study. Details regarding ethical approval, eligibility criteria, informed consent, and study conduct are available in the previous chapter. Most participants were recruited from hyper-acute/acute stroke units, with a minority recruited from a TIA clinic. The baseline visit was completed as soon as possible from recent stroke or TIA. The subsequent follow-ups were completed at approximately one and six months following the baseline visit, with a two-week window permitted either side of the follow-up date.

I followed the Standards for Reporting Diagnostic Accuracy in dementia (STARDdem) checklist (270) for the conduct and reporting of the study. The checklist is available in Appendix 5.

6.2.2 Index tests

The CSIs were administered by researchers working on the APPLE study (PhD students, research nurses and investigators) across 11 UK hospital sites. All researchers received an instruction/training manual, detailing how to administer and score the CSIs used in the study.

As detailed in the previous chapter, the questions included in the 'AMT-plus', along with sections from the CAM-ICU (245) and the NIHSS (244) were used to

score eight different CSIs: CDT, AMT-4 (149), AMT-10 (61), 6-CIT (255), 4AT *(www.the4AT.com)*, Cog-4 (252), Abbreviated MoCA (147), NINDS-CSN 5-min MoCA (110). These CSIs were chosen as they shared a number of test items in common and have some supportive evidence in other settings (271-273).

The relevant pages of the case report form (CRF) are available in Appendix 12. All index tests except the Cog-4 were completed at the same time-point. The Cog-4 was based on admission NIHSS and therefore from an earlier timepoint than the other CSIs.

At the one-month visit, the AMT-plus and CAM-ICU were completed again but not the NIHSS, therefore all CSIs except the Cog-4 could be scored. At six months, the AMT-plus was completed in the face-to-face follow-up, and all tests except Cog-4 and 4AT could be scored. For those participants completing the 6-month follow-up via telephone, only the AMT-4 could be scored.

Researchers were asked to document whether physical or verbal assistance was required to complete the cognitive assessment, for example if a participant had limb weakness, aphasia, hearing impairment, but these were not formally operationalised. Each CSI was administered and scored as close to the scale's published guidelines as possible (deviations are detailed in the previous chapter). Published recommended threshold scores were used to classify a participant as cognitively impaired (Table 6-1). Some CSIs provide more than one threshold for different severities of cognitive impairment, for example the 6-CIT has a threshold for 'MCI' and another for 'significant cognitive impairment'. In these cases, the threshold to detect the milder form of cognitive impairment was chosen.

Table 6-1 CSI threshold scores

Test	Direction of score reflecting better cognition	Threshold score/total score
Clock-drawing test	High score	<3/3
Abbreviated MoCA	High score	<4/8
Cog-4	Low score	>0/9
AMT-4	High score	<4/4
AMT-10	High score	<7/10
4AT	Low score	>0/12
6-CIT	Low score	>7/28
NINDS-CSN 5-min MoCA	High score	<10/11

Table 6-2 Scoring algorithm of each brief CSI

CSI	AMT-plus:																		NIHSS:				CAM-ICU:	
	Age	Time	Day	Month	Year	Place	City	Two-person recognition	Date of Birth	WW1	Prime minister	Count 20-1	Delayed Recall	Clock Face	Clock Numbers	Clock Hands	News items	Months backwards	Letter fluency	1b level of concsiousness	1c level of consciousness	9 Best language	11 Extinction & Inattention	Feature 1	Feature 3
CDT														x	x	x									
Abbrev. MoCA													x	x	x	x									
AMT-4	x				x	x			x																
4AT	x				x	x			x									x						x	x
6-CIT		x		x	x							x	x					x							
10-AMT	x	x			x	x		x	x	x	x	x	x												
NINDS-CSN 5-min MoCA			x	x	x	x	x						x						x						
Cog-4																				x	x	x	x		

Abbreviations: AMT, Abbreviated mental test; CDT, Clock drawing test; NINDS-CSN, National Institute of Neurological Disorders and Stroke-Canadian Stroke Network; 6-CIT, Six-item cognitive impairment test.

6.2.3 Reference standards

The first reference standard I used was a diagnosis of pre-existing cognitive impairment. These data were obtained through the patient's medical record and captured either a dementia or MCI diagnosis. Although the focus of this work is post-stroke impairment, it is important to also examine whether the CSIs identify pre-existing issues too. The same person collecting data from the medical record carried out the index tests and was therefore not blinded.

I used the Oxford Cognitive Screen (OCS) as the post-stroke reference standard (completed at one month). There is no perfect choice of reference standard for the acute period following a stroke and clinical diagnosis of MCI/dementia is not recommended until at least six months post-stroke (264). Although the OCS is also a screening test, it is more comprehensive than all the index tests (greater coverage of cognitive domains) and designed specifically for a stroke population. When compared to the MoCA in previous research, the OCS had higher sensitivity (274). A full description of this CSI can be found in Chapter 1.

Participants were scored on each of the 13 components of the OCS. Threshold scores are provided by the authors for each subtest and these were used to define impairment in a particular task (66) (Table 6-3). Each subtest sits within one of five OCS cognitive domains: attention, memory, language, praxis, and number processing. If a participant is impaired in a subtest, they are considered to be impaired within that domain (66). I used two categorisations of cognitive impairment: impaired in ≥1 domain (to capture single and multi-domain) and impaired in ≥2 domains (multi-domain only). All untestable data on the OCS were excluded to ensure that the categorisation of cognitive impairment was in fact related to cognition and not due to another reason, such as hemiparesis. Participants with a partially completed OCS were included and scored based on the tasks completed. Researchers carrying out the one-month assessment had access to the baseline CRF and therefore were potentially not blinded.

Table 6-3 OCS tasks and threshold scores

Domain	Task	Cut-off
Memory	Orientation	<4/4
	Recall & recognition	<3/4
	Episodic	<3/4
Language	Picture naming	<3/4
	Semantics	<3/4
	Sentence reading	<14/15
Number	Number writing	<3/3
	Calculation	<3/4
Attention	Broken hearts total	<42
	Space asymmetry	<-2 or >3
	Object asymmetry	<-2 or >1
Praxis	Gesture imitation	<8
Executive functioning	Executive score	>4*

*The executive score is calculated using the scores from three different trail tasks: circles, triangles and mixed

6.2.4 Analysis

At the time of writing this chapter, the database was not locked for the APPLE study as the host Clinical Trial Unit and biostatistics centre had not completed data queries and internal quality control checks following double data entry, therefore results may differ from subsequent publications.

I summarised descriptive statistics for the demographic and clinical characteristics of the sample; variables with normally distributed data are summarised by mean and standard deviation, whereas skewed data were summarised by median and interquartile range (IQR). I used SPSS version 27 (IBM, Armonk, NY, USA).

As discussed in the previous chapter, the APPLE database was managed by a biostatistics centre, and I had access to the data for the variables required for my PhD. To score each index test I created a scoring spreadsheet (Table 6-2; full version available in Appendix 14) for the biostatistics centre, detailing items and scoring for each CSI. I performed quality checks by independently scoring the CSIs for a subsample of participants to compare against the RCB derived scores. The subsample checked were chosen on a stratified basis, with a focus on checking the scoring of those with partially and fully untestable CSI data. Any issues found were resolved through discussion with the biostatistics centre and scoring was re-calculated again where necessary.

I made some assumptions for scoring the CSIs. For three questions (months of the year backwards, count backwards from 20 and 5-word delayed recall), I required the number of errors made by the participant. In the CRF, researchers were asked to document the number of errors made by the participant, as well as indicate 'yes' or 'no' according to certain rules: yes for delayed recall = ≥4 words correct; yes for count backwards = all correct; yes for months of the year backwards = ≥7 months correct). For months of the year backwards, count backwards from 20 and 5-word delayed recall, if yes was ticked, and number of mistakes missing, I assumed these cases were 0 mistakes. Where no was ticked and number of errors missing, I considered their possible score range, for example a 'no' for delayed recall could mean the participant scored between 0-3 correct. If this score range would alter their categorisation of being impaired the participant was excluded.

To address the multiple research objectives the following analyses were carried out for each CSI, which are detailed in the following subsections:

- Percentages of participants falling below the threshold score and classified as test positive;

- Test accuracy of baseline CSI compared to a pre-stroke diagnosis of MCI/dementia;

- Prognostic accuracy of baseline CSI compared to one-month OCS;

- Test accuracy of one month CSI compared to one-month OCS;
- Completion rates and reasons for incompletion;
- Floor/ceiling effects.

6.2.4.1 Dealing with incomplete and untestable data

To build on the work completed in Chapter 2, I anticipated data to be missing for varying reasons. In the CRF, a range of pre-specified untestable reasons were provided for researchers to choose where assessment could not be fully completed: aphasia, limb weakness, confused, drowsy/reduced consciousness, deaf, motor problem, visual problem, unwell, refused, other (specify).

For two CSIs (4AT and Cog-4), there are ways for dealing with untestable data built into the scoring. For the other CSIs, methods for missing data were planned a priori. For this particular study, I considered data to be missing under the following conditions: missing CRF, participants declined individual questions or the full test; participants were discharged; a task could not be completed due to positioning of the participant; a task was not completed due to a researcher error (e.g., missed); missing data with reason unknown.

Participants who could not be tested for other reasons, for example aphasia, limb weakness, confusion, reduced consciousness were classified as impaired for that item (and assigned the corresponding impaired score). This approach mirrors that taken in the 4AT and is supported by previous research (275) as it reduces the number of type two errors (false negatives).

6.2.4.2 CSI results and test accuracy

At all three time-points (baseline, one and six months), I calculated the percentage of participants categorised as ‘test positive’ on each of the eight CSIs using recommended thresholds.

To estimate sample sizes required in test accuracy studies, simple nomograms can be used (276). There are four elements to the nomogram (prevalence, sensitivity/specificity, confidence intervals and number of patients) and if a

researcher knows any three, the fourth can be calculated. In the APPLE protocol, an estimate of 400 participants was made, based on a prevalence of 40% cognitive impairment at one month (a=0.05).

The results of each CSI were evaluated against the two reference standard assessments. When the baseline brief CSIs were compared to the OCS at one month, I refer to this as prognostic test accuracy, since the index test and reference standard were completed at different timepoints, whereas when the one-month brief CSI results were compared to the OCS I refer to this as test accuracy (completed at same timepoint).

Sensitivity, specificity, positive and negative predictive values (PPV/NPV) and area under the receiver operating characteristic curve (AUC) were calculated using the Delong et al. (1988) method (277). An AUC value of 0.5 indicates the test is no better than chance at detecting the desired outcome. Values of 0.7-0.8 are considered acceptable, 0.8-0.9 excellent and 0.9-1.00 outstanding (278). Based on the reasons discussed in the introduction, CSIs with a high sensitivity (>0.7) were required to be recommended, rather than those with high specificity. The optimal threshold score, determined by the Youden index method, was also determined. The Youden index however gives equal weight to sensitivity and specificity. I used MedCalc version 19.5.3 (MedCalc Software Ltd, Ostend, Belgium) for these analyses.

For the primary analyses, participants who were untestable were included. I then ran sensitivity analyses excluding these participants to examine if accuracy altered. This was done for all CSIs apart from the 4AT and Cog-4, which already incorporate scoring for these scenarios.

6.2.4.3 Completion rates, floor/ceiling effects, errors

The number of participants with a complete data set (no missing or untestable data) for each CSI was recorded. Floor and ceiling effects were calculated for both the individual items of the AMT-plus and of the eight named CSIs, through calculating the percentage of participants scoring zero (floor effect) or full marks (ceiling effect) on each item/CSI. Three of the CSIs have the opposite scoring structure, in that a high score indicates poorer cognition (4AT, Cog-4 and

6-CIT), so for these instruments a score of zero was captured as the ceiling effect. A criterion of >15% was applied to determine if floor/ceiling effects were present (116).

I also recorded the types of errors made by researchers in administrating/scoring the AMT-plus items. Throughout the APPLE study, CRFs from other hospital sites were sent to me and the rest of the team at the host clinical site. We checked the scoring of tasks which could be scored retrospectively (e.g., clock draw).

6.3 Results

There were 354 participants recruited into the APPLE study (97% recruited from a HASU). Follow-up data were available for 268 participants at one month and 220 participants at six months. The participants at each timepoint are detailed in the flow diagram in Figure 6-1.

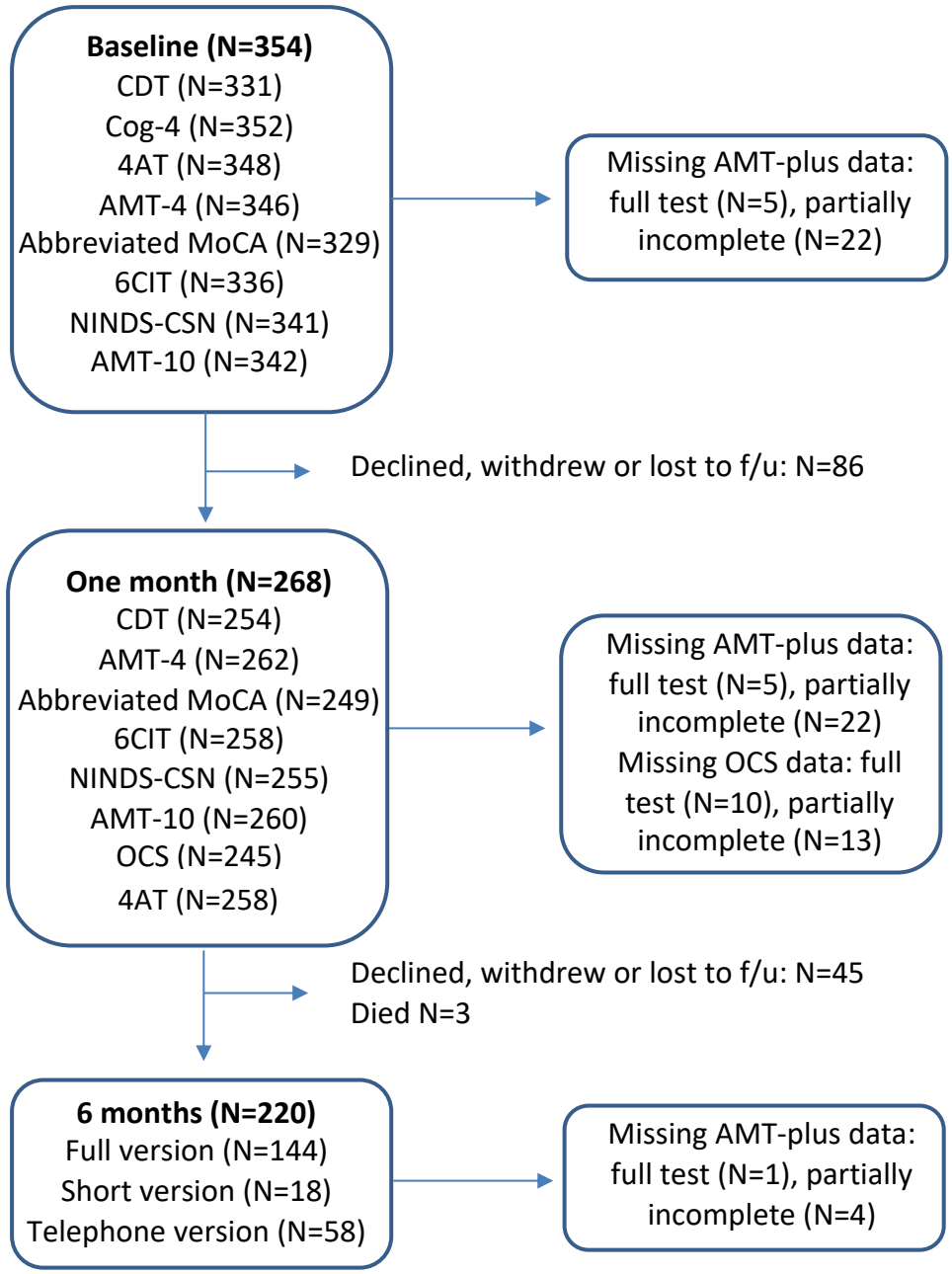


Figure 6-1 Flow diagram: Fully completed CSIs at each timepoint

Table 6-4 Characteristics of the sample (N=354)

Variable	Summary statistic
Age: Mean (SD)	69.1 (12.8)
Sex: N Male (%)	197 (56%)
Stroke Classification (includes IS & ICH) Missing data = 2	TACS: 33 (9%) PACS: 118 (34%) LACS: 85 (24%) POCS: 68 (20%) TIA: 47 (13%)
Side of brain affected by index stroke (includes TACS/PACS/LACS/TIA) Missing data = 2	Right: 146 (41%) Left: 131 (37%) Bilateral: 5 (1%)
NIHSS: Median (IQR) Missing data = 2	2 (1-4)
Pre-morbid mRS: Median (IQR) Missing data = 3	0 (0-2)
Capacity to consent themselves: N (%)	332 (93%)
Previous stroke: N (%)	87 (25%)
Years in education: Mean (SD) Missing data = 31	12.0 (3.4)
Pre-stroke dementia or MCI: N (%)	26 (7%)
Presence of aphasia on admission NIHSS Q.9 (>0): N (%)	48 (14%)
Limb weakness on admission NIHSS Q.5 (>0): N (%)	Right: 60 (17%) Left: 90 (26%)
Hearing impairment: N (%) Missing data = 2	63 (18%)
Visual impairment: N (%) Missing data = 2	96 (27%)
Recruited from TIA clinic	9 (3%)

Abbreviations: HASU, hyper acute stroke unit; ICH, Intracerebral haemorrhage; IS, Ischaemic stroke; IQR, Interquartile range; LACS, lacunar stroke; NIHSS, National Institute of Health Stroke Scale; PACS, Partial anterior circulation stroke; POCS, posterior circulation stroke; TACS, Total anterior circulation stroke; TIA, Transient ischaemic attack

6.3.1 Baseline

The baseline assessment took place at a median of 6 days (IQR 4-9) post stroke and a median of 4 days (IQR 3-7) post admission to the stroke unit or TIA clinic. Five participants were CAM-ICU positive, indicating a potential presence of delirium.

6.3.1.1 AMT-plus completion rates

Five participants were fully untestable on the AMT-plus (did not complete any questions) due to aphasia, confusion, and reduced consciousness. There were 22 participants with partially completed AMT-plus; 10 of these participants were classified partially untestable due to motor problems, aphasia, deafness, visual impairment, confusion and 12 participants had incomplete assessments but were not classified as untestable, with the following reasons: discharged, declined, taken from medical records, completed via telephone, and missing with unknown reason. For completion of the AMT-plus: 11 participants required physical assistance to complete the AMT-plus and 46 required verbal assistance.

6.3.1.2 CSI results and accuracy

At baseline, the percentage of participants classified as impaired across the 8 CSIs ranged 12% (AMT-10) to 69% (CDT), using recommended thresholds. Sensitivity of each CSI when comparing the results of the CSIs to a pre-stroke cognitive impairment syndrome ranged 0.46 (Cog-4) to 0.84 (NINDS-CSN 5-min MoCA). Specificity ranged 0.32 (CDT) to 0.89 (AMT-4). AUC ranged 0.63-0.81 and was highest for AMT-10 and 6-CIT. In sensitivity analyses, where untestable participants were removed, sensitivity decreased/specificity increased slightly across all CSIs (Table 6-5).

When comparing the accuracy of each CSI to detect any level of cognitive impairment (single and multi-domain) at 1 month, sensitivity ranged 0.14 (AMT-10) to 0.76 (CDT). Specificity ranged 0.45 (CDT) to 0.96 (AMT-4). AUC ranged 0.56-0.70 and was highest for the NINDS-CSN 5-min MoCA. In sensitivity analyses, where untestable participants were removed, sensitivity decreased/specificity increased slightly or stayed the same across all CSIs (Table 6-6).

When comparing the accuracy of each CSI to detect multi-domain cognitive impairment at 1 month, sensitivity ranged 0.25 (AMT-10) to 0.87 (CDT). Specificity ranged 0.45 (CDT) to 0.96 (AMT-4). AUC ranged 0.62-0.77 and was highest for the NINDS-CSN 5-min MoCA. In sensitivity analyses, where untestable participants were removed, sensitivity decreased/specificity increased slightly or stayed the same across all CSIs (Table 6-7).

Table 6-5 Accuracy of each CSI at baseline to identify pre-stroke diagnosis of dementia or MCI

	Main analyses (including untestable)							Sensitivity analyses					
CSI (threshold)	N	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)	Optimal threshold	N	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)
CDT (<3)	344	0.75 (0.53-0.90)	0.32 (0.27-0.37)	0.08 (0.06-0.10)	0.94 (0.89-0.97)	0.63 (0.58-0.68)	<2	331	0.71 (0.48-0.89)	0.33 (0.27-0.38)	0.07 (0.05-0.09)	0.94 (0.89-0.97)	0.60 (0.55-0.65)
AMT-4 (<4)	351	0.64 (0.43-0.82)	0.89 (0.85-0.92)	0.31 (0.23-0.41)	0.97 (0.95-0.98)	0.77 (0.72-0.81)	<4	346	0.61 (0.39-0.80)	0.90 (0.86-0.93)	0.30 (0.22-0.41)	0.97 (0.95-0.98)	0.76 (0.71-0.80)
AMT-10 (<7)	349	0.52 (0.31-0.73)	0.91 (0.87-0.94)	0.29 (0.19-0.40)	0.96 (0.95-0.98)	0.81 (0.76-0.85)	<8	334	0.45 (0.23-0.69)	0.94 (0.90-0.96)	0.31 (0.19-0.46)	0.96 (0.95-0.98)	0.80 (0.75-0.84)
6-CIT (>7)	343	0.71 (0.49-0.87)	0.75 (0.70-0.80)	0.18 (0.13-0.23)	0.97 (0.95-0.99)	0.81 (0.77-0.85)	>8	336	0.68 (0.45-0.86)	0.76 (0.71-0.81)	0.17 (0.13-0.22)	0.97 (0.95-0.98)	0.81 (0.76-0.85)
Abbreviated MoCA (<4)	341	0.54 (0.33-0.74)	0.83 (0.78-0.87)	0.19 (0.13-0.27)	0.96 (0.94-0.97)	0.75 (0.70-0.80)	<5*	322	0.45 (0.23-0.69)	0.86 (0.82-0.90)	0.18 (0.11-0.27)	0.96 (0.94-0.97)	0.70 (0.65-0.75)
NINDS-CSN 5-min MoCA (<10)	350	0.84 (0.64-0.96)	0.38 (0.33-0.44)	0.10 (0.08-0.11)	0.97 (0.93-0.99)	0.77 (0.72-0.81)	<7	334	0.82 (0.60-0.95)	0.40 (0.35-0.46)	0.09 (0.07-0.11)	0.97 (0.93-0.99)	0.75 (0.70-0.79)
4AT (>0)	348	0.65 (0.43-0.84)	0.77 (0.72-0.81)	0.17 (0.12-0.22)	0.97 (0.95-0.98)	0.72 (0.67-0.76)	>0						
Cog-4 (>0)	352	0.46 (0.27-0.67)	0.78 (0.73-0.83)	0.15 (0.10-0.21)	0.95 (0.93-0.96)	0.63 (0.58-0.68)	>0						

*<4 in sensitivity analysis. Abbreviations: AMT, Abbreviated mental test; AUC, Area under curve; CDT, Clock drawing test; MoCA, Montreal Cognitive Assessment; NINDS-CSN, National Institute of Neurological Disorders and Stroke-Canadian Stroke Network; NPV, Negative predictive value; PPV, Positive predictive value; 4AT, 4 A's test; 6-CIT, Six-item cognitive impairment test.

Table 6-6 Prognostic accuracy of each CSI at baseline to identify post-stroke cognitive impairment (single and multi-domain)

	Main analyses (including untestable)							Sensitivity analyses				
CSI	N	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)	Optimal threshold	N	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)
CDT (<3)	254	0.74 (0.66-0.81)	0.45 (0.35-0.54)	0.64 (0.59-0.68)	0.57 (0.48-0.65)	0.65 (0.59-0.71)	<2	247	0.73 (0.65-0.80)	0.45 (0.35-0.55)	0.63 (0.58-0.67)	0.57 (0.48-0.65)	0.64 (0.58-0.70)
AMT-4 (<4)	258	0.16 (0.11-0.23)	0.96 (0.91-0.99)	0.86 (0.68-0.94)	0.47 (0.45-0.49)	0.56 (0.50-0.63)	<4	257	0.16 (0.10-0.23)	0.96 (0.91-0.99)	0.85 (0.67-0.94)	0.47 (0.45-0.49)	0.56 (0.50-0.62)
AMT-10 (<7)	255	0.14 (0.09-0.20	0.97 (0.92-0.99)	0.87 (0.67-0.96)	0.46 (0.44-0.48)	0.66 (0.60-0.72)	<9	249	0.09 (0.05-0.16)	0.97 (0.92-0.99)	0.81 (0.56-0.94)	0.46 (0.45-0.48)	0.65 (0.58-0.71)
6-CIT (>7)	254	0.36 (0.28-0.46)	0.93 (0.86-0.97)	0.87 (0.76-0.93)	0.53 (0.49-0.56)	0.69 (0.63-0.75)	>6	252	0.35 (0.27-0.43)	0.93 (0.86-0.97)	0.86 (0.75-0.93)	0.52 (0.49-0.56)	0.69 (0.62-0.74)
Abbreviated MoCA (<4)	253	0.25 (0.18-0.32)	0.93 (0.86-0.97)	0.81 (0.68-0.90)	0.49 (0.46-0.51)	0.67 (0.61-0.73)	<6	242	0.21 (0.14-0.29)	0.94 (0.88-0.98)	0.82 (0.67-0.92)	0.49 (0.47-0.52)	0.67 (0.60-0.73)
NINDS-CSN 5-min MoCA (<10)	257	0.73 (0.65-0.80)	0.55 (0.45-0.64)	0.68 (0.63-0.73)	0.61 (0.53-0.68)	0.70 (0.64-0.75)	<10	249	0.72 (0.64-0.79)	0.55 (0.46-0.65)	0.67 (0.62-0.72)	0.61 (0.53-0.68)	0.69 (0.62-0.74)
4AT (>0)	257	0.30 (0.23-0.38)	0.91 (0.84-0.96)	0.82 (0.70-0.89)	0.50 (0.47-0.53)	0.61 (0.54-0.67)	>0						
Cog-4 (>0)	258	0.29 0.21-0.37)	0.86 (0.78-0.92)	0.72 (0.61-0.82)	0.48 (0.44-0.51)	0.57 (0.51-0.63)	>0						

Abbreviations AMT, Abbreviated mental test; AUC, Area under curve; CDT, Clock drawing test; MoCA, Montreal Cognitive Assessment; NINDS-CSN, National Institute of Neurological Disorders and Stroke-Canadian Stroke Network; NPV, Negative predictive value; PPV, Positive predictive value; 4AT, 4 A's test; 6-CIT, Six-item cognitive impairment test.

Table 6-7 Prognostic accuracy of each CSI at baseline to identify post-stroke cognitive impairment (multi-domain)

	Main analyses (including untestable)							Sensitivity analyses				
CSI (threshold)	N	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)	Optimal threshold	N	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)
CDT (<3)	253	0.87 (0.77-0.94)	0.41 (0.34-0.48)	0.33 (0.30-0.36)	0.91 (0.83-0.95)	0.74 (0.68-0.79)	<2	247	0.87 (0.75-0.94)	0.42 (0.35-0.49)	0.32 (0.29-0.36)	0.91 (0.83-0.95)	0.74 (0.68-0.79)
AMT-4 (<4)	257	0.29 (0.19-0.42)	0.96 (0.92-0.98)	0.70 (0.52-0.84)	0.80 (0.77-0.82)	0.63 (0.56-0.69)	<4	257	0.29 (0.19-0.42)	0.96 (0.92-0.98)	0.70 (0.52-0.84)	0.80 (0.77-0.82)	0.63 (0.56-0.69)
AMT-10 (<7)	256	0.25 (0.15-0.37)	0.97 (0.93-0.99)	0.73 (0.52-0.87)	0.80 (0.77-0.82)	0.72 (0.66-0.78)	<9	249	0.19 (0.10-0.31)	0.97 (0.94-0.99)	0.69 (0.44-0.86)	0.80 (0.78-0.82)	0.70 (0.64-0.76)
6-CIT (>7)	253	0.52 (0.39-0.64)	0.86 (0.81-0.91)	0.56 (0.45-0.66)	0.84 (0.80-0.87)	0.73 (0.67-0.78)	>6	252	0.50 (0.37-0.63)	0.86 (0.81-0.91)	0.54 (0.44-0.65)	0.84 (0.80-0.87)	0.72 (0.66-0.78)
Abbreviated MoCA (<4)	252	0.38 (0.26-51)	0.90 (0.85-0.94)	0.57 (0.44-0.70)	0.81 (0.78-0.84)	0.74 (0.68-0.79)	<6	242	0.32 (0.20-0.45)	0.91 (0.86-0.95)	0.53 (0.38-0.67)	0.81 (0.78-0.84)	0.73 (0.66-0.78)
NINDS-CSN 5-min MoCA (<10)	256	0.86 (0.75-0.93)	0.47 (0.40-0.55)	0.35 (0.32-0.39)	0.91 (0.84-0.95)	0.77 (0.71-0.82)	<8	249	0.85 (0.73-0.93)	0.48 (0.41-0.55)	0.34 (0.30-0.38)	0.91 (0.85-0.95)	0.75 (0.69-0.80)
4AT (>0)	256	0.50 (0.37-0.63)	0.89 (0.84-0.93)	0.60 (0.49-0.71)	0.84 (0.81-0.87)	0.70 (0.64-0.76)	>0						
Cog-4 (>0)	257	0.40 (0.28-0.53)	0.84 (0.78-0.89)	0.46 (0.35-0.57)	0.81 (0.77-0.84)	0.62 (0.56-0.68)	>0						

Abbreviations: AMT, Abbreviated mental test; AUC, Area under curve; CDT, Clock drawing test; MoCA, Montreal Cognitive Assessment; NINDS-CSN, National Institute of Neurological Disorders and Stroke-Canadian Stroke Network; NPV, Negative predictive value; PPV, Positive predictive value; 4AT, 4 A's test; 6-CIT, Six-item cognitive impairment test.

6.3.2 One Month

The one-month assessment took place at a mean of 42.3 days (SD 19.5) post-stroke. Between the baseline and one-month assessment, three participants had a further stroke.

6.3.2.1 AMT-plus completion rates

Five participants were fully untestable on the AMT-plus (did not complete any questions) due to aphasia, unwell, reduced consciousness, confusion. This included two participants who had been previously deemed testable at baseline. There were 12 participants with partially completed AMT-plus; 10 of these classified partially untestable due to motor problems, aphasia, deafness, visual impairment, confusion, language barrier and two participants had incomplete assessments due to declining the assessment but were not classified as untestable. For completion of the AMT-plus: 9 participants required physical assistance and 35 required verbal assistance.

6.3.2.2 CSI results and accuracy

At one month the percentage of participants classified as impaired across the CSIs ranged 8% (AMT-10) to 64% (CDT). Sensitivity to detect any level of cognitive impairment (as detected by the OCS) ranged 0.17 (AMT-4) to 0.75 (CDT). Specificity ranged 0.54 (CDT) to 1.00 (AMT-10, Abbreviated MoCA). AUC ranged 0.58-0.70 and was highest for the CDT and Abbreviated MoCA (Table 6-8). In the sensitivity analyses, where untestable participants were removed, sensitivity decreased/specificity increased slightly across all CSIs (Table 6-8).

When comparing the accuracy of each CSI to detect multi-domain cognitive impairment, sensitivity ranged 0.18 (AMT-10) to 0.86 (CDT). Specificity ranged 0.45 (CDT) to 0.99 (AMT-10). AUC ranged 0.63-0.72 and was highest for the NINDS-CSN 5-min MoCA. In sensitivity analyses, where untestable participants were removed, sensitivity decreased/specificity increased slightly or stayed the same across all CSIs.

Table 6-8 Test accuracy of each CSI to identify post-stroke cognitive impairment (single and multi-domain)

	Main analyses (including untestable)							Sensitivity analyses					
CSI (threshold)	N	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)	Optimal threshold	N	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)
CDT (<3)	256	0.75 (0.67-0.81)	0.54 (0.44-0.64)	0.68 (0.63-0.73)	0.62 (0.54-0.69)	0.70 (0.64-0.76)	<2	252	0.74 (0.66-0.81)	0.54 (0.44-0.64)	0.67 (0.62-0.72)	0.62 (0.54-0.69)	0.69 (0.63-0.75)
AMT-4 (<4)	258	0.17 (0.11-0.24)	0.98 (0.94-1.00)	0.93 (0.75-0.98)	0.47 (0.45-0.49)	0.58 (0.51-0.64)	<4	256	0.16 (0.10-0.23)	0.98 (0.94-1.00)	0.92 (0.74-0.98)	0.47 (0.45-0.49)	0.57 (0.51-0.63)
AMT-10 (<7)	258	0.10 (0.05-0.16)	1.00 (0.97-1.00)	1.00	0.46 (0.44-0.47)	0.68 (0.61-0.73)	<10	253	0.07 (0.03-0.13)	1.00 (0.97-1.00)	1.00	0.46 (0.45-0.47)	0.67 (0.60-0.72)
6-CIT (>7)	257	0.30 (0.23-0.38)	0.98 (0.94-1.00)	0.96 (0.85-0.99)	0.51 (0.49-0.54)	0.69 (0.63-0.75)	>2	254	0.27 (0.20-0.35)	0.98 (0.94-1.00)	0.95 (0.83-0.99)	0.51 (0.48-0.53)	0.68 (0.62-0.74)
Abbreviated MoCA (<4)	253	0.19 (0.13-0.26)	1.00 (0.97-1.00)	1.00	0.49 (0.47-0.51)	0.70 (0.64-0.75)	<6	248	0.16 (0.10-0.23)	1.00 (0.97-1.00)	1.00	0.48 (0.47-0.51)	0.69 (0.63-0.75)
NINDS-CSN 5-min MoCA (<10)	258	0.63 (0.54-0.70)	0.60 (0.51-0.70)	0.68 (0.62-0.73)	0.55 (0.49-0.61)	0.68 (0.62-0.74)	<9	251	0.61 (0.52-0.69)	0.61 (0.51-0.70)	0.67 (0.61-0.72)	0.55 (0.49-0.61)	0.67 (0.61-0.73)
4AT (>0)	258	0.29 (0.22-0.37)	0.95 (0.89-0.98)	0.88 (0.76-0.94)	0.50 (0.47-0.53)	0.62 (0.56-0.68)	>0						

Abbreviations: AMT, Abbreviated mental test; AUC, Area under curve; CDT, Clock drawing test; MoCA, Montreal Cognitive Assessment; NINDS-CSN, National Institute of Neurological Disorders and Stroke-Canadian Stroke Network; NPV, Negative predictive value; PPV, Positive predictive value; 4AT, 4 A's test; 6-CIT, Six-item cognitive impairment test.

Table 6-9 Test accuracy metrics of each CSI against the OCS (multi-domain CI)

	Main analyses (including untestable)							Sensitivity analyses				
CSI (threshold)	N	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)	Optimal threshold	N	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)
CDT (<3)	255	0.86 (0.75-0.93)	0.46 (0.39-0.53)	0.34 (0.31-0.38)	0.91 (0.84-0.95)	0.75 (0.69-0.80)	<2	252	0.85 (0.74-0.93)	0.46 (0.39-0.53)	0.34 (0.30-0.37)	0.91 (0.84-0.95)	0.74 (0.68-0.79)
AMT-4 (<4)	257	0.29 (0.19-0.42)	0.96 (0.93-0.99)	0.73 (0.55-0.86)	0.80 (0.77-0.83)	0.63 (0.57-0.69)	<4	256	0.28 (0.18-0.41)	0.96 (0.93-0.99)	0.72 (0.53-0.86)	0.80 (0.78-0.83)	0.62 (0.56-0.68)
AMT-10 (<7)	257	0.18 (0.10-0.30)	0.99 (0.97-1.00)	0.92 (0.61-0.99)	0.78 (0.76-0.80)	0.72 (0.67-0.78)	<10	253	0.15 (0.07-0.26)	0.99 (0.97-1.00)	0.90 (0.54-0.99)	0.79 (0.77-0.80)	0.71 (0.65-0.77)
6-CIT (>7)	256	0.43 (0.31-0.56)	0.91 (0.86-0.95)	0.62 (0.49-0.74)	0.83 (0.79-0.85)	0.72 (0.66-0.77)	>5	254	0.40 (0.28-0.53)	0.92 (0.87-0.95)	0.61 (0.47-0.73)	0.82 (0.79-0.85)	0.71 (0.65-0.76)
Abbreviated MoCA (<4)	252	0.28 (0.17-0.41)	0.95 (0.91-0.98)	0.65 (0.47-0.80)	0.81 (0.78-0.83)	0.72 (0.66-0.78)	<6	248	0.24 (0.14-0.37)	0.96 (0.92-0.98)	0.64 (0.44-0.80)	0.81 (0.78-0.83)	0.71 (0.65-0.77)
NINDS-CSN 5-min MoCA (<10)	257	0.71 (0.58-0.81)	0.54 (0.46-0.61)	0.34 (0.29-0.39)	0.84 (0.78-0.89)	0.72 (0.66-0.78)	<8	251	0.68 (0.55-0.80)	0.54 (0.47-0.61)	0.32 (0.27-0.37)	0.84 (0.79-0.89)	0.70 (0.64-0.76)
4AT (>0)	257	0.45 (0.32-0.58)	0.90 (0.85-0.94)	0.60 (0.48-0.72)	0.83 (0.79-0.86)	0.68 (0.62-0.74)	>0						

Abbreviations: AMT, Abbreviated mental test; AUC, Area under curve; CDT, Clock drawing test; MoCA, Montreal Cognitive Assessment; NINDS-CSN, National Institute of Neurological Disorders and Stroke-Canadian Stroke Network; NPV, Negative predictive value; PPV, Positive predictive value; 4AT, 4 A's test; 6-CIT, Six-item cognitive impairment test.

6.3.3 Six months

The six-month assessment took place at a mean of 195.7 days (SD 21.3) post-stroke. 144 participants completed the full follow-up, 18 completed the shortened version, and 58 completed the telephone-based version.

6.3.3.1 AMT-plus completion rates

One participant was fully untestable on the AMT-plus (did not complete any questions) due to aphasia. This participant was also untestable at the previous timepoints. There were four participants with partially completed AMT-plus; 3 of these classified partially untestable due to motor problems and one participant had an incomplete assessment without a documented reason so were not classified as untestable. For completion of the AMT-plus: 6 participants required physical assistance to complete the AMT-plus and 23 required verbal assistance.

6.3.3.2 CSI results

The percentage of participants classified as impaired ranged 7% (AMT-10) to 64% (CDT).

6.3.4 All timepoints

6.3.4.1 CSI results

Figure 6-2 illustrates the percentage of participants classified as impaired across the three timepoints. In this sample of participants, across all timepoints, the CDT resulted in the greatest number of participants classified as cognitively impaired and the AMT-10 resulted in the least. Between the first and second timepoints, the percentage of participants classified as impaired decreased across all CSIs. Between the second and third timepoints, the percentage of those classified as impaired decreased or plateaued for all but one test: the 5-min MoCA recommended by NINDS-CSN which increased from 54-59%. The largest decrease was 9% between baseline and one month using the NINDS-CSN 5-min MoCA.

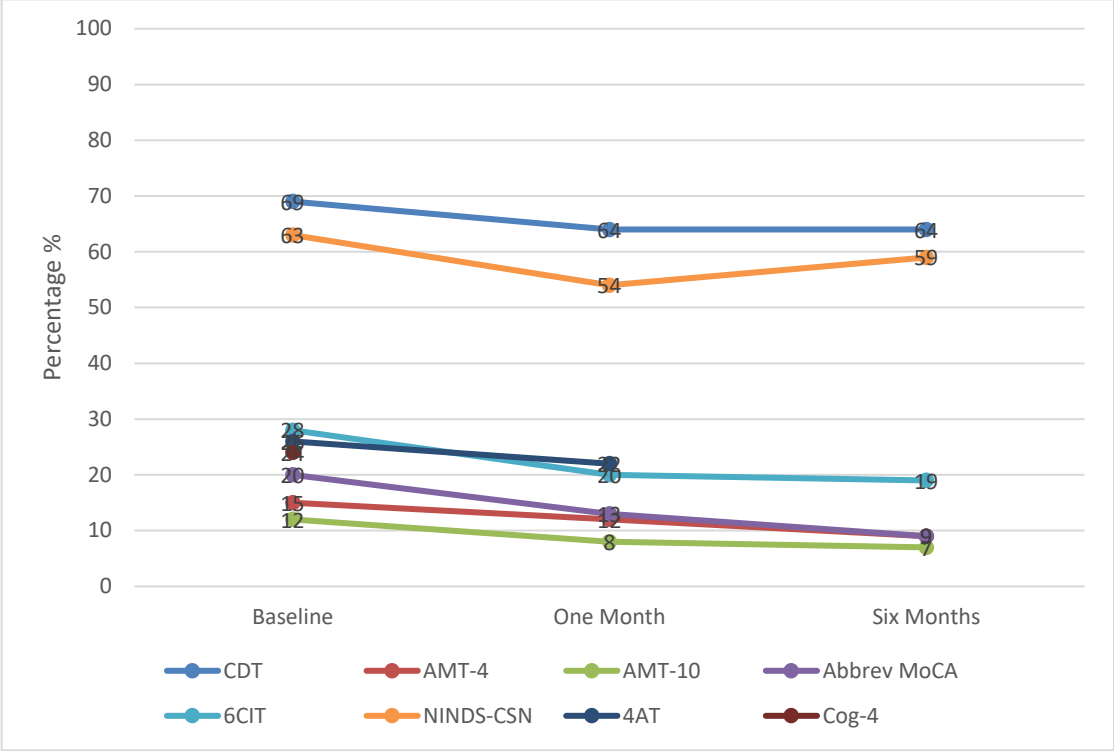


Figure 6-2 Percentage of participants scoring below the threshold for each test (classified as cognitively impaired) at each time-point using each CSI

6.3.4.2 Completion rates, floor/ceiling effects, errors

Completion of the different test items included in the AMT-plus ranged 94-98% at baseline, 95-98% at one month, and 98-100% at six months. Across all timepoints, the clock-draw had the lowest completion rate. The percentage of items correct ranged 39-99% at baseline, 44-99% at one month, and 43-100% at 6 months. At baseline and one month, the date of birth (DOB) item had the highest percentage correct and clock draw hands had the lowest percentage correct. At 6 months, all participants got the two-person recognition item correct and verbal fluency had the lowest percentage correct (Table 6-11).

Across the eight CSIs at baseline, the Cog-4 had the highest completion rate (99%) and the Abbreviated MoCA had the lowest (93%). All CSIs demonstrated ceiling effects (>15%) but effects were highest for the Cog-4 (99%), 4AT (98%) and AMT-4 (87%). Across the seven CSIs at one month, the 4AT had the highest completion rate (100%) and the Abbreviated MoCA had the lowest (93%). All CSIs demonstrated ceiling effects, but effects were highest for the 4AT (78%) and AMT-4 (90%). Across the six CSIs at 6 months, completion rates were similar (98-

99%). Four out of six CSIs demonstrated ceiling effects: 6-CIT (25%), AMT-4 (92%), AMT-10 (45%), CDT (37%) (Table 6-12).

On reviewing the CRFs from other sites, some errors were identified for three of the AMT-plus items: verbal fluency, the CDT, and months of the year backwards (Table 6-10). In some cases, the task could be re-scored, whereas no action could be taken if it resulted from an administration error. For verbal fluency, it was decided that if the administrator did not specify that proper nouns/numbers were not permitted, it was unfair to penalise the participant. Additional training materials clarifying these items were provided to the sites (Appendix 15).

Table 6-10 Errors identified in administration and scoring of AMT-plus

AMT-plus item	Error type	Action
Fluency	Different letter to F used	None
	Proper nouns or numbers scored as correct	None
Clock draw	Different time used to 11:10	None
	Scoring errors e.g. hands drawn as the same length marked as correct	Re-scored
Months of the year backwards	Scoring and recording ≥7 months as correct	Re-scored (if the participant's responses were recorded)

Table 6-11 Data completeness and percentage correct for each AMT-plus question

Item	Baseline (N=354)		One month (N=268)		Six months* (N=162)	
	Completeness N (%)	N (%) correct	Completeness N (%)	N (%) correct	Completeness N (%)	N (%) correct
Age	346 (98%)	330 (95%)	262 (98%)	255 (97%)	162 (100%)	152 (94%)
Time	346 (98%)	299 (86%)	261 (97%)	243 (93%)	162 (100%)	156 (96%)
Date	346 (98%)	232 (67%)	261 (97%)	188 (72%)	162 (100%)	113 (70%)
Month	346 (98%)	313 (90%)	262 (98%)	245 (94%)	161 (99%)	149 (93%)
Year	343 (97%)	316 (92%)	262 (98%)	241 (92%)	161 (99%)	152 (94%)
Place	346 (98%)	329 (95%)	262 (98%)	252 (96%)	162 (100%)	160 (99%)
City	345 (97%)	331 (96%)	262 (98%)	255 (97%)	162 (100%)	161 (99%)
Two-person recognition	343 (97%)	316 (92%)	262 (98%)	252 (96%)	161 (99%)	161 (100%)
Date of Birth	345 (97%)	342 (99%)	262 (98%)	259 (99%)	162 (100%)	161 (99%)
Date of WW1	344 (97%)	242 (70%)	261 (97%)	188 (72%)	162 (100%)	114 (70%)
Name of Prime Minister	343 (97%)	284 (83%)	261 (97%)	234 (90%)	160 (99%)	144 (90%)
Count backwards from 20[a]	341 (96%)	309 (91%)	261 (97%)	243 (93%)	161 (99%)	147 (91%)
5-word recall[b]	342 (97%)	196 (57%)	260 (97%)	184 (71%)	161 (99%)	111 (69%)
Clock-draw (face)	331 (94%)	298 (90%)	254 (95%)	237 (93%)	158 (98%)	144 (91%)
Clock-draw (Numbers)	331 (94%)	214 (65%)	254 (95%)	186 (73%)	158 (98%)	120 (76%)
Clock-draw (Hands)	331 (94%)	128 (39%)	254 (95%)	111 (44%)	158 (98%)	73 (46%)
News item	342 (97%)	280 (82%)	261 (97%)	228 (87%)	162 (100%)	137 (85%)
Months of the Year backwards[c]	341 (96%)	281 (82%)	261 (97%)	227 (87%)	162 (100%)	143 (88%)
Letter F fluency[d]	343 (97%)	162 (47%)	260 (97%)	125 (48%)	162 (100%)	70 (43%)

*Includes only those participants who completed the face-to-face follow-up.

[a]Correct = all numbers correct; [b]Correct = ≥4 words correct; [c]Correct = ≥7 months correct; [d]Correct = ≥11 words correct.

Table 6-12 Floor and ceiling effects of each CSI across the three time-points (only those with full completion of each CSI included)

	Baseline (N=354)			One Month (N=268)			Six Months (Face to face N=162; Telephone N=58)		
	Fully completed	Floor	Ceiling	Fully completed	Floor	Ceiling	Fully completed	Floor	Ceiling
CDT	331 (94%)	25 (8%)	107 (32%)	254 (95%)	6 (2%)	97 (38%)	158 (98%)	5 (3%)	58 (37%)
AMT-4	346 (98%)	0 (0%)	300 (87%)	262 (98%)	3 (1%)	235 (90%)	219 (99%)*	0 (0%)	201 (92%)
AMT-10	342 (97%)	0 (0%)	121 (35%)	260 (97%)	0 (0%)	123 (46%)	161 (99%)	0 (0%)	72 (45%)
6-CIT	336 (95%)	2 (1%)	91 (27%)	258 (96%)	89 (34%)	0 (0%)	161 (99%)	41 (25%)	0 (0%)
Abb. MoCA	329 (93%)	4 (1%)	55 (17%)	249 (93%)	0 (0%)	50 (20%)	158 (98%)	1 (1%)	23 (15%)
NINDS-CSN	341 (96%)	2 (1%)	59 (17%)	255 (95%)	0 (0%)	59 (23%)	159 (98%)	0 (0%)	20 (13%)
4AT	348 (98%)	0 (0%)	257 (74%)	268 (100%)	210 (78%)	1 (<1%)	-	-	-
Cog-4	352 (99%)	2 (1%)	269 (76%)	-	-	-	-	-	-

*denominator N=220, denominator of all other CSIs N=162

6.4 Discussion

The absolute numbers of participants who screened positive for cognitive impairment varied widely at all three timepoints based on the CSI used. Consequently, test accuracy metrics of each CSI varied. To detect pre- and post-stroke cognitive syndromes, brief CSIs had high specificity but low sensitivity. The two exceptions to this were the CDT and NINDS-CSN 5-min MoCA which had higher sensitivity than specificity. Generally, brief CSIs completed early after a stroke were more sensitive in detecting a pre-stroke diagnosis of MCI/dementia than detecting a post-stroke cognitive syndrome. Completion rates were good for all eight CSIs, but they all demonstrated ceiling effects (these were particularly high for the AMT-4, Cog-4 and 4AT).

The accuracy of each CSI to detect post-stroke cognitive syndromes were similar when the index tests and reference standard were carried out at the same timepoint vs. when they were a month apart, with very slightly higher AUC when completed at the same timepoint. In sensitivity analyses, accuracy metrics were also similar when untestable patients were included and classified as impaired vs. excluded from analyses. The percentage of participants classified as impaired on each CSI did not decrease much over the three timepoints (largest percentage decrease was 9% for the NINDS-CSN 5-min MoCA between baseline and 6 months). This could be due to the brief CSIs having poor responsiveness to change, since they have a small number of test items. However, this comparison was based on those participants who had completed the follow-up at each timepoint rather than the exact same sample (different numbers of participants completed baseline, one month and six months).

The NINDS-CSN and CDT had a different test accuracy profile to the other CSIs. Both test content and a high threshold score should be factored into the interpretation of this. Of the included CSIs, the NINDS-CSN 5-min MoCA was the only test to include a verbal fluency task. This could be considered one of the more difficult tasks, since only 47% of participants got this item correct (≥11 words) at baseline, which dropped to 43% at 6 months. A high proportion of participants were also classified as impaired on the CDT. The scoring of this task, taken from the MoCA, is stricter than other CDT scoring methods in that it requires the hour hand to be clearly shorter than the minute hand. At baseline

only 39% participants got a point for the clock hands, rising to 46% at the 6-month visit. Many clocks had to be re-scored since many researchers tended to score the hands as correct when they were the same length. While we did not calculate inter-rater reliability, it is likely for this component it would be low. Finally, verbal fluency and the CDT are also the only task components examining executive functioning, which could explain the higher sensitivity.

To identify pre-stroke cognitive impairment, the 6-CIT and NINDS-CSN were the only CSIs to have both sensitivity and AUC ≥0.7. There were 26 participants with a pre-stroke diagnosis of MCI or dementia. Since the breakdown of dementia vs. MCI was not recorded in the CRF, I was unable to examine these groups separately. Using clinical diagnosis from the medical record as the choice of reference standard has limitations as it is not systematic; some individuals may not seek help and have undiagnosed cognitive impairment.

The pattern of low sensitivity and high specificity found in this study mirrors previous research completed in stroke using the AMT-4, AMT-10, 4AT and CDT (267). The Cog-4 has previously been criticised due to having limited ability to detect cognitive impairment when compared to the MoCA (279). The Cog-4 had lower sensitivity in this study when compared to the OCS. Across other settings and purposes, the CSIs used in this chapter have a different test accuracy profile, for example the AMT-4, AMT-10 and 4AT have higher sensitivity to detect delirium (103), the 6-CIT, AMT-10 have higher sensitivity to identify dementia (271) and MCI (280).

The rates of participants who were untestable were lower than Chapter 2. This is not surprising since participants entering research studies are often those with less disability compared to all patients admitted to a stroke unit. The rates of those untestable were similar across baseline and one month. As mentioned in Chapter 2, the 4AT is unique in the fact that it provides scoring for untestable patients, and for this reason has an advantage over other CSIs. The Cog-4 should in theory be similar, since it is scored from the NIHSS and there is guidance to score the NIHSS in patients with aphasia, comatose, intubated etc.

6.4.1 Strengths and Limitations

Strengths of this work include a comparison of multiple CSIs and the use of a stroke specific CSI as reference standard. This study contributes to the limited literature of brief CSIs in a stroke population and is the first to provide psychometric data on the 6-CIT in stroke. This study is also unique to previous studies using brief CSIs as it provides longitudinal data across three timepoints. I followed best practice guidelines in the conduct and reporting of this study.

In this study I used one set of questions to score the different CSIs, with the aim of reducing test burden for the participant. While this has a strength for feasibility in acute settings, it comes with limitations. As detailed in the previous chapter, there were some differences between the published version and the version used in this study for three CSIs: the 6-CIT, AMT-10 and NINDS-CSN 5 min MoCA. The delayed recall component in two CSIs (name and address in 6-CIT; address in AMT-10) was replaced by the 5 words used in the MoCA and it could be argued that the difficulty is not comparable. In addition, the NINDS-CSN 5-min MoCA used in this study was 1-point lower as we did not ask participants the day of the week.

No alternate versions of test items were used in this longitudinal study, so the same words were used for delayed recall at baseline, one and six months. I therefore should acknowledge the possibility of practice effects. There is no easy solution for this in longitudinal studies as while using alternate versions can attenuate practice effects, alternate forms may not be equivalent in difficulty (281).

The research team took a pragmatic approach to assessment; however different assessors likely took different approaches if a patient had aphasia, limb weakness or hearing impairment, since the training materials did not provide specific instructions on how to approach assessment and scoring for those with impairments which would affect the assessment. With this being a multi-site study there were several different assessors. This is not necessarily a limitation since it mirrors clinical practice, but for future studies it is worth providing very specific examples for training purposes. To take the example of assessing a patient with severe hearing impairment, one may take the approach of other

studies and alter administration of items from verbal to visual (282) (e.g., presenting the words for delayed recall rather than reading them out), whereas another researcher may have recorded them as untestable. Both approaches are sensible, and the preferred approach should depend on the study objectives. Previous work has found that the modality of presentation of the MoCA (verbal vs. visual) does not affect overall performance but for the delayed recall item, those with visual presentation of the words scored higher (283).

6.4.2 Implications

Firstly, it is worth emphasising that the context of use for brief CSIs is routine, clinical use, rather than a clinical trial setting. All the CSIs demonstrated ceiling effects, meaning they are less sensitive to change and poorer at discriminating the milder end of the spectrum of cognitive impairment.

The implications of these results for clinical practice depend on whether all patients testing 'positive' for cognitive impairment will be followed up, since the choice of CSI would therefore have an impact on resources. If CSIs are to be used as intended, for initial triaging, then those with high sensitivity are preferred. To detect pre-stroke cognitive impairment the NINDS-CSN 5-min MoCA and 6-CIT had some supportive evidence to be used for this purpose. For post-stroke cognitive impairment, the NINDS-CSN 5-min MoCA and the CDT had supportive evidence to detect multi-domain CI (as detected by the OCS). When single domain cognitive impairment was included in the reference standard, performance was less optimal, so brief CSIs would not be recommended for this purpose. These results are based on the published recommended threshold scores, however in this sample, the thresholds recommended in the CSI development paper were not always the best threshold to optimise the AUC.

In a CSI with high sensitivity, but lower specificity, a negative result can confidently rule out cognitive impairment. However, this comes at a cost of a greater number of false positives. Whereas CSIs with the opposite pattern, low sensitivity but high specificity, a positive result can confidently rule in cognitive impairment. This comes at a cost of a greater number of false negatives.

A way to illustrate this is to use a theoretical example. In a stroke unit admitting 1000 patients, 260 would have persisting multi-domain cognitive impairment at one month (according to the OCS). Using the NINDS-CSN 5-min MoCA in the first week after a stroke (which has higher sensitivity than specificity), 224/260 people would be correctly identified for having persisting impairment. However, 392 people without persisting cognitive impairment would also screen positive and could receive additional unnecessary follow-up investigations. On the other hand, using the AMT-4, which has high specificity but low sensitivity, only 42/260 would be correctly identified with cognitive impairment, but false positives would be reduced to 30.

Another limitation to highlight regarding brief CSIs is that they do not adjust for years of education, however it has been argued that the 6-CIT is not sensitive to varying educational levels (284).

6.4.3 Conclusion

Using eight different CSIs resulted in varying proportions of participants being categorised as test positive for cognitive impairment. Most of these CSIs had low sensitivity, high specificity so would not be recommended for clinical use. The NINDS-CSN 5-min MoCA and the CDT however had some supportive data to detect both pre-stroke cognitive impairment and post-stroke multi-domain cognitive impairment. Considering accuracy in combination with completion rates, the NINDS-CSN 5-min MoCA had higher completion rates than the CDT, therefore out of the brief CSIs examined in this study, this would be the best option for the acute stroke context.

7 The Oxford Cognitive Screen (OCS): Feasibility, floor/ceiling effects and associations with later outcomes

The previous chapters have focused on generic CSIs, however there is rationale to use a CSI which has been developed specifically for stroke, especially if the purpose is to detect cognitive impairments which are more common in this population, such as apraxia and visuospatial neglect. CSIs designed for other disease areas often do not account for physical or speech problems, therefore it is easy to lose points for these reasons and be classified as impaired. Chapter 2 illustrated that limb weakness and speech impairment were the two main reasons for partially incomplete cognitive screens. Therefore, a CSI that takes these factors into account is favourable.

In the previous chapter the brief CSIs demonstrated low sensitivity to identify cognitive impairment when compared to the OCS at one month. In this chapter I focus solely on the OCS and provide data on feasibility, floor/ceiling effects and prognostic ability.

7.1 Introduction

A running theme throughout this thesis is the purpose of post-stroke cognitive screening, bearing in mind that many CSIs were developed for different contexts of use. Results from screening can provide cross-sectional information but also potentially offer insight into a person's future condition. If results can aid understanding of prognosis, clinicians can provide targeted support and better manage expectations for patients and their families.

Thinking about prognosis, one should consider different timepoints after a stroke and the outcomes of interest. Prognosis of cognition itself is of interest since acute cognitive impairment after stroke can improve over time. Being able to differentiate between those who may not improve will help tailor/plan services accordingly. Beyond cognition, one may want to know whether results from early cognitive screening can predict other outcomes which are important to patients,

for example independence in activities of daily living (ADLs), mood and health related quality of life (HRQoL).

It is well recognised that admission NIHSS is associated with later outcomes, with previous studies reporting associations with later morbidity, disability (285, 286) and function (287). The NIHSS however poorly represents cognition and therefore does not address the impact of early cognitive impairment. There has been increasing interest in examining whether early cognitive screening predicts later activities and participation. This topic was addressed in a systematic review which found 14 eligible studies (288). The CSIs varied between studies; a mixture of global dementia screens (MoCA, MMSE) and domain-specific tests (e.g., trail making test, line bisection test) were used within 6 weeks of stroke. A range of outcome measures were used, including the Barthel index (BI), Frenchay Activity Index (FAI), mRS and Functional Independence Measure (FIM) at 6-12 months. The authors concluded that the relationship of acute cognitive impairment and later functional impairment was more consistent when domain-specific tests were used. One reason for this could be that CSIs like MoCA and MMSE, have poorer sensitivity to stroke specific impairments (68). Domain-specific results found impairments in visuospatial abilities, visual memory, visual neglect and executive functioning independently associated with activities, but there were no high-quality studies addressing participation. This review highlighted gaps in the literature to be addressed in future research, such as use of a stroke-specific CSI, and controlling for important confounders, such as education, to reduce risk of bias. Additional studies have also been published on this topic more recently, with varying timepoints for baseline completion and later follow-up, for example one study completed the MoCA at 36-48hrs after stroke and found the results to be associated with functional dependence at 3 months, defined by the mRS dichotomised at <3 (289).

Although global CSIs have also demonstrated ability to predict later outcomes, without carrying out domain-specific tests, one is unable to determine which specific areas of cognition contribute towards later functioning and whether domains are differentially associated, since there are no normative data or threshold scores provided for individual tasks. No study to date has examined early assessment using the OCS or another stroke specific CSI to address later

functional, mood and QoL outcomes. Studies either have studied cognitive impairment as a unitary concept or focused on one specific impairment such as aphasia, neglect with later outcomes.

As described in previous chapters, the OCS is becoming a popular choice for clinical use (53), yet it is yet to be used as an endpoint in an interventional clinical trial. Traditionally, cognitive screening has only given an aggregate total score. The OCS however does not take this approach; it was purposefully designed to avoid an overall pass/fail, with threshold scores provided for each task. Researchers and trial methodologists would therefore benefit from guidance on ways to best utilise the data from this screen and have data to justify analysis plans. I will compare domain-specific results vs. a single global score to assess the utility of this approach.

The primary aim was to examine associations between OCS individual cognitive domains and OCS global score at one-month post-stroke with cognitive, functional, mood and QoL outcomes at 6 months. Secondary aims were to provide details on completion rates and floor/ceiling effects of the OCS.

7.2 Methods

This study used data from the APPLE study: an observational, longitudinal cohort study described in Chapter 5. For this study I used data collected in the study at baseline, one month and six-month time-points.

7.2.1 Cognition

At the one-month follow-up in the APPLE study, the OCS (66) was completed face-to-face. A two-week window either side of the one-month date was permitted, so data were collected up to 6 weeks following study entry. Study entry however was not the same as stroke onset.

The OCS provides domain-specific data across five broad cognitive domains: memory, attention, language, number processing and praxis. Within each domain are subtests which have individual threshold scores to determine impairment (66) (thresholds provided in the previous chapter). Based on this a

visual snapshot of the patient's strengths and weaknesses can be provided (Figure 7-1). A full description of the scale is given in Chapter 1.

I split the attention domain into its two subdomains (spatial attention and executive functioning) based on previous literature addressing these domains separately. I categorised participants as impaired across each of the six cognitive domains, if they scored below the recommended threshold score in any of the 12 subtests detailed in the inner circle in Figure 7-1. A global OCS score was also calculated for all participants who had completed at least 1 subtest, based on the number of domains impaired (range 0-6). While the OCS was not designed to make an overall 'global' categorisation of impairment, the scale's authors have used a global score in previous publications (290) and advise on their website that being impaired in at least one task would be outside of the population norm.

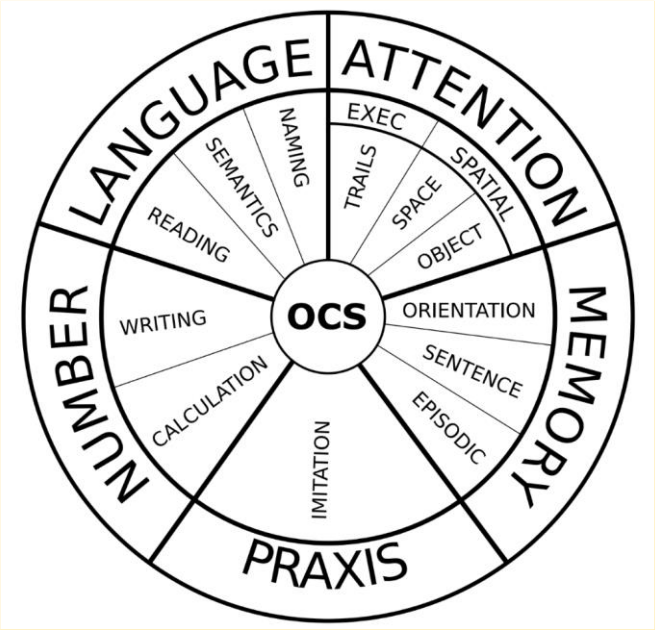


Figure 7-1 OCS domains and tasks

7.2.2 Six-month outcome measures

At six months, outcomes were chosen to reflect different aspects of stroke recovery. These data were collected either face-to-face or over the telephone. A generic HRQoL measure was used: the EQ-5D 3-level (259). The index scores range from -0.594 to 1 (higher scores reflecting better QoL).

The Center for Epidemiologic Studies Depression Scale Revised (CESD-R) (262) was used to measure symptoms of depression (score range 0-80; higher scores indicate more depressive symptoms). The Barthel index (BI) (247) was used to measure independence in basic ADLs (score range 0-20; higher scores indicate greater independence). The Lawton scale was used to measure independence in instrumental ADLs (score range 0-14; higher scores indicate greater independence). The mRS was used to measure global disability (score range 0-6; higher scores indicate greater disability).

The BI, Lawton, and CESD-R scales were usually completed directly by the participant, whereas the mRS and OCS were administered and scored by the researcher. The BI, Lawton and EQ-5D were also completed by a family member/informant, where available, and these data were used in this study where patient self-report data were not available. All outcome measures, apart from the mRS, were not developed for the stroke setting (they are disease-generic scales). Full descriptions of all measures are available in Chapter 5.

For participants where the 6-month follow-up was not completed but had died by the time of the 12-month follow-up, a score of 5 was assigned for the mRS to reflect deteriorating health. No other assumptions were made.

7.2.3 Analysis

As detailed in the previous chapter, at the time of writing this chapter, the database was not locked for the APPLE study, therefore results may differ from subsequent publications.

7.2.3.1 Completion rates, floor/ceiling effects, errors

Consistent with previous chapters, I examined the rates of item-level and scale-level missing data (completion rates) across the full data set and recorded the reasons for incompletion or untestable.

I calculated the percentage of participants scoring the lowest score (0) or full marks (floor and ceiling effects) for each OCS item. The OCS does not have one overall total score range, so scale level floor/ceiling effects were calculated by the number of participants scoring the lowest/highest possible score across all

12 tasks. A criterion of >15% was applied to determine if floor/ceiling effects were present (116).

I checked the case report forms from other hospital sites and checked the scoring of tasks which could be scored retrospectively (e.g., broken hearts, trails). In addition, I recorded other errors in OCS administration/scoring which were identified based on queries from the research nurses.

7.2.3.2 Associations with 6-month outcomes

Dealing with untestable data and time since stroke

For models including each of the six cognitive domains, all missing (incomplete and untestable) data were excluded, since this would prevent interpretation of specific cognitive impairments. For models including a global OCS score, participants who were partially untestable were included (score reflected tasks completed); those who were fully untestable were excluded. Any data collected more than 7 weeks post stroke were excluded from regression analyses, to ensure that the heterogeneity in time from stroke to OCS assessment was minimised. A time window of 6 weeks is often used in the literature (288). In this study the choice of 7 weeks was a compromise between minimising heterogeneity but maximising the participants that could be included in the analysis. Due to a small sample size, I could not restrict the time window further.

I summarised descriptive statistics for the demographic and clinical characteristics of the sample; variables with normally distributed data are summarised by mean and standard deviation, whereas skewed data were summarised by median and interquartile range (IQR).

I compared the participants with 6-month follow-up data to those without on a range of demographic and clinical variables. Group comparisons were made using Pearson's chi^2 (categorical data), independent t test (continuous variables) or Mann-Whitney U test (where data were not normally distributed).

Covariates

Based on previous literature (7, 288, 291, 292) I identified variables that could act as confounders: Age at time of study entry; Sex; Years in education; Stroke severity measured using the NIHSS on admission (score range 0-42); Pre-stroke disability measured using the mRS (score range 0-5; 0 indicating no symptoms and 5 indicating severe disability). Full description of the NIHSS and mRS can be found in Chapter 2.

The NIHSS scores had a positive skew so were placed into five categories, as defined by the following score ranges: 0, 1-4, 5-15, 16-20, 21-42 and treated as ordinal. This is an approach used in other studies (293). Years in education had a positive skew and was transformed with a natural log transform ('LN'). Pre-morbid mRS had a positive skew and was dichotomised using >1 to define disability (294).

Outcomes

I conducted a series of regression models to investigate the associations of impairments in memory, language, praxis, number processing, spatial attention, executive functioning, or a global OCS score, with a range of stroke outcomes (each cognitive domain or global score were examined in a separate model).

For the 6-month outcome variables, the mRS, EQ-5D TTO and CESD-R data were treated as continuous. The CESD-R data were positively skewed and transformed with a natural log transformation. As 0 is a possible score on the CESD-R, I preceded the transformation by adding a value of 1 to all participant's scores.

The BI and Lawton scores were negatively skewed, and the normal probability plots indicated that the residuals were not normally distributed (an assumption of linear regression). Data transformations did not improve the distributions, so the BI and Lawton IADL were dichotomised at <20 and <14, respectively. These threshold scores were chosen to capture any level of disability in ADLs.

Linear regression was therefore used for three outcomes (mRS, EQ-5D, CESD-R) and logistic regression for two outcomes (BI and Lawton ADL). I first ran univariate/unadjusted models which did not control for any potential confounders. I then ran multivariate/adjusted models including the five

covariates mentioned above. For linear regression I checked normality of residuals (normal probability plots are provided in Appendix 16) and multicollinearity of continuous variables (lack of multicollinearity indicated by variance inflation factor <2). To describe the associations with the mRS, EQ-5D, CESD-R I used standardised betas. For associations with the BI and Lawton, I reported odds ratios (ORs).

Sample size calculation

I used G-Power, version 3.1 to inform a sample size calculation for multiple linear regression. This calculation was performed after the APPLE study had started but before this sub-study data analysis was performed. Based on inclusion of 6 independent variables, a moderate effect size of F^2= 0.15, power of 0.95 and a statistical significance level of 0.05, I required a sample size of 146 participants. For logistic regression sample size, I used the rule of 10 outcome events per covariate. To include 6 covariates, I therefore required 60 cases.

I conducted all statistical analyses using IBM SPSS version 26 (NY: IMB Corp). For primary analyses, I used a significance level of 5%. As a sensitivity analysis, I applied a Bonferroni correction, to account for multiple testing and the increased risk of a Type 1 error, accepting differences as significant at p<0.007.

7.3 Results

Administration of the OCS was attempted in 268 participants at a mean of 42.3 (SD 19.5) days post-stroke. Of those that had completed one-month, six-month follow-up data for at least one outcome variable were available for N=202. For those participants without follow-up data, 66 (25%) were either lost to follow-up or declined assessment, and two participants had died (Figure 7-2).

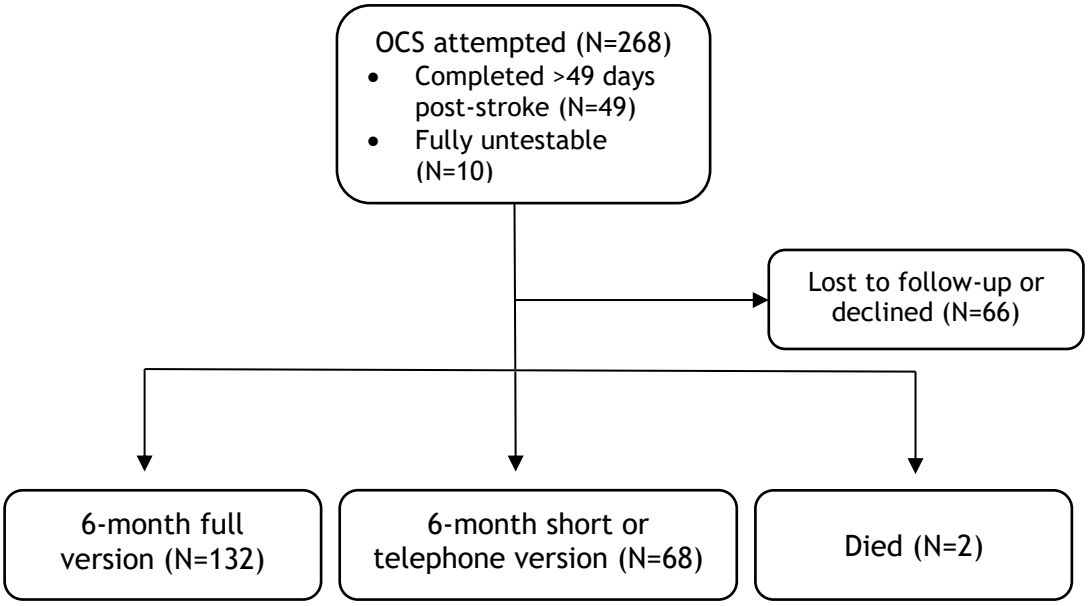


Figure 7-2 Consort flow diagram

7.3.1 Completion rates, floor/ceiling effects, errors

A breakdown of data completeness, floor and ceiling effects for each task are provided in Table 7-1. Data completeness was good for all items (>90%). The heart cancellation test had the lowest completion (93%) and orientation, picture naming and semantics had the highest completion (96%). On an item level, floor effects were low across all tasks (all ≤2%). Ceiling effects were lowest for broken hearts (17%) and highest for the semantic task (98%). On a scale level there was no evidence of a floor or ceiling effect.

Table 7-1 Completeness, floor/ceiling effects of OCS tasks (N=268)

Domain	Task	Completeness N (%)	Floor N (%)	Ceiling N (%)
Memory	Orientation (0-4)	258 (96%)	1 (<1%)	240 (93%)
	Recall & recognition (0-3)	255 (95%)	1 (<1%)	185 (73%)
	Episodic recognition (0-4)	254 (94%)	1 (<1%)	202 (80%)
Language	Picture naming (0-4)	258 (96%)	2 (1%)	173 (67%)
	Semantics (0-4)	258 (96%)	1 (<1%)	253 (98%)
	Sentence reading (0-15)	257 (96%)	2 (1%)	195 (76%)
Number	Number writing (0-3)	256 (96%)	5 (2%)	217 (85%)
	Calculation (0-4)	257 (96%)	1 (<1%)	198 (77%)
Spatial Attention	Broken hearts total (0-50)	248 (93%)	0 (0%)	42 (17%)
Praxis	Gesture imitation (0-12)	257 (96%)	0 (0%)	182 (71%)
Executive functioning	Trails (circles) (0-6)	251 (94%)	5 (2%)	207 (82%)
	Trails (triangles) (0-6)	251 (94%)	2 (1%)	216 (86%)
	Trails (mixed) (0-13)	251 (94%)	5 (2%)	139 (55%)
Full test	-	245 (91%)	0 (0%)	15 (6%)

Overall, there were missing OCS data (partially and fully) for 23 (9%) participants (Table 7-2). Ten participants (4%) were fully untestable; 9 (3%) were untestable on the full test due to stroke-related impairments or reduced consciousness, and one (<1%) declined the full test. Thirteen participants (5%) were partially untestable; 8 (3%) due to stroke-related impairments, 2 (1%) declined certain items and 3 (1%) participants had incomplete data due to positioning of the participant and a page printing error.

Table 7-2 Incomplete OCS tasks with reasons (N=23)

Participant	OCS task	Reason given by researcher for missing data
1	Heart Cancellation	Motor problem
2	Full test	Aphasia
3	Heart Cancellation	Unable to understand tasks
	Trails	
4	Full test	Confused
5	Sentence reading	Aphasia and poor understanding
	Writing numbers & calculation	
	Heart Cancellation	
	Recall & recognition, episodic recognition	
	Trails	
6	Heart cancellation	Language barrier – could not understand
	Recall & recognition, episodic recognition	
	Mixed trails	
7	Trails	Confused
8	Full test	Aphasia
9	Heart cancellation	Visual problem
10	Heart cancellation	Too tired to complete test
	Gesture imitation	
	Recall & recognition, episodic recognition	
	Trails	
11	Full test	Medically unwell
12	Trails	Declined, said it was too difficult
13	Heart cancellation	Lying flat
	Trails	
14	Full test	Drowsy
15	Heart Cancellation	Couldn’t position patient properly

16	Full test	Mix of aphasia, tired and motor problems
17	Full test	Pre-existing cognitive impairment and at limit
18	Trails (circles and triangle)	Missed by researcher
19	Full test	Medically unwell
20	Full test	Medically unwell
21	Heart Cancellation	Page printing error*
22	Full test	Declined
23	Number	Declined
	Heart cancellation	

* heart cancellation page did not print out landscape

In addition to missing data, there were cases of administration/scoring errors where data were available. Items were re-scored for tasks which could be scored retrospectively. Other errors in administration which were identified but not linked to a specific patient are also documented in Table 7-3. Additional training materials clarifying these items were provided to the sites (Appendix 15).

Table 7-3 Errors identified in administration and scoring of OCS

OCS item	**Error or query type**	**Action**
Orientation	Not showing multiple choice options following an incorrect response	None
Gesture imitation	Scoring of finger positions	None
Trails	Scoring errors	Re-scored
	Mixed task - started at circle rather than triangle	None
Broken hearts	Grid lines drawn onto paper	None
	Page printed smaller	None

	Incorrect totals	Re-scored
Delayed recall	Scoring of free recall and recognition combined – scored out of 8 rather than 4	Re-scored

Across the domains assessed by the OCS, impairments in spatial attention were most common (N=65/248; 26%), while impairments in praxis were least common (N=20/257; 8%). The percentages of those classified impaired in each domain (both excluding untestable participants and including them as impaired) are provided in Figure 7-3.

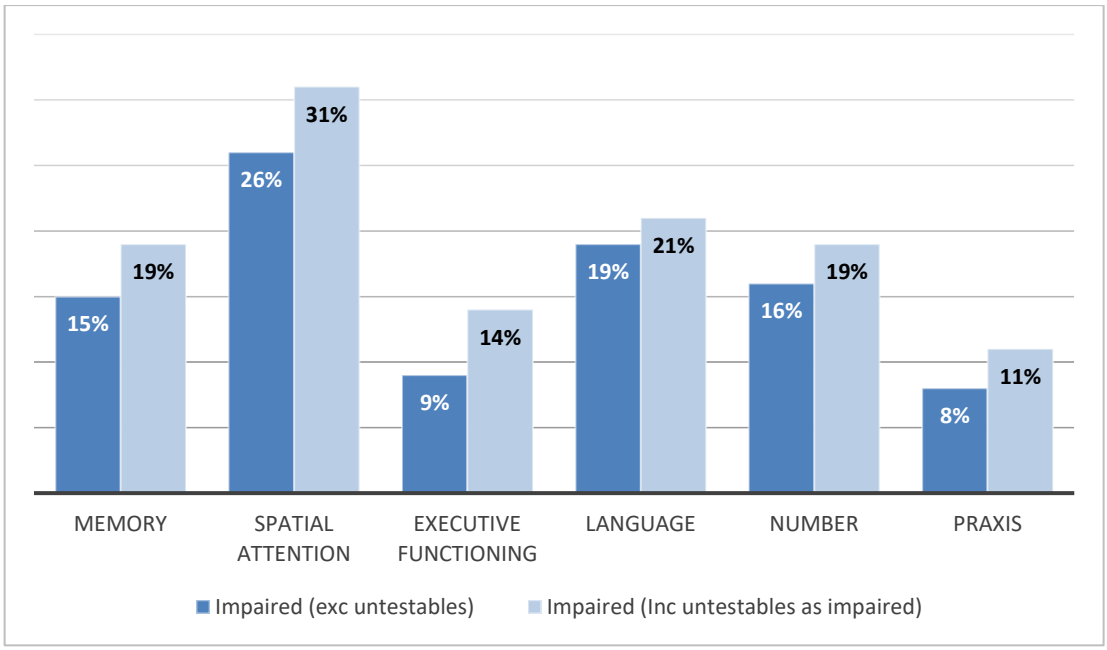


Figure 7-3 Percentage of patients impaired in each cognitive domain

There were 140 (54%) participants who were impaired in at least one cognitive domain. There were 65 participants (26%) with multi-domain cognitive impairment (≥2 or more domains). The median global OCS score was 1 (IQR 0-2) (Figure 7-4).

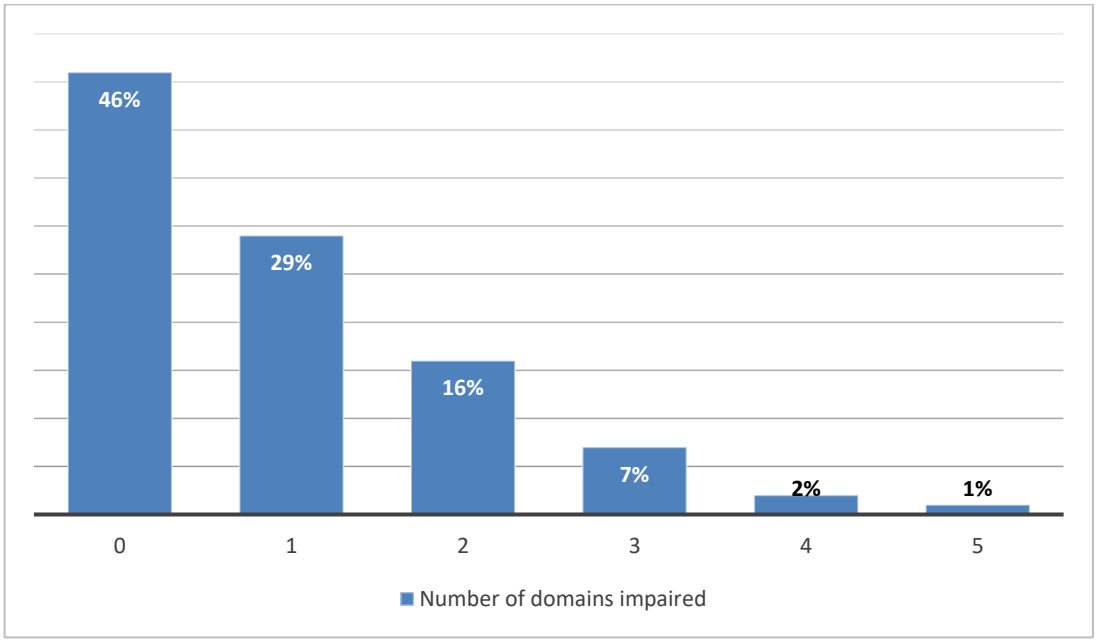


Figure 7-4 Percentage of participants with each global OCS score (excluding untestable participants) (N=258)

7.3.2 Associations with 6-month outcomes

Restricting OCS assessment to within seven weeks of stroke onset excluded 49 participants from analysis. In total, there were 164 participants that received an OCS assessment within this timeframe (mean: 37.8 days, SD 5.6, range 18-49 days), and had data for at least one 6-month outcome, which were included in regression analyses.

The 55 participants without 6-month follow-up data did not differ from those with follow-up data with regards to age, sex, NIHSS, years in education, pre-morbid mRS, previous stroke, impairment in executive functioning, spatial attention, and praxis. However, a greater proportion of those without follow-up data had a history of dementia/cognitive impairment and were impaired across the domains of memory, language, and number processing (Table 7-4). Using a threshold score of <20 on the BI, 41 participants were classified as impaired in basic ADLs. Using a threshold score of <14 on the Lawton, 62 participants were classified as impaired in instrumental ADLs.

Table 7-4 Comparison of participants with/without 6-month f/u

	With 6-month follow-up data (N=164)	Without 6-month follow-up data (N=55)	p value
Age (Mean ± SD)	68.8 (13.2) Missing: 1	69.2 (10.9) Missing: 1	0.73
Male, N (%)	99 (61%)	28 (50%)	0.22
Baseline NIHSS (Median, IQR)	2 (0-4) Missing: 1	2 (1-5) Missing: 1	0.19
Previous stroke, N (%)	38 (23%)	13 (23%)	0.94
Pre-morbid mRS (Median, IQR)	0 (0-2) Missing: 2	0 (0-2)	0.49
Years in Education (Mean, SD)	11 (10-13) Missing: 12	11 (10-12) Missing: 4	0.10
Pre-stroke diagnosis of dementia/MCI (n, %)	7 (4%)	7 (13%)	0.03
Impaired in Memory (yes) N (%)	19 (12%) Missing: 4	12 (24%) Missing: 7	0.03
Impaired in Language (yes) N (%)	23 (14%) Missing: 3	15 (27%) Missing: 6	0.01
Impaired in Executive functioning (yes) N (%)	12 (7%) Missing: 5	5 (9%) Missing: 11	0.42
Impaired in Spatial Attention (yes) N (%)	37 (23%) Missing: 7	11 (20%) Missing: 10	0.90
Impaired in Praxis (yes) N (%)	10 (7%) Missing: 10	6 (11%) Missing: 6	0.17
Impaired in Number processing (yes) N (%)	15 (11%) Missing: 3	15 (27%) Missing: 7	0.00

Abbreviations: IQR, interquartile range; MCI, mild cognitive impairment; mRS, modified Rankin scale; N, number; NIHSS, National Institute of Health Stroke Scale.

HRQoL (EQ-5D TTO)

In unadjusted models, impairments in spatial attention, executive functioning and global OCS score were significantly associated with lower EQ-5D scores (Table 7-5). In adjusted models, executive dysfunction was the only type of cognitive impairment that was independently associated with lower EQ-5D (standardised beta = -0.21 (95% CI: -0.41 to 0.07); this was significant at $p<0.007$ (Table 7-6).

Depressive symptoms (log CESD-R)

In unadjusted models, praxis impairment was significantly associated with lower scores on the CESD-R (Table 7-5). In adjusted models, no cognitive data were independently associated with symptoms of depression (Table 7-6).

Global disability (mRS)

In unadjusted models, impairments in number processing, language and global OCS score were significantly associated with higher 6-month mRS (Table 7-5). In adjusted models, impairment in number processing was the only type of cognitive impairment that was independently associated with a higher 6-month mRS (standardised beta = 0.15 (95% CI: 0.03 to 1.25); this was significant at $p<0.05$ but not at $p<0.007$ (Table 7-6).

Independence in Basic ADLs (BI)

In unadjusted models, impairments in spatial attention, executive functioning, number processing and global OCS were significantly associated with impairment on the BI (Table 7-5). In adjusted models, no cognitive data were independently associated (Table 7-6).

Independence in instrumental ADLs (Lawton)

In unadjusted and adjusted models, no cognitive domains were significantly associated with impairment on the Lawton (Tables 7-5 and 7-6).

Table 7-5 Unadjusted models: Associations between cognitive domains and six-month outcomes

	Six-month outcomes				
	Log CESD-R	**EQ-5D**	**mRS**	**Barthel**	**Lawton**
	Standardised beta (95% CI)			**Odds ratio (95% CI)**	
Memory	0.04 (-0.41, 0.71)	-0.09 (-0.24, 0.06)	0.02 (-0.52, 0.70)	3.20 (0.97, 10.55)	3.01 (0.78, 11.61)
	N=149	N=155	N=160	N=112	N=112
Executive functioning	0.00 (-0.74, 0.75)	-0.25 (-0.47, -0.11)**	0.15 (-0.04, 1.43)	6.01 (1.17, 31.77)*	2.62 (0.51, 13.59)
	N=149	N=154	N=159	N=111	N=111
Spatial attention	0.02 (-0.38, 0.46)	-0.19 (-0.25, -0.02)*	0.12 (-0.11, 0.83)	2.90 (1.19, 7.08)*	1.51 (0.62, 3.68)
	N=147	N=152	N=157	N=110	N=110
Number processing	0.10 (-0.31, 1.26)	-0.08 (-0.29, 0.09)	0.30 (0.67, 1.96)**	16.97 (2.04, 141.33)*	3.06 (0.61, 15.41)
	N=149	N=155	N=161	N=112	N=112
Language	0.09 (-0.23, 0.83)	-0.10 (-0.24, 0.05)	0.24 (0.31, 1.40)**	1.48 (0.53, 4.11)	2.39 (0.79, 7.23)
	N=149	N=155	N=161	N=112	N=112
Praxis	-0.16 (-1.56, -0.01)*	0.09 (-0.09, 0.33)	-0.08 (-1.21, 0.42)	0.23 (0.03, 1.93)	0.10 (0.01, 0.85)*
	N=149	N=155	N=160	N=112	N=112
Global	0.06 (-0.14, 0.28)	-0.22 (-0.13, -0.02)*	0.26 (0.15, 0.57)**	2.30 (1.40, 3.77)**	0.09 (0.94, 2.34)
	N=149	N=155	N=161	N=112	N=112

*p<0.05; **p<0.007. Abbreviations: CESD-R, Center for Epidemiologic Studies Depression Scale Revised; EQ-5D, Euro-QoL 5-dimensions; mRS, modified Rankin scale.

Table 7-6 Adjusted models: Associations between cognitive domains and six-month outcomes

	Six-month outcomes				
	Log CESD-R	**EQ-5D**	**mRS**	**Barthel**	**Lawton**
	Standardised beta (95% CI)			**Odds ratio (95% CI)**	
Model 1 (Memory)	0.03 (-0.52, 0.74)	-0.08 (-0.24, 0.06)	-0.05 (-0.79, 0.36)	3.23 (0.60, 17.32)	4.21 (0.73, 24.25)
	N=134	N=139	N=142	N=104	N=104
Model 2 (Executive functioning)	0.00 (-0.74, 0.71)	-0.21 (-0.41, 0.07)**	0.04 (-0.47, 0.87)	4.47 (0.73, 27.52)	0.96 (0.15, 6.28)
	N=134	N=138	N=142	N=102	N=103
Model 3 (Spatial attention)	-0.08 (-0.65, 0.24)	-0.04 (-0.15, 0.08)	0.00 (-0.43, 0.43)	2.21 (0.71, 6.86)	0.80 (0.25, 2.58)
	N=132	N=136	N=139	N=103	N=102
Model 4 (Number processing)	0.06 (-0.47, 1.06)	0.05 (-0.12, 0.23)	0.15 (0.03, 1.25)*	8.71 (0.92, 82.15)	0.83 (0.12, 5.84)
	N=134	N=139	N=144	N=104	N=104
Model 5 (Language)	0.09 (-0.24, 0.87)	0.02 (-0.12, 0.16)	0.12 (-0.07, 0.92)	0.96 (0.24, 3.86)	1.30 (0.31, 4.33)
	N=134	N=139	N=144	N=104	N=104
Model 6 (Praxis)	-0.11 (-1.26, 0.26)	0.01 (-0.18, 0.19)	-0.01 (-0.73, 0.62)	0.21 (0.02, 2.28)	0.11 (0.01, 1.17)
	N=134	N=139	N=143	N=104	N=104
Model 7 (global)	0.00 (-0.22, 0.23)	-0.09 (-0.09, 0.02)	0.08 (-0.09, 0.31)	1.75 (0.97, 3.17)	0.91 (0.51, 1.61)
	N=134	N=139	N=144	N=104	N=104

All models included age (years), sex, NIHSS , pre-morbid mRS (dichotomised) and years of education as covariates. Results of covariates available in Appendix 16.
*p<0.05. **p<0.007.

There were 12 participants with executive functioning impairment included in regression analyses. The executive function score is calculated using three trails tasks: two simple (circles, triangles) and one complex (mixed). In these 12 participants, the median scores on the first two trails tasks were 6/6 (range 5-6 on circles, 4-6 on triangles), and on the mixed task 3/13 (range 0-7). The median executive score was 8.5 (range 5-12).

There were 15 participants with number processing impairment included in regression analyses. There are two tasks in this domain: number writing and calculations. All participants were impaired on the number writing task. There were different types of errors for this task (Figure 7-5). The main error type (73%) was adding extra zeros, for example '1500200' instead of '15,200'. One participant was impaired on the calculation task.

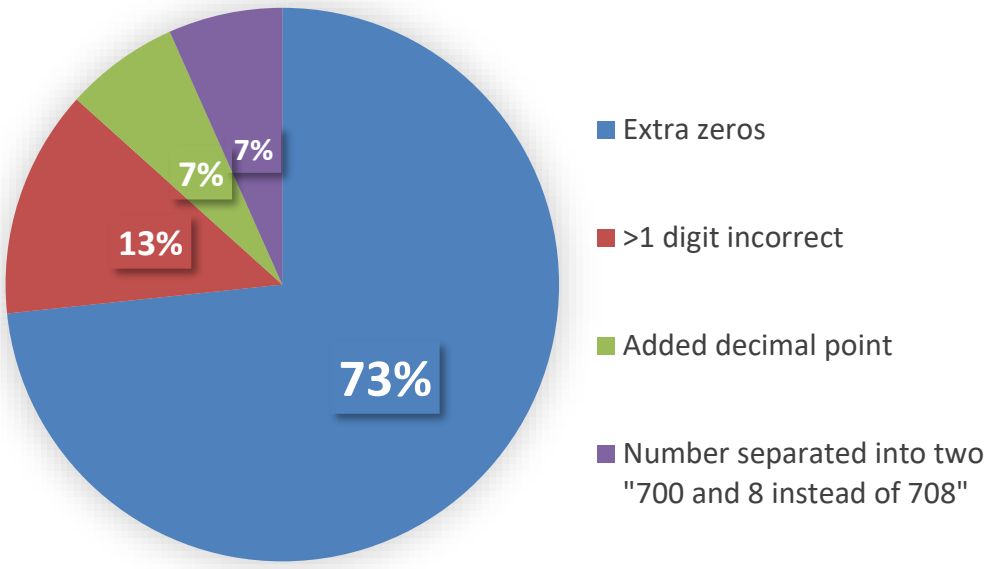


Figure 7-5 Number encoding errors of participants with 6-month f/u (N=15)

7.4 Discussion

The OCS had good completion rates and no scale level floor/ceiling effects. Most OCS domains were associated with at least one outcome at six months in unadjusted models, but only two domains remained significant when controlling for potential confounders; number processing impairment was associated with greater disability on the mRS and executive dysfunction was associated with

reduced QoL on the EQ-5D. However, the magnitude of association was modest and only executive dysfunction remained significant when adjusting for multiple testing. No independent associations were found using the OCS global score.

Early executive dysfunction has previously been found to be independently associated with various later outcomes, including HR-QoL as found in this study (288). In one study executive functioning was evaluated using six different tests as part of a NPB within the first three weeks post-stroke, followed up with a stroke specific QoL scale (14). In another study, an independent association was found between the trail making test (TMT) part B at 72 days post-stroke and QoL at 10 months, as assessed by the sickness impact profile (295). I used the EQ-5D, which covers the following areas: walking, self-care, usual activities, pain/discomfort, anxiety/depression. It therefore covers a wide range of concepts, some of which may be more affected by impairments in executive functioning than others. Previous research has also found associations of executive dysfunction with the Barthel index (7) and depression (14) which were not replicated in this study, although the EQ-5D captures both functional and mood aspects.

While previous research has found associations with other types of cognitive impairment with later outcomes, number processing impairments have not been amongst these findings. This impairment was found to be associated with greater disability as measured by the mRS. To explore this result, I examined the types of errors made by participants. Transcoding zeros were the most common number processing error for this task, as previously reported in another study (296). In this study the researchers hypothesised that zero takes a special role in the process of number transcoding, which requires a higher cognitive load and attentional demand. This could therefore reflect an impairment in complex attention, consistent with previous findings of attentional disorders being associated with later function and participation (13, 288, 297). Since cognitive screening tasks involve a number of processes, there is an argument that domain ‘purity’ does not exist, which can mean that some apparently different results could in fact involve the same processes. This association however could also have been due to a type 1 error since it was not significant when a Bonferroni correction was applied.

Over half of the sample (54%) were impaired in at least one OCS domain which is lower than the incidence found in previous studies using the OCS; one study reported 86% (274) and another 92% (91). Across the different OCS cognitive domains, spatial neglect was the most common type of cognitive impairment in this sample. Previous research has also found the OCS attention domain to be commonly impaired (274, 290). The number of tasks within each domain in the OCS differs, for example there are three subtests for language and only one for praxis, therefore it could be argued that it is easier to be classified as impaired in a domain which has a greater number of tasks. This matches my data, since only 8% were impaired on praxis.

Most OCS domains at one-month post-stroke were not independently associated with the range of outcomes selected in this study. This analysis was of a predictive nature rather than cross-sectional. Results should therefore not be interpreted to mean that cognitive impairment does not affect these areas of one's life. These results may illustrate that domain-specific impairments had improved by the time of the 6-month follow-up.

Considering these results in context with previous findings raises the following areas of discussion in terms of reproducibility of findings: which cognitive screens and outcome measures are used, the timing of assessment post-stroke and differing ways of data management/analysis. Different studies have included both different CSIs and outcome measures and therefore may not be comparable. Results are likely to be specific to the measures used, therefore, although evidence was not found for impairments in certain domains, this does not mean that they are not associated with later functional abilities. In terms of previous studies which found associations of impaired MoCA/MMSE with later outcomes (289, 298, 299), there is evidence that patients with aphasia (274, 300) and limb weakness can perform poorer on these tests, so it may not necessarily be cognitive impairment driving these associations.

Previous studies use the term 'early' cognitive impairment, but how early is early? While I restricted analyses to data collected within seven weeks, this is still a considerably wide range, considering that the prevalence of cognitive impairment will be highest early on after a stroke (7, 20, 301). My sample size however prevented me restricting this time window further. Some previous

studies completed the CSI within the first few days after stroke (302) which could reflect the differences in findings. To compare this with research concerning the NIHSS predicting later outcomes (285, 286, 303), data is usually taken from admission, when the score is at its highest.

Finally, differing results across the literature may be due to differing methods of analysing data. Different studies treat the same outcomes as continuous, ordinal, or binary and results could differ because of this. Dichotomising continuous data is not recommended, however this is usually only applied when dealing with a skewed distribution, such as in this study.

7.4.1 Strengths and weaknesses

Strengths of this work include using a stroke-specific CSI and controlling for important confounding variables recommended by previous work (288). Having informants participate in the study enabled us to obtain outcome data for participants who were unwell.

Limitations of this work include a limited sample size and not having follow-up data for all patients. The data available for the two outcomes (BI and Lawton) were particularly underpowered as reflected in the wide CI's. Lost to follow-up or declining follow-up could be due to various reasons, for example it could reflect greater disability, or it could reflect no disability as participants are back at work/too busy to take part. I had knowledge that some participants without 6-month follow-up data had poor health outcomes, including being admitted to a care home, significant health deterioration or hospitalised. However, with this information alone, I was unable to infer scores on outcome measures. Participants without follow-up data also included a greater number of people who were impaired across three cognitive domains.

Finally, although APPLE was an all-inclusive study, the sample included in this chapter generally consisted of participants with mild strokes and I also included those with a TIA diagnosis, therefore these results are not necessarily representative of patients with severe strokes and functional impairment. When comparing the sample in this study to an unselected cohort, such as SSNAP (2020 results: https://www.strokeaudit.org/results/Clinical-audit/National-

Results.aspx), the median NIHSS was two points lower. Results may differ if future samples include participants with severe strokes, since they are at a heightened risk of cognitive impairment. However, as demonstrated in chapter 2, participants with a higher NIHSS are also more likely to be untestable on CSIs.

7.4.2 Implications

There was insufficient evidence of results from the OCS at one month being associated with a range of patient outcomes at six months, but some evidence that executive dysfunction was independently associated with reduced QoL. Further studies are necessary to understand the prognostic ability of the OCS.

The OCS had strengths when compared to the other CSIs covered in this thesis. The OCS firstly did not demonstrate floor/ceiling effects. Another strength of using a domain-specific CSI rather than a global one is that data from partially incomplete assessments, can be included rather than excluding the participant entirely, since there are cut-off scores for individual questions.

Weaknesses of the OCS are that more resources and training are required. In terms of materials needed to administer, each participant requires a 14-page document, along with a 30-page test booklet which can be re-used. More administration/scoring error types were observed for this screen, in comparison to the brief CSIs and completion rates were no better than the brief CSIs described in the previous chapter.

7.4.3 Conclusion

The OCS performed well in terms of completion rates and floor/ceiling effects. When completed one-month post-stroke it had limited prognostic utility for a range of functional and mood outcomes at six months, when controlling for confounders and adjusting for multiple testing; only executive dysfunction was associated with reduced QoL. Further studies are required to understand the associations of early cognitive impairments with later patient outcomes.

8 Discussion

8.1 General overview

Cognitive screening after a stroke is recommended by clinical guidelines, specialist societies and as part of national audit programs. However, due to vague recommendations, different cognitive syndromes and differing opinions regarding CSI choice and timing, a range of CSIs are being used in clinical practice and research. For clinical use, the decision is left to the individual clinical team and with hundreds of CSIs available, stroke-specific and generic, there can be uncertainty in deciding which CSI is fit for purpose.

In this thesis I undertook a series of studies to provide empirical data to contribute to our understanding of the performance of CSIs at varying timepoints after stroke. I focused on feasibility, accuracy, floor/ceiling effects and prognostic ability of a range of different CSIs.

Understanding psychometric properties of CSIs is essential to provide the foundation for many other areas of research into cognition. Without knowing these properties, results cannot be fully interpreted and their ability to work as clinical outcome assessments (COAs) is limited. These methodological considerations are important since limitations with CSIs can result in misleading or incorrect conclusions. For example, in clinical practice, cases of cognitive impairment may be missed and in research an interventional clinical trial could fail to meet its endpoint based on the CSI chosen, rather than lack of treatment efficacy, or a longitudinal study may find limited change over time due to the CSI lacking responsiveness.

8.2 Summary of key findings

A summary of the thesis objectives and key findings are provided in Table 8-1.

Chapter 2 provided valuable insight into completion rates of CSIs in a real-world setting, with a quarter of patients being untestable for varying reasons. Across the individual items, the CDT had the lowest completion rate, whereas age had the highest. Across the different named CSIs, the Abbreviated MoCA, Mini-Cog

and GP-COG had the lowest completion, whereas there was no missing data for the 4AT, due to scoring options for those who are untestable. In the context of acute stroke, incomplete assessments and missing data are common, although the reasons are not always reported in research studies, since these participants are often excluded. In addition to stroke related impairments, older adults can have other comorbidities or hearing/visual impairments which can impact administration and interpretation of a CSI.

In Chapter 3, through undertaking a systematic review, I identified 13 shortened versions of the MoCA. I reviewed the accuracy of the SF-MoCAs in the published literature to identify varying cognitive syndromes and then validated the SF-MoCA versions in two independent data sets. Across the published literature and in the external validation, the performance of the short forms varied but demonstrated high sensitivity to detect multi-domain cognitive impairment, according to different reference standards.

In Chapter 4, I carried out another test accuracy systematic review; this time focusing on multi-domain telephone-based CSIs to identify MCI or dementia. Across 34 studies, I identified 15 CSIs, but only four of these CSIs were used in participants post-stroke (TICS, TICS-m, T-MoCA, T-MoCA short). Of the limited data available in stroke, the telephone CSIs demonstrated high sensitivity to detect multi-domain cognitive impairment, but more work should be undertaken before they can be recommended for clinical use. In non-stroke settings, the TICS and TICS-m had the greatest supportive evidence base to screen for dementia.

In Chapter 5, I described the methods and conduct of a prospective multi-centre post-stroke cohort study of cognition: APPLE. This study had a more inclusive inclusion criteria than other stroke cohorts, with the intention to include groups often excluded from research (e.g., those with aphasia and pre-existing cognitive impairment). The methods outlined in this chapter relate to Chapters 6 and 7.

In Chapter 6, I evaluated the use of brief CSIs in stroke. I examined completion rates, floor/ceiling effects and accuracy to detect both pre- and post-stroke cognitive impairment. Completion rates were higher than Chapter 2 (all >90%).

Three tests (AMT-4, Cog-4, 4AT) exhibited high ceiling effects. The pattern of accuracy for pre- and post-stroke cognitive syndromes was generally high specificity, low sensitivity, apart from the CDT and NINDS-CSN 5-min MoCA which had the opposite pattern. The two SF-MoCAs used in this chapter demonstrated lower sensitivity than the work completed in Chapter 3, likely due to the choice of reference standard (OCS).

In Chapter 7, I investigated whether domain-specific results from the OCS completed at one month after stroke admission were associated with patient outcomes at six months. I also provided data on OCS completion rates, and floor/ceiling effects. In unadjusted models, all domains apart from memory were significantly associated with at least one outcome. However, when controlling for confounding variables (such as age, education, pre-stroke disability and stroke severity), and adjusting for multiple testing, only one domain remained significant with one outcome: executive dysfunction had a modest association with reduced quality of life (EQ-5D). The OCS had good completion rates, but fewer participants fully completed it in comparison to the brief CSIs and there were more types of administrator errors. There were no issues of floor/ceiling effects.

8.3 Limitations

Limitations of each chapter have been discussed throughout the thesis. In addition to those already detailed, there are some broader limitations to take into account.

Across all chapters with prospective data collection (chapters 2, 6 and 7), most of the participants were recruited from Glasgow. Therefore, it is important to take into account that the samples included in these studies are not necessarily representative of all patients post-stroke, for example with regards to socioeconomic status (304).

Another important limitation and broader discussion point concerns the choice of reference standard in diagnostic test accuracy of CSIs, seeing as there is no consensus on a gold standard for assessing cognition. For example, if dementia is the cognitive syndrome of interest, clinical diagnosis would be considered the

gold standard. However, there are different diagnostic criteria available, which may result in different outcomes (305, 306).

8.4 Implications for clinical practice and research

8.4.1 Choice of CSI

First and foremost, purpose must drive CSI choice and there are various purposes of post-stroke cognitive screening. There is no perfect CSI; each have different strengths and limitations. These trade-offs need to be weighed up within the specific context of use before recommendations are made. To choose appropriately, one should define the specific purpose of screening at a particular point of time, including the cognitive syndrome of interest and what action will be taken based on the results. Differences in opinion regarding CSI choice are generally down to viewing the purpose of screening, especially at an early-stage post-stroke, through different paradigms, e.g., triaging for those at risk of dementia vs. wanting to detect milder cognitive impairment that reveal more about a person at that particular point in time.

Screening should not be undertaken purely to satisfy guidelines/local protocols, without any plans to act on the results. In the absence of evidence-based treatments for post stroke cognitive issues, other actions could include communicating the results back to the patient and MDT, informing other rehabilitation areas (physiotherapy, speech and language therapy, occupational therapy) and making plans for follow-up and relevant support. I recognise that recommendations for CSI choice need to be pragmatic, bearing in mind that resources are quite different across hospitals; CSI recommendation cannot be universal for all hospitals and all patients. Results from all CSIs applied in acute stroke settings should be interpreted in the context of potential pre-stroke cognitive impairment, delirium, depression, and aphasia.

Based on the work completed in this thesis, completion rates vary across different CSIs. If you were choosing a CSI based on this aspect alone, you would choose one which incorporates a scoring option for being untestable, such as the 4AT. Stroke-specific CSIs do not necessarily result in higher completion in comparison to brief generic CSIs; at the one-month follow-up in the APPLE study,

all CSIs had good full completion rates, but the OCS had the lowest (91%) and the 4AT had the highest (100%). However, it should be noted that the OCS was administered after the AMT-plus. A strength of having cut-off scores for each task (like in the OCS) is that in the scenario of a partially incomplete assessment, you can score and interpret the other complete tasks. This is unlike most other CSIs which use sum scores, and normative data is not available for subtasks. As the reasons vary widely for incompletion (stroke related and non-stroke factors), you may choose a CSI to address one aspect of non-completion (e.g., use of non-verbal tasks for those with aphasia) but it may not address other areas (e.g., writing tasks for those with limb weakness). Of note in this thesis, drawing tasks such as the CDT, trails, broken hearts, had lower completion rates and there were cases of examiner errors in scoring, therefore they can be problematic for both patients and those administering the CSI.

Test completion is just one aspect to consider with regards to CSI choice. Another important aspect is validity. If MCI or dementia are present prior to the index stroke, the NINDS-CSN 5-min MoCA, CDT and 6-CIT carried out after the stroke had the highest sensitivity to identify it. Many of the brief CSIs examined in this thesis could not be recommended for clinical practice for the purpose of identifying post-stroke single or multi-domain cognitive impairment (as defined by the OCS) due to low sensitivity. High sensitivity is preferred in acute screening, as tests are often used to identify those patients who require more detailed assessment and a CSI with low sensitivity will miss many patients who have cognitive impairment. The NINDS-CSN 5-min MoCA and the CDT were the only two CSIs to have high sensitivity (above 0.8); these could be used as an initial screen for multi-domain cognitive impairment and followed up with a more detailed assessment at a later timepoint. Many of the brief CSIs exhibited high ceiling effects, which indicates they would have poor responsiveness if serial assessment is planned.

CSI choice would differ for other screening purposes. To identify delirium post-stroke, the 4AT has a different pattern of accuracy (very high sensitivity) (307), which is not surprising, since this is the syndrome it was developed for. Likewise, to detect milder single-domain impairments, CSI choice would differ. Stroke-specific instruments, such as the OCS, are particularly advantageous to identify

cognitive syndromes more prevalent in this population, for example neglect and apraxia. For clinical use, there are a number of benefits of having more detailed domain-specific information as opposed to CSIs which provide one overall score. I would argue the results are more beneficial to the therapy team (as opposed to medical) to ascertain the nature of cognitive impairment, to consider the impact of the specific cognitive profile on various aspects of one's life and to provide targeted rehabilitation. If we do not identify and recognise specific types of cognitive impairment, we cannot manage them. Domain-specific measures are also valuable to examine at trajectories since impairments in different domains can progress differently (301). In terms of prognostic ability beyond cognition, the results from Chapter 7 suggested that when completed at one month, results from the OCS provided limited prognostic information for later functional abilities/disability, when variables such as age, education, stroke severity and previous disability were controlled for.

In light of the recent pandemic, CSIs which can be administered remotely are required. Based on the review I undertook in Chapter 4, telephone-based CSIs, such as T-MoCA and TICS, are an alternative to screen for multi-domain cognitive impairment but would not be appropriate to identify single-domain or milder forms of cognitive impairment. They have limitations as visuospatial abilities cannot be assessed and they are likely an inappropriate format for those with aphasia. Video-based options may address some of these limitations, but more research is needed so we understand the feasibility and validity of this approach.

Finally, there are additional aspects not covered in this thesis that also contribute to the choice of CSI. These include cost/copyright restrictions, availability of alternate versions and cross-cultural translations. Digital CSIs are also alternatives to traditional pencil and paper CSIs. Automated scoring and reporting will likely improve acceptability for clinicians and improve standardisation (308), since we know that there can be subjectivity with scoring and they also allow for results to be available in real time. Having a record of real-time performance also allows for improved quality control and training.

8.4.2 Training

CSIs are carried out by healthcare professionals and researchers from different specialities, with differing training and experience. While instruction materials exist for CSIs, administration and scoring errors occur frequently. Training is essential for use of CSIs and should not be underestimated or overlooked. In the scenario where CSIs are not regularly used, refresher training should be undertaken. Recent studies carried out with OTs working in stroke services, also highlighted the importance of training in this area (53, 309).

In APPLE we had unique insight into how research teams are completing CSIs. Case report forms and the anonymised OCS patient pack were sent directly to the host clinical site for quality checking before being sent to the biostatistics centre for data input. While having an independent biostatistics centre has many strengths, it is important to remember that they will not identify scoring errors, unless they happen to be out of range. In non-commercial, observational studies, data monitoring (including source data verification) rarely happens. In the APPLE study, if the case report forms had been sent directly to the biostatistics centre, a number of errors would have gone unnoticed as without seeing the source data and knowing the assessments, you are blind to these issues. This process had time implications and is likely not practical for large clinical studies. It could be argued that quality may be further compromised when using electronic case report forms, since there are often no spaces for the researcher to record notes or explanations. Electronic forms do, however, have other strengths.

While I documented the types of errors identified in cognitive screening throughout the APPLE study, I did not record the incidence of them. Only some items can be checked and re-scored. Therefore, there are likely many more errors being made in administrating/scoring CSIs than we are aware of, without being present at the time of screening. It is fair to say that of the brief CSIs and the OCS used in the APPLE study, more training and support were needed for the OCS, since there were more types of administration and scoring errors.

CSI errors are not often recorded or mentioned in published studies, although they are likely to be prevalent. One study using routine clinical data from community older adult's mental health teams, reported that 78% of assessments

using the Addenbrooke's cognitive evaluation (ACE-III) had scoring or arithmetical errors (310). Another study completed in primary care also highlighted issues with scoring and reporting of results from CSIs (311). Test scores were ambiguous, incorrect or incomplete in 26% of cases using the 6-CIT and in 32% of cases using the GPCOG. I believe that errors reveal that training has been insufficient, but also a possibility that the CSI may be poorly acceptable to all clinicians. CSIs which have numerous different scoring guidelines, for example the CDT, are especially problematic and likely to be at a higher risk of scoring errors.

CSI authors need to provide greater guidance, including examples of common errors, so there is less subjectivity surrounding cognitive screening. Different formats of training are available for different CSIs. In addition to written instructions, some CSIs provide instructional videos. However, it is rare to finding training materials demonstrating the different incorrect ways in which a participant could respond, with the corresponding scoring.

Research teams can often have high turnover of staff, so it is important to check in with sites and ensure that new staff are trained appropriately. Based on the experience in this the APPLE study, I would recommend checking scoring of CSIs for a sample of participants from each site at an early stage of the study so any issues can be rectified. All participant responses on CSI items should be recorded so scoring can be checked later (for example record the responses for months of the year backwards and all words mentioned in a fluency task).

The authors of the MoCA have moved towards a mandatory paid 1-hour training and certification. Their justification for introducing this is to ensure consistency and accuracy. Dr. Nasreddine states some common mistakes made on the MoCA in an email advertising the training; Informing patients of the words they missed, missing the 2nd learning trial of words for immediate recall, scoring the clock incorrectly and over explaining tasks. While I support mandatory training, it should not incur a cost. Free training materials should be available for all clinicians and researchers using them. I would also support the requirement of a certification exam for usage.

8.4.3 Recommendations for future studies in cognition

Carrying out work for this thesis has highlighted a number of areas of improvement for the conduct of studies using CSIs. The first topic concerns providing sufficient details about the CSI and the threshold score used. While undertaking the two systematic reviews, some studies poorly described these aspects in the methods. This is particularly a concern where shortened versions of tests are used and the test items are not specified, for example 'mini-MoCA' is not an informative title. It also should not be assumed that shortened versions will perform as well as the full scale. Scales should not be altered, shortened, or combined to create a new COA, without appropriate supportive validation work. Researchers should also clearly describe the threshold score used to define impairment, for example stating a threshold score of 10, could be interpreted by different people as <10 or ≤10.

I recommend to always rely on the original development paper for a copy of the scale, and instructions for administration/scoring. While this may sound obvious, CSIs and other COAs can be altered by third parties, and then incorrect versions can be carried forward or important details lost. An example of this was the GPCOG, where one resource online (https://www.alz.org/media/documents/gpcog-screening-test-english.pdf) does not specify details for scoring the hands of the clock, whereas the development paper details that the hand length is not important.

Finally, studies should specify the profession of the person administering the CSI in the study, as well as detailing any training they have received. It is important to know this as acceptability/feasibility from the point of view of the assessor may be different across professions. Many of these aspects are captured in STARDdem but reporting guidance for CSIs could be created for use in any type of study into cognition (outside of diagnostic accuracy studies).

8.4.4 Missing or untestable data

Missing CSI data in the context of stroke is not the same as missing data in PRO questionnaires. The data are rarely missing at random, as the participants who are untestable are often those who are at a high risk of dementia (123). There

are a range of reasons for incomplete CSIs and they should not always be grouped together. The reasons for missing data should be recorded so the data can be handled most appropriately and interpreted in a meaningful way. Statistical analysis plans should set a priori rules based on the specific objectives of the study. It would however be beneficial to have some evidence-based consensus and guidelines on how to deal with missing CSI data, with specific examples for researchers to refer to (e.g., if a participant declines a particular item).

8.4.5 Future topics

Completion rates were reported throughout this thesis, but other aspects of feasibility/acceptability should be explored in future work, including use of qualitative methods, to understand participant's and clinician's experience of cognitive screening post-stroke. The published literature in this area appears to be limited, even outside the context of stroke. Work in this area would move us towards a more person-centred approach to cognitive screening. I identified one non-stroke study investigating older adults' attitudes towards cognitive screening through using a questionnaire (312). Some participants indicated preference to be assessed at home rather than a doctor's office (35% vs. 10%), some indicated preference for a particular modality: computer/mobile device was preferred to paper and pencil (29% vs. 4%) and 63% indicated preference to complete the CSI without any company. Qualitative methodology could also be used to evaluate content validity of CSIs; Are CSIs measuring all aspects of cognition which are important to patients? Future work could also focus on understanding what constitutes meaningful change in each CSI.

Since starting this PhD, a number of new stroke specific CSIs have been published, which are listed in Chapter 1. More work is needed to understand their performance beyond the original development paper and to examine how they perform against the OCS, which has now become the most popular stroke-specific CSI. In this thesis I did not have a gold standard of which I could compare the OCS to. It would be beneficial to have accuracy data of the OCS compared to a recommended NPB, such as that recommended by NINDS-CSN. Future work from the APPLE study can compare one month OCS to NINDS-CSN 30-

and 60-minute protocols at 6/12/18 months for a measure of prognostic accuracy.

Future work should produce data to provide recommendations concerning screening individuals with hearing or visual impairment and interpreting results in those with low education levels. I encourage future research to report on issues and errors arising using CSIs and to share experiences with the wider community, so issues can be recognised and addressed. Without action, there is a concern that the same mistakes are being repeated across studies. The number of errors made could also be used as a proxy measure of rater acceptability.

Finally, this thesis included classical test theory (CTT) approaches to psychometrics. Future work would benefit from modern psychometric approaches (item response theory (IRT), Rasch measurement theory (RMT)) as these approaches provide additional valuable information, including information at an item-level. CTT assumptions are easy to meet, so many instruments can appear to perform well under this framework.

8.5 Conclusion

I have investigated feasibility, accuracy and prognostic ability of different CSIs used at varying time points post-stroke. Data from this thesis contribute to the growing literature concerning cognitive screening post-stroke and can be used to justify using (or not using) a particular CSI. Recommendations for CSI choice differ depending on the purpose of screening, including plans for following up those with identified cognitive impairment.

I am keen to help improve the quality of both cognitive screening and research conduct in this area. Based on the work completed in this thesis, I have made a set of recommendations for those carrying out cognitive screening and for those designing and leading on studies into cognition. These include defining the purpose of screening, the training resources required, reporting guidance for studies using CSIs, and planning for participants being untestable.

Table 8-1 Summary of main findings

Thesis aim	Chapter	Main findings
To evaluate the feasibility (completion rates and reasons for missing data) of brief CSIs in acute stroke and to examine associations with being untestable.	2	A quarter of participants were untestable on at least one test item. Across all individual items, the CDT had the lowest completion rate and age had the highest. Across the different named CSIs, the abbreviated MoCA, Mini-cog and GP-COG had the lowest completion rates, and the 4AT had the highest due to scoring for untestable being incorporated.
To identify the number of shortened versions of the MoCA in the literature, to review their accuracy and independently validate them in stroke and non-stroke contexts.	3	Thirteen SF-MoCAs were identified across the literature. Sensitivity was high across the context of stroke and non-stroke but based on varying reference standards. In the stroke validation, sensitivity was high against the MMSE.
To identify telephone-based CSIs and to review their accuracy in stroke and non-stroke contexts.	4	I identified 15 telephone-based CSIs, four of which were used in participants post-stroke (TICS, TICS-m, T-MoCA, T-MoCA short). Of the limited data available in stroke, the telephone CSIs demonstrated high sensitivity to detect multi-domain cognitive impairment. Outside of stroke the TICS and TICS-m had supportive evidence to screen for dementia.
To examine the feasibility (completion rate and reasons for missing data) and floor/ceiling effects of brief CSIs across three timepoints.	6	Completion rates were all above 90% at all three timepoints but there was fewer missing data at 6 months compared to baseline/one month. Ceiling effects were highest for the AMT-4, Cog-4 and 4AT.
To examine the test accuracy of brief CSIs to identify pre-stroke MCI/dementia and post stroke cognitive impairment.	6	The pattern of accuracy for pre- and post-stroke cognitive syndromes was generally high specificity, low sensitivity, apart from the CDT and NINDS-CSN 5-min MoCA which had the opposite pattern.
To evaluate the feasibility (completion rate and reasons for missing data) and floor/ceiling effects of the OCS at one month.	7	Full completion was 91%. Reasons for missing data included both participant factors and administrator errors. No floor/ceiling effects were observed.
To examine whether the results from the OCS at one month are associated with patient outcomes at 6 months.	7	When controlling for variables known to impact functional and mood outcomes (such as age, education and stroke severity), and adjusting for multiple testing, impairment in only one domain (executive functioning) had a modest association with lower scores on the EQ-5D (reduced QoL).

Appendices

Appendix 1: Chapter 2 Ethics approval

WoSRES
West of Scotland Research Ethics Service

NHS Greater Glasgow and Clyde

West of Scotland REC 5
West of Scotland Research Ethics Service
West Glasgow Ambulatory Care Hospital
Dalnair Street
Glasgow
G3 8SW

Dr Terence J Quinn
NHS Greater Glasgow and Clyde
c/o ward 17/31
Glasgow Royal Infirmary
G4 0SF

Date 04 February 2016

Direct line 0141 232 1809
E-mail WoSREC5@ggc.scot.nhs.uk

Dear Dr Quinn

Title of the Database: **Stroke Assessment Area Database**
REC reference: **16/WS/0001**
IRAS project ID: **188817**

Thank you for your letter of 1 February 2016. I can confirm the REC has received the documents listed below and that these comply with the approval conditions detailed in our letter dated 26 January 2016.

Documents received

The documents received were as follows:

Document	*Version*	*Date*
Covering letter on headed paper		01 February 2016
Protocol for management of the database	1.2	01 February 2016

Approved documents

The final list of approved documentation for the study is therefore as follows:

Document	*Version*	*Date*
Covering letter on headed paper [cover letter]		31 December 2015
Covering letter on headed paper		01 February 2016
Other [Stroke admission proforma]		
Other [SOPs]	V1.0	25 December 2015
Other [CV]		31 December 2015
Protocol for management of the database	1.2	01 February 2016

Appendix 2: Database proforma

Patient ID:_____
Admission date:_____
Date and time of symptom onset:
_____
Date of stroke: _____ Age: ____

Cognitive Assessment	
1 Age (in years)	/1
2 Time (to nearest hour)	/1
3 Date (day, month, year)	/1
4 Place	/1
5 Recognition (2 people)	/1
6 Date of Birth	/1
7 Start of WW1 (1914)	/1
8 Prime Minister	/1
9 Months Backwards (Dec-Jan; see AMT-4 for scoring)	/1
10 Remember face, velvet, church, daisy, red	
11 Clock draw time of 11:10 (face, numbers, hands)	/3
12 Recent news item	/1
14 Recall face, velvet, church, daisy, red	/1
15 Words Beginning with "F" >11 in 1 min	/1
Total Score	

Yale Single Q:
Do you often feel sad or depressed?
Yes ☐ No ☐

Have you been more forgetful in the past year, to the extent that it interferes with daily life?
Yes ☐ No ☐

GP-Cog Informant Interview
Initials and date:
Informant:

Q1 remembering recent things	
Q2 recalling conversations	
Q3 struggle to find right words	
Q4 able to manage finances	
Q5 able to manage medication	
Q6 help with public or private transport	
Total (score > suggests pre-stroke cognitive decline)	/6

4AT **TOTAL SCORE** _____

Acute change mental status: Y ☐ N ☐

Fluctuating mental status: Y ☐ N ☐

Alertness
- Normal (0) ☐
- Mild sleepiness (0) ☐
- Clearly abnormal (4) ☐

Attention (MOYB Dec-Jan)
- >7 correct (0) ☐
- < 7 correct (0) ☐
- Untestable (2) ☐

AMT-4
- no mistakes (0) ☐
- 1 mistake (1) ☐
- >1 mistake/untestable (2) ☐

Pre-stroke functional status
Pre-stroke mRS (0-5) _____
Care-home resident Yes ¯ No ¯
No. of carers _____
Glasses _ Hearing Aids _
Worn on ward Worn during assess

ASU Cognitive Assessment Questionnaire 2015

Patient ID_____
Admission date _____
Date and time of symptom onset
_____
Date of stroke: _____ Age: ___

Do they have or have had any known mood disorders: Y ☐ N ☐

Is this ongoing: Y ☐ N ☐

Specify:_____

Relevant medication:

Are they taking any drugs form the Medication list? If so specify:

_____
_____
_____
_____

STROKE RISK FACTORS
Hypertension
Atrial fibrillation
Previous STROKE/TIA
Diabetes Mellitus
Vascular disease (IHD/PVD)
Hyperlipidaemia
Heart failure
Oral contraceptive
Family Hx of premature vascular disease
Other

Stroke Checklist

	Yes	No	Uncertain
Stroke	☐	☐	☐
Antiplt	☐	☐	☐
Swallow	☐	☐	☐
IPC	☐	☐	☐
Pre-stroke Cog issues	☐	☐	☐
Delirium	☐	☐	☐
AWI	☐	☐	☐

NIHSS

1 LOC a)	
1 LOC b)	
1 LOC c)	
2 gaze	
3 visual	
4 facial	
5 motor arm L	
5 motor arm R	
6 motor leg L	
6 motor leg R	
7 ataxia	
8 sensory	
9 language	
10 dysarthria	
11 inattention	
Total	

If unable to score why:

Temp:
Blood Glucose:
AVPU rating:
BP:
HR:

TACS☐ PACS☐ LACS ☐ POCS☐ TIA ☐
Other ☐
Sex: M ☐ F ☐
Dominant Hand: L ☐ R ☐ Stroke: R L
Care home resident: Y ☐ N ☐
Previous stroke: Y ☐ N ☐
Alcohol/drug dependence: Y ☐ N ☐
Known dementia diagnosis: Y ☐ N ☐
Known depression diagnosis: Y ☐ N ☐
Sensory issue (deaf /blind): Y ☐ N ☐
GCS E: V: M: Score: _____
NEWS _____ Meds (n): _____

SHx
Smoking: Current/ex/never (no. of pack yrs_____)
Alcohol: Units per week ____
Illicit drug use: yes (specify)/no
_____

Pre-stroke functional status
Care-home resident Yes No
Mobility I Independent mobility aid
_ direct assistance
Continent bladder Yes No
Continent bowel Yes I No I

Modified Rankin Scale Criteria
0 No symptoms at all
1 No significant disability despite symptoms; able to carry out all usual duties and activities
2 Slight disability; unable to carry out all previous activities, but able to look after own affairs without assistance
3 Moderate disability; requiring some help, but able to walk without assistance
4 Moderately severe disability; unable to walk without assistance and unable to attend to own bodily needs without assistance
5 Severe disability; bedridden, incontinent and requiring constant nursing care and attention
6 Dead

TACS: Big stroke, Motor and 2 cortical issues
PACS: Motor and 1 cortical issue
LACS: Face, arm and leg (only motor issues)
POCS: Sensory issues (hemianopia)
TIA: Smaller stroke <24h

Appendix 3: PRISMA checklist (SF-MoCA)

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
TITLE / ABSTRACT			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	81
Abstract	2	Abstract: See PRISMA-DTA for abstracts.	Not in thesis but in publication
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	81-82
Clinical role of index test	D1	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	81
Objectives	4	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	82
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	82
Eligibility criteria	6	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	83
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	83
Search	8	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	Appendix 4
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	83

Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	83-84
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	84
Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	84
Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	84
Synthesis of results	14	Describe methods of handling data, combining results of studies and describing variability between studies. This could include, but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity, c) handling multiple index test readers, d) handling of indeterminate test results, e) grouping and comparing tests, f) handling of different reference standards	84

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
Meta-analysis	D2	Report the statistical methods used for meta-analyses, if performed.	N/A
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	87
Study characteristics	18	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	94-96
Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	97
Results of individual studies	20	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	94-96

Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	92
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	106
Limitations	25	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	107-108
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	108-110
FUNDING			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	Not in thesis but in publication

Appendix 4: Chapter 3 SF-MoCA search strategy

1. Montreal Cognitive Assessment.ti,ab,kf
2. Montreal Cognitive*.ti,ab,kf
3. MoCA.ti,ab,kf
4. Mini adj3 Montreal Cognitive.ti,ab,kf
5. Mini adj3 MoCA.ti,ab,kf
6. Mini-MoCA.ti,ab,kf
7. miniMoCA.ti,ab,kf
8. Short adj3 Montreal Cognitive.ti,ab,kf
9. Short adj3 MoCA.ti,ab,kf
10. Montreal Cognitive adj3 5-minute protocol.ti,ab,kf
11. MoCA adj3 5-minute protocol.ti,ab,kf
12. 1 OR 2 OR 3 OR 4 or 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11
13. Short* form*.ti,ab,kf
14. Abbreviate*.ti,ab,kf
15. item reduction*.ti,ab,kf
16. minimum dataset. ti,ab,kf
17. Rasch*ti,ab,kf
18. Principal component analys*.ti,ab,kf
19. 13 OR 14 OR 15 OR 16 OR 17 OR 18
20. 12 AND 19

Appendix 5: STARDdem checklist

Table 3 STARDdem checklist for the reporting of diagnostic accuracy studies in dementia

Section, topic, and item no.	STARD checklist item	Points of particular relevance to dementia
Title/abstract/ keywords		
1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading "sensitivity and specificity")	Studies reporting a sensitivity/specificity or 2 × 2 data derivable fall within the scope of STARDdem and should be indexed accordingly
Introduction		
2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups	Some studies describing aims related to "prognosis" or "prediction" may also fall within the remit of STARDdem. Report test purpose: "stand-alone" test or as an addition to other tests or clinical criteria
Methods		
Participants		
3	The study population: the inclusion and exclusion criteria, setting and locations where data were collected. See also item 4 on recruitment and item 5 on sampling	Key inclusion criteria: (1) demographic, especially age; (2) cognition- or disease-related criteria. Accurate description of the target sample is required including reporting criteria used to define the study population. Report referral pathways, precise locations of patient recruitment, where index test and reference standard were performed. For secondary/ tertiary settings, helpful to report the medical subspecialty or hospital department (e.g., psychiatry, neurology). Diagnostic accuracy studies in dementia are often nested within larger cohort studies. If this is the case, then the targeted population for the cohort study and the method of selection into the cohort should be described and/or the parent study cited
4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard? See also item 5 on sampling and item 16 on participant loss at each stage of the study	Report whether those in intermediate categories (e.g., possible AD or possible DLB) were excluded
5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in items 3 and 4? If not, specify how participants were further selected. See also item 4 on recruitment and item 16 on participant loss	Planned analyses showing how characteristics of the subgroup entering the study differ from the eligible population are strongly recommended (i.e., if a convenience sample has been used because of the invasive nature of the test or tests)
6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	Authors should report the timing of the analysis plan regarding data collection: Was the analysis plan set out in a protocol before index and reference standards were performed? If not, when was the analysis plan created?
Test methods		
7	The reference standard and its rationale	For neuropathologic and clinical reference standards, the diagnostic criteria used should be specified. Where relevant, reference should be made to studies validating the criteria. Report whether standard consensus clinical criteria incorporate the index test (incorporation bias rendering blinding of index test impossible)
8	Technical specifications of materials and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard. See also item 10 concerning the person(s) executing the tests	Use of scales: specify details of administration, which version. Clinical diagnostic criteria: what information was available to inform the diagnoses; how the criteria were applied (e.g., by individual clinicians, by consensus conference, by semiautomated algorithm). Imaging and laboratory tests: specify materials and instruments, including sample handling and concordance with any harmonization criteria. In new assays, describe all steps in detail. Any particular preparation of participants should be described
9	Definition of and rationale for the units, cutoffs, and/or categories of the results of the index tests and the reference standard	Justify any cutoffs used, because these may vary with clinical context

Section, topic, and item no.	STARD checklist item	Points of particular relevance to dementia
10	The number, training, and expertise of the persons executing and reading the index tests and the reference standard. See also item 8	Especially where subjective judgments are involved, e.g., the interpretation of neuroimaging results. Report inter- and intrarater agreement. Reference or describe the content of training materials used. Reference or describe details of lab certification and harmonized biomarker assays
11	Whether or not the readers of the index tests and reference standard were blinded (masked) to the results of the other test and describe any other clinical information available to the readers. See also item 7	Also, the index test may form a part of the reference standard. This is often referred to as incorporation bias and renders blinding of the index test impossible
Statistical methods		
12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g., 95% confidence intervals)	
13	Methods for calculating test reproducibility, if done	Applies to the reference standard as well as to the index test. Both should be reported/ adequately referenced. Report interrater and test-retest reliability of reference standard as established in the study being reported, rather than simply referring to other studies in which reproducibility has been established. The training that image readers receive should be carefully described. Studies in which the accuracy of "majority" judgments are reported should also report data for the minority judgments. Reports of the impact of training should clearly describe the characteristics of the sample used for training and whether it is representative of the group to which the test will be applied
Results		
Participants		
14	When study was performed, including beginning and end dates of recruitment	Pertinent particularly to longitudinal (delayed verification) studies, authors should report recruitment dates of the study (not to be confused with recruitment dates of the wider cohort study from which it might be drawn), and the beginning (first participant) and end (last participant) dates of the periods during which the index test(s) and reference standard were performed. Report the period for the index test and period for the reference standard separately if it is not clear
15	Clinical and demographic characteristics of the study population (at least information on age, sex, spectrum of presenting symptoms). See also item 18	Report key demographic variables: age, sex, and education. Report age distribution of sample in detail. Ethnicity and genetic factors (e.g., *APOE* genotype) may also be particularly important. The cognitive characteristics are covered in item 18
16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended). See also items 3–5	
Test results		
17	Time interval between the index tests and the reference standard, and any treatment administered in between	Specify the follow-up period for all subjects in relation to their outcomes. It should be specified whether participants had received any treatments that might affect disease progression
18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition	Include a description of the severity of the target condition at the time the index test is performed. Usually captured by a cognitive score and/or duration of symptoms. For delayed verification studies, report distribution of severity of disease and the degree of certainty (such as probable/possible) about the diagnosis at the time of case ascertainment. Report other diagnoses (not target condition). Report relationship of test to other diagnoses

Table 3 Continued

Section, topic, and item no.	STARD checklist item	Points of particular relevance to dementia
19	A cross-tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard	
20	Any adverse events from performing the index tests or the reference standard	Report all adverse events, even if unlikely to be related to the diagnostic test performed
Estimates		
21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g., 95% confidence intervals). See also item 12	
22	How indeterminate results, missing data, and outliers of the index tests were handled	
23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers, or centers, if done	
24	Estimates of test reproducibility, if done. See also item 13	
Discussion		
25	Discuss the clinical applicability of the study findings	Discuss differences in age and comorbidity between the study population and the patients typically seen in clinical practice. Discuss whether the reported data demonstrate "added" or "incremental" value of the index test over and above other routine diagnostic tests. Identify stage of development of the test (e.g., proof of concept; defining accuracy in a typical spectrum of patients). Discuss the further research needed to be done to make test applicable to population in whom likely to be applied in practice

Abbreviations: AD = Alzheimer disease; DLB = dementia with Lewy bodies; MeSH = Medical Subject Headings; STARD = Standards for Reporting of Diagnostic Accuracy.

Appendix 6: QUADAS-2

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection:

- Was a consecutive or random sample of patients enrolled? Yes/No/Unclear
- Was a case-control design avoided? Yes/No/Unclear
- Did the study avoid inappropriate exclusions? Yes/No/Unclear

Could the selection of patients have introduced bias? **RISK: LOW/HIGH/UNCLEAR**

B. Concerns regarding applicability

Describe included patients (prior testing, presentation, intended use of index test and setting):

Is there concern that the included patients do not match the review question? **CONCERN: LOW/HIGH/UNCLEAR**

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted:

- Were the index test results interpreted without knowledge of the results of the reference standard? Yes/No/Unclear
- If a threshold was used, was it pre-specified? Yes/No/Unclear

Could the conduct or interpretation of the index test have introduced bias? **RISK: LOW /HIGH/UNCLEAR**

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question? **CONCERN: LOW /HIGH/UNCLEAR**

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted:

- Is the reference standard likely to correctly classify the target condition? Yes/No/Unclear
- Were the reference standard results interpreted without knowledge of the results of the index test? Yes/No/Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **RISK: LOW /HIGH/UNCLEAR**

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question? **CONCERN: LOW /HIGH/UNCLEAR**

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram):

Describe the time interval and any interventions between index test(s) and reference standard:

- Was there an appropriate interval between index test(s) and reference standard? Yes/No/Unclear
- Did all patients receive a reference standard? Yes/No/Unclear
- Did patients receive the same reference standard? Yes/No/Unclear
- Were all patients included in the analysis? Yes/No/Unclear

Could the patient flow have introduced bias? **RISK: LOW /HIGH/UNCLEAR**

Appendix 7: Chapter 4 Telephone CSIs search strategy

Search syntax used across electronic databases:
Key search terms: telephone interview, telephone screening, cognitive screening and cognitive status.

Concept 1 = “Dementia”

1. exp Dementia/
2. Delirium, Dementia, Amnestic, Cognitive Disorders/
3. dement*.ti,ab.
4. alzheimer*.ti,ab.
5. AD.ti,ab.
6. ("lewy bod*" or DLB or LBD).ti,ab.
7. "cognit* impair*".ti,ab.
8. (cognit* adj4 (disorder* or declin* or fail* or function*)).ti,ab.
9. (memory adj3 (complain* or declin* or function*)).ti,ab.

Concept 2 = “Cognition”

1. (cognition or cognitive).ti,ab.
2. Cognition/
3. Cognition Disorders/
4. Memory/

Concept 3 = “Telephone Assessment”

1. telephon*.ti,ab.
2. (tele* adj5 (screen* or interview* or study* or question* or assess*)). ti,ab.
3. phone*.ti,ab.
4. “telephone administered”.ti,ab.

5. “telephone-administered”. ti,ab.
6. “testing by telephone” . ti,ab.
7. “telephone test”. ti,ab.
8. “telephone-test”. ti,ab.
9. Concept 4= Specific Telephone Administered Cognitive Screening Tests
10. “Telephone Interview for cognitive status”. ti,ab.
11. TICS-m.ti,ab.
12. “Telephone Interview for cognitive status- modified”.ti,ab.
13. “The Brief Test of Adult Cognition by Telephone”.ti,ab.
14. BTACT.ti,ab.
15. “Telephone Dementia Questionnaire”.ti,ab.
16. TDQ.ti,ab.
17. “Brief Screen for Cognition Impairment”.ti,ab.
18. “Memory and Ageing Telephone Screen”.ti,ab.
19. “Telephone Cognitive Assessment Battery”.ti,ab.
20. “Memory Impairment Screen- Telephone”.ti,ab.
21. MIS-T.ti,ab.
22. “Short Portable Mental Status Questionnaire”.ti,ab.
23. SPMSQ.ti,ab.
24. “Telephone Modified Mini- Mental state exam”.ti,ab.
25. T3MS.ti,ab.
26. “Telephone administered Minnesota Cognitive Acuity Screen”.ti,ab.
27. MCAS.ti,ab.
28. “Blessed Telephone Information Memory Concentration Test”.ti,ab.

29. BTIMC.ti,ab.

30. "Structured telephone interview for dementia assessment".ti,ab

31. STIDA.ti,ab.

Dementia OR Cognition AND Telephone Assessment OR Specific Telephone

CINAHL

1. TI ((tele* N5 (screen* or interview* or study or question* or assess))) OR AB ((tele* N5 (screen* or interview* or study or question* or assess)))

2. (MH "Telephone Consultation (Iowa NIC)") OR (MH "Telephone+") OR (MH "Telecommunications") OR (MH "Telehealth+")

3. (MH "Remote Consultation")

4. (MH "Delirium, Dementia, Amnestic, Cognitive Disorders") OR (MH "Amnesia+") OR (MH "Cognition Disorders+") OR (MH "Consciousness Disorders") OR (MH "Dementia+")

5. TI dement* OR AB dement*

6. TI cognit* OR AB cognit*

7. Tl memory OR AB memory

PsycINFO

1. TI ((tele* N5 (screen* or interview* or study or question* or assess))) OR AB ((tele* N5 (screen* or interview* or study or question* or assess)))
2. DE "Telemedicine"
3. 2 or 3
4. DE "Dementia" OR DE "AIDS Dementia Complex" OR DE "Dementia with Lewy Bodies" OR DE "Presenile Dementia" OR DE "Semantic Dementia" OR DE "Senile Dementia" OR DE "Vascular Dementia" OR DE "Alzheimer's Disease" OR DE "Cognitive Impairment" OR DE "Neurodegenerative Diseases" OR DE "Alzheimer's Disease" OR DE "Amyotrophic Lateral Sclerosis" OR DE "Corticobasal Degeneration" OR DE "Dementia with Lewy Bodies" OR DE "Multiple System Atrophy" OR DE "Parkinson's Disease" OR DE "Semantic Dementia"
5. DE "Cognitive Impairment" OR DE "Cognition" OR DE "Animal Cognition" OR DE "Mental Lexicon" OR DE "Cognitive Ability" OR DE "Brain Training" OR DE "Mathematical Ability" OR DE "Reading Ability" OR DE "Spatial Ability" OR DE "Verbal Ability" OR DE "Memory Disorders" OR DE "Amnesia"
6. TI dement* OR AB dement*
7. TI cognit* OR AB cognit*
8. TI memory AND AB memory
9. 4 or 5 or 6 or 7 or 8
10. 9 and 3

Appendix 8: PRISMA checklist (telephone-based CSI)

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
TITLE / ABSTRACT			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	111
Abstract	2	Abstract: See PRISMA-DTA for abstracts.	Not in thesis but in publication
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	111-112
Clinical role of index test	D1	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	111
Objectives	4	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	112
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	112
Eligibility criteria	6	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	113
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	112-113
Search	8	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	Appendix 7
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	113

Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	112
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	114
Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	114
Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	114
Synthesis of results	14	Describe methods of handling data, combining results of studies and describing variability between studies. This could include, but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity, c) handling multiple index test readers, d) handling of indeterminate test results, e) grouping and comparing tests, f) handling of different reference standards	114

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
Meta-analysis	D2	Report the statistical methods used for meta-analyses, if performed.	114
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	115
RESULTS			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	116
Study characteristics	18	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	122-123, 126, 127
Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	120
Results of individual studies	20	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	122-123, 125, 126, 127, 130

Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	124, 128
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	129, 131-132
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	133
Limitations	25	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	135
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	135-138
FUNDING			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	Not in thesis but in publication

Appendix 9: APPLE study protocol

Improving our assessment and understanding of the short, medium and longer term neuropsychological consequences of stroke

Running Title:	**Assessing Post-stroke Psychology Longitudinal Evaluation (APPLE)**
Lay Title:	**Understanding the emotional, thinking and memory problems that can follow a stroke**
Protocol Version:	**1.5**
Date:	**08.12.17**
REC Reference Number:	**16/SS/0105**
Sponsors Protocol Number:	**GN14NE496**
Protocol registration:	**researchregistry1018**
Sponsor:	**NHS Greater Glasgow and Clyde**
Funders Reference:	**PPA 2015/01_CSO**
Funder:	**Joint Stroke Association and Chief Scientist Office Programme Grant**

Amendment number	Date	Protocol version
AM01	19.09.16	1.3 (GN14NE496)
AM02	22.05.17	1.4 (GN14NE496)
AM03	08.12.17	1.5 (GN14NE496)

This study will be performed according to the Research Governance Framework for Health and Community Care (Second edition, 2006) and WORLD MEDICAL ASSOCIATION DECLARATION OF HELSINKI Ethical Principles for Medical Research Involving Human Subjects 1964 (as amended).

Contacts

Chief Investigator

Dr Terry Quinn
Stroke Association / Chief Scientist Office Senior Clinical Lecturer
College of Medicine, Veterinary & Life Sciences
Institute of Cardiovascular and Medical Sciences
University of Glasgow
Room 2.44 New Lister Building
Glasgow Royal Infirmary
Glasgow G4 0SF
Tel: 0141 201 8510
Email: terry.quinn@glasgow.ac.uk

Co-Investigators

Professor Peter Langhorne
Professor in Stroke Care
College of Medicine, Veterinary & Life Sciences
Institute of Cardiovascular and Medical Sciences
Level 2, New Lister Building
Glasgow Royal Infirmary
Glasgow G4 0SF
Tel: 0141 201 8510
Email: peter.langhorne@glasgow.ac.uk

Professor David Stott
Professor in Geriatric Medicine
College of Medicine, Veterinary & Life Sciences
Institute of Cardiovascular and Medical Sciences
Level 2, New Lister Building
Glasgow Royal Infirmary
Glasgow G4 0SF
Tel: 0141 201 8510
Email: david.j.stott@glasgow.ac.uk

Professor Matthew Walters
Professor in Clinical Pharmacology
College of Medicine, Veterinary & Life Sciences
Institute of Cardiovascular and Medical Sciences
Queen Elizabeth University Hospital
1345 Govan Road
Glasgow G51 4TF
Tel: 0141 201 1100
Email: matthew.walters@glasgow.gla.ac.uk

Professor Kennedy Lees
Professor of Cerebrovascular Medicine
College of Medicine, Veterinary & Life Sciences
Institute of Cardiovascular and Medical Sciences
Queen Elizabeth University Hospital
1345 Govan Road
Glasgow G51 4TF
Tel: 0141 201 1100

Email: kennedy.lees@glasgow.gla.ac.uk

Dr Jesse Dawson
Reader in Clinical Pharmacology
College of Medicine, Veterinary & Life Sciences
Institute of Cardiovascular and Medical Sciences
Queen Elizabeth University Hospital
1345 Govan Road
Glasgow G51 4TF
Tel: 0141 201 1100
Email: jesse.dawson@glasgow.ac.uk

Dr Niall Broomfield
Honorary Clinical Teacher
Institute of Health and Wellbeing
Primary Services HQ
Gartnavel Royal Hospital
Glasgow G12 0XH
Tel: 211 2134
Email: Niall.Broomfield@glasgow.ac.uk

Professor Jonathan Evans
Professor of Applied Neuropsychology
Institute of Health and Wellbeing
R212 Level 2, Mental Health & Wellbeing,
Gartnavel Royal Hospital,
Glasgow G12 0XH
Tel: 0141 211 3978
Email: Jonathan.Evans@glasgow.ac.uk

Professor Gillian Mead
Professor of Geriatric Medicine
University of Edinburgh
Room S1642, Royal Infirmary of Edinburgh,
Little France Crescent,
Edinburgh EH16 4SA
Tel: 0131 242 6481
Email: gmead@staffmail.ed.ac.uk

Professor Sarah T Pendlebury
Associate professor in Medicine and Old Age Neuroscience
Nuffield Department of Clinical Neurosciences
University of Oxford
Level 6 West Wing John Radcliffe Hospital,
Oxford, OX2 7RX
Tel: 01865 231606
Email: Sarah.pendlebury@ndcn.ox.ac.uk

Dr Christine McAlpine
Consultant in Stroke and Managed Clinical Network Lead
Centre Block
Glasgow Royal Infirmary

Glasgow G4 0SF
Tel: 0141 211 4843
Email: Christine.mcalpine@ggc.scot.nhs.uk

Trial Statistical Lead
Professor Ian Ford
Professor of Medical Statistics
Robertson Centre for Biostatistics
University of Glasgow, Boyd Orr Building
University Avenue,
Glasgow, G12 8QQ
Tel: 0141 330 4744
Email: Ian.Ford@glasgow.ac.uk

Data Centre
Robertson Centre for Biostatistics
University of Glasgow
Boyd Orr Building
University Avenue
Glasgow G12 8QQ
Tel: 0141 330 4744

Sponsor:

The sponsors is NHS Greater Glasgow and Clyde.

Contact: **Emma-Jane Gault**
Research Governance Officer
University of Glasgow
Clinical Research and Development
West Glasgow Ambulatory Care Hospital
Dalnair Street
Glasgow G3 8SW
Tel: 0141 232 1819
Email: emmajane.gault@glasgow.ac.uk

Funding Body

Funded via a Stroke Association / Chief Scientist Office Joint Programme Grant.
Administered by: The Stroke Association
Research Department
Stroke House
240 City Road
London, EC1V 2PR

Protocol Approval
Improving our assessment and understanding of the short, medium and longer term neuropsychological consequences of stroke

Chief Investigator: **Dr Terry Quinn**
Clinical Senior Lecturer
College of Medicine, Veterinary & Life Sciences
Room 2.44 New Lister Building, Glasgow Royal Infirmary
Glasgow G4 0SF

Signature:

Date:

Sponsor Representative: **Emma-Jane Gault**
Research Governance Officer
University of Glasgow
Clinical Research and Development
West Glasgow Ambulatory Care Hospital
Dalnair Street
Glasgow G3 8SW

Signature:

Date:

TABLE OF CONTENTS

ABBREVIATIONS

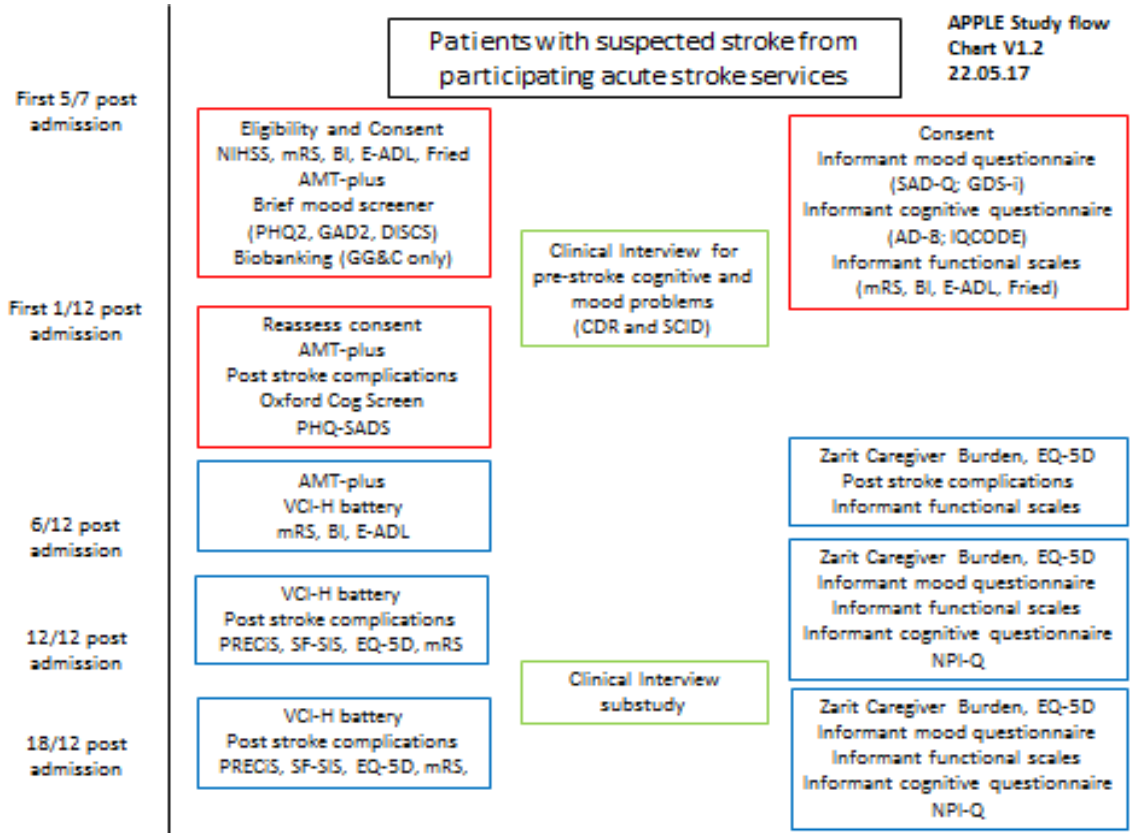
AD8	Ascertaining Dementia 8 Question Screener
AE	Adverse Event
AMT	Abbreviated Mental Testing
ASU	Acute Stroke Unit
BI	Barthel Index
CAM	Confusion Assessment Method
CDR	Clinical Dementia Rating
CRF	Case Report Form
DISCS	Depression Intensity Scale Circles
DSM	Diagnostic and Statistical Manual of Mental Disorders
E-ADL	Extended Activities of daily living
EQ-5D	Euro-QOL 5 Dimensions
GAD-2	Generalised anxiety disorder (2 question screener)
GCP	Good Clinical Practice
GDS	Geriatric Depression Scale
IQCODE	Informant Questionnaire Cognitive Decline in the Elderly
MCN	Managed Clinical Network
mRS	Modified Rankin Scale
NE-ADL	Nottingham Extended Activities of Daily Living
NIHSS	National Institutes of Health Stroke Scale
NPI-Q	Neuropsychiatric Inventory Questionnaire
OCS	Oxford Cognitive Screen
PHQ-2/SADS	Patient Health Questionnaire (2 question screener) (structured assessment)
PI	Principal Investigator
PIS	Patient Information Sheet
PRECiS	Patient Reported Evaluation Cognitive Impairment Scale
QEUH	Queen Elizabeth University Hospital
RAH	Royal Alexendra Hospital
SADQ-10	Stroke Aphasia Depression Questionnaire
SCID	Structured Clinical Interview for DSM
SF-SIS	Short Form Stroke Impact Scale
SOP	Standard Operating Procedure
SRN	Stroke Research Nurse
SSRN	Scottish Stroke Research Network
STARD	Strengthening Transparency and Reporting in Diagnostic Studies
TIA	Transient ischaemic attack
VCI-H	National Institute of Neurological Disorders and Stroke–Canadian Stroke Network Vascular Cognitive Impairment Harmonization Standards
WP	Work Package

1 STUDY SYNOPSIS

Title of Study:	Improving our assessment and understanding of the short, medium and longer term neuropsychological consequences of stroke
Study Centre:	Glasgow Royal Infirmary and associated hospitals of the Managed Clinical Network for Stroke or Stroke Research Network.
Duration of Study:	4 years
Objectives:	To establish a prospective inception cohort, recruited early after stroke and followed for up to 18 months with a focus on psychological outcomes.
Primary Objective:	There are three distinct work packages (WP). WP 1. To assess the prevalence of psychological problems that pre-date stroke. (A separate complementary study will describe test accuracy of short questionnaires for assessing pre-stroke psychological problems). WP 2. To assess test accuracy and utility of brief cognitive and mood tests short for assessment of short and longer term psychological outcomes. WP 3. To describe change in cognition and mood over time following a stroke, with assessments at one, six, twelve and eighteen months.
Secondary Objectives:	The secondary objective is to create a resource that can be used for future studies of psychological impact of stroke. To this end we will ask participants if they wish to have blood taken for biobanking; if we can hold their anonymised data (clinical, laboratory and radiological) in a secure database and if we can access de-identified data from electronic health records.
Main Study Endpoints	Pre-stroke cognitive and physical function (based on CDR and SCID structured interviews). Change in cognition or mood symptoms based on repeated neuropsychological assessment (using VCI Harmonization Standard). Development of incident cognitive or mood disorder (consensus agreement based on collected materials).
Rationale:	National stroke guidance recommends early cognitive and mood screening but this policy lacks evidence-base. Building on previous work, we will create a programme of research designed to inform practice and policy. We will major on themes of "natural history" of neuropsychological problems; screening test accuracy/feasability; prognosis and user experience.
Methodology:	Prospective, observational cohort with nested test accuracy studies.

Sample Size:	500 participants recruited to primary study, with plans for pooled analyses with other studies. Attrition is expected and we have based sample size on 200 participants completing 18 month assessments. The pre-stroke assessment diagnostic study is based on a separate sample size calculation and requires 100 informant interviews and diagnostic assessments.
Screening	Case note review of in-patient / outpatient attendees to the Acute Stroke Services by clinicians. A full log will be maintained.
Inclusion Criteria	1. Clinical diagnosis of stroke or transient ischaemic attack (TIA) at time of assessment. 2. Age greater than 18 years. 3. Treating clinician happy that the patient would have some form of psychological assessment as part of usual care.
Exclusion Criteria	1. Non-stroke diagnosis at time of assessment. 2 Unable to consent and no suitable proxy available. 4. No spoken English pre-stroke. 4. Prisoners.
Statistical Analysis	WP 1,2: Accuracy of screening tools will be described in terms of usual test accuracy metrics against a reference standard of semi-structured baseline clinical assessment (WP1) or prospective assessment with neuropsychological battery (WP2). We will employ an "intention to diagnose" approach. WP 3: Outcomes of interest are change in scores on neuropsychological battery and incident clinical mood disorder or cognitive impairment. We will use generalized linear models for prospective data to describe associations of baseline characteristics with change across repeated neuropsychological measures and use varying competing risk survival models. We will describe univariate and adjusted independent predictors of "outcomes" using odds-ratios for binary "outcomes" at chosen time-points. We will create prognostic models and if data allow predictive risk scores for outcomes, describing calibration; discrimination and validation using bootstrapping.

STUDY FLOW CHART



NB. All aspects of the study are optional, participants can chose to contribute to all or only one part of the study. Some sites may not be able to offer biobanking.

Key: *Red boxes: short screening assessments; blue boxes: detailed screening assessments; green: structured psychology assessment with clinician input.*

Details of all the neuropsychological battery assessments in appendix.

PARTICIPANT SCHEDULE

Detailed Participant Schedule					
	ASU	**Follow up**			
	Week 1	1 month	6 months	12 months	18 months
Review Eligibility	√				
Consent	√				
Blood / Urine for Biobanking (separate optional study)	√	√	√	√	√
Patient assessment	√				
Informant assessment	√				
Structured clinical interview study (separate optional study)	√				
Consent re-assessed		√			
Patient psychological screen		√			
Patient neuropsychological battery			√	√	√
Informant questionnaires			√	√	√
Clinical assessment (separate optional study)				√	
Consensus assessment					√

NB. All aspects of the study are optional, participants can choose to contribute to all or only one part of the study. Some sites may not be able to offer biobanking.

1 INTRODUCTION

1.1 Background

We propose a programme of work designed to improve our understanding of neuropsychological effects of stroke. We will focus on themes of assessment, prognosis and natural history. Outputs will have immediate relevance and impact, providing an evidence base to policy and practice around early cognitive and mood screening and informing the design and conduct of future studies. The prospective cohort and biobank/big data resources created through this work will act as foundation for an ongoing portfolio, creating cross institutional research synergy; encouraging new researchers and providing the "substrate" for ongoing interdisciplinary work.

People affected by stroke have consistently highlighted the importance of neuropsychological issues.[1] However, the field remains relatively under researched. Important evidence gaps collated at a Stroke Association convened priority setting workshop, were around the "natural history" of neuropsychological change after stroke; utility of early assessments and predicting who will require later specialist input.[2] Our proposed body of research is designed to address these priority areas. Specifically, we will create a "real world" acute stroke inception cohort, offering prospective cognitive and mood testing to progress inter-related themes (Work Packages). We will also offer related complementary, optional studies looking at pre-stroke assessment and facilitating biobanking and 'big data' approaches.

Despite the importance of psychological issues, memory, thinking and mood have not received as much attention in stroke research as other areas.[3,4] For this reason there are still fundamental questions that we don't know the answer to. These include:

- How do memory thinking and mood change after stroke?
- What happens to memory, thinking and mood in the longer term after stroke?
- Can we predict which people will have problems with memory, thinking and mood?
- What is the best way to look for problems with memory, thinking and mood?
- When should we perform tests of memory, thinking and mood?

These are the questions we wish to answer with this programme of research.

1.2 Pilot data to support the creation of a cohort

The proposed programme of work builds on our previous systematic review and original research. Our national questionnaire and literature review has shown inconsistency in neuropsychological assessment strategies both in clinical practice and in research.[3,4] Subject responses from stroke units across Scotland suggest that clinical teams are looking for guidance around method and timing of neuropsychological assessment and around prognosis.

Our systematic review work has highlighted a lack of data around cognitive and mood screening tools in acute care, albeit this is where the majority of initial assessment is performed.[5]

With the Standards for Reporting of Diagnostic accuracy in Dementia (STARDdem)working group, we have creating guidance for conduct and reporting of diagnostic test accuracy studies and have used this to inform the proposed work.[6]

Importantly, our pilot work has shown that studies of early neuropsychological assessment with prospective follow up can recruit rapidly and efficiently.[7,8]

1.3 Involvement of stroke-survivors and others affected by stroke

This body of work has been created with input from stroke survivors and others affected by stroke. Input from stroke survivors and those affected by stroke will continue for the lifespan of the study.

The researchers involved in this application were part of a national research priority setting group that collated feedback from various groups including strong representation from stroke survivors and care-givers. The number one research priority identified through this work was around problems with memory and thinking that can occur after stroke. This feedback was the inspiration behind this work.[1]

The Stroke Association ran their own workshop around memory and mood problems, the lead applicant in this work was part of this group, that also included stroke survivors and representation from various professional groups. The conversations and experiences shared as part of this workshop and our daily clinical work in stroke units helped us create a body of work that we believe is relevant to stroke survivors and is in keeping with the issues that they feel are important. The project was further reviewed by Stroke Association lay members as part of the grant review process.

The research plan outlined in this application has been previously assessed by panel members of the UK Stroke Research Network CSG (acute and rehabilitation groups). This group includes clinicians from various disciplines and representation from those affected by stroke. The insightful suggestions and comments we received, particularly around the conduct and reporting of the work, have improved the proposal considerably.

In designing a study, researchers always need to balance their desire to collect detailed information with not over-burdening the person taking part in the study. With a stroke group based in Edinburgh we have shared the various tests of memory, thinking and mood that we propose to use. Feedback from the group helped us refine our set of tests to a selection that should be acceptable to stroke survivors while still giving us the necessary information we need to answer our research questions.

As part of the study we will create an advisory group, who will meet once yearly. The advisory group will include two stroke survivors as well as representation from doctors, nurses and therapists. The group will act as a forum for stroke survivors and others to comment on the design of the study; the progress of the study; the "meaning" of the results obtained and how to share these results with the wider stroke community including stroke survivors. (*advisory group members detailed in appendix*)

1.4 Principal research questions

This application is towards a programme of work supported by the Stroke Association and Chief Scientist Office Scotland.

Within the programme are three distinct work packages (WP) designed to offer rich data that answer a number of important questions in stroke care.

WP 1. The primary aim is to assess the prevalence of memory and thinking (cognitive) and mood problems that pre-date the stroke. A complementary (optional) study seeks to describe the test accuracy of short questionnaires for assessing pre-stroke psychological problems.

WP 2. The primary aim is to assess how useful short tests are for detecting cognitive and mood problems immediately after a stroke and for detecting persisting cognitive and mood problems.

WP 3. The primary aim is to describe change in cognition and mood over time following a stroke, with assessments at around one month, six months, twelve months and eighteen months.

A further important objective is to create a resource that can be used for future studies of psychological impact of stroke. To this end we will ask participants if they wish to have blood taken for biobanking (optional); if we can hold their anonymised clinical, laboratory and imaging data in a secure database (optional) and if we can access de-identified data from electronic health records (optional). All these aspects are optional and may not be available in certain centres.

1.5 Summary of Risk Assessment

We recognise the potential issues associated with this project; we have worked with patient groups, lay representatives and clinical study advisory groups to create a methodology that minimises issues while maintaining the research potential of the programme of research. We have listed the potential issues and steps taken to minimise their impact.

Test burden: The project involves cognitive and mood testing of patients and informants at various stages in the stroke journey. We recognise the importance of minimising test burden. Our pilot work suggests that patients struggle with standard, multidomain cognitive tests in the first days post stroke. In this study we will concentrate on very brief tests. As the brief tests share a number of questions, we can assess the performance of several tests at once by simply adding some questions to the short cognitive assessment that is used as standard in our clinical service.

As part of our preparatory work we asked a stroke group based in Edinburgh to look at the tests we proposed for the acute study and they were happy that the tests were not overly burdensome. Piloting the acute test battery with an Edinburgh research group suggests that completion should take around 20 minutes at most. Patients are not required to complete all the tests and they can ask to stop testing at any time. Testing can be performed in two sessions or more depending on patient preference. Feasibility of using brief tests is an important metric of this work and we will record how many patients attempt and complete tests. If a participant becomes distressed or frustrated and it is clear that they are unable to complete testing, testing will be stopped. Any distress will be handled through reassurance and ending the assessment.

Informants (family, friends, carers) will also be asked to complete paper based questionnaires. We have chosen brief assessments that should take around 20 minutes and can be completed at a time that suits the participant.

The prospective arm of the study will use a longer test battery. Completing the study follow-up will involve four assessment visits over 18 months (one month; six months; twelve months; eighteen months). We have chosen cognitive and mood tests recommended for stroke cohorts and which we use in clinical practice. There is considerable experience of using these tests with stroke survivors. The first session using these longer test batteries will not begin before six months post stroke to allow time for recovery. Again testing can be performed in split sessions if the patient prefers. Completion of the tests is not mandatory and the patient can request to stop testing at any time. Where completion of the full assessment is not possible, we have specified a short form assessment protocols for use in person or over telephone.

Opportunity cost: We recognise that while a patient is working on cognitive assessment they will not be able to work with ward staff / allied health professionals on other rehabilitation tasks. We will work with the ward team to minimise disruption. We will be performing an activity, cognitive and mood testing screen, that is a recommended part of routine care. We will share the inpatient test results with ward staff on request and this should release their time for other activities.

Disclosure of sensitive information: We will be assessing mood (emotions and feelings) we recognise that this can be a sensitive area. If we detect probable depression, or other mood disorder we would advise the clinical team to refer to the Stroke Psychology service. In the event that suicidal thoughts/ideations are disclosed assessment will be stopped and a member of the treating clinical team will be informed immediately. This action will also be documented in the patient's case notes. The study

has input from the local Clinical Stroke Psychology service and they are happy to be contacted in the event of suicidal ideation or any other disclosures that may require clinical input.

Patients may ask for their scores on the cognitive or mood tests. We will share these data with the patient but we will also explain that these tests in and of themselves are not diagnostic of dementia / depression or other serious psychological problems. Rather they are part of an assessment that will be shared with the treating clinical team. If there is concern regarding a patients cognitive function or mood, the research team can access the stroke clinical psychology services and referrals can be made to Memory Clinic services.

Informed consent: We want our study to produce results that reflect "real world" stroke care. Previous studies of cognition and mood in stroke have limited themselves to consenting patients. This gives a biased sample and produces results that lack external validity and clinical utility. We propose a more generalizable approach, where we potentially include all patients with stroke unless the clinical team feel that any form of testing is inappropriate. There will be a proportion of patients who may struggle to provide informed consent to research. For a study with a cognitive focus, it is important that these patients are included. In this instance we will seek consent from a suitable proxy (family, friend, carer). For those patients who are included in the study with proxy consent; we will reassess capacity to consent and seek informed consent at one month follow-up visit.

Test environment: For follow-up testing we will recommend that testing is performed within one of the clinical research facilities of the participating hospitals. We have a budget to cover patient travel by taxi to allow this. Some patients may be unable to attend the research facility or may for any reason choose to be assessed at home. Telephone based assessment is possible if required. For home assessments, we will follow NHS GG&C and GU lone working procedures for safety.

Use of participant data: Our cohort will provide a unique resource for understanding post-stroke psychological problems. We wish to maximise the potential of the data collected, so that it can be used to answer clinically important questions beyond those outlined in the primary study. Certain centres will invite participants to give a blood samples that will be stored for future analyses. Participants can decline this biobanking aspect at any stage of the study and still help with other aspects. The Robertson Centre for Biostatistics will hold anonymised patient data within a secure resource. Our study follow-up is limited to eighteen months. We will ask participants permission to link their study data to anonymised data from electronic health-records (clinical, laboratory, imaging). This will allow future studies to look at longer term outcomes. Again, participants can chose to decline this aspect of the study but contribute to the other aspects.

2 STUDY OBJECTIVES AND PURPOSE

We propose a study that has been designed to answer pressing clinical questions. National and international stroke guidelines recommend early cognitive and mood screening but this policy is based on expert opinion and lacks evidence-base. Building on our previous pilot work, we will create a programme of research designed to inform practice and policy. We will major on themes of "natural history" of neuropsychological problems following stroke; screening test accuracy/feasibility and prognosis.

We anticipate that at study completion we will be able to offer:

- Guidance on the optimal methods to assess for pre-stroke cognitive and mood problems.
- Guidance on the optimal methods to assess for cognitive and mood problems in the acute stroke setting and in early follow-up.
- Descriptions on the natural history of cognitive and mod symptoms following stroke.
- An understanding of clinical, demographic features that predict poor and good psychological outcomes following stroke.

STUDY DESIGN

This study will be performed according to the Research Governance Framework for Health and Community Care (Second edition, 2006) and WORLD MEDICAL ASSOCIATION DECLARATION OF HELSINKI Ethical Principles for Medical Research Involving Human Subjects 1964 (as amended). All investigators and key study personnel will undergo biennial GCP training.

The study design is detailed below and is summarised in the flow chart and schedule.

The programme of work is based on a prospective observational cohort, recruited at time of stroke and followed up with assessments focussing on neuropsychological aspects. The cohort will allow for studies of test properties, studies of prognosis and epidemiology. Inclusion of a biobank and consent to future electronic data linkage increases the research potential of the cohort.

3.1 Study Population

The study will involve participants aged over 18 years with clinical diagnosis of stroke or transient ischaemic attack (TIA) who meet the inclusion criteria and have none of the specified exclusion criteria. All will give full informed consent or have consent provided by appropriate proxy.

Participants will be consecutive, stroke patients over 18 month recruitment. Primary sites will be Glasgow Royal Infirmary (GRI); Royal Alexandra Hospital (RAH) and Queen Elizabeth University Hospital (QEUH), with additional recruitment from other Scottish Stroke Research Network (SSRN) sites or research active stroke centres in other parts of the UK. To allow descriptions of generalizability and feasibility we will adopt an inclusive policy, offering testing to all adult (over 18 years) stroke survivors except where clinical team feel that any form of testing is inappropriate (for example end of life care). We will define stroke using World Health Organisation criteria. Our stroke rubric will include TIA and minor stroke and recruitment from outpatient clinics will be possible. Co-recruitment with other observational or investigational trial will be possible.

We will include patients with varying levels of communication problems. Our national ethics application will allow us to seek proxy consent for assessment and follow up where participant is unable to give direct consent at time of recruitment. The context is disease orientated and so will include new TIA/minor stroke seen at clinics as well as inpatients. Eligibility screening and recruitment will be performed by stroke research nurses or trained researchers. Initial assessment of capacity and willingness to be approached will be determined by the clinical team.

3.2 Main Study Inclusion Criteria

1. Clinical diagnosis of stroke or TIA at time of assessment

2. Age greater than 18 years.

3. Clinical team happy that patient is suitable for some form of psychological testing.

Stroke will be diagnosed by a stroke specialist, defined as a focal neurological event of presumed vascular cause. We will operate no time or imaging based inclusion criteria.

3.3 Main Study Exclusion Criteria

1. Non-stroke diagnosis at time of assessment.

2. Unable to consent and no suitable proxy available.

4. No spoken English pre-stroke.

4. Prisoners.

3.4 Description of the work packages

We propose a programme of work themed around improving cognitive and mood assessment.

The portfolio is described as interlinked work-packages each with distinct aims and objectives. In addition we offer optional, complementary studies.

Work package one: Assessing pre-stroke psychological problems.

- To describe prevalence of pre-stroke psychological problems (specifically, cognitive decline and depression) in an acute stroke cohort.
- A separate (optional) study will assess the feasibility of using informant based screening tools for pre-stroke depression (GDS-informant scale; SADQ-10) and cognitive decline (IQCODE, AD-8) in an acute stroke setting.
- A separate (optional) study will assess the accuracy of informant based screening tools for pre-stroke depression (GDS-informant scale; SADQ-10) and cognitive decline (IQCODE, AD-8) against a reference standard of semi-structured clinical assessment (using the Structured Clinical Interview [SCID] for DSM mood disorder and the clinical dementia rating [CDR] for cognitive assessment).

Published research describing cognitive and mood problems following stroke assumes that the person had no problems prior to the stroke event. This is overly reductionist approach fails to appreciate the complex relationship between psychological symptoms and cerebrovascular disease. Stroke is predominantly a disease of older age and older people will show varying degrees of cognitive decline and mood problems. These may be sufficient to warrant a diagnostic label, albeit often a diagnosis of dementia or mood disorder is not made in the community.[9] Both cognitive decline and mood disorder seem to be associated with increased risk of stroke.[10]

To understand the psychological picture seen after stroke we need robust methods of capturing the pre-stroke state. A common approach is to conduct a questionnaire based interview with informants (family, friends, carers) and use the description of past cognitive and mood symptoms to assign a retrospective label. Scales are available and are used in stroke care, for example the Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE). Our recent systematic review has shown that while test properties of informant scales are good in community dwelling older adults, no informant questionnaire has been validated in a stroke population.[11,12]

We will use a classical test accuracy study design to describe the properties of informant tools in acute stroke. Stroke research nurses (SRN) or trained researchers will interview informants with short questionnaires looking to describe pre-stroke depression and cognition. Within one month of this assessment, a trained member of the research team will conduct a semi-structured interview (based on standardised questionnaires of SCID and CDR (sum of boxes scoring https://www.alz.washington.edu/cdrnacc.html) with patient and family and formulate a clinical assessment of pre-stroke problems. Following discussion with a clinician, results will be operationalised as pre-stroke dementia or depression probable; possible; unlikely; unable to assess.

Accuracy of screening tools will be described in terms of sensitivity; specificity; predictive value; receiver operating space analyses. Index test questionnaires will be compared against each other and against a reference standard of semi-structured clinical assessment. To describe feasibility we will collate numbers completing each test fully and partially; time and assistance required for completion. To incorporate feasibility into analyses, we will employ an "intention to diagnose" approach, including those unable to complete tests.[13]

Work package two: Test accuracy and prognostic utility of brief screening tools

- To describe feasibility of using brief screening tools for diagnosis of cognitive and mood problems in acute stroke.
- To describe accuracy of brief screening tools for diagnosis of cognitive and mood problems in acute stroke; comparing to each other and to a one month multi-domain assessment.
- To describe prognostic accuracy of a one month multi-domain cognitive and mood assessments against detailed assessment at six, twelve, eighteen months.
- To describe neuropsychological "case-mix" with reference to incident/prevalent delirium and impairments that may complicate cognitive and mood testing.

The first step to management of neuropsychological problems is recognition and diagnosis. At present we have no agreed method on how or when to assess for these problems. Our pilot data suggests that standard multi-domain assessment tools are not feasible as a universal screen in the first days post stroke.[5,8] Thus, we suggest a neuropsychological assessment paradigm where brief assessments are used in the hyperacute period with increasingly detailed assessment at later time period.

Various brief (less than five minutes) assessment tools for cognition and mood are available. Such tools are suited to acute settings and indeed are often used in the ASU, however data on test properties are limited.[5] Many of these brief assessments have shared items. We have created an instrument that combines elements from popular brief tests in a single assessment, allowing derivation of various scores while minimising test burden. Our brief mood testing includes a depression and anxiety questionnaire; pictorial assessment and single question. Tests for delirium are also included. We have not modified assessments for those with communication problems, as describing feasibility of tests across a range of stroke related impairment is an important outcome of our work. However, the tests used should be feasible for those with mild to moderate aphasia. At one month, a longer test battery will include multi-domain screening tool. (Assessments described in appendix).

Our methodology is based on best practice in conduct and reporting guidance for dementia test accuracy studies (STARDdem).[6] Index test will be brief screening tools (acute assessment) and multi-domain screening tools (one month and beyond). Given the dynamic early changes in cognition and mood seen early after stroke, purpose of early testing should be to predict later problems. Thus our reference (gold) standard comparator will be mood disorder and multi-domain cognitive impairment as described by our neuropsychological battery at six, twelve and eighteen months with expert consensus diagnosis based on all collated materials at end of study. We recognise that these assessments are not diagnostic, rather they offer a suitable compromise between validity of assessment and suitability post stroke where formal diagnosis of dementia or mood disorder can be challenging. As an optional study, at 12/12 and 12/18 follow-up a random selection of participants, will be offered additional face to face clinical assessment with a senior stroke neuropsychologist or clinician blinded to other assessment scores. At completion all 6,12,18 month study materials will be reviewed by the senior investigators (TQ, NB, JD, DJS) and a consensus diagnosis assigned for incident mood disorder and/or incident cognitive disorder, using descriptors of:probable, possible, unlikely.

Accuracy of screening tools will be described in terms of sensitivity; specificity; predictive value; receiver operating space analyses. Index tests will be acute and one month assessments and will be compared against each other and reference standard of follow up assessment data. From the acute assessments, we will describe the accuracy of brief screening tests used in isolation and combined with Boolean operators of "OR"/"AND". To describe feasibility we will collate numbers completing each test fully and partially; time and assistance required for completion. To incorporate feasibility into analyses, we will employ an "intention to diagnose" approach, including those unable to complete tests.[13]

Work package three: Describing and predicting neuropsychological prognosis

- To describe serial change in cognition/mood test scores and to describe prevalence of cognitive and mood diagnoses at time points of one month; six month; twelve months and eighteen months.
- To describe univariate and adjusted independent predictors of both post stroke cognitive decline and post stroke mood disorder.
- To develop, calibrate and validate predictive models for post stroke neuropsychological factors.
- To estimate likely recruitment, "event rates" and loss to follow up for future cognitive/mood studies.

Systematic reviews suggest substantial post stroke neuropsychological burden, however these data may have limited generalizability to acute settings.[14] Problems include selection bias; non-acute sampling and lack of data on important comorbidities such as delirium and prevalent dementia. Our pilot data describes a high incidence of cognitive/mood problems in first days post stroke with trajectories of improvement, stabilisation and decline.[8] We need "real world" data on baseline and natural history of neuropsychological change to inform practice, research and policy in this regard.

Follow up will be at six, twelve and eighteen months, time-points chosen to reflect common clinical and study assessment times. Assessments will be face-to-face and performed in study centres or in participant's home as required/requested. There will be opportunity for telephone assessment if required. The six/twelve/eighteen month assessments will be performed by trained members of the research team. We make no assumptions around the pathology underlying post stroke cognitive change and so we have devised a battery of assessment that will allow derivation of scores for "vascular" dementia and Alzheimer's Disease dementia.[15,16] While our principal mood interest is depression we have chosen a mood assessment that screens for various other disorders using structured clinical interview.[17] (see appendix for full details of all assessments) After 12 and 18 month follow-up, a proportion of participants will be asked if they wish to take part in an optional study, where they are assessed by a clinician and assigned a clinical label. These results will be compared to our standard assessments.

The work is modelled around the "fundamental" prognosis research paradigm as described by MRC PROGRESS prognosis research group.[18] Taking acute stroke as start-point, we will create an inception cohort, collecting clinical, demographic and neuropsychological "phenotyping" data at baseline and then prospectively following up with serial cognitive and mood assessments.

For prospective follow up, outcomes of interest are change in scores on cognitive and mood screening tools and incident clinical mood disorder or multi-domain cognitive impairment. Multi-domain tools will be analysed as ordinal data and dichotomised at varying thresholds. Neuropsychological battery data will be transformed into z scores, with impairment defined as greater than 1.5 standard deviations below age and sex based norms. We will collect data on recurrent stroke, complications (falls, seizure, infection) hospitalisation/institutionalisation and death.

We will explore repeated measures analyses adjusting for baseline covariates and describe temporal change in test scores. We will create prognostic models and if data allow predictive risk scores for the various cognitive and mood outcomes, describing calibration; discrimination and validation using bootstrapping.

3.5 Identification of Participants and Consent

Potential participants will be identified (by clinical or case note review by a member of the clinical team or attending Doctor) whilst in-patients or in a cerebrovascular out-patient clinic. If the patient asks not to be approached no further action will be taken. The clinical team will make an assessment of capacity to consent to inclusion in the study. The principal criterion for entry into the study is that the treating team believe

an attempt at cognitive and mood assessment is appropriate. We have used this approach in previous pilot studies and it has worked well.

Following identification, potential participants will be approached in person and asked whether they would wish to consider taking part in the trial. Those who are willing to hear more will be given the participant information sheet (PIS) and a date (at least 24 hours later) arranged for further discussion with a member of the research team. Eligibility will be confirmed by an investigator.

At this second meeting, subjects will be asked if they have any questions and those who wish to participate will be asked to sign the consent form. Two copies will be signed (one each for the participant and the site file) and a copy of the signed consent form will be inserted into the casenotes.

Consent will be taken by one of the investigators, research nurse or trained researcher.

Consent will be staged to ensure that participation in the study is always voluntary and fully informed. At all points we will stress that taking part in the study is voluntary and if patients wish to terminate the cognitive testing early we will respect this wish which will not impact on the clinical care that they receive.

For patients unable to provide informed consent, we will seek consent from a legal proxy or family, carer, friend. We have outlined the details of this approach in the section on adults lacking capacity (see below).

We offer additional complementary studies looking at informant assessment; blood taking for biobanking; prospective follow-up; clinical diagnostic study; data storage and linkage. Participants will be given the option to consent to all aspects of the study or to limit their participation to certain aspects only. In centres where biobanking is not possible this will not be included in consent form.

We recognise that cognition can change over time. Our pilot data suggests that immediately after stroke patients can have cognitive impairments that improve over the first weeks.[8] At early follow-up (around 4 weeks post stroke) the participant's capacity to consent will be reassessed.

3.5.1 Including participants unable to provide informed consent

We wish to include a representative sample of stroke survivors. For a study that is concerned with post stroke psychological problems we need to include a spectrum of cognitive abilities and impairments. Previous work in this area has been limited by including non-representative populations and so results have lacked real world validity. To ensure our results have clinical utility, we will be maximally inclusive in our recruitment strategy.

Patients may have cognitive problems, problems with communication/language or physical impairments. Some may have severe communication or cognitive difficulties. The assessment battery we propose, while not specific to aphasia, should be suitable for those with mild to moderate communication problems. We will only assess those patients where the clinical team feel that an attempt at assessment of mood or cognition is appropriate.

We do not wish to deny stroke survivors involvement in a study that might lead to benefit for those like them. We believe the risk of participation in this observational study is minimal.

Decisions on patient capacity to consent will be made by the Consultant/senior members of the Acute Stroke team at daily ward rounds or on first assessment. This is a standard part of usual clinical practice for stroke clinical teams.

Where the ward clinical team determine a patient does not have capacity to consent, we would seek informed consent from a close relative/welfare guardian. We would still include the patient in decision making around the study as possible. Choice of proxy will be made by the patient, either at the time of testing or based on previously expressed wishes.

We will involve the nearest relative/guardian/welfare attorney in the study, regardless of patient ability to consent as some of our measures require to be completed by an informant that knows the patient well. We have developed a specific information leaflet (PIS) for this purpose.

Capacity to consent will be re-assessed at one month follow-up. If a patient has been included using proxy consent but it is felt the patient now has capacity, consent will be rechecked at the follow-up visit. In this scenario, if the participant does not give consent the participant would be withdrawn from the study. No further data or tissue would be collected or any other research procedures carried out on or in relation to the participant. We would ask if those identifiable data or tissue already collected with consent could be retained and used in the study. If the participant does not agree to this, the data and biobank samples will be removed from study registers.

If the patient is felt to no longer have capacity to consent, the assessor will follow procedures outlined for including a patient that lacks capacity. In this scenario, if a relevant proxy does not give consent the participant would be withdrawn from the study. Identifiable data or tissue already collected with consent would be retained and used in the study. No further data or tissue would be collected or any other research procedures carried out on or in relation to the participant.

At subsequent follow-up visits, capacity to consent will not be formally reassessed but we will check that the participant is still happy to continue with the study and emphasise that the participant can withdraw at any time and not give a reason

3.5.2 Withdrawal of subjects

Participants will be told that they can withdraw their consent at any time without giving a reason and that this will not affect their care in any way. Participants will be informed that they can participate in any or all of the follow up assessments.

3.6 Assessment Schedule

The study will comprise a maximum of seven patient assessments. A short baseline assessment; (optional) semi-structured clinical interview within first month; one month follow-up with short screening tests; then six, twelve, eighteen month follow-ups with multi-domain assessments with an optional clinical diagnostic assessment. Following the baseline assessments, each visit has a two week time window either side of the scheduled date during which it can be completed. Other than baseline assessment, assessments will be preferentially performed in the Clinical Research Facility of the participating hospital. There is the option for home assessment or for telephone assessment if required.

3.6.1 Baseline assessment

This will be completed as soon as possible following index stroke but not before 24 hours to allow participants sufficient time to read study materials. Initial assessment will confirm eligibility and consent. Clinical and demographic details will be extracted from case-notes. Clinical assessment will include National Institute of Health Stroke Scale (NIHSS) score, modified Rankin Scale (mRS); Barthel Index (BI): five question assessment for frailty (Fried), Lawton Extended Activities of Daily Living (E-ADL) a short questionnaire around physical activity (Brief Physical Activity Assessment [BPAA] and a measure of social inclusion (Medical Outcome Study Social Support Scale[MOSS-SSS] 4 item).

The cognitive assessment (AMT-plus) will comprise the 10 point abbreviated mental test and clock drawing test, supplemented by a recall question, one letter fluency test and naming months of the year backwards. This battery allows us to derive the score from 9 different screening tests without performing each test individually. We will assess for delirium using Confusion Assessment Method (CAM-ICU). We will assess for mood symptoms using Depression Intensity Scale Circles (DISCS) and the short forms of Patient Health Questionnaire PHQ-2/GAD-2. If patient agrees and facility is available, bloods and urine will be taken for biobanking.

Informants will be chosen by the stroke patient or ward staff if stroke patient unable to make this decision. Informants will complete brief questionnaires describing the patient's mood and cognition pre-stroke. Questionnaires will comprise the Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE); the Ascertain Dementia screener (AD-8). The Geriatric Depression Scale informant version (GDS-i) and Stroke Aphasia Depression Questionnaire (SAD-Q). Patients pre-stroke functional ability will be assessed using the BI, Fried and E-ADL. The baseline visit will confirm a suitable time to organise the semi-structured clinical interview.

3.6.1.1 Semi-structured clinical interview

This optional study interview will be performed within one month of baseline assessment. A trained member of the research team will interview the patient and informant. Interview will cover diagnostic criteria necessary to assign a label of major neurocognitive disorder; delirium and major depression. The content will be based on the operationalised structured clinical interview for DSM-5 (SCID) and the Clinical Dementia Rating (CDR – sum of boxes scoring https://www.alz.washington.edu/cdrnacc.html). The interviewer will not have access to previous cognitive and mood screening assessment results. Results of the interview will be discussed with the study team and a final consensus label will be operationalised as: probable cognitive/mood disorder pre-stroke; possible disorder; unlikely disorder; unable to assign a label. We will emphasise that the assessments are not diagnostic but will share the information with the treating clinical team on request.

3.6.2 One month assessment

The one month assessment will be performed at a time convenient for the patient and informant. One month assessments will comprise a repeat of the short patient cognitive

battery performed at baseline (AMT-plus, CAM-ICU), the Oxford Cognitive Screen (OCS) and the complete Patient Health Questionnaire (PHQ-SADS). We will collect information on post stroke complications (stroke, cardiac, seizure, infection, falls, fatigue [using brief fatigue inventory]) and any change in medication. If the patient is agreeable and if available then further samples for biobanking will be taken.

3.6.3 Six, twelve, eighteen month visit

Assessments at six, twelve and eighteen months will be performed by researchers trained in the various assessments. Patients will be assessed according to Vascular National Institute of Neurological Disorders and Stroke-Canadian Stroke Network vascular cognitive impairment harmonization standards (VCI-H).[15,16] Function will be assessed with mRS, BI, EADL and BPAA, MOSS-SSS at 12 and 18 month. The patient will be asked about specific stroke complications of interest. List of medication will be updated.

At the six month assessment the assessor will use the 30 minute version of the VCI-H If the patient struggles with this assessment, does not wish such a lengthy assessment or the assessment is not possible for any other reason, we have proposed a shorter assessment based on the VCI-H five minute battery. For twelve and eighteen month assessments the patient will be offered the choice of full VCI-H (around 45 minutes) or shorter assessments. Choice of assessment used will be at the discretion of the researcher in discussion with participant and informant.

In addition at the twelve and eighteen month visits the patient will complete generic and stroke specific quality of life measures: Euro-Qol 5 domains (EQ-5D); Short Form of the Stroke Impact Scale (SIS) and Patient Reported Evaluation of Cognitive Status (PRECiS).

The informant will complete a caregiver burden scale (Zarit Caregiver Burden) and will complete the generic quality of life EQ-5D. At 12 and 18 months the informant will complete the cognitive and mood questionnaires employed at baseline (IQCODE,Yes include GDS-i) and will complete the neuropsychiatric inventory questionnaire (NPI-Q).

Completion of the eighteen month visit marks the end of the study.

3.7 Biobanking

Urine and blood samples will be obtained as outlined in the appendix and then will be stored in the NHS GG&C biorepository; all aspects of collection and storage will be in line with NHS Greater Glasgow and Clyde policies. Biobanking samples will be from GG&C participants only.

Venepuncture will be performed from the antecubital fossa where possible (using a ~ 19G (green needle) vacutainer (or similar) system). Three lavender top EDTA tube (or similar), a gold top clot activator (or similar) for serum chemistry measures and two grey tube (or similar) for glucose determination will be collected (ca 40 mls in total)

3.8 Team Expertise and Project Management

NHS Greater Glasgow and Clyde have agreed to act as sponsor. All protocols will be stored in publically accessible registers. Creation of case report forms (CRF), data management, archiving and analyses will be supported by Robertson Centre for Biostatistics.

Terry Quinn (Glasgow) will lead the work and act as principal investigator (PI). He has particular expertise in stroke study methodology; test accuracy and cognitive/functional assessments. The core research team will include stroke research nurses at both sites; new researcher posts, designed to allow study towards PhD and dedicated statistical support. The multifaceted nature of the topic requires knowledge and skills in various areas and our collaborators bring this multidisciplinary expertise. Our experienced site leads have international reputations for excellent multicentre, prospective research:

Peter Langhorne (GRI); Kennedy Lees (QEUH). Ian Ford (Glasgow) will support all aspects of statistical analysis. Niall Broomfield (clinical lead for Glasgow stroke psychology services) will provide training for research nurses and doctoral students and will facilitate clinical assessments.

We will form an advisory group who will provide oversight and guidance, the group will have representation from stroke survivors (x2); primary care; research networks (SSRN, SCDRN); neuropsychology (Jonathan Evans, Glasgow); the local stroke managed clinical network lead (Christine McAlpine) and an external expert on neuropsychological outcomes in stroke (Sarah Pendlebury, Oxford).

4 Rater training

We propose assessments using a battery of differing neuropsychological and functional tests. We have extensive experience of training researchers in use of assessment scales. Our previous work around outcomes assessments for large clinical trials has shown the importance of offering training, standardisation and quality control, even for those tests considered "routine" in stroke research.[19]

We will use training materials produced for use with the assessments of interest. Online training resources will be available for functional outcomes (NIHSS, mRS. BI). For the neuropsychological tests we will offer face-to-face training. Educational materials will be complemented by an investigator work book and Standard operating Procedures (SOPs) for all of the assessments required in the study. To accompany the SOPs we will create study-specific case report forms to facilitate standardised assessment and scoring. For PhD student assessors, the first three assessments will be supervised. There is scope for further direct assessment and training as required. Contact details of the principal investigator and research team will be made available to all the sites should issues arise.

4 PHARMACOVIGILENCE

We propose an observational study with no intervention or change to usual care. There are no pharmacovigilance issues specific to this work.

6 STATISTICS AND DATA ANALYSIS

6.1 Primary Outcomes

We propose a programme of inter-related projects themed around improving cognitive and mood assessment.

The portfolio is described as work-packages and optional studies each with distinct aims and objectives. The outcomes and analysis plan for each will be described in turn.

WP 1: Assessing pre-stroke psychological problems.

- To describe prevalence (n, [%]) of pre-stroke psychological problems (specifically, cognitive decline and depression) in an acute stroke cohort.
- As part of an optional, separate study, to assess the feasibility (n, [%] return rate, items complete, time for testing) of using informant based screening tools for pre-stroke depression (GDS-informant scale; SADQ-10) and cognitive decline (IQCODE, AD-8) in an acute stroke setting.
- As part of an optional, separate study, to assess the accuracy (sensitivity, specificity, positive/negative predictive value) of informant based screening tools for pre-stroke depression (GDS-informant scale; SADQ-10) and cognitive decline (IQCODE, AD-8) against a reference standard of semi-structured clinical assessment.

WP 2: Test accuracy and prognostic utility of brief screening tools

- To describe feasibility (n, [%] items complete, time for testing) of using brief screening tools for diagnosis of cognitive and mood problems in acute stroke.
- To describe accuracy (sensitivity, specificity, positive/negative predictive value) of brief screening tools for diagnosis of cognitive and mood problems in acute stroke; comparing to each other and to a one month multi-domain assessment.
- To describe prognostic accuracy (sensitivity, specificity, positive/negative predictive value, ROC analyses) of a one month multi-domain cognitive and mood assessments against detailed assessment at six, twelve, eighteen months.
- To describe neuropsychological "case-mix" with reference to (n, [%]) prevalence of pre-stroke cognitive decline; pre-stroke mood disorder (depression) and incident/prevalent delirium.

Work package three: Describing and predicting neuropsychological prognosis

- To describe the natural history (rates of outcomes; change over time) of post stroke neuropsychological problems at time points of one month; six month; twelve months and eighteen months.
- To describe univariate and adjusted independent predictors of both post stroke cognitive decline and post stroke mood disorder (odds ratios, with corresponding 95% confidence intervals).
- To develop, calibrate and validate predictive models for post stroke neuropsychological factors.
- To estimate likely recruitment, "event rates" and loss to follow up for future cognitive/mood studies.

6.2 Statistical Analysis Plan

The study will have a comprehensive Statistical Analysis Plan (SAP), which will govern all statistical aspects of the study, and will be authored by the Trial Statistician. Full details of all statistical issues and planned statistical analyses will be specified in the SAP which will be agreed before analyses begin.

6.3 Overview of statistical analysis

6.3.1 WP1: Assessing pre-stroke psychological problems

Accuracy of screening tools will be described in terms of sensitivity; specificity; predictive value; receiver operating space analyses. Index test questionnaires will be compared against each other and against a reference standard of semi-structured clinical assessment. From the acute assessments, we will describe the accuracy of brief screening tests used in isolation and combined with Boolean operators of "OR"/"AND". To describe feasibility we will collate numbers completing each test fully and partially; time and assistance required for completion. To incorporate feasibility into analyses, we will employ an "intention to diagnose" approach, including those unable to complete tests.

6.3.2 WP2: Test accuracy and prognostic utility of brief screening tools

Accuracy of screening tools will be described in terms of sensitivity; specificity; predictive value; receiver operating space analyses. Index tests will be acute and one month assessments and will be compared against each other and reference standard of follow up assessment data. From the acute assessments, we will describe the accuracy of brief screening tests used in isolation and combined with Boolean operators of "OR"/"AND". To describe feasibility we will collate numbers completing each test fully and partially; time and assistance required for completion. To incorporate feasibility into analyses, we will employ an "intention to diagnose" approach, including those unable to complete tests.

6.3.3 Work Package three: Describing and predicting neuropsychological prognosis

For prospective follow up, outcomes of interest are change in scores on cognitive and mood screening tools and incident clinical mood disorder or multi-domain cognitive impairment. Multi-domain tools will be analysed as ordinal data and dichotomised at varying thresholds. Neuropsychological battery data will be transformed into z scores, with impairment defined as greater than 1.5 standard deviations below age and sex based norms. We will collect data on recurrent stroke, complications (falls, seizure, infection) hospitalisation/institutionalisation and death. All data from 6,12,18 month assessments will be assessed by a panel of the senior investigators and a consensus assessment for incident mood disorder and incident cognitive disorder made.

We will use generalized linear models for prospective data to describe associations of baseline characteristics with change across repeated neuropsychological measures. With our statistician we will use varying competing risk survival models to account for events that may precede our neuropsychological outcomes of interest (mortality). We will describe univariate and adjusted independent predictors of "outcomes". We will describe odds-ratios for binary "outcomes" at chosen time-points, using multivariate Poisson regression.
We will explore repeated measures analyses adjusting for baseline covariates and describe temporal change in test scores. We will create prognostic models and if data allow predictive risk scores for the various cognitive and mood outcomes, describing calibration; discrimination and validation using bootstrapping.

6.4 General Considerations

In general we will apply parametric statistical methods; any variable not suitable for parametric analysis will be analysed using non-parametric methods.
Descriptive statistics by study centre will be provided. A summary and listing of patients with protocol violations will be produced.

6.5 Software for Statistical Analysis

All statistical analysis will be performed using SAS version 9.1 or later.

6.6 Sample Size

We anticipate recruiting n=500 participants across the three sites over 18 months recruitment. We expect substantial attrition (death, loss to follow-up, development of cognitive problems that preclude further assessment) and anticipate n=400 one month; n=350 six month; n=300 twelve month and n=200 eighteen month follow up data.

Data to allow sample size calculations for future studies is an intended output of this work. Recognising the uncertainty, we do not offer definitive "power" calculation per se, but our recruitment estimates suggest we will have sufficient patients to achieve our research aims.

Scottish Stroke Care Audit reports over 1500 stroke discharges per annum across our three Glasgow sites. Our pilot data suggest that over 18 month recruitment, at a conservative estimate 500 will be suitable and agree to early assessment and follow up. Based on Information Services Division stroke data, we project estimates of n=400 one month; n=350 six month; n=300 twelve month and n=200 eighteen month follow up data. These numbers make our study equivalent to or larger than other international neuropsychological focused studies. By using research nurses for initial assessments and three full time PhD student assessors for follow up, daily maximum number of assessments per team member would be two.

For the optional study describing accuracy of informant questionnaires we have a separate power calculation. Using a nomogram approach [20] describing test properties of informant questionnaires, based on estimated prevalence of pre-stroke problems of 20% and anticipated specificity of around 0.8, recruiting n=100 gives sufficient power to assess the scales.

WP1 and WP2. Our recruitment is designed mindful of potential attrition. For the test accuracy work, using a nomogram [20] based on prevalence of 40% cognitive impairment at one month, (α=0.05); our estimate of 400 participants would allow description of accuracy across a full range of plausible sensitivity/specificity.

WP3. Based on published data on mood we would anticipate annual rates of outcomes at around 30% with n=125 "outcomes" in survivors at end of follow up (although our data suggests rates of cognitive/mood disorder may be considerably higher in unselected cohorts). This gives sufficient power for the prospective models we have planned. Based on our anticipated recruitment and retention, prognostic models will have power to describe multiple covariates.

The optional subgroup study where results on neuropsychological assessment are compared to clinical assessment will be performed on n=25 in the first instance. This is a pragmatic sample size. Recruitment will be of sequential consenting participants from the Glasgow sites.

6.7 Procedures for Accounting for Missing Data

There will be no imputation of missing data for the primary or secondary endpoints in the first instance. As part of the analyses we will explore the effects of various approaches to handling missing data.

6.8 Procedures for Reporting Deviations from the Original Statistical Plan

A detailed statistical analysis plan (SAP) will be agreed before analyses begin. Any deviations from this plan will be documented and justified in the final study report.

6.9 Selection of Subjects to be Included in the Analyses

We will run analyses including those with full test data and those with missing data, using intention to diagnose approaches.

7 STUDY Closure / DEFINITION OF END OF STUDY

The study will end when the last patient has their last study visit.

8 SOURCE DATA/DOCUMENTS

8.1 Case Report Forms / Electronic Data Record

Primary data collection will use paper based case report form (CRF). Inpatient assessment scores will be shared with the hospital team on request. For out-patient/community assessments, screening test summary results will not be shared with the General Practitioner (GP). This approach was suggested by the Scotland A Research Ethics Committee and recognises that the screening tests are not diagnostic If assessment suggests a serious cognitive or mood disorder that requires urgent treatment results will be shared with the appropriate team.

All participant data will be identified by the participant study identification number. CRF data will be securely transferred to the Robertson Centre for Biostatistics (RCB) for electronic entry. Data will be validated at the point of entry into and at regular intervals during the study. Data discrepancies will be flagged to the study site by the statistician and any data changes will be recorded in order to maintain a complete audit trail (reason for change, date change made, who made change).

8.1.1 Data Handling and Record Keeping

All CRF data will be held in the RCB. The RCB manages all studies to the highest standards in accordance with its internal Standard Operating Procedures, ICH Good Clinical Practice, the European Union Clinical Trials Directive 2001/20/EC, the ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trials E9 and all other industry legal and regulatory guidelines. It has extensive experience of managing data in the context of privacy and data protection legislature, including the Data Protection Act 1998 and EU Data Protection Directive 95/46/EC. The Centre is certified for ISO 9001:2008 for its quality systems, has TickIT accreditation for its software development and is BS7799 compliant.

Only the study investigators will have access to participant identifiable data. We will permit trial-related monitoring, audits and regulatory inspections and will provide direct access to source data and documents.

8.1.3 Data Security

The RCB systems are fully validated in accordance with industry and regulatory standards, and incorporate controlled access security. High volume servers are firewall protected and preventative system maintenance policies are in place to ensure no loss of service. Web servers are secured by digital certificates. Data integrity is assured by strictly controlled procedures, including secure data transfer procedures.

8.1.4 Database Software

Data will be stored in MS SQL Server.

8.1.5 Record Retention

To enable evaluations and/or audits from regulatory authorities, the investigator agrees to keep records, including the identity of all participating subjects (sufficient information to link records), all original signed informed consent forms, source document in accordance with ICH GCP, local regulations, or as specified in the Clinical Study Agreement, whichever is longer. Data will be retained at the Data Centre for a minimum of 10 years.

8.1.6 Archiving

CRF data will be stored by the Robertson Centre for Biostatistics for 10 years after completion of the study.

9 STUDY MANAGEMENT

The trial management teams will be in place before recruitment begins.

9.1 Routine Management of Study

The study will be co-ordinated from the Glasgow Royal Infirmary, Glasgow by the PI. The study will be subject to review at any time by the West Glasgow Local Research Ethics Committee.

9.2 Trial Management Committee (TMC)

There will be no DSMC for this observational trial. Independent oversite will be provided by the study advisory group.

9.3 Data Safety Monitoring Committee (DSMC)

There will be no DSMC for this observational trial.

10 STUDY MONITORING AND AUDITING

Study monitoring visits will be conducted according to a study-specific monitoring plan devised by NHS Greater Glasgow and Clyde and subsequent monitoring reports will be reviewed by NHS Greater Glasgow and Clyde. The Sponsor, NHS Greater Glasgow and Clyde, audit a randomly selected 10% of studies conducted under the Research Governance Framework per annum, as well as those identified using a risk assessment tool as specifically requiring assessment. Investigators and site staff will notified in advance of any audit and/or monitoring visits.

11 PROTOCOL AMENDMENTS

Any change in the study protocol will require an amendment. Any proposed protocol amendments will be initiated by the Chief Investigator and any required amendment forms will be submitted to the regulatory authority, ethics committee and sponsor. The Sponsor will determine whether an amendment is non-substantial or substantial. All amended versions of the protocol will be signed by the Chief Investigator and sponsor representative. Before the amended protocol can be implemented (or sent to other participating sites) favourable opinion/approval must be sought from the original reviewing REC and Greater Glasgow & Clyde Health Board Research and Development (R & D) office. The Chief Investigator will sign any amended versions of the protocol. All protocol versions and their amendments must be notified to the study team and to the data centre.

12 ETHICAL CONSIDERATIONS

12.1 Ethical Conduct of Study

Study will be carried on accordance with the World Medical Association Declaration of Helsinki (1964) and it revisions (Tokyo (1975), Venice (1983), Hong Kong (1989), South Africa (1996), Edinburgh (2000), Seoul (2008) and Fortaleza (2013).

There are no special ethical considerations pertaining to this study. Favourable ethical opinion will be sought before patients are entered into this study. The Chief Investigator will update the ethics committee of any new information related to the study.

12.2 Informed Consent

The clinical team will assess study participant's ability to provide informed consent. Where possible we will obtain written informed consent from both study patient and informant.

Where a patient is unable to provide informed consent but clinical team are still happy for the person to participate in the study, informed consent will be sought from a suitable proxy. Choice of proxy will be guided by patient preference expressed at time of assessment or expressed pre-stroke.

The research nurse or trained member of the research team will explain the exact nature of the study in writing, provide patient and carer information sheets, and verbal information. Study participants will be informed that they are free to withdraw their consent from the study or study treatment at any time.

13 INSURANCE AND INDEMNITY

The study is sponsored by NHS Greater Glasgow and Clyde. The sponsors will be liable for negligent harm caused by the design of the trial. NHS Indemnity is provided under the Clinical Negligence and Other Risks Indemnity Scheme (CNORIS). As the substantive employer of the CI, The University of Glasgow also has insurance with Newline. It will be confirmed prior to the study starting that insurance cover will be provided automatically under the current policy. The insurance cover will be subject to NHS indemnity being in place and Ethics Committee approval being obtained.

The NHS has a duty of care to patients treated, whether or not the patient is taking part in a clinical trial, and the NHS remains liable for clinical negligence and other negligent harm to patients under this duty of care.

As this is a clinician-led study there are no arrangements for no-fault compensation.

14 FUNDING

The study is funded by a Chief Scientist Office / Stroke Association Programme grant.

15 ANNUAL REPORTS

The funders mandate progress report and outputs to be submitted electronically via the Researchfish resource; these will be updated in real time and reviewed annually. Annual reports will be submitted to the ethics committee, regulatory authority and sponsor with the first submitted one year after the date that all trial related approvals are in place.

16 Dissemination of Findings

Study results will be submitted to an International Conference and will be submitted for publication in a peer review journal. No personal data will be used when publishing the results. A lay summary and other material as appropriate will be offered to those participants who wish to receive it. Participants will be asked at their last study visit if they are happy to be contacted and the preferred method for contact. These data will be held securely in the CRF in a password protected file that is separate from the main study archive.

18 REFERENCES

1. Pollock A, St George B, Fenton M, Firkins L. Top ten research priorities relating to life after stroke. Lancet Neurol.2012,11:209.
2. Research Gaps and priorities Around Psychological Consequences of Stroke. Stroke Association 2014 http://www.stroke.org.uk/research/priority-programme-awards last accessed Jan 2015.
3. Lees RA Fearon P, Harrison JK, Broomfield NM, Quinn TJ. Cognitive and mood research in stroke research: focussed review of contemporary studies. Stroke. 2012; 43:1678-80.
4. Lees RA, Broomfield NM, Quinn TJ. Questionnaire assessment of usual practice in mood and cognitive assessment in Scottish stroke units. Disabil Rehabil. 2014;36:339-343.
5. Lees R, Selvarajah J, Fenton C et al. Test accuracy of cognitive screening tests for diagnosis of dementia and multidomain cognitive impairment in stroke. Stroke. 2014;45:3008-3018
6. Noel-Storr AH, McCleery JM, Richard E, et al. Reporting standards for studies of diagnostic test accuracy in dementia: the STARDdem initiative. Neurology.2104;83:364-373.
7. Lees R, Corbet S, Johnston C, Moffitt E, Shaw G, Quinn TJ. Test accuracy of short screening tests for diagnosis of delirium or cognitive impairment in acute stroke unit. Stroke. 2013;44:3078-83.
8. Lees R, Stott DJ, Quinn TJ, Bromfield NM. Feasibility and diagnostic accuracy of early mood screening to diagnose persisting clinical depression/anxiety disorder after stroke. Cerebrovasc Dis. 2014; 37:323-9.
9. Brainin, M., Tuomilehto, J., Heiss, W.-D., Bornstein, N. M., Bath, P. M. W., Teuschl, Y., Richard, E., Guekht, A., Quinn, TJ. and the Post Stroke Cognition Study Group. Post-stroke cognitive decline: an update and perspectives for clinical research. Euro J Neurol. doi: 10.1111/ene.12626
10. Appelros P, Nydevik I, Seiger A, Terent A.Predictors of severe stroke: influence of preexisting dementia and cardiac disorders.Stroke.2002,33:2357–2362.
11. Harrison JK, Fearon P, Noel-Storr AH, McShane R, Stott DJ, Quinn TJ. IQCODE for diagnosis dementia in secondary care settings. Cochrane Database Systematic Reviews. 2015; 3:CD010772.
12. McGovern A, Pendlebury ST, Mishra N, Fan Y, Quinn TJ. Test Accuracy of Informant-Based Cognitive Screening Tests for Diagnosis of Dementia and Multidomain Cognitive Impairment in Stroke. Stroke. 2016 published online ahead of print.
13. Schuetz Georg M, Schlattmann Peter, Dewey Marc. Use of 3×2 tables with an intention to diagnose approach to assess clinical performance of diagnostic tests: meta-analytical evaluation of coronary CT angiography studiesBMJ 2012; 345:e6717
14. Pendlebury ST, Rothwell PM. Prevalence, incidence, and factors associated with pre-stroke and post-stroke dementia: a systematic review and meta-analysis. Lancet Neurol.2009,8:1006-18
15. Hachinski V, Iadecola C, Petersen RC et al National Institute Neurological Disorders and Canadian Stroke Network vascular cognitive impairment harmonization standards.Stroke.2006;37:2220-41.
16. Chandler MJ, Lacritz LH, Hynan et al A total score for the CERAD neuropsychological battery. Neurology.2005;65:102-6
17. Michael BF, Williams JBW, Spitzer RL, Gibbon M. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Clinical Trials Version (SCID-CT). New York: Biometrics Research, New York State Psychiatric Institute, 2007.
18. Hemingway H, Croft P, Perel P, et al. Prognosis research strategy (PROGRESS) 1: framework for researching clinical outcomes. BMJ.2013;346:e5595
19. Quinn TJ, Lees KR, Hardemark HG, Dawson J. Walters MR. Initial experience of a digital training resource for modified Rankin scale assessment in clinical trials. Stroke.2007;38:2257-61.
20. Carley S, Dosman S, jones SR, Harrison M. Simple nomogrmas to calculate sample size in diagnostic studies. Emerg Med J 2005; 22:180-1.

Appendix 10: APPLE Patient Information Sheet

Assessing for memory, thinking & mood problems following a stroke

Information sheet and consent form for potential participants

Contact information: If you wish any further information about the study please contact Dr Terry Quinn, Email:terry.quinn@glasgow.ac.uk, Tel:0141-201-8510

Title of project: Improving our assessment and understanding of the short, medium and longer term neuropsychological consequences of stroke

Assessing Post-stroke Psychology Longitudinal Evaluation – the APPLE study

You are being invited to take part in a clinical research study. The study will look at memory, thinking and mood changes that can happen after a stroke. We are particularly interested in how well questionnaires and pencil and paper tests detect any problems. You will be asked to complete some questionnaires and some pencil and paper tests looking at memory, thinking and mood. If you agree, as an optional study, we will also ask someone that knows you well about any memory, thinking, mood problems that they may have noticed.

Before you make a decision, it is important that you fully understand why the research is being done and what will be involved. This study is part of a program of work supported by the Stroke Association. This project is part of the PhD work for three students based at the University of Glasgow.

This study is part of a program of work supported by the Stroke Association. This project is part of the PhD work for three students based at the University of Glasgow.

Please take time to read the following information carefully and discuss it with others if you wish. Please ask me if there is anything that is not clear or if you would like more information.

Thank you for taking the time to read and consider this.

Summary

The aim of the study is to describe changes in memory, thinking and mood that occur after a stroke. We are not testing a new drug or method of assessment and there will be no change to the usual clinical care. We provide a summary of the research in this section, with more detail later in this information sheet.

Participants will be recruited at the time of their stroke. The clinical team will identify suitable patients and if they agree, then we will approach them with details of the study.

The study has a number of components and potential participants can choose to take part in all, some or none of these. We will check that the person who has had a stroke and their family member/friend are both agreeable to the study and which parts they wish to assist with. An optional part of the study involves asking a friend, family member or carer about their perception of the participant's memory, thinking and mood.

If you agree to take part you will be asked some brief questions looking at memory, thinking and mood. If agreeable, your family member/friend will also be asked to complete

questionnaires on your mood and memory. As an optional step a member of the research team will then interview you and your family member/friend together. They will use a structured questionnaire. The purpose of this more detailed interview is to assess whether you had memory, thinking, mood problems before the stroke.

At around one month later, a research nurse or member of the research team will assess you again. They will use short questionnaires and pencil and paper tests to describe memory, thinking, mood and any complications from stroke that have happened. At six, twelve and eighteen months you will be asked about their recovery, their quality of life and will be reassessed for memory, thinking and mood problems using a more detailed questionnaire. Your family member/friend will be asked about your recovery, their perception of your memory, thinking and mood and whether they feel under strain as a carer.
If the participant finds the assessments too long or tiring; a shorter assessment is available. If the participant prefers a telephone assessment rather than face to face interview is available. Completing all the study assessments is not mandatory. You can choose to participate in all the assessments or chose to only help with some of them.

The assessments performed at one, six, twelve and eighteen months are not diagnostic. A proportion of participants will be asked if they can be assessed by a trained psychologist who will try and make a clinical assessment of whether the person has important cognitive or mood problems.

If you agree, you will have some blood taken and will be asked to provide a urine sample at the first and subsequent assessments. This step is not mandatory. If you agree the information collected will be stored and used for future studies. Some of these studies may involve linking the research information to other sources of information such as hospital records, x-ray files or national records of admissions. All the information held will be anonymous.

What is the purpose of study?
After a stroke many people develop changes in their memory, thinking and mood. Our study is looking at questionnaires and pencil and paper based tests of thinking. We wish to see if these tests can be used to detect problems in patients who have had a stroke. This is important as early detection of memory and thinking problems may allow treatment of these problems. We also wish to describe how memory, thinking and mood change over time.

Why have I been chosen?
We wish to include a range of people who have had stroke. We are particularly interested in whether certain pencil and paper or questionnaire tests are suitable for use in acute stroke units. We are approaching patients with stroke if they wish to participate. Your hospital consultant and team are aware of the study and have suggested that you may be suitable. The final decision on participation is up to you.

Do I have to take part?
No. It is up to you to decide whether or not you take part. If you do decide to take part you will be given this information sheet to keep and will be asked to sign a consent form.

You are free to withdraw at any time and do not have to give a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you will receive. You can choose to participate in all aspects of the project or only part of the project.
The researchers are not involved with your general care and will not be involved with treatment of any memory, thinking or mood problems. The initial results of the tests of memory, thinking and mood can be shared with the hospital team looking after you.

What will happen if I take part?
The project involves different steps. You can choose to participate in all aspects or to restrict your participation to certain areas only.

Main study: A researcher will ask you to complete a series of pencil and paper tests of memory, thinking and mood. We want to assess how easy it is for people with stroke to complete the tests, so it is fine if you are unable to complete all or any of the tests. The tests should take around 20 minutes to complete and the researcher can help. As part of the initial assessment the research team may access the clinical notes, laboratory reports or x-rays relating to your admission. We will also ask a relative or someone who knows you well to comment on any memory, thinking or mood problems that they may have noticed.
As an optional study, within one month of the first assessment, a psychology graduate studying for a PhD will chat to you and your relative about memory, thinking and mood. This interview will be at a time and date that suits you both. The total interview should last no more than 60 minutes and can be in the hospital or at another place that suits you.
The nurse or psychology student will arrange to see you both again at around four weeks, six months, twelve months and eighteen months after the first interview. They will complete some other pencil and paper tests around memory, thinking and mood with you and your relative. These assessments should last less than 60 minutes. Assessments will be at a date and time that suits you and can be in the hospital (we will pay for transport), in your house or over the telephone. You can choose to help with all of these assessments or only a selection.

Additional study A: After this study is complete, we wish to keep the results of the (anonymised) questionnaires and pencil and paper tests so that we can use these, with other researchers, to answer new questions on memory, thinking and mood problems after stroke. We will use your hospital identifier (CHI number) to link the research results to other electronic databases, for example the other hospital records, the national record of hospital admissions or national record of drug prescribing. Any matching of information between databases will be performed in such a way that participants' data will be kept anonymous. The information will be held securely in the Robertson Centre for Biostatistics, Glasgow. As we wish to look at how problems now influence future health, we have set no time limit on how long we will keep the information for.

Additional study B: While in the acute stroke unit, and at each study visit, a research nurse will ask if a blood and/or urine sample can be taken. This will involve one needle for the blood sample at each visit. This step is optional and at any visit you can decline the blood sample. We will store the blood and urine to allow for future studies looking at ill-health in stroke. Storage will be in the secure facilities of the NHS Greater Glasgow and Clyde biobank. The samples will allow us to look at blood cells and other molecules in

blood and proteins in urine. We may use the materials to look at genetic factors. The genetic studies may involve looking at genes associated with certain disease or looking at all the genes. We would not use samples for stem cell work.

What do I have to do?
If you agree to participate you will be asked to sign a consent form. We are also interested in the thoughts and assessment of a family member/friend on your memory, thinking and mood. If you agree, we would also like them to complete some short questionnaires. There is a separate information sheet and consent form describing this study. You will have at least a day to decide if you want to participate.

What are the possible benefits to me from taking part?
There are no direct benefits to you or your family member/friend from taking part. Participants in the study will get a detailed assessment that can be shared with the clinical team. By taking part you will help us decide on the best way to test for memory, thinking and mood problems after stroke. The study will also help us understand how memory, thinking and mood change over time.

What are the possible disadvantages and risks from part?
From being part of the study you will get more detailed, and longer, assessments of memory, thinking and mood than would happen in standard care.

What if something goes wrong?
If you have a concern about any aspect of this study please contact the research team (details at end of leaflet) who will do their best to answer your questions.

You have the right to withdraw from assessment at any time without providing a reason and with no impact to the care you receive. If you are unhappy about any aspect of the study and wish to make a formal complaint the normal NHS complaints mechanism is also available to you.

What happens when the study is finished?
We will collate all the information from participants and look to see which tests are best at picking up memory, thinking and mood problems and how these issues change over time. The information collected will be securely stored for an indefinite time. Other researchers may access anonymised information to answer new research questions. The blood samples taken will be stored securely in the NHS Greater Glasgow & Clyde Biobank facility for ten years. At the end of the study, if you agree, we will send you the results of the various research projects.

What will happen to the results of the study?
The results of the tests may be shared with the clinical team working within the stroke unit and with your GP. We hope to publish the final results of our study in a scientific journal and discuss the results at professional meetings. Personal details will not be available in any of these materials. If you are interested in the results when the study is complete, details can be sent to you.

Who will see my information? (confidentiality)
All the information we collect during the course of the research will be kept confidential and there are strict laws which safeguard the privacy of the patient at every stage. Initial

scores on the questionnaires and pencil and paper tests may be shared you're your hospital stroke team. If during testing we detect any new diagnoses, we will share this with your hospital team. All information collected by the research team will be anonymised and stored in a secure way. The information will be held securely in the Robertson Centre for Biostatistics, Glasgow. If you agree we will keep the results of the anonymised questionnaires to allow us and other researchers to use them in future projects. The information collected as part of the study may be looked at by representatives of the study Sponsor, NHS GG&C, for audit purposes.
Part of our questions on mood includes asking about low mood (depression). If we suspect severe depression or suicidal thoughts, questioning will be stopped and the treating physician contacted immediately. In this case you may also be referred to appropriate specialised help.

Some additional information on how we use your information
NHS Greater Glasgow and Clyde is the sponsor for this study based in UK. We will be using information from and/or your medical records in order to undertake this study and will act as the data controller for this study. This means that we are responsible for looking after your information and using it properly. NHS Greater Glasgow and Clyde will keep identifiable information about you for 10 years after the study has finished.

Your rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. If you withdraw from the study, we will keep the information about you that we have already obtained. To safeguard your rights, we will use the minimum personally-identifiable information possible.

The NHS hospitals taking part in the study will collect information from you and/or your medical records for this research study in accordance with our instructions. The NHS hospitals will use your name and contact details to contact you about the research study, and make sure that relevant information about the study is recorded for your care, and to oversee the quality of the study. Individuals from NHS Greater Glasgow and Clyde and regulatory organisations may look at your medical and research records to check the accuracy of the research study. The participating NHS hospitals will pass these details to NHS Greater Glasgow and Clyde/University of Glasgow along with the information collected from you and/or your medical records. The only people in NHS Greater Glasgow and Clyde/University of Glasgow who will have access to information that identifies you will be people who need to contact you about research or audit the data collection process. The people who analyse the information will not be able to identify you and will not be able to find out your name or contact details. The participating NHS hospitals will keep identifiable information about you from this study for 10 years after the study has finished.

When you agree to take part in a research study, the information about your health and care may be provided to researchers running other research studies in this organisation and in other organisations. These organisations may be universities, NHS organisations or companies involved in health and care research in this country or abroad. Your information will only be used by organisations and researchers to conduct research in accordance with the UK Policy Framework for Health and Social Care Research.

Your information could be used for research in any aspect of health or care, and could be combined with information about you from other sources held by researchers, the NHS or government.

Where this information could identify you, the information will be held securely with strict arrangements about who can access the information. The information will only be used for the purpose of health and care research, or to contact you about future opportunities to participate in research. It will not be used to make decisions about future services available to you, such as insurance.

Where there is a risk that you can be identified your data will only be used in research that has been independently reviewed by an ethics committee.

You can find out more about how we use your information by contacting the Principal Investigator: terry.quinn@glasgow.ac.uk

Who is organising and funding the study?
This study is being organised by the Institute of Cardiovascular and Medical Sciences, University of Glasgow. The study is funded by the Stroke Association. The researchers will receive no remuneration for including you in the study.

Who has reviewed this study?
This study has been reviewed and approved by Scotland A Research Ethics Committee. All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee. A favourable ethical opinion has been obtained from Scotland A REC. NHS management approval has also been obtained.

SUMMARY
If you agree to participate you will be asked to complete some tests assessing memory, thinking skills and mood. Tests will be performed during this admission and at four future visits.

Name of Lead Researcher
Dr Terry Quinn, Senior Clinical Lecturer, University of Glasgow.

Name of sponsor	**Name of Funder**
NHS Greater Glasgow and Clyde	**The Stroke Association**

Appendix 11: APPLE Ethics and local R&D approval letters

Scotland A Research Ethics Committee

Research Ethics Service
2nd Floor Waverley Gate
2-4 Waterloo Place
Edinburgh
EH1 3EG
Telephone: 0131 465 5680
www.hra.nhs.uk

Scotland A REC
2nd Floor Waverley Gate
2 - 4 Waterloo Place
Edinburgh
EH1 3EG
Tel: 0131-465-5679

28 July 2016

Dr Terence J Quinn
Room 2.44 New Lister Building
Institute of Cardiovascular and Medical Sciences
Glasgow Royal Infirmary, Alexandra Parade
G4 0SF

Dear Dr Quinn

Study title:	**Improving our assessment and understanding of the short, medium and longer term neuropsychological consequences of stroke**
REC reference:	**16/SS/0105**
Protocol number:	**GN14NE496**
IRAS project ID:	**199099**

Thank you for your letter, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to make a request to postpone publication, please contact the REC Manager, Miss Manx Neill, manx.neill@nhslothian.scot.nhs.uk.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Adults with Incapacity (Scotland) Act 2000

I confirm that the Committee has approved this research project for the purposes of the Adults with Incapacity (Scotland) Act 2000. The Committee is satisfied that the requirements of section 51 of the Act will

Scotland A Research Ethics Committee

Research Ethics Service
2nd Floor Waverley Gate
2-4 Waterloo Place
Edinburgh
EH1 3EG
Telephone: 0131 465 5680
www.hra.nhs.uk

Scotland A REC
2nd Floor Waverley Gate
2 - 4 Waterloo Place
Edinburgh
EH1 3EG
Tel: 0131 465 5678

26 July 2017

Dr Terence J Quinn
Room 2.44 New Lister Building
Institute of Cardiovascular and Medical Sciences
Glasgow Royal Infirmary, Alexandra Parade
Glasgow, G4 0SF

Dear Dr Quinn,

Study title:	**Improving our assessment and understanding of the short, medium and longer term neuropsychological consequences of stroke**
REC reference:	**16/SS/0105**
Protocol number:	**GN14NE496**
Amendment number:	**AM02 (REC Ref 16/SS/0105/AM02)**
Amendment date:	**09 June 2017**
IRAS project ID:	**199099**

The above amendment was reviewed on 14 July 2017 by the Sub-Committee in correspondence.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

The Committee had no ethical concerns regarding this amendment.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Notice of Substantial Amendment (non-CTIMP)		09 June 2017
Other [Bogna Drozdowska Curriculum Vitae]		
Other [Emma Elliot Curriculum Vitae]		
Participant information sheet (PIS) [APPLE PIS Assent]	1.5	22 May 2017

Coordinator/administrator:Maureen Travers
Telephone Number: 0141 232 1813
E-Mail: Maureen.Travers@ggc.scot.nhs.uk
website www.nhsggc.org.uk/r&d

Clinical Research & Development
West Glasgow ACH
Dalnair Street
Glasgow G3 8SJ
Scotland, UK

02/08/2016

Dr Terence Quinn
University of Glasgow
Institute of Cardiovascular and Medical Sciences
New Lister Building
Glasgow Royal Infirmary
Glasgow
G4 0SF
Scotland

NHS GG&C Board Approval

Dear Dr Terence Quinn

Study Title:	Improving our assessment and understanding of the short, medium and longer term neuropsychological consequences of stroke
Principal Investigator:	Dr Terence Quinn
GG&C HB site	Glasgow Royal Infirmary
Sponsor	NHS Greater Glasgow and Clyde
R&D reference:	GN14NE496
REC reference:	
Protocol no: (including version and date)	version 1.2 (05.07.2016)

I am pleased to confirm that Greater Glasgow & Clyde Health Board is now able to grant **Approval** for the above study.

Conditions of Approval

1. **For Clinical Trials** as defined by the Medicines for Human Use Clinical Trial Regulations, 2004
   a. During the life span of the study GGHB requires the following information relating to this site
      i. Notification of any potential serious breaches.
      ii. Notification of any regulatory inspections.

It is your responsibility to ensure that all staff involved in the study at this site have the appropriate GCP training according to the GGHB GCP policy (www.nhsggc.org.uk/content/default.asp?page=s1411), evidence of such training to be filed in the site file.

2. **For all studies** the following information is required during their lifespan.
   a. Recruitment Numbers on a quarterly basis
   b. Any change of staff named on the original SSI form

Appendix 12: Case report form pages

APPLE Study
Protocol Version 1.5
Version 4.0 (26 Feb 2018)

Baseline Participant Assessments
AMT-plus

Site Number	Participant Number	Initials	Date of Visit
01	02	03	D D M M Y Y Y Y 04

	Score 1 Point			Untestable - Reason *(enter appropriate number from table on page 13)*	Number of mistakes
1. Age		Yes ☐ 1	No ☐ 2 05	UN ⌊⌋ 06 If other (9), specify: _____ 07	N/A
2. Time		Yes ☐ 1	No ☐ 2 08	UN ⌊⌋ 09 If other (9), specify: _____ 10	N/A
3. Date	Day Yes ☐ 1 No ☐ 2 11	Month Yes ☐ 1 No ☐ 2 12	Year Yes ☐ 1 No ☐ 2 13	UN ⌊⌋ 14 If other (9), specify: _____ 15	N/A
4. Place		Place Yes ☐ 1 No ☐ 2 16	City Yes ☐ 1 No ☐ 2 17	UN ⌊⌋ 18 If other (9), specify: _____ 19	N/A
5. Two person recognition		Yes ☐ 1	No ☐ 2 20	UN ⌊⌋ 21 If other (9), specify: _____ 22	N/A
Give recall items at this point: face, velvet, church, daisy, red					N/A
6. Date of birth		Yes ☐ 1	No ☐ 2 23	UN ⌊⌋ 24 If other (9), specify: _____ 25	N/A
7. World War 1		Yes ☐ 1	No ☐ 2 26	UN ⌊⌋ 27 If other (9), specify: _____ 28	N/A
8. Prime Minister		Yes ☐ 1	No ☐ 2 29	UN ⌊⌋ 30 If other (9), specify: _____ 31	N/A
9. Count 20-1		Yes ☐ 1	No ☐ 2 32	UN ⌊⌋ 33 If other (9), specify: _____ 34	⌊⌋ 35
10. Recall (items)		Yes ☐ 1	No ☐ 2 36	UN ⌊⌋ 37 If other (9), specify: _____ 38	⌊⌋ 39
11. Clock Draw	Face Yes ☐ 1 No ☐ 2 40	Numbers Yes ☐ 1 No ☐ 2 41	Hands Yes ☐ 1 No ☐ 2 42	UN ⌊⌋ 43 If other (9), specify: _____ 44	N/A
12. News item		Yes ☐ 1	No ☐ 2 45	UN ⌊⌋ 46 If other (9), specify: _____ 47	N/A
13. Months backwards		Yes ☐ 1	No ☐ 2 48	UN ⌊⌋ 49 If other (9), specify: _____ 50	⌊⌋ 51
14. One letter fluency		Yes ☐ 1	No ☐ 2 52	UN ⌊⌋ 53 If other (9), specify: _____ 54	**Number of words** ⌊⌋ 55

1. Time taken to complete ⌊⌋ 56 seconds
2. Verbal assistance required to complete Yes ☐ 1 No ☐ 2 57
3. Hands on assistance required to complete Yes ☐ 1 No ☐ 2 58

APPLE Study
Protocol Version 1.3
Version 3.0 (22 Jun 2017)

Baseline Participant Assessments
Confusion Assessment Method for Intensive Care (CAM-ICU)

Site Number	Participant Number	Initials	Date of Visit
01	02	03	D D M M Y Y Y Y 04

1. Untestable - Reason ⌊_⌋ 05 If other (9), specify: _____ 06
(enter appropriate number fromtable on page 13)

Feature 1: Acute Onset or Fluctuating Course	
Is the patient different than his/her baseline mental status? OR has the patient had any fluctuation in mental status in the past 24 hours as evidenced by fluctuation on a sedation scale (i.e. RASS), GCS, or previous delirium assessment?	Either Question Yes Yes ☐ 1 No ☐ 2 07
Feature 2: Inattention	
Letters Attention Test (see training manual for alternate **Pictures**) Directions: Say to the patient, "I am going to read you a series of 10 letters. Whenever you hear the letter 'A', indicate by squeezing my hand." Read letters from the following letter list in a normal tone 3 seconds apart. **SAVEAHAART** **Errors are counted when patient fails to squeeze on the letter "A" and when the patient squeezes on any letter other than "A".**	Number of Errors > 2 Yes ☐ 1 No ☐ 2 08
Feature 3: Altered Level of Consciousness	RASS anything other than zero
Present if the Actual RASS score is anything other than alert and calm (zero)	Yes ☐ 1 No ☐ 2 09
Feature 4: Disorganized Thinking	
Yes/No Questions (see training manual for alternate **set of questions)** 1. Will a stone float on water? 2. Are there fish in the sea? 3. Does one pound weigh more than two pounds? 4. Can you use a hammer to pound a nail? **Errors are counted when patient incorrectly answers a question.** **Command** Say to patient: "Hold up this many fingers" (Hold 2 fingers in front of patient) "Now do the same thing with the other hand" (Do not repeat number of fingers) *if pt is unable to move both arms for 2nd part of command ask patient to "Add one more finger" **An error is countedif patient is unable to complete the entire command.**	Combined number of errors > 1 Yes ☐ 1 No ☐ 2 10

Overall CAM-ICU **Feature 1 plus 2 and either 3 or 4 present = CAM-ICU positive**	Criteria Met	☐ 1 **CAM-ICU Positive** (Delerium Present)
	Criteria Not Met	☐ 0 **CAM-ICU Negative** (No Delerium) 11

Site Number	Participant Number	Initials	Date of Visit
01	02	03	D D M M Y Y Y Y 04

PRE-STROKE FUNCTION

1. Bowels
   - Incontinent (or needs to be given enemata) ☐ 0
   - Occasional accident (once a week) ☐ 1
   - Continent ☐ 2 (05)

2. Bladder
   - Incontinent or catheter ☐ 0
   - Occasional accident (maximum once per 24 hours) ☐ 1
   - Continent ☐ 2 (06)

3. Grooming
   - Needs help ☐ 0
   - Independent (face, hair, teeth, shaving) ☐ 1 (07)

4. Toilet use
   - Dependent ☐ 0
   - Needs some help, but can do something alone ☐ 1
   - Independent (on and off, dressing, wiping) ☐ 2 (08)

5. Feeding
   - Unable ☐ 0
   - Needs help cutting, spreading butter, etc ☐ 1
   - Independent ☐ 2 (09)

6. Transfer (bed to chair and back)
   - Unable, no sitting balance ☐ 0
   - Major help (one or two people, physical), can sit ☐ 1
   - Minor help (verbal or physical) ☐ 2
   - Independent ☐ 3 (10)

7. Mobility
   - Immobile ☐ 0
   - Wheelchair independent, including corners ☐ 1
   - Walks with help of one person (verbal or physical) ☐ 2
   - Independent (but may use aid; for example, stick) ☐ 3 (11)

8. Dressing
   - Dependent ☐ 0
   - Needs help but can do about half unaided ☐ 1
   - Independent (including buttons, zips, laces, etc) ☐ 2 (12)

9. Stairs
   - Unable ☐ 0
   - Needs help (verbal, physical, carrying aid) ☐ 1
   - Independent ☐ 2 (13)

10. Bathing
    - Dependent ☐ 0
    - Independent (or in shower) ☐ 1 (14)

Site Number	Participant Number	Initials	Date of Visit
01	02	03	D D M M Y Y Y Y 04

PRE-STROKE FUNCTION

No symptoms at all	☐ 0
No significant disability despite symptoms; able to carry out all usual duties and activities	☐ 1
Slight disability; unable to carry out all previous activities; but able to look after own affairs without assistance	☐ 2
Moderate disability; requires some help, but able to walk without assistance	☐ 3
Moderately severe disability; unable to walk without assistance, unable to attend to own bodily needs without assistance	☐ 4
Severe disability; bedridden, incontinent, requiring constant nursing care	☐ 5 05

Site Number	Participant Number	Initials	Date of Visit
01	02	03	D D M M Y Y Y Y 04

PRE-STROKE FUNCTION

1. Can you use the telephone:
   - Without help, including looking up numbers and dialing. ☐ 2
   - With some help (can answer phone or dial operator in emergency but need a special phone or help in getting the number or dialling). ☐ 1
   - Completely unable to answer the telephone. ☐ 0 05

2. Can you get to places out of walking distance:
   - Without help (i.e. drive your own car, travel alone on buses or taxis). ☐ 2
   - With some help (need someone to help you or go with you when travelling). ☐ 1
   - Unable to travel unless emergency arrangements are made for a specialized vehicle like an ambulance. ☐ 0 06

3. Can you go shopping for groceries or clothes (assuming has transportation):
   - Without help (taking care of all shopping needs yourself). ☐ 2
   - With some help (need someone to go with you on shopping trips). ☐ 1
   - Completely unable to do any shopping. ☐ 0 07

4. Can you prepare your own meals:
   - Without help (plan and cook full meals yourself). ☐ 2
   - With some help (can prepare some things but unable to cook full meals yourself). ☐ 1
   - Completely unable to prepare any meals. ☐ 0 08

5. Can you do your housework:
   - Without help (can clean floors etc.). ☐ 2
   - With some help (can do light housework but need help with heavy work). ☐ 1
   - Completely unable to do any housework. ☐ 0 09

6. Can you take your own medicine
   - Without help (in the right doses at the right time). ☐ 2
   - With some help (able to take medicine if someone prepares it for you, reminds you to take it). ☐ 1
   - Completely unable to take medicines. ☐ 0 10

7. Can you handle your own money
   - Without help (write cheques, pay bills etc). ☐ 2
   - With some help (manage day to day buying but needs help with managing chequebook and paying bills etc.). ☐ 1
   - Completely unable to handle money. ☐ 0 11

Appendix 13: The Oxford Cognitive Screen

OCS

Oxford Cognitive Screen

ID: _____

Date: __ / __ / __

1. Picture Naming

Present the drawings from the Test Booklet (4 pages)

1.	hippo / hippopotamus	☐
2.	melon / watermelon	☐
3.	filing cabinet - chest of drawers	☐
4.	pear	☐

Total: /4

2. Semantics / Picture Pointing

Present the drawings from the Test Booklet (1 page)

1.	Point to the fruit	☐
2.	Point to the animal	☐
3.	Point to the tool	☐

Total: /3

3. Orientation

MCQ (in Test Booklet) if expressive problems or after errors (no penalties for mcq)

		FR	MCQ
1.	City / town	☐	☐
2.	Part of the day	☐	☐
3.	Month	☐	☐
4.	Year	☐	☐

Total: /4

4. Visual Field Test

Patient asked to look at examiner's nose and point to the moving hand

	LEFT	RIGHT
Upper quadrant	☐	☐
Lower quadrant	☐	☐
Total:	/2	/2

Total: /4

5. Sentence Reading

See Test Booklet for printed sentence, tick underneath words when read correctly. After reading (attempt): read the correct sentence out loud to the participant. ***Please try to remember this sentence as I will ask you about it later on.***

Have any of the islands got a quay, thought the colonel sitting on his yacht.

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Total: /15

6. Number Writing

Read out numbers for writing on lined page

708	☐
15 200	☐
400	☐

Total: /3

Calculations

See Test Booklet. Use MCQ if expressive problems

	FR	MCQ
6 + 3 = 9	☐	☐
7 + 9 = 16	☐	☐
8 – 5 = 3	☐	☐
36 – 17 = 19	☐	☐

Total: /4

7. Broken Hearts Test

Explain task and practice up to 2 times, then allow a ***MAXIMUM of 3 min*** *for full page task*

Box 1	Box 3	Box 5	Box 7	Box 9
/5 /5 /5	/5 /5 /5	/5 /5 /5	/5 /5 /5	/5 /5 /5
Box 2	**Box 4**	**Box 6**	**Box 8**	**Box 10**
/5 /5 /5	/5 /5 /5	/5 /5 /5	/5 /5 /5	/5 /5 /5

Total Left gap: /50

Total Right gap: /50

Object Asymmetry:

(total Left opening minus total Right opening hearts)

Total Time:

Total Correct: /50

Space Asymmetry:

(total in boxes 7,8,9,10 minus total correct in 1,2,3,4)

OCS

Oxford Cognitive Screen

ID: _____

Date: __ / __ / __

8. Meaningless Gesture Imitation

Which is your best hand now ? Left ☐ Right ☐

*This is a **mirror task**: use patient's best hand (e.g. right) and your opposite hand (e.g. left) to demonstrate the action sequence (2 sec per gesture). After practice: a **MAXIMUM of 2** demonstrations per sequence.*

Hand sequences:

	1st	2nd
 1st presentation	☐	☐
2nd presentation	☐	☐
	/3	
 1st presentation	☐	☐
2nd presentation	☐	☐
	/3	

Finger positions:

 1st presentation	☐
2nd presentation	☐
	/3
 1st presentation	☐
2nd presentation	☐
	/3

Total: /12

9. Delayed Recall & Recognition

Free recall first – only give MCQ if no response or incorrect answers

A. Verbal memory: Free recall

Have any of the islands got a quay, thought the colonel sitting on his yacht.

☐ ☐ ☐ ☐ **Total:** /4

B. Verbal memory: Recognition

Which word was in the sentence ?

1.	ocean	islands	hearts	bicycle
2.	harbour	restaurant	quay	dock
3.	pirate	priest	major	colonel
4.	yacht	hammock	boat	car

Total score (recall + recognition): /4

C. Episodic memory: Recognition

1.	Which picture did you see before?	banana	hippo	cow	
2.	Which picture did you see before?	beaver	carrot	broccoli	
3.	What did you cross out?	hearts	stars	crosses	faces
4.	What did you write?	words	letters	numbers	prices

Total score (episodic recognition): /4

10. Executive task

Circles (large to small)	/6	Mixed trail connections	/13
TIME	sec	TIME	sec
Triangles (large to small)	/6	**Executive score**	
TIME	sec	(circles + triangles) *minus* mixed	

Appendix 14: Brief CSIs scoring algorithm

	AMT-plus:																			NIHSS:				CAM-ICU:	
Test and total score	**1. Age**	**2. Time**	**3. Day**	**3. Month**	**3. Year**	**4. Place**	**4. City**	**5. Two-person recognition**	**6. Date of Birth**	**7. WW1**	**8. Prime minister**	**9. Count 20-1**	**10. Delayed Recall**	**11. Face**	**11. Numbers**	**11. Hands**	**12. News items**	**13. Months backwards**	**14. letter fluency**	**1-b level of concsiousness**	**1-c level of consciousness**	**9 Best language**	**11 Extinction & Inattention**	**Feature 1 (acute onset or fluctuating course)**	**Feature 3 (altered level of consciousness)**
Clock-drawing test														x	x	x									
Total: 3														/1	/1	/1									
Abbreviated MoCA													x	x	x	x									
Total: 8													/5	/1	/1	/1									
AMT-4	x				x	x			x																
Total: 4	/1				/1	/1			/1																
4AT	x				x	x			x									x						x	x
Total: 12	if AMT-4 is 4/4 = 0 points. If AMT is 3/4 = 1 point, if there are ≥2 errors or if they are untestable = 2 points																	0 points if ≥7 correct, 1 point if <7 OR refusal, 2 points if untestable						4 points If yes, 0 if no	4 points if yes, 0 if no
6-CIT		x		x	x							x	x					x							
Total: 28		correct: 0, incorrect: 3		correct: 0, incorrect: 3	correct: 0, incorrect: 4							correct: 0, 1 error: 2 points, >1 error: 4 points	points = no. of errors x 2					correct: 0, 1 error: 2 points, >1 error: 4 points							
10-AMT	x	x			x	x		x	x	x	x	x	x												
Total: 10	1	1			/1	/1		/1	/1	/1	/1	/1	1 point if ≥4 words correct, 0 points otherwise												
NINDS-CSN 5-min MoCA			x	x	x	x	x						x						x						
Total: 11			1	1	/1	/1	/1						/5						/1						
Cog-4																				x	x	x	x		
Total: 9																				/2	/2	/3	/2		

Appendix 15: Guidance document sent to APPLE sites

CLARIFICATION OF AMBIGUOUS TEST ITEMS – ADMINISTRATION AND SCORING

Oxford Cognitive Screen

Question	Clarification
3. Orientation	If the patient doesn't get any of these questions right by free recall, show them the multiple-choice answers
5. Sentence reading	The patient must correctly pronounce each word to get full marks (e.g. quay needs to read as 'key'). Self-correction is allowed. Don't penalise for dysarthria.
6. Number writing	The patient needs to write out the number numerically (708), not 'seven hundred and eight'
7. Broken Hearts	• The maximum time for this task is **3 minutes**. Please stop the task after this time. • Keep the page centred with the triangle in the patient's midline. The page should not be turned or moved from this position. • Please do not draw lines across the heart cancellation page (like how the scoring template has) • Space asymmetry "total in boxes 7,8,9,10 minus total correct in 1,2,3,4" – this means only add up the full hearts which have been cancelled out in these boxes, don't include any that have cancelled out which have gaps
8. Meaningless gesture imitation	If the patient correctly copies the 2 gestures after 1 demonstration they receive the full 3 points and you don't need to repeat them a second time. Scoring: On the two gestures: **3** if both correct first time **2** if both correct on second go **1** if one correct on second go **0** if neither correct on second go

	For the single finger positioning: **3** if correct first time **2** if correct second time **1** if incorrect but recognisable second time (e.g orientation error – this means it is the correct shape but not the mirror image) **0** if completely wrong (not recognizable)
9. Delayed recall and recognition	Verbal memory - If the patient gets all 4 words correct by free recall then you do not need to show them the multiple choice options at all. If they get 1 correct, then just show them the multiple-choice pages for the other 3 words.
10. Executive task	If an error is made at some point, but subsequent performance is correct, the correct connections are acknowledged. Self-correction is allowed.

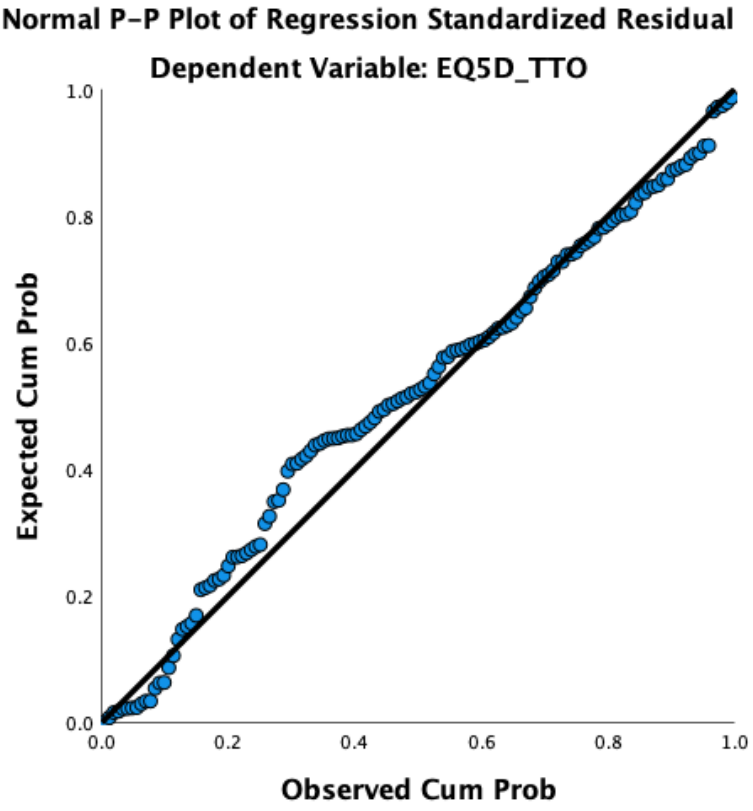
AMT

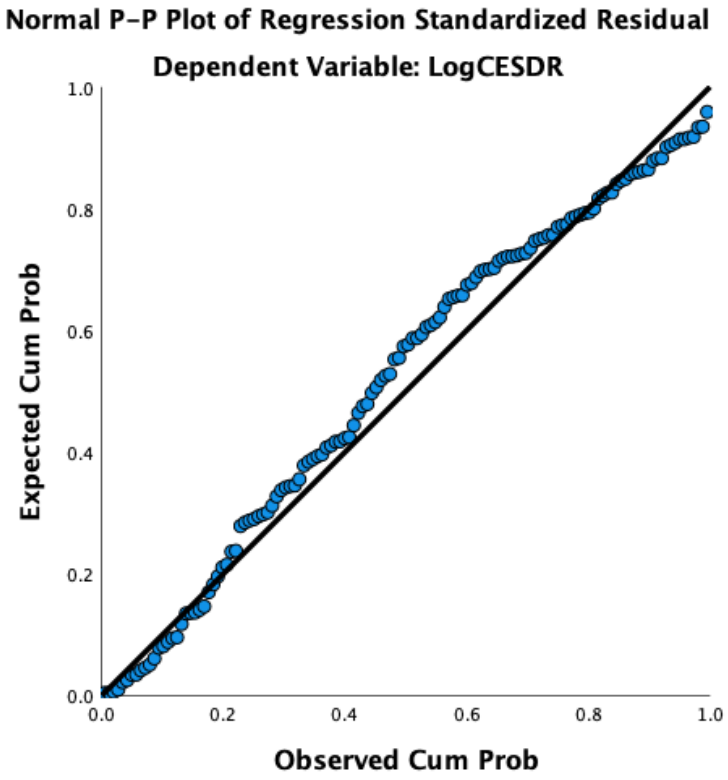
Question	Clarification
3. Date	In the current CRF version this item is split into 3 subpoints. The "Day" item relates to the day of the month e.g. "16th", not the day of the week e.g. "Tuesday". If the patient says the day of the week, prompt them for the exact date.
10. Recall	The patient must correctly recall a minimum of 4 words (in any order) to pass. "Number of mistakes" in the last column refers to the number of recall items not mentioned by the patient. For example, if a patient says "church, daisy, face", "number of mistakes" should be 2. If a patient says "face, rose, church, silk, arm, purple", the number of mistakes is 3.
11. Clock draw	Please ask the patient to set the time to 11:10. If a different time is used please write down what time you asked them to complete. We are following the scoring guidelines from the MoCA. In the current CRF version this item is split into 3 subpoints: • Face: the clock face must be a circle with only minor distortion acceptable (e.g. Slight imperfection on closing the circle) • Numbers: all clock numbers must be present with no additional numbers; numbers must be in the correct order and placed in the approximate quadrants on the clock face; Roman numerals are acceptable; numbers can be placed outside the circle contour • Hands: there must be two hands jointly indicating the correct time; the hour hand must be clearly shorter than the minute hand; hands must be centred within the clock face with their junction close to the clock

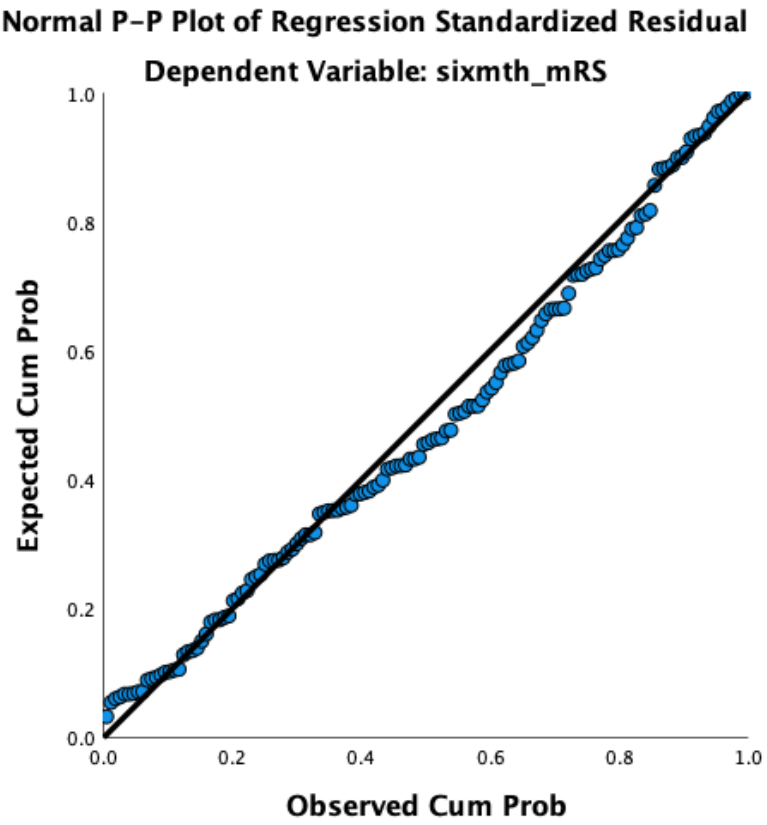
	centre
13. Months backwards	In order to pass, the patient must say a string of minimum 7 consecutive months correctly. A mistake is recorded if a patient does not mention a particular month(s) at all and/or lists a month(s) in an incorrect order. If a patient mentions a particular month more than once, and it is not an attempt at self-correction, this will also mean that at least on one occasion the month appeared in the wrong order. This should be however treated as a single error, not to penalise the patient for the same mistake twice. Overall, to simplify scoring, we suggest a rule that 1 particular month can account for 1 error maximum, meaning a highest possible total of 12 mistakes. Correctness of order should be assessed based on the preceding month. For example, if a patient says: "December, September, August, October, September, July, June, May, April, March, February, January", it's a pass (a correct sequence including 7 months, from July to January), with 4 errors (omitting November, September mentioned after December, October mentioned after August, July mentioned after September). Self-correction is allowed. We find that it's always easier to write down what the patient says and then score it later. If you're struggling with scoring we can do this if you have written down what the patient said.
14. One letter fluency	As mentioned in earlier instructions, if a patient says words with the same core but different suffixes, only 1 point is given – for the first word, e.g. 1 point for love, but 0 for lover and loving. However, if the first part of the words is the same, but the second part makes the words unrelated in terms of meaning, a point is awarded for each of these words e.g. for saying aircraft, airway and airtight a patient should receive 3 points. A score of 11 or more is a 'pass'.

Appendix 16: Chapter 7 linear regression plots

Normal probability plots for EQ-5D, log CESD-R, mRS outcomes. All three models include age, sex, NIHSS, pre-morbid mRS, years in education (log) and impairment in executive functioning.






Normal P-P Plot of Regression Standardized Residual
Dependent Variable: sixmth_mRS
Expected Cum Prob
Observed Cum Prob
0.0
0.2
0.4
0.6
0.8
1.0

Appendix 17: Chapter 7 full regression results

All models including memory

	Log CESD-R	EQ-5D	mRS	Barthel	Lawton
	Standardised beta (95% CI)			Odds ratio (95% CI)	
Age	-0.30 (-0.04, -0.01)**	0.16 (0.00, 0.01)*	0.02 (-0.01, 0.02)	1.01 (0.97, 1.05)	1.02 (0.98, 1.06)
Sex (male)	-0.09 (-0.57, 0.19)	0.10 (-0.03, 0.16)	-0.05 (-0.49, 0.24)	1.05 (0.35, 3.18)	1.78 (0.60, 5.26)
NIHSS	0.08 (-0.13, 0.38)	-0.17 (-0.14, -0.01)*	0.17 (0.05, 0.55)*	5.63 (2.26, 14.04)**	3.05 (1.37, 6.82)**
Pre-morbid mRS (>1)	0.15 (-0.06, 0.83)	-0.38 (-0.39, -0.17)**	0.51 (1.05, 1.86)**	4.47 (1.52, 13.18)*	10.17 (2.49, 41.51)**
Years in education (log)	-0.20 (-1.75, -0.14)*	0.23 (0.52, 0.94)**	-0.13 (-1.49, 0.09)	0.23 (0.02, 2.58)	0.02 (0.00, 0.29)**
Memory	0.03 (-0.52, 0.74)	-0.08 (-0.24, 0.06)	-0.05 (-0.79, 0.36)	3.23 (0.60, 17.32)	4.21 (0.73, 24.25)

*P<0.05, **p<0.007

All models including language

	Six-month outcomes				
	Log CESD-R	**EQ-5D**	**mRS**	**Barthel**	**Lawton**
	Standardised beta (95% CI)			**Odds ratio (95% CI)**	
Age	-0.30 (-0.04, -0.01)**	0.15 (0.00, 0.01)	0.02 (-0.01, 0.02)	1.01 (0.97, 1.05)	1.02 (0.98, 1.06)
Sex (male)	-0.09 (-0.59, 0.17)	0.11 (-0.03, 0.16)	-0.05 (-0.48, 0.24)	0.89 (0.31, 2.57)	1.48 (0.52, 4.21)
NIHSS	0.08 (-0.13, 0.39)	-0.17 (-0.14, -0.01)*	0.17 (0.07, 0.56)*	5.14 (2.15, 12.29)**	2.75 (1.28, 5.93)*
Pre-morbid mRS (>1)	0.13 (-0.09, 0.80)	-0.39 (-0.40, -0.18)**	0.48 (0.98, 1.79)*	4.72 (1.59, 14.01)**	9.79 (2.47, 38.81)**
Years in education (log)	-0.18 (-1.69, -0.08)*	0.23 (0.11, 0.52)**	-0.12 (-1.44, 0.12)	0.32 (0.03, 3.20	0.04 (0.00, 0.42)*
Language	0.09 (-0.24, 0.87)	0.02 (-0.12, 0.16)	0.12 (-0.07, 0.92)	0.96 (0.24, 3.86)	1.30 (0.31, 4.33)

*P<0.05, **p<0.007

All models including spatial attention

	Six-month outcomes				
	Log CESD-R	**EQ-5D**	**mRS**	**Barthel**	**Lawton**
	Standardised beta (95% CI)			**Odds ratio (95% CI)**	
Age	-0.29 (-0.04, -0.01)**	0.14 (0.00, 0.01)	0.01 (-0.01, 0.02)	1.01 (0.97, 1.05)	1.02 (0.98, 1.06)
Sex (male)	-0.07 (-0.55, 0.22)	0.12 (-0.02, 0.17)	-0.04 (-0.46, 0.28)	0.79 (0.27, 2.35)	1.60 (0.56, 4.58)
NIHSS	0.09 (-0.12, 0.40)	-0.19 (-0.15, -0.02)*	0.17 (0.04, 0.55)*	5.10 (2.10, 12.38)**	2.66 (1.23, 5.76)*
Pre-morbid mRS (>1)	0.17 (0.00, 0.90)*	-0.38 (-0.39, -0.17)**	0.50 (1.00, 1.84)**	4.21 (1.43, 12.38)*	9.87 (2.44, 39.96)**
Years in education (log)	-0.20 (-1.76, -0.15)*	0.21 (0.49, 0.93)**	-0.13 (-1.52, 0.09)	0.37 (0.04, 3.74)	0.03 (0.00, 0.39)*
Spatial attention	-0.08 (-0.65, 0.24)	-0.04 (-0.15, 0.08)	0.00 (-0.43, 0.43)	2.21 (0.71, 6.86)	0.80 (0.25, 2.58)

*P<0.05, **p<0.007

All models including executive functioning

	Six-month outcomes				
	Log CESD-R	**EQ-5D**	**mRS**	**Barthel**	**Lawton**
	Standardised beta (95% CI)			**Odds ratio (95% CI)**	
Age	-0.30 (-0.04, -0.01)**	0.16 (0.00, 0.01)*	0.01 (-0.01, 0.02)	1.00 (0.96, 1.05)	1.02 (0.98, 1.06)
Sex (male)	-0.09 (-0.58, 0.18)	0.10 (-0.03, 0.16)	-0.04 (-0.46, 0.27)	0.87 (0.29, 2.58)	1.52 (0.54, 4.33)
NIHSS	0.08 (-0.13, 0.38)	-0.18 (-0.14, -0.01)*	0.17 (0.05, 0.55)*	5.20 (2.10, 12.88)**	2.74 (1.26, 5.93)*
Pre-morbid mRS (>1)	0.15 (-0.06, 0.84)	-0.36 (-0.37, -0.16)**	0.49 (1.00, 1.82)**	4.26 (1.44, 12.60)*	9.68 (2.43, 38.55)**
Years in education (log)	-0.19 (-1.74, -0.13)*	0.20 (0.08, 0.47)*	-0.12 (-1.48, 0.11)	4.47 (0.73, 27.52)	0.04 (0.00, 0.41)*
Executive functioning	0.00 (-0.74, 0.71)	-0.21 (-0.41, 0.07)**	0.04 (-0.47, 0.87)	4.47 (0.73, 27.52)	0.96 (0.15, 6.28)

*P<0.05, **p<0.007

All models including number processing

	Six-month outcomes				
	Log CESD-R	**EQ-5D**	**mRS**	**Barthel**	**Lawton**
	Standardised beta (95% CI)			**Odds ratio (95% CI)**	
Age	-0.30 (-0.04, -0.01)**	0.15 (0.00, 0.01)	0.01 (-0.01, 0.01)	1.00 (0.96, 1.05)	1.02 (0.98, 1.06)
Sex (male)	-0.09 (-0.57, 0.18)	0.11 (-0.03, 0.17)	-0.04 (-0.45, 0.26)	0.84 (0.28, 2.49)	1.51 (0.53, 4.28)
NIHSS	0.07 (-0.15, 0.37)	-0.18 (-0.15, -0.02)*	0.16 (0.03, 0.53)*	4.68 (1.93, 11.39)**	2.78 (1.28, 6.04)*
Pre-morbid mRS (>1)	0.14 (-0.06, 0.82)	-0.40 (-0.40, -0.18)**	0.47 (0.95, 1.76)**	4.16 (1.37, 12.61)*	10.11 (2.55, 40.12)**
Years in education (log)	-0.19 (-1.71, -0.10)*	0.23 (0.11, 0.52)**	-0.12 (-1.44, 0.11)	0.41 (0.04, 4.06)	0.03 (0.00, 0.40)*
Number processing	0.06 (-0.47, 1.06)	0.05 (-0.12, 0.23)	0.15 (0.03, 1.25)*	8.71 (0.92, 82.15)	0.83 (0.12, 5.84)

*P<0.05, **p<0.007

All models including praxis

	Six-month outcomes				
	Log CESD-R	**EQ-5D**	**mRS**	**Barthel**	**Lawton**
	Standardised beta (95% CI)			**Odds ratio (95% CI)**	
Age	-0.29 (-0.04, -0.01)**	0.15 (0.00, 0.01)	0.01 (-0.01, 0.02)	1.01 (0.97, 1.05)	1.02 (0.98, 1.06)
Sex (male)	-0.08 (-0.56, 0.19)	0.11 (-0.03, 0.17)	-0.04 (-0.47, 0.25)	0.98 (0.33, 2.90)	1.75 (0.59, 5.17)
NIHSS	0.09 (-0.12, 0.40)	-0.18 (-0.14, -0.01)*	0.17 (0.05, 0.56)*	5.86 (2.31, 14.85)**	3.29 (1.43, 7.55)**
Pre-morbid mRS (>1)	0.13 (-0.09, 0.80)	-0.39 (-0.40, -0.18)**	0.50 (1.03, 1.85)**	4.15 (1.41, 12.20)*	8.37 (2.11, 33.26)**
Years in education (log)	-0.18 (-1.68, -0.07)*	0.22 (0.10, 0.51)**	-0.13 (-1.50, 0.08)	0.39 (0.04, 3.64)	0.04 (0.00, 0.52)*
Praxis	-0.11 (-1.26, 0.26)	0.01 (-0.18, 0.19)	-0.01 (-0.73, 0.62)	0.21 (0.02, 2.28)	0.11 (0.01, 1.17)

*P<0.05, **p<0.007

All models including global OCS score

	Six-month outcomes				
	Log CESD-R	**EQ-5D**	**mRS**	**Barthel**	**Lawton**
	Standardised beta (95% CI)			**Odds ratio (95% CI)**	
Age	-0.30 (-0.04, -0.01)**	0.16 (0.00, 0.01)*	0.01 (-0.01, 0.01)	1.00 (0.96, 1.04)	1.02 (0.98, 1.06)
Sex (male)	-0.09 (-0.58, 0.18)	0.12 (-0.02, 0.17)	-0.05 (-0.48, 0.24)	0.75 (0.25, 2.24)	1.54 (0.54, 4.39)
NIHSS	0.08 (-0.13, 0.38)	-0.17 (-0.14, -0.01)*	0.17 (0.06, 0.55)*	5.11 (2.06, 12.66)**	2.79 (1.29, 6.03)*
Pre-morbid mRS (>1)	0.15 (-0.06, 0.84)	-0.37 (-0.39, -0.17)**	0.48 (0.97, 1.79)**	4.12 (1.38, 12.32)*	10.30 (2.58, 41.18)**
Years in education (log)	-0.19 (-1.74, -0.13)*	0.21 (0.09, 0.50)**	-0.13 (-1.48, 0.08)	0.38 (0.04, 4.09)	0.04 (0.00, 0.40)*
Global OCS	0.00 (-0.22, 0.23)	-0.09 (-0.09, 0.02)	0.08 (-0.09, 0.31)	1.75 (0.97, 3.17)	0.91 (0.51, 1.61)

*P<0.05, **p<0.007

List of References

1. Stroke Association. A complete guide to cognitive problems after stroke. 2018.
2. Harvey PD. Domains of cognition and their assessment. Dialogues in clinical neuroscience. 2019;21(3):227-37.
3. American Psychiatric Association., American Psychiatric Association. DSM-5 Task Force. Diagnostic and statistical manual of mental disorders : DSM-5. 5th ed. Washington, D.C.: American Psychiatric Association; 2013. xliv, 947 p. p.
4. Sachdev PS, Blacker D, Blazer DG, Ganguli M, Jeste DV, Paulsen JS, et al. Classifying neurocognitive disorders: the DSM-5 approach. Nature Reviews Neurology. 2014;10(11):634-42.
5. Stroke Association. State of the Nation: Stroke Statistics January 2017 [
6. Renjen PN, Gauba C, Chaudhari D. Cognitive Impairment After Stroke. Cureus. 2015;7(9):e335.
7. Leśniak M, Bak T, Czepiel W, Seniów J, Członkowska A. Frequency and prognostic value of cognitive disorders in stroke patients. Dement Geriatr Cogn Disord. 2008;26(4):356-63.
8. van Zandvoort MJ, Kessels RP, Nys GM, de Haan EH, Kappelle LJ. Early neuropsychological evaluation in patients with ischaemic stroke provides valid information. Clin Neurol Neurosurg. 2005;107(5):385-92.
9. Sexton E, McLoughlin A, Williams DJ, Merriman NA, Donnelly N, Rohde D, 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):160-71.
10. Wentzel C, Rockwood K, MacKnight C, Hachinski V, Hogan DB, Feldman H, et al. Progression of impairment in patients with vascular cognitive impairment without dementia. Neurology. 2001;57(4):714-6.
11. Pendlebury ST, Rothwell PM. Prevalence, incidence, and factors associated with pre-stroke and post-stroke dementia: a systematic review and meta-analysis. Lancet Neurol. 2009;8(11):1006-18.
12. Galski T, Bruno RL, Zorowitz R, Walker J. Predicting length of stay, functional outcome, and aftercare in the rehabilitation of stroke patients. The dominant role of higher-order cognition. Stroke. 1993;24(12):1794-800.
13. Nys GM, van Zandvoort MJ, de Kort PL, van der Worp HB, Jansen BP, Algra A, et al. The prognostic value of domain-specific cognitive abilities in acute first-ever stroke. Neurology. 2005;64(5):821-7.
14. Nys GM, van Zandvoort MJ, van der Worp HB, de Haan EH, de Kort PL, Jansen BP, et al. Early cognitive impairment predicts long-term depressive symptoms and quality of life after stroke. J Neurol Sci. 2006;247(2):149-56.
15. Towfighi A, Ovbiagele B, El Husseini N, Hackett ML, Jorge RE, Kissela BM, et al. Poststroke Depression: A Scientific Statement for Healthcare Professionals From the American Heart Association/American Stroke Association. Stroke. 2017;48(2):e30-e43.
16. Sinha P, Bharath S, Chandra S. DSM-5 in Vascular Dementia - Comparison with other diagnostic criteria in a retrospective study. EC Neurology. 2015;2.3:135-43.
17. Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. Clin Geriatr Med. 2013;29(4):737-52.
18. Mijajlović MD, Pavlović A, Brainin M, Heiss WD, Quinn TJ, Ihle-Hansen HB, et al. Post-stroke dementia - a comprehensive review. BMC Med. 2017;15(1):11.

19. Dichgans M, Leys D. Vascular Cognitive Impairment. Circ Res. 2017;120(3):573-91.
20. Nys GM, van Zandvoort MJ, de Kort PL, Jansen BP, de Haan EH, Kappelle LJ. Cognitive disorders in acute stroke: prevalence and clinical determinants. Cerebrovasc Dis. 2007;23(5-6):408-16.
21. Blackburn DJ, Bafadhel L, Randall M, Harkness KA. Cognitive screening in the acute stroke setting. Age Ageing. 2013;42(1):113-6.
22. Sachdev P, Looi J. Neuropsychological differentiation of Alzheimer's disease and vascular dementia. In: Bowler J, Hachinski V, editors. Vascular Cognitive Impairment: Reversible Dementia. Oxford: Oxford University Press; 2003. p. 153-75.
23. Brookes RL, Hannesdottir K, Lawrence R, Morris RG, Markus HS. Brief Memory and Executive Test: evaluation of a new screening test for cognitive impairment due to small vessel disease. Age and Ageing. 2012;41(2):212-8.
24. Hoffmann M. Higher cortical function deficits after stroke: an analysis of 1,000 patients from a dedicated cognitive stroke registry. Neurorehabil Neural Repair. 2001;15(2):113-27.
25. Stamenova V, Roy E, Black S. Poster 29: Limb Apraxia Recovery After Stroke. Archives of Physical Medicine and Rehabilitation. 2008;89(10):e35-e6.
26. Farnè A, Buxbaum LJ, Ferraro M, Frassinetti F, Whyte J, Veramonti T, et al. Patterns of spontaneous recovery of neglect and associated disorders in acute right brain-damaged patients. J Neurol Neurosurg Psychiatry. 2004;75(10):1401-10.
27. Merriman NA, Sexton E, McCabe G, Walsh ME, Rohde D, Gorman A, 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-e.
28. Larner A. Cerebrovascular disease: vascular dementia and vascular cognitive impairment. Neuropsychological Neurology. 2nd ed: Cambridge University Press; 2013.
29. Rasquin SM, Lodder J, Verhey FR. Predictors of reversible mild cognitive impairment after stroke: a 2-year follow-up study. J Neurol Sci. 2005;229-230:21-5.
30. Ganguli M, Jia Y, Hughes TF, Snitz BE, Chang CH, Berman SB, et al. Mild Cognitive Impairment that Does Not Progress to Dementia: A Population-Based Study. Journal of the American Geriatrics Society. 2019;67(2):232-8.
31. Al-Qazzaz NK, Ali SH, Ahmad SA, Islam S, Mohamad K. Cognitive impairment and memory dysfunction after a stroke diagnosis: a post-stroke memory assessment. Neuropsychiatr Dis Treat. 2014;10:1677-91.
32. Santisteban MM, Iadecola C. Hypertension, dietary salt and cognitive impairment. J Cereb Blood Flow Metab. 2018;38(12):2112-28.
33. Iadecola C. The overlap between neurodegenerative and vascular factors in the pathogenesis of dementia. Acta Neuropathol. 2010;120(3):287-96.
34. Chui HC, Victoroff JI, Margolin D, Jagust W, Shankle R, Katzman R. Criteria for the diagnosis of ischemic vascular dementia proposed by the State of California Alzheimer's Disease Diagnostic and Treatment Centers. Neurology. 1992;42(3 Pt 1):473-80.
35. Shaw RC, Walker G, Elliott E, Quinn TJ. Occurrence Rate of Delirium in Acute Stroke Settings: Systematic Review and Meta-Analysis. Stroke. 2019;50(11):3028-36.

36. Mohd Zulkifly MF, Ghazali SE, Che Din N, Singh DKA, Subramaniam P. A Review of Risk Factors for Cognitive Impairment in Stroke Survivors. The Scientific World Journal. 2016;2016:3456943.
37. Levine DA, Wadley VG, Langa KM, Unverzagt FW, Kabeto MU, Giordani B, et al. Risk Factors for Poststroke Cognitive Decline: The REGARDS Study (Reasons for Geographic and Racial Differences in Stroke). Stroke. 2018;49(4):987-94.
38. Sun JH, Tan L, Yu JT. Post-stroke cognitive impairment: epidemiology, mechanisms and management. Ann Transl Med. 2014;2(8):80.
39. Zhao L, Biesbroek JM, Shi L, Liu W, Kuijf HJ, Chu WW, et al. Strategic infarct location for post-stroke cognitive impairment: A multivariate lesion-symptom mapping study. J Cereb Blood Flow Metab. 2018;38(8):1299-311.
40. Biesbroek JM, Weaver NA, Biessels GJ. Lesion location and cognitive impact of cerebral small vessel disease. Clin Sci (Lond). 2017;131(8):715-28.
41. Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. Lancet Neurol. 2010;9(7):689-701.
42. Wardlaw JM, Smith EE, Biessels GJ, Cordonnier C, Fazekas F, Frayne R, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. The Lancet Neurology. 2013;12(8):822-38.
43. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet. 2020;396(10248):413-46.
44. Rosenich E, Hordacre B, Paquet C, Koblar SA, Hillier SL. Cognitive Reserve as an Emerging Concept in Stroke Recovery. Neurorehabilitation and Neural Repair. 2020;34(3):187-99.
45. Larsson SC, Traylor M, Malik R, Dichgans M, Burgess S, Markus HS. Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. BMJ. 2017;359:j5375.
46. Yeh S-CJ, Liu Y-Y. Influence of social support on cognitive function in the elderly. BMC Health Serv Res. 2003;3(1):9-.
47. Ahlskog JE, Geda YE, Graff-Radford NR, Petersen RC. Physical exercise as a preventive or disease-modifying treatment of dementia and brain aging. Mayo Clin Proc. 2011;86(9):876-84.
48. Umarova RM, Sperber C, Kaller CP, Schmidt CSM, Urbach H, Klöppel S, et al. Cognitive reserve impacts on disability and cognitive deficits in acute stroke. Journal of Neurology. 2019;266(10):2495-504.
49. Rattanabannakit C, Risacher SL, Gao S, Lane KA, Brown SA, McDonald BC, et al. The Cognitive Change Index as a Measure of Self and Informant Perception of Cognitive Decline: Relation to Neuropsychological Tests. Journal of Alzheimer's disease : JAD. 2016;51(4):1145-55.
50. Patchick E, Vail A, Wood A, Bowen A. PRECiS (Patient Reported Evaluation of Cognitive State): psychometric evaluation of a new patient reported outcome measure of the impact of stroke. Clinical Rehabilitation. 2016;30(12):1229-41.
51. Teasdale G, Jennett B. ASSESSMENT OF COMA AND IMPAIRED CONSCIOUSNESS: A Practical Scale. The Lancet. 1974;304(7872):81-4.
52. Swartz RH, Cayley ML, Lanctôt KL, Murray BJ, Smith EE, Sahlas DJ, et al. Validating a Pragmatic Approach to Cognitive Screening in Stroke Prevention Clinics Using the Montreal Cognitive Assessment. Stroke. 2016;47(3):807-13.
53. Ablewhite J, Geraghty J, das Nair R, Lincoln N, Drummond A. Cognitive Management Pathways in Stroke Services (COMPASS): A qualitative investigation of key issues in relation to community stroke teams undertaking cognitive assessments. British Journal of Occupational Therapy. 2019;82(7):404-11.

54. Bowen A, Hazelton C, Pollock A, Lincoln NB. Cognitive rehabilitation for spatial neglect following stroke. Cochrane Database of Systematic Reviews. 2013(7).
55. das Nair R, Cogger H, Worthington E, Lincoln NB. Cognitive rehabilitation for memory deficits after stroke. Cochrane Database of Systematic Reviews. 2016(9).
56. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. Journal of the American Geriatrics Society. 2005;53(4):695-9.
57. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research. 1975;12(3):189-98.
58. Stolwyk RJ, O'Neill MH, McKay AJ, Wong DK. Are cognitive screening tools sensitive and specific enough for use after stroke? A systematic literature review. Stroke. 2014;45(10):3129-34.
59. Rodrigues JdC, Becker N, Beckenkamp CL, Miná CS, de Salles JF, Bandeira DR. Psychometric properties of cognitive screening for patients with cerebrovascular diseases A systematic review. Dementia & neuropsychologia. 2019;13(1):31-43.
60. Molnar FJ, Benjamin S, Hawkins SA, Briscoe M, Ehsan S. One Size Does Not Fit All: Choosing Practical Cognitive Screening Tools for Your Practice. Journal of the American Geriatrics Society. 2020;n/a(n/a).
61. Hodkinson HM. Evaluation of a mental test score for assessment of mental impairment in the elderly. Age Ageing. 1972;1(4):233-8.
62. Brooke P, Bullock R. Validation of a 6 item cognitive impairment test with a view to primary care usage. Int J Geriatr Psychiatry. 1999;14(11):936-40.
63. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The mini-cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. Int J Geriatr Psychiatry. 2000;15(11):1021-7.
64. Brodaty H, Pond D, Kemp NM, Luscombe G, Harding L, Berman K, et al. The GPCOG: a new screening test for dementia designed for general practice. Journal of the American Geriatrics Society. 2002;50(3):530-4.
65. Mathuranath PS, Nestor PJ, Berrios GE, Rakowicz W, Hodges JR. A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. Neurology. 2000;55(11):1613-20.
66. Demeyere N, Riddoch MJ, Slavkova ED, Bickerton WL, Humphreys GW. The Oxford Cognitive Screen (OCS): validation of a stroke-specific short cognitive screening tool. Psychol Assess. 2015;27(3):883-94.
67. Holmes J, Gilbody S. Differences in use of abbreviated mental test score by geriatricians and psychiatrists. BMJ. 1996;313(7055):465.
68. Chan E, Khan S, Oliver R, Gill SK, Werring DJ, Cipolotti L. Underestimation of cognitive impairments by the Montreal Cognitive Assessment (MoCA) in an acute stroke unit population. J Neurol Sci. 2014;343(1-2):176-9.
69. Nicholson K, Martelli MF, Zasler ND. Does pain confound interpretation of neuropsychological test results? NeuroRehabilitation. 2001;16(4):225-30.
70. Dagenais E, Rouleau I, Demers M, Jobin C, Roger E, Chamelian L, et al. Value of the MoCA test as a screening instrument in multiple sclerosis. Can J Neurol Sci. 2013;40(3):410-5.
71. Hoops S, Nazem S, Siderowf AD, Duda JE, Xie SX, Stern MB, et al. Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. Neurology. 2009;73(21):1738-45.

72. Rambeau A, Beauplet B, Laviec H, Licaj I, Leconte A, Chatel C, et al. Prospective comparison of the Montreal Cognitive Assessment (MoCA) and the Mini Mental State Examination (MMSE) in geriatric oncology. J Geriatr Oncol. 2019;10(2):235-40.
73. Gluhm S, Goldstein J, Brown D, Van Liew C, Gilbert PE, Corey-Bloom J. Usefulness of the Montreal Cognitive Assessment (MoCA) in Huntington's disease. Mov Disord. 2013;28(12):1744-7.
74. Chiti G, Pantoni L. Use of Montreal Cognitive Assessment in patients with stroke. Stroke. 2014;45(10):3135-40.
75. Jaywant A, Toglia J, Gunning FM, O'Dell MW. The diagnostic accuracy of the Montreal Cognitive Assessment in inpatient stroke rehabilitation. Neuropsychological Rehabilitation. 2019;29(8):1163-76.
76. Koski L. Validity and applications of the Montreal cognitive assessment for the assessment of vascular cognitive impairment. Cerebrovasc Dis. 2013;36(1):6-18.
77. Nasreddine ZS. MoCA Test Mandatory Training and Certification: What Is the Purpose? Journal of the American Geriatrics Society. 2020;68(2):444-5.
78. Dong Y, Sharma VK, Chan BP, Venketasubramanian N, Teoh HL, Seet RC, et al. The Montreal Cognitive Assessment (MoCA) is superior to the Mini-Mental State Examination (MMSE) for the detection of vascular cognitive impairment after acute stroke. J Neurol Sci. 2010;299(1-2):15-8.
79. Pendlebury ST, Cuthbertson FC, Welch SJ, Mehta Z, Rothwell PM. Underestimation of cognitive impairment by Mini-Mental State Examination versus the Montreal Cognitive Assessment in patients with transient ischemic attack and stroke: a population-based study. Stroke. 2010;41(6):1290-3.
80. Chan E, Altendorff S, Healy C, Werring DJ, Cipolotti L. The test accuracy of the Montreal Cognitive Assessment (MoCA) by stroke lateralisation. J Neurol Sci. 2017;373:100-4.
81. Bour A, Rasquin S, Boreas A, Limburg M, Verhey F. How predictive is the MMSE for cognitive performance after stroke? Journal of neurology. 2010;257(4):630-7.
82. Beatty WW, Goodkin DE. Screening for cognitive impairment in multiple sclerosis. An evaluation of the Mini-Mental State Examination. Arch Neurol. 1990;47(3):297-301.
83. Nys GMS, van Zandvoort MJE, de Kort PLM, Jansen BPW, Kappelle LJ, de Haan EHF. Restrictions of the Mini-Mental State Examination in acute stroke. Archives of Clinical Neuropsychology. 2005;20(5):623-9.
84. Saar K, Nyrkkö H, Tolvanen A, Kuikka P, Poutiainen E, Aro T. Validation of a New Cognitive Screening Method for Stroke Patients. Behavioural Neurology. 2019;2019:2943603.
85. Lunardelli A, Mengotti P, Pesavento A, Sverzut A, Zadini A. The Brief Neuropsychological Screening (BNS): valuation of its clinical validity. Eur J Phys Rehabil Med. 2009;45(1):85-91.
86. Williams PM, Johnson C, Swan S, Barber C, Murphy PK, Devine J, et al. The northwick park examination of cognition: A brief cognitive assessment tool for use in acute stroke services. International journal of therapy and rehabilitation. 2016;23:314-22.
87. Brookes RL, Hollocks MJ, Khan U, Morris RG, Markus HS. The Brief Memory and Executive Test (BMET) for detecting vascular cognitive impairment in small vessel disease: a validation study. BMC Medicine. 2015;13(1):51.

88. Barnay JL, Wauquiez G, Rousseaux M, Bonnin-Koang HY, Dischler F, De Boissezon X, et al. Presentation of the cognitive assessment scale for stroke patient (CASP). Ann Phys Rehabil Med. 2012;55:e188.
89. Pasotti F, Magnani FG, Gallucci M, Salvato G, Ovadia D, Scotto M, et al. Neuropsychological assessment in acute stroke patients. Neurol Sci. 2020;41(5):1259-66.
90. Oh HS, Kim JS, Shim EB, Seo WS. Development and Clinical Validity of a Mild Vascular Cognitive Impairment Assessment Tool for Korean Stroke Patients. Asian Nursing Research. 2015;9(3):226-34.
91. Mancuso M, Varalta V, Sardella L, Capitani D, Zoccolotti P, Antonucci G. Italian normative data for a stroke specific cognitive screening tool: the Oxford Cognitive Screen (OCS). Neurol Sci. 2016;37(10):1713-21.
92. Huygelier H, Schraepen B, Demeyere N, Gillebert CR. 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. 2019:1-22.
93. Valera-Gran D, Lopez-Roig S, Hurtado-Pomares M, Peral-Gomez P, Garcia-Manzanares M, Sunyer Catlla 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):724-36.
94. Shendyapina M, Kuzmina E, Kazymaev S, Petrova A, Demeyere N, Weekes BS. The Russian version of the Oxford Cognitive Screen: Validation study on stroke survivors. Neuropsychology. 2019;33(1):77-92.
95. 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.
96. Hong WJ, Tao J, Wong AWK, Yang SL, Leung MT, Lee TMC, 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.
97. Kong AP-H, Lam PH-P, Ho DW-L, Lau JK, Humphreys GW, Riddoch J, et al. The Hong Kong version of the Oxford Cognitive Screen (HK-OCS): validation study for Cantonese-speaking chronic stroke survivors. Aging, Neuropsychology, and Cognition. 2016;23(5):530-48.
98. Ramos CCF, Amado DK, Borges CR, Bergamaschi E, Nitrini R, Brucki SMD. Oxford Cognitive Screen - Brazilian Portuguese version (OCS-Br) A pilot study. Dementia & neuropsychologia. 2018;12(4):427-31.
99. Rodrigues JdC, Bandeira DR, de Salles JF. Cognitive Screening (TRIACOG) for adults with cerebrovascular diseases: Construction process and validity evidence. Psychology & Neuroscience. 2020:No Pagination Specified-No Pagination Specified.
100. Brandt J, Spencer M, Folstein M. The Telephone Interview for Cognitive Status. Neuropsychiatry Neuropsychol Behavioral Neurol. 1988(1):111-7.
101. Welsh KA, Breitner JCS, Magruderhabib KM. Detection of Dementia in the Elderly Using Telephone Screening of Cognitive Status. Neuropsy Neuropsy Be. 1993;6(2):103-10.
102. Pendlebury ST, Welch SJ, Cuthbertson FC, Mariz J, Mehta Z, Rothwell PM. 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.

103. Hendry K, Quinn TJ, Evans J, Scortichini V, Miller H, Burns J, et al. Evaluation of delirium screening tools in geriatric medical inpatients: a diagnostic test accuracy study. Age Ageing. 2016;45(6):832-7.
104. Carin-Levy G, Mead GE, Nicol K, Rush R, van Wijck F. Delirium in acute stroke: screening tools, incidence rates and predictors: a systematic review. J Neurol. 2012;259(8):1590-9.
105. Jorm AF, Korten AE. Assessment of cognitive decline in the elderly by informant interview. Br J Psychiatry. 1988;152:209-13.
106. McGovern A, Pendlebury ST, Mishra NK, Fan Y, Quinn TJ. Test Accuracy of Informant-Based Cognitive Screening Tests for Diagnosis of Dementia and Multidomain Cognitive Impairment in Stroke. Stroke. 2016;47(2):329-35.
107. Winstein CJ, Stein J, Arena R, Bates B, Cherney LR, Cramer SC, et al. Guidelines for Adult Stroke Rehabilitation and Recovery: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. Stroke. 2016;47(6):e98-e169.
108. National Institute for Health and Clinical Excellence. Psychological care after stroke. Improving stroke services for people with cognitive and mood disorders. London: NICE; 2012.
109. McGurn B, Starr JM, Topfer JA, Pattie A, Whiteman MC, Lemmon HA, et al. Pronunciation of irregular words is preserved in dementia, validating premorbid IQ estimation. Neurology. 2004;62(7):1184-6.
110. Hachinski V, Iadecola C, Petersen RC, Breteler MM, Nyenhuis DL, Black SE, et al. National Institute of Neurological Disorders and Stroke-Canadian Stroke Network vascular cognitive impairment harmonization standards. Stroke. 2006;37(9):2220-41.
111. Intercollegiate Stroke Working Party. National Clinical Guideline for Stroke, 5th ed. London: Royal College of Physicians; 2016.
112. European Stroke Organisation (ESO) Executive Committee, ESO Writing Committee. Guidelines for management of ischaemic stroke and transient ischaemic attack 2008. Cerebrovasc Dis. 2008;25(5):457-507.
113. Stroke Foundation. Clinical Guidelines for Stroke Management. Available at https://informme.org.au/en/Guidelines/Clinical-Guidelinesfor-Stroke-Management [
114. Lanctôt KL, Lindsay MP, Smith EE, Sahlas DJ, Foley N, Gubitz G, et al. Canadian Stroke Best Practice Recommendations: Mood, Cognition and Fatigue following Stroke, 6th edition update 2019. International Journal of Stroke. 2019;15(6):668-88.
115. Souza AC, Alexandre NMC, Guirardello EB. Psychometric properties in instruments evaluation of reliability and validity. Epidemiol Serv Saude. 2017;26(3):649-59.
116. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42.
117. Larner AJ. What is test accuracy? Comparing unitary accuracy metrics for cognitive screening instruments. Neurodegener Dis Manag. 2019;9(5):277-81.
118. Takwoingi Y, Quinn TJ. Review of Diagnostic Test Accuracy (DTA) studies in older people. Age and Ageing. 2018;47(3):349-55.
119. Hemmy LS, Linskens EJ, Silverman PC, Miller MA, Talley KMC, Taylor BC, et al. Brief Cognitive Tests for Distinguishing Clinical Alzheimer-Type Dementia From Mild Cognitive Impairment or Normal Cognition in Older Adults With Suspected Cognitive Impairment. Ann Intern Med. 2020;172(10):678-87.

120. Lees R, Selvarajah J, Fenton C, Pendlebury ST, Langhorne P, Stott DJ, et al. Test accuracy of cognitive screening tests for diagnosis of dementia and multidomain cognitive impairment in stroke. Stroke. 2014;45(10):3008-18.
121. Burton L, Tyson SF. Screening for cognitive impairment after stroke: A systematic review of psychometric properties and clinical utility. Journal of rehabilitation medicine. 2015;47(3):193-203.
122. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess. 1998;2(14):i-iv, 1-74.
123. Pendlebury ST, Klaus SP, Thomson RJ, Mehta Z, Wharton RM, Rothwell PM, et al. Methodological Factors in Determining Risk of Dementia After Transient Ischemic Attack and Stroke: (III) Applicability of Cognitive Tests. Stroke. 2015;46(11):3067-73.
124. Lees RA, Hendry Ba K, Broomfield N, Stott D, Larner AJ, Quinn TJ. Cognitive assessment in stroke: feasibility and test properties using differing approaches to scoring of incomplete items. Int J Geriatr Psychiatry. 2016.
125. Pasi M, Salvadori E, Poggesi A, Inzitari D, Pantoni L. Factors predicting the Montreal cognitive assessment (MoCA) applicability and performances in a stroke unit. J Neurol. 2013;260(6):1518-26.
126. Pendlebury ST, Klaus SP, Thomson RJ, Mehta Z, Wharton RM, Rothwell PM. Methodological Factors in Determining Risk of Dementia After Transient Ischemic Attack and Stroke: (III) Applicability of Cognitive Tests. Stroke. 2015;46(11):3067-73.
127. Horstmann S, Rizos T, Rauch G, Arden C, Veltkamp R. Feasibility of the Montreal Cognitive Assessment in acute stroke patients. Eur J Neurol. 2014;21(11):1387-93.
128. Pendlebury ST, Chen PJ, Bull L, Silver L, Mehta Z, Rothwell PM. Methodological factors in determining rates of dementia in transient ischemic attack and stroke: (I) impact of baseline selection bias. Stroke. 2015;46(3):641-6.
129. Alderman S, F V, A B, S S. Feasibility of cognitive testing in patients with acute ischemic stroke. Stroke. 2013;44(2 MeetingAbstract):no pagination.
130. Collas D. Detecting Cognitive Impairment in Acute Stroke Using the Oxford Cognitive Screen: Characteristics of Left and Right Hemisphere Strokes, with More Severe Impairment in AF Strokes (P1.190). Neurology. 2016;86(16 Supplement):P1.190.
131. van Zandvoort MJE, Kessels RPC, Nys GMS, de Haan EHF, Kappelle LJ. Early neuropsychological evaluation in patients with ischaemic stroke provides valid information. Clinical Neurology and Neurosurgery. 2005;107(5):385-92.
132. Barnay JL, Wauquiez G, Bonnin-Koang HY, Anquetil C, Perennou D, Piscicelli C, et al. Feasibility of the cognitive assessment scale for stroke patients (CASP) vs. MMSE and MoCA in aphasic left hemispheric stroke patients. Annals of physical and rehabilitation medicine. 2014;57(6-7):422-35.
133. Benaim C, Barnay JL, Wauquiez G, Bonnin-Koang HY, Anquetil C, Perennou D, et al. The Cognitive Assessment scale for Stroke Patients (CASP) vs. MMSE and MoCA in non-aphasic hemispheric stroke patients. Annals of physical and rehabilitation medicine. 2015;58(2):78-85.
134. Cumming TB, Bernhardt J, Linden T. The montreal cognitive assessment: short cognitive evaluation in a large stroke trial. Stroke. 2011;42(9):2642-4.
135. Kwa VIH, Limburg M, Voogel AJ, Teunisse S, Derix MMA, Hijdra A. Feasibility of cognitive screening of patients with ischaemic stroke using the CAMCOG A hospital-based study. Journal of Neurology. 1996;243(5):405-9.

136. Mancuso M, Demeyere N, Abbruzzese L, Damora A, Varalta V, Pirrotta F, 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.
137. Lees RA, Hendry Ba K, Broomfield N, Stott D, Larner AJ, Quinn TJ. Cognitive assessment in stroke: feasibility and test properties using differing approaches to scoring of incomplete items. Int J Geriatr Psychiatry. 2017;32(10):1072-8.
138. Party ISW. Sentinel Stroke National Audit Programme (SSNAP). Clinical audit April 2017 – July 2017 Public Report. 2017.
139. Bamford J, Sandercock P, Dennis M, Warlow C, Burn J. Classification and natural history of clinically identifiable subtypes of cerebral infarction. The Lancet. 1991;337(8756):1521-6.
140. Kasner SE, Chalela JA, Luciano JM, Cucchiara BL, Raps EC, McGarvey ML, et al. Reliability and validity of estimating the NIH stroke scale score from medical records. Stroke. 1999;30(8):1534-7.
141. Lyden P. Using the National Institutes of Health Stroke Scale: A Cautionary Tale. Stroke. 2017;48(2):513-9.
142. Lyden P, Claesson L, Havstad S, Ashwood T, Lu M. Factor analysis of the National Institutes of Health Stroke Scale in patients with large strokes. Arch Neurol. 2004;61(11):1677-80.
143. Meyer BC, Lyden PD. The modified National Institutes of Health Stroke Scale: its time has come. International journal of stroke : official journal of the International Stroke Society. 2009;4(4):267-73.
144. Banks JL, Marotta CA. Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis. Stroke. 2007;38(3):1091-6.
145. Quinn TJ, Dawson J, Walters MR, Lees KR. Reliability of the modified Rankin Scale: a systematic review. Stroke. 2009;40(10):3393-5.
146. Fearon P, McArthur KS, Garrity K, Graham LJ, McGroarty G, Vincent S, et al. Prestroke modified rankin stroke scale has moderate interobserver reliability and validity in an acute stroke setting. Stroke. 2012;43(12):3184-8.
147. Panenková E, Kopecek M, Lukavsky J. Item analysis and possibility to abbreviate the Montreal Cognitive Assessment2016. 63-9 p.
148. Ismail Z, Rajji TK, Shulman KI. Brief cognitive screening instruments: an update. International journal of geriatric psychiatry. 2010;25(2):111-20.
149. Swain DG, Nightingale PG. Evaluation of a shortened version of the Abbreviated Mental Test in a series of elderly patients. Clin Rehabil. 1997;11(3):243-8.
150. O'Regan C, Meagher D. A comparison of the lighthouse (LH) test with other bedside tests of attention in elderly medical inpatients. Irish Journal of Medical Science. 2016;5):S210.
151. Champod AS, Gubitz GJ, Phillips SJ, Christian C, Reidy Y, Radu LM, et al. Clock Drawing Test in acute stroke and its relationship with long-term functional and cognitive outcomes. The Clinical neuropsychologist. 2019;33(5):817-30.
152. Renovanz M, Reitzug L, Messing L, Scheurich A, Grüninger S, Ringel F, et al. Patient reported feasibility and acceptance of Montreal Cognitive Assessment (MoCA) screening pre- and postoperatively in brain tumour patients. J Clin Neurosci. 2018;53:79-84.
153. Lorentz WJ, Scanlan JM, Borson S. Brief screening tests for dementia. Can J Psychiatry. 2002;47(8):723-33.

154. Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. Journal of the American Geriatrics Society. 2005;53(4):695-9.
155. Beck AT, Beck RW. Screening depressed patients in family practice. A rapid technic. Postgrad Med. 1972;52(6):81-5.
156. Jenkinson C, Fitzpatrick R, Crocker H, Peters M. The Stroke Impact Scale: validation in a UK setting and development of a SIS short form and SIS index. Stroke. 2013;44(9):2532-5.
157. MacIsaac RL, Ali M, Taylor-Rowan M, Rodgers H, Lees KR, Quinn TJ. Use of a 3-Item Short-Form Version of the Barthel Index for Use in Stroke: Systematic Review and External Validation. Stroke. 2017;48(3):618-23.
158. Roalf DR, Moore TM, Wolk DA, Arnold SE, Mechanic-Hamilton D, Rick J, et al. Defining and validating a short form Montreal Cognitive Assessment (s-MoCA) for use in neurodegenerative disease. J Neurol Neurosurg Psychiatry. 2016;87(12):1303-10.
159. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. Jama. 2018;319(4):388-96.
160. Julayanont P, Tangwongchai S, Hemrungrojn S, Tunvirachaisakul C, Phanthumchinda K, Hongsawat J, et al. The Montreal Cognitive Assessment-Basic: A Screening Tool for Mild Cognitive Impairment in Illiterate and Low-Educated Elderly Adults. Journal of the American Geriatrics Society. 2015;63(12):2550-4.
161. Takwoingi Y, Quinn TJ. Review of Diagnostic Test Accuracy (DTA) studies in older people. Age and Ageing. 2018.
162. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of internal medicine. 2011;155(8):529-36.
163. Noel-Storr AH, McCleery JM, Richard E, Ritchie CW, Flicker L, Cullum SJ, et al. Reporting standards for studies of diagnostic test accuracy in dementia: The STARDdem Initiative. Neurology. 2014;83(4):364-73.
164. Ali M, Bath P, Brady M, Davis S, Diener HC, Donnan G, et al. Development, expansion, and use of a stroke clinical trials resource for novel exploratory analyses. International journal of stroke : official journal of the International Stroke Society. 2012;7(2):133-8.
165. Larner AJ. Screening utility of the Montreal Cognitive Assessment (MoCA): in place of--or as well as--the MMSE? International psychogeriatrics. 2012;24(3):391-6.
166. Larner AJ. Short Montreal Cognitive Assessment: Validation and Reproducibility. Journal of Geriatric Psychiatry and Neurology. 2016;30(2):104-8.
167. Wong A, Nyenhuis D, Black SE, Law LS, Lo ES, Kwan PW, et al. Montreal Cognitive Assessment 5-minute protocol is a brief, valid, reliable, and feasible cognitive screen for telephone administration. Stroke. 2015;46(4):1059-64.
168. Horton DK, Hynan LS, Lacritz LH, Rossetti HC, Weiner MF, Cullum CM. An Abbreviated Montreal Cognitive Assessment (MoCA) for Dementia Screening. The Clinical neuropsychologist. 2015;29(4):413-25.
169. Lim JS, Oh MS, Lee JH, Jung S, Kim C, Jang MU, et al. Prediction of post-stroke dementia using NINDS-CSN 5-minute neuropsychology protocol in acute stroke. International psychogeriatrics. 2017;29(5):777-84.

170. Mai LM, Oczkowski W, Mackenzie G, Shuster A, Wasielesky L, Franchetto A, et al. Screening for cognitive impairment in a stroke prevention clinic using the MoCA. Can J Neurol Sci. 2013;40(2):192-7.
171. Bocti C, Legault V, Leblanc N, Berger L, Nasreddine Z, Beaulieu-Boire I, et al. Vascular cognitive impairment: most useful subtests of the Montreal Cognitive Assessment in minor stroke and transient ischemic attack. Dement Geriatr Cogn Disord. 2013;36(3-4):154-62.
172. Dong Y, Koay WI, Yeo LL, Chen CL, Xu J, Seet RC, et al. Rapid Screening for Cognitive Impairment in Parkinson's Disease: A Pilot Study. Parkinsons Dis. 2015;2015:348063.
173. Cecato JF, Martinelli JE, Izbicki R, Yassuda MS, Aprahamian I. A subtest analysis of the Montreal cognitive assessment (MoCA): which subtests can best discriminate between healthy controls, mild cognitive impairment and Alzheimer's disease? Int Psychogeriatr. 2016;28(5):825-32.
174. Campbell N, Rice D, Friedman L, Speechley M, Teasell RW. Screening and facilitating further assessment for cognitive impairment after stroke: application of a shortened Montreal Cognitive Assessment (miniMoCA). Disabil Rehabil. 2016;38(6):601-4.
175. Freitas S, Batista S, Afonso AC, Simões MR, de Sousa L, Cunha L, et al. The Montreal Cognitive Assessment (MoCA) as a screening test for cognitive dysfunction in multiple sclerosis. Applied Neuropsychology: Adult. 2018;25(1):57-70.
176. Wittich W, Phillips N, Nasreddine ZS, Chertkow H. Sensitivity and Specificity of the Montreal Cognitive Assessment Modified for Individuals who are Visually Impaired. Journal of Visual Impairment & Blindness. 2010;104(6):360-8.
177. Kaur D, Kumar G, Singh AK. Quick screening of cognitive function in Indian multiple sclerosis patients using Montreal cognitive assessment test-short version. Ann Indian Acad Neurol. 2013;16(4):585-9.
178. Xu X, Chan QL, Hilal S, Ikram MK, Venketasubramanian N, Tan BY, et al. The Diagnostic Utility of the NINDS-CSN Neuropsychological Battery in Memory Clinics. Dement Geriatr Cogn Dis Extra. 2016;6(2):276-82.
179. Lin HF, Chern CM, Chen HM, Yeh YC, Yao SC, Huang MF, et al. Validation of NINDS-VCI Neuropsychology Protocols for Vascular Cognitive Impairment in Taiwan. PLoS One. 2016;11(6):e0156404.
180. Cameron JD, Gallagher R, Pressler SJ, McLennan SN, Ski CF, Tofler G, et al. Sensitivity and Specificity of a Five-Minute Cognitive Screening Test in Patients With Heart Failure. J Card Fail. 2016;22(2):99-107.
181. Bocti C, Cossette P, Fulop T, Lacombe G, Létourneau C. P3-226: Designing an abridged Montreal Cognitive Assessment for more efficient cognitive screening. Alzheimer's & Dementia. 2012;8(4S_Part_14):P539-P.
182. Davies LM, Oczkowski W, MacKenzie G, Sahlas DJ. Detection of cognitive impairment in a stroke prevention clinic using the moca: rationale for prospective validation of a mini-MoCA. P.066. 2nd Canadian Stroke Congress; Canada: Stroke; 2011.
183. Dong C, Xu K, Chen C. Abstract 18752: An Ultra-short Cognitive Screening Test to Detect Significant Post-stroke Vascular Cognitive Impairment. Circulation. 2015;132(suppl_3):A18752-A.
184. Wong SS, Fong KN. Reliability and validity of the telephone version of the Cantonese Mini-mental State Examination (T-CMMSE) when used with elderly patients with and without dementia in Hong Kong. International psychogeriatrics. 2009;21(2):345-53.

185. Goetz C, Coste J, Lemetayer F, Rat AC, Montel S, Recchia S, et al. Item reduction based on rigorous methodological guidelines is necessary to maintain validity when shortening composite measurement scales. J Clin Epidemiol. 2013;66(7):710-8.
186. Liew TM. The Optimal Short Version of Montreal Cognitive Assessment in Diagnosing Mild Cognitive Impairment and Dementia. Journal of the American Medical Directors Association. 2019;20(8):1055.e1-.e8.
187. Bezdicek O, Červenková M, Moore TM, Stepankova Georgi H, Sulc Z, Wolk DA, et al. Determining a Short Form Montreal Cognitive Assessment (s-MoCA) Czech Version: Validity in Mild Cognitive Impairment Parkinson's Disease and Cross-Cultural Comparison. Assessment. 2018:1073191118778896.
188. Dawes P, Pye A, Reeves D, Yeung WK, Sheikh S, Thodi C, et al. Protocol for the development of versions of the Montreal Cognitive Assessment (MoCA) for people with hearing or vision impairment. BMJ Open. 2019;9(3):e026246.
189. Freitas S, Prieto G, Simões MR, Santana I. Psychometric properties of the Montreal Cognitive Assessment (MoCA): an analysis using the Rasch model. Clin Neuropsychol. 2014;28(1):65-83.
190. Freitas S, Prieto G, Simões MR, Santana I. Scaling Cognitive Domains of the Montreal Cognitive Assessment: An Analysis Using the Partial Credit Model. Archives of Clinical Neuropsychology. 2015;30(5):435-47.
191. Luo H, Andersson B, Tang JYM, Wong GHY. Applying Item Response Theory Analysis to the Montreal Cognitive Assessment in a Low-Education Older Population. Assessment. 2019:1073191118821733.
192. Davis DH, Creavin ST, Yip JL, Noel-Storr AH, Brayne C, Cullum S. Montreal Cognitive Assessment for the diagnosis of Alzheimer's disease and other dementias. Cochrane Database Syst Rev. 2015(10):Cd010775.
193. Larner AJ. Speed versus accuracy in cognitive assessment when using CSIs. Progress in Neurology and Psychiatry. 2015;19(1):21-4.
194. Wouters H, van Campen J, Appels B, Lindeboom R, Buiter M, de Haan RJ, et al. Does adaptive cognitive testing combine efficiency with precision? Prospective findings. J Alzheimers Dis. 2011;25(4):595-603.
195. van Rooij FG, Schaapsmeerders P, Maaijwee NA, van Duijnhoven DA, de Leeuw FE, Kessels RP, et al. Persistent cognitive impairment after transient ischemic attack. Stroke. 2014;45(8):2270-4.
196. Fisher GG, Ryan LH. Overview of the Health and Retirement Study and Introduction to the Special Issue. Work, Aging and Retirement. 2017;4(1):1-9.
197. Herr M, Ankri J. A Critical Review of the Use of Telephone Tests to Identify Cognitive Impairment in Epidemiology and Clinical Research. Journal of Telemedicine and Telecare. 2013;19(1):45-54.
198. Berthier ML. Poststroke Aphasia. Drugs & Aging. 2005;22(2):163-82.
199. Martin-Khan M, Wootton R, Gray L. A systematic review of the reliability of screening for cognitive impairment in older adults by use of standardised assessment tools administered via the telephone. J Telemed Telecare. 2010;16(8):422-8.
200. Castanho TC, Amorim L, Zihl J, Palha JA, Sousa N, Santos NC. Telephone-based screening tools for mild cognitive impairment and dementia in aging studies: a review of validated instruments. Frontiers in aging neuroscience. 2014;6:16-.
201. Lewis BE, Mills CS, Mohs RC, Hill J, Fillit H. Improving early recognition of Alzheimer's disease: a review of telephonic screening tools. JCOM-WAYNE PA-. 2001;8(8):41-6.

202. Davis DHJ, Creavin ST, Noel-Storr A, Quinn TJ, Smailagic N, Hyde C, et al. Neuropsychological tests for the diagnosis of Alzheimer's disease dementia and other dementias: a generic protocol for cross-sectional and delayed-verification studies. Cochrane Database of Systematic Reviews. 2013(3).
203. Covidence systematic review software. Melbourne, Australia: Veritas Health Innovation. Available at www.covidence.org
204. Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. Neurology. 1993;43(2):250-60.
205. Steinberg DM, Fine J, Chappell R. Sample size for positive and negative predictive value in diagnostic research using case-control designs. Biostatistics. 2009;10(1):94-105.
206. Takwoingi Y, Riley RD, Deeks JJ. Meta-analysis of diagnostic accuracy studies in mental health. Evidence Based Mental Health. 2015;18(4):103-9.
207. Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. Stat Methods Med Res. 2017;26(4):1896-911.
208. Review Manager (RevMan). 5.3 ed. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration; 2014.
209. Freeman SC, Kerby CR, Patel A, Cooper NJ, Quinn T, Sutton AJ. Development of an interactive web-based tool to conduct and interrogate meta-analysis of diagnostic test accuracy studies: MetaDTA. BMC Med Res Methodol. 2019;19(1):81.
210. Dal Forno G, Chiovenda P, Bressi F, Ferreri F, Grossi E, Brandt J, et al. Use of an Italian version of the telephone interview for cognitive status in Alzheimer's disease. Int J Geriatr Psych. 2006;21(2):126-33.
211. Kempen GI, Meier AJ, Bouwens SF, van Deursen J, Verhey FR. Telefonisch Interview Cognitieve Status (TICS): psychometrische aspecten. Tijdschr Gerontol Geriatr. 2007;38(1):38-45.
212. Konagaya Y, Washimi Y, Hattori H, Takeda A, Watanabe T, Ohta T. Validation of the Telephone Interview for Cognitive Status (TICS) in Japanese. Int J Geriatr Psych. 2007;22(7):695-700.
213. Knopman DS, Roberts RO, Geda YE, Pankratz VS, Christianson TJ, Petersen RC, et al. Validation of the telephone interview for cognitive status-modified in subjects with normal cognition, mild cognitive impairment, or dementia. Neuroepidemiology. 2010;34(1):34-42.
214. Seo EH, Lee DY, Kim SG, Kim KW, Kim DH, Kim BJ, et al. Validity of the telephone interview for cognitive status (TICS) and modified TICS (TICSm) for mild cognitive imparment (MCI) and dementia screening. Arch Gerontol Geriatr. 2011;52(1):e26-30.
215. Cook SE, Marsiske M, McCoy KJ. The use of the Modified Telephone Interview for Cognitive Status (TICS-M) in the detection of amnestic mild cognitive impairment. J Geriatr Psychiatry Neurol. 2009;22(2):103-9.
216. Lipton RB, Katz MJ, Kuslansky G, Sliwinski MJ, Stewart WF, Verghese J, et al. Screening for dementia by telephone using the memory impairment screen. Journal of the American Geriatrics Society. 2003;51(10):1382-90.
217. Manly JJ, Schupf N, Stern Y, Brickman AM, Tang MX, Mayeux R. Telephone-based identification of mild cognitive impairment and dementia in a multicultural cohort. Arch Neurol. 2011;68(5):607-14.
218. Jarvenpaa T, Rinne JO, Raiha I, Koskenvuo M, Lopponen M, Hinkka S, et al. Characteristics of two telephone screens for cognitive impairment. Dement Geriatr Cogn Disord. 2002;13(3):149-55.

219. Desmond DW, Tatemichi TK, Hanzawa L. The Telephone Interview for Cognitive Status (Tics) - Reliability and Validity in a Stroke Sample. Int J Geriatr Psych. 1994;9(10):803-7.
220. Brandt J, Spencer M, Folstein M. The Telephone Interview for Cognitive Status1988.
221. Johnston D, Samus QM, Morrison A, Leoutsakos JS, Hicks K, Handel S, et al. Identification of community-residing individuals with dementia and their unmet needs for care. Int J Geriatr Psychiatry. 2011;26(3):292-8.
222. Go RC, Duke LW, Harrell LE, Cody H, Bassett SS, Folstein MF, et al. Development and validation of a Structured Telephone Interview for Dementia Assessment (STIDA): the NIMH Genetics Initiative. J Geriatr Psychiatry Neurol. 1997;10(4):161-7.
223. Crooks VC, Clark L, Petitti DB, Chui H, Chiu V. Validation of multi-stage telephone-based identification of cognitive impairment and dementia. BMC Neurol. 2005;5(1):8.
224. Plassman BL, Newman TT, Welsh KA, Helms M, Breitner JCS. Properties of the telephone interview for cognitive status: Application in epidemiological and longitudinal studies. Neuropsychiatry, Neuropsychology and Behavioral Neurology. 1994;7(3):235-41.
225. Zietemann V, Kopczak A, Muller C, Wollenweber FA, Dichgans M. Validation of the Telephone Interview of Cognitive Status and Telephone Montreal Cognitive Assessment Against Detailed Cognitive Testing and Clinical Diagnosis of Mild Cognitive Impairment After Stroke. Stroke. 2017;48(11):2952-7.
226. Meng C, Zhang X, Zhou J. The telephone interview for cognitive status-modified used in screening cognitive impairment. Chinese Mental Health Journal. 2005;19:34-7.
227. Alexopoulos P, Perneczky R, Cramer B, Grimmer T, Kurz A. [Validation of a Short Telephone Test (T3MS) for the Diagnosis of Cognitive Impairment]. Fortschritte der Neurologie-Psychiatrie. 2006;74(6):329-36.
228. Roccaforte WH, Burke WJ, Bayer BL, Wengel SP. Reliability and validity of the Short Portable Mental Status Questionnaire administered by telephone. J Geriatr Psychiatry Neurol. 1994;7(1):33-8.
229. Gatz M, Reynolds C, Nikolic J, Lowe B, Karel M, Pedersen N. An empirical test of telephone screening to identify potential dementia cases. International psychogeriatrics. 1995;7(3):429-38.
230. Van Mierlo LD, Wouters H, Sikkes SAM, Van Der Flier WM, Prins ND, Bremer JAE, et al. Screening for mild cognitive impairment and dementia with automated, anonymous online and telephone cognitive self-tests. Journal of Alzheimer's Disease. 2017;56(1):249-59.
231. Zhou J, Zhang X, Mundt JC, Wang L, Meng C, Chu C, et al. A Comparison of Three Dementia Screening Instruments Administered by Telephone in China. Dementia. 2004;3(1):69-81.
232. McDicken JA, Elliott E, Blayney G, Makin S, Ali M, Larner AJ, et al. Accuracy of the short-form Montreal Cognitive Assessment: Systematic review and validation. Int J Geriatr Psychiatry. 2019;34(10):1515-25.
233. Sachs-Ericsson N, Blazer DG. The new DSM-5 diagnosis of mild neurocognitive disorder and its relation to research in mild cognitive impairment. Aging Ment Health. 2015;19(1):2-12.
234. Tsoi KK, Chan JY, Hirai HW, Wong SY, Kwok TC. Cognitive Tests to Detect Dementia: A Systematic Review and Meta-analysis. JAMA internal medicine. 2015;175(9):1450-8.

235. Breton A, Casey D, Arnaoutoglou NA. Cognitive tests for the detection of mild cognitive impairment (MCI), the prodromal stage of dementia: Meta-analysis of diagnostic accuracy studies. Int J Geriatr Psychiatry. 2018.
236. Jorm AF, Jacomb PA. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): socio-demographic correlates, reliability, validity and some norms. Psychol Med. 1989;19(4):1015-22.
237. Li K, Malhotra PA. Spatial neglect. Pract Neurol. 2015;15(5):333-9.
238. Pena MCS, Sobreira EST, Souza CP, Oliveira GN, Tumas V, do Vale FdAC. Visuospatial cognitive tests for the evaluation of patients with Parkinson's disease. Dementia & neuropsychologia. 2008;2(3):201-5.
239. Salimi S, Irish M, Foxe D, Hodges JR, Piguet O, Burrell JR. Can visuospatial measures improve the diagnosis of Alzheimer's disease? Alzheimers Dement (Amst). 2017;10:66-74.
240. Chen P, Ratcliff G, Belle SH, Cauley JA, DeKosky ST, Ganguli M. Patterns of cognitive decline in presymptomatic Alzheimer disease: a prospective community study. Arch Gen Psychiatry. 2001;58(9):853-8.
241. Dupuis K, Pichora-Fuller MK, Chasteen AL, Marchuk V, Singh G, Smith SL. Effects of hearing and vision impairments on the Montreal Cognitive Assessment. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn. 2015;22(4):413-37.
242. Brearly TW, Shura RD, Martindale SL, Lazowski RA, Luxton DD, Shenal BV, et al. Neuropsychological Test Administration by Videoconference: A Systematic Review and Meta-Analysis. Neuropsychol Rev. 2017;27(2):174-86.
243. Chapman JE, Cadilhac DA, Gardner B, Ponsford J, Bhalla R, Stolwyk RJ. Comparing face-to-face and videoconference completion of the Montreal Cognitive Assessment (MoCA) in community-based survivors of stroke. J Telemed Telecare. 2019:1357633x19890788.
244. Brott T, Adams HP, Jr., Olinger CP, Marler JR, Barsan WG, Biller J, et al. Measurements of acute cerebral infarction: a clinical examination scale. Stroke. 1989;20(7):864-70.
245. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. Ann Intern Med. 1990;113(12):941-8.
246. Rankin J. Cerebral vascular accidents in patients over the age of 60. I. General considerations. Scott Med J. 1957;2(4):127-36.
247. Mahoney FI, Barthel DW. Functional evaluation: The Barthel Index: A simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill. Maryland State Medical Journal. 1965;14:61-5.
248. Lawton MP, Brody EM. Assessment of Older People: Self-Maintaining and Instrumental Activities of Daily Living1. The Gerontologist. 1969;9(3_Part_1):179-86.
249. Aprahamian I, Martinelli JE, Neri AL, Yassuda MS. The Clock Drawing Test: A review of its accuracy in screening for dementia. Dementia & neuropsychologia. 2009;3(2):74-81.
250. AGRELL B, DEHLIN O. The clock-drawing test. Age and Ageing. 1998;27(3):399-403.
251. Nishiwaki Y, Breeze E, Smeeth L, Bulpitt CJ, Peters R, Fletcher AE. Validity of the Clock-Drawing Test as a screening tool for cognitive impairment in the elderly. Am J Epidemiol. 2004;160(8):797-807.
252. Cumming TB, Blomstrand C, Bernhardt J, Linden T. The NIH Stroke Scale Can Establish Cognitive Function after Stroke. Cerebrovascular Diseases. 2010;30(1):7-14.

253. Lees R, Lua J, Melling E, Miao Y, Tan J, Quinn TJ. Cog-4 has limited diagnostic test accuracy and validity for cognitive assessment in stroke survivors. J Stroke Cerebrovasc Dis. 2014;23(6):1604-10.
254. Ankolekar S, Renton C, Sprigg N, Bath PMW. The cog-4 subset of the national institutes of health stroke scale as a measure of cognition: relationship with baseline factors and functional outcome after stroke using data from the virtual international stroke trials archive. Stroke Res Treat. 2013;2013:562506-.
255. Katzman R, Brown T, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short Orientation-Memory-Concentration Test of cognitive impairment. Am J Psychiatry. 1983;140(6):734-9.
256. Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. Br J Psychiatry. 1968;114(512):797-811.
257. O'Sullivan D, O'Regan NA, Timmons S. Validity and Reliability of the 6-Item Cognitive Impairment Test for Screening Cognitive Impairment: A Review. Dement Geriatr Cogn Disord. 2016;42(1-2):42-9.
258. Schofield I, Stott DJ, Tolson D, McFadyen A, Monaghan J, Nelson D. Screening for cognitive impairment in older people attending accident and emergency using the 4-item Abbreviated Mental Test. Eur J Emerg Med. 2010;17(6):340-2.
259. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001;33(5):337-43.
260. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199-208.
261. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-108.
262. Center for Epidemiologic Studies Depression Scale: Review and Revision (CESD and CESD-R) [press release]. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers2004.
263. Dyer AH, Briggs R, Nabeel S, O'Neill D, Kennelly SP. The Abbreviated Mental Test 4 for cognitive screening of older adults presenting to the Emergency Department. Eur J Emerg Med. 2017;24(6):417-22.
264. Quinn TJ, Elliott E, Langhorne P. Cognitive and Mood Assessment Tools for Use in Stroke. Stroke. 2018;49(2):483-90.
265. Johnson-Greene D, Touradji P, Emmerson LC. The Three Cities Test: Preliminary Validation of a Short Bedside Memory Test in Persons with Acute Stroke. Topics in Stroke Rehabilitation. 2009;16(5):321-9.
266. Wong A, Mok VCT, Yim P, Fu M, Lam WWM, Yau C, et al. The executive clock drawing task (CLOX) is a poor screening test for executive dysfunction in Chinese elderly patients with subcortical ischemic vascular disease. Journal of Clinical Neuroscience. 2004;11(5):493-7.
267. Lees R, Corbet S, Johnston C, Moffitt E, Shaw G, Quinn TJ. Test accuracy of short screening tests for diagnosis of delirium or cognitive impairment in an acute stroke unit setting. Stroke. 2013;44(11):3078-83.
268. Mok VCT, Wong A, Lam WWM, Fan YH, Tang WK, Kwok T, et al. Cognitive impairment and functional outcome after stroke associated with small vessel disease. Journal of Neurology, Neurosurgery & Psychiatry. 2004;75(4):560-6.
269. Jehkonen M, Laihosalo M, Kettunen JE. Impact of neglect on functional outcome after stroke: a review of methodological issues and recent research findings. Restor Neurol Neurosci. 2006;24(4-6):209-15.

270. Noel-Storr AH, McCleery JM, Richard E, Ritchie CW, Flicker L, Cullum SJ, et al. Reporting standards for studies of diagnostic test accuracy in dementia: The STARDdem Initiative. Neurology. 2014;83(4):364-73.
271. Velayudhan L, Ryu S-H, Raczek M, Philpot M, Lindesay J, Critchfield M, et al. Review of brief cognitive tests for patients with suspected dementia. International psychogeriatrics. 2014;26(8):1247-62.
272. Woodford HJ, George J. Cognitive assessment in the elderly: a review of clinical methods. QJM: An International Journal of Medicine. 2007;100(8):469-84.
273. Hwang AB, Boes S, Nyffeler T, Schuepfer G. Validity of screening instruments for the detection of dementia and mild cognitive impairment in hospital inpatients: A systematic review of diagnostic accuracy studies. PLOS ONE. 2019;14(7):e0219569.
274. Demeyere N, Riddoch MJ, Slavkova ED, Jones K, Reckless I, Mathieson P, et al. Domain-specific versus generalized cognitive screening in acute stroke. J Neurol. 2016;263(2):306-15.
275. DeCrane SK, Sands LP, Young KM, DePalma G, Leung JM. Impact of missing data on analysis of postoperative cognitive decline (POCD). Appl Nurs Res. 2013;26(2):71-5.
276. Carley S, Dosman S, Jones SR, Harrison M. Simple nomograms to calculate sample size in diagnostic studies. Emergency Medicine Journal. 2005;22(3):180-1.
277. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. Biometrics. 1988;44(3):837-45.
278. Mandrekar JN. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. Journal of Thoracic Oncology. 2010;5(9):1315-6.
279. Abzhandadze T, Reinholdsson M, Stibrant Sunnerhagen K. NIHSS is not enough for cognitive screening in acute stroke: A cross-sectional, retrospective study. Scientific Reports. 2020;10(1):534.
280. Larner AJ. Cognitive screening instruments for the diagnosis of mild cognitive impairment. Progress in Neurology and Psychiatry. 2016;20(2):21-6.
281. Goldberg TE, Harvey PD, Wesnes KA, Snyder PJ, Schneider LS. Practice effects due to serial cognitive assessment: Implications for preclinical Alzheimer's disease randomized controlled trials. Alzheimers Dement (Amst). 2015;1(1):103-11.
282. Lin VY, Chung J, Callahan BL, Smith L, Gritters N, Chen JM, et al. Development of cognitive screening test for the severely hearing impaired: Hearing-impaired MoCA. Laryngoscope. 2017;127 Suppl 1:S4-s11.
283. Parada JC, Hillyer J, Parbery-Clark A. Performance on the standard and hearing-impaired Montreal Cognitive Assessment in cochlear implant users. Int J Geriatr Psych. 2020;35(4):338-47.
284. Tuijl JP, Scholte EM, de Craen AJ, van der Mast RC. Screening for cognitive impairment in older general hospital patients: comparison of the Six-Item Cognitive Impairment Test with the Mini-Mental State Examination. Int J Geriatr Psychiatry. 2012;27(7):755-62.
285. Adams HP, Jr., Davis PH, Leira EC, Chang KC, Bendixen BH, Clarke WR, et al. Baseline NIH Stroke Scale score strongly predicts outcome after stroke: A report of the Trial of Org 10172 in Acute Stroke Treatment (TOAST). Neurology. 1999;53(1):126-31.
286. Wouters A, Nysten C, Thijs V, Lemmens R. Prediction of Outcome in Patients With Acute Ischemic Stroke Based on Initial Severity and Improvement in the First 24 h. Frontiers in Neurology. 2018;9(308).

287. Ahmed R, Zuberi BF, Afsar S. Stroke scale score and early prediction of outcome after stroke. J Coll Physicians Surg Pak. 2004;14(5):267-9.
288. 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):346-70.
289. Abzhandadze T, Rafsten L, Lundgren Nilsson Å, Palstam A, Sunnerhagen KS. Very Early MoCA Can Predict Functional Dependence at 3 Months After Stroke: A Longitudinal, Cohort Study. Front Neurol. 2019;10:1051.
290. Demeyere N, Sun S, Milosevich E, Vancleef K. Post-stroke cognition with the Oxford Cognitive Screen vs Montreal Cognitive Assessment: a multi-site randomized controlled study (OCS-CARE) [version 1; peer review: 1 approved, 1 approved with reservations]. AMRC Open Research. 2019;1(12).
291. Haacke C, Althaus A, Spottke A, Siebert U, Back T, Dodel R. Long-term outcome after stroke: evaluating health-related quality of life using utility measurements. Stroke. 2006;37(1):193-8.
292. Predicting outcome after acute ischemic stroke: an external validation of prognostic models. Neurology. 2004;62(4):581-5.
293. Kogan E, Twyman K, Heap J, Milentijevic D, Lin JH, Alberts M. Assessing stroke severity using electronic health record data: a machine learning approach. BMC Medical Informatics and Decision Making. 2020;20(1):8.
294. Weisscher N, Vermeulen M, Roos YB, de Haan RJ. What should be defined as good outcome in stroke trials; a modified Rankin score of 0-1 or 0-2? J Neurol. 2008;255(6):867-74.
295. Hochstenbach JB, Anderson PG, van Limbeek J, Mulder TT. Is there a relation between neuropsychologic variables and quality of life after stroke? Archives of Physical Medicine and Rehabilitation. 2001;82(10):1360-6.
296. Haupt M, Gillebert CR, Demeyere N. The zero effect: voxel-based lesion symptom mapping of number transcoding errors following stroke. Sci Rep. 2017;7(1):9242.
297. de Graaf JA, Nijsse B, Schepers VPM, van Heugten CM, Post MWM, Visser-Meily JMA. Which approach to measure cognitive functioning should be preferred when exploring the association between cognitive functioning and participation after stroke? Neuropsychol Rehabil. 2020:1-17.
298. Zietemann V, Georgakis MK, Dondaine T, Müller C, Mendyk AM, Kopczak A, et al. Early MoCA predicts long-term cognitive and functional outcome and mortality after stroke. Neurology. 2018;91(20):e1838-e50.
299. Dong Y, Slavin MJ, Chan BP, Venketasubramanian N, Sharma VK, Crawford JD, et al. Cognitive screening improves the predictive value of stroke severity scores for functional outcome 3-6 months after mild stroke and transient ischaemic attack: an observational study. BMJ Open. 2013;3(9):e003105.
300. Fonseca J, Raposo A, Martins IP. Cognitive performance and aphasia recovery. Top Stroke Rehabil. 2018;25(2):131-6.
301. Hurford R, Charidimou A, Fox Z, Cipolotti L, Werring DJ. Domain-specific trends in cognitive impairment after acute ischaemic stroke. J Neurol. 2013;260(1):237-41.
302. Abzhandadze T, Rafsten L, Lundgren Nilsson Å, Palstam A, Sunnerhagen KS. Very Early MoCA Can Predict Functional Dependence at 3 Months After Stroke: A Longitudinal, Cohort Study. Frontiers in Neurology. 2019;10(1051).
303. Mahdy ME, Ghonimi NA, Elserafy TS, Mahmoud W. The NIHSS score can predict the outcome of patients with primary intracerebral hemorrhage. The Egyptian Journal of Neurology, Psychiatry and Neurosurgery. 2019;55(1):21.

304. Fraser SD, George S. Perspectives on differing health outcomes by city: accounting for Glasgow's excess mortality. Risk Manag Healthc Policy. 2015;8:99-110.
305. Wetterling T, Kanitz RD, Borgis KJ. Comparison of different diagnostic criteria for vascular dementia (ADDTC, DSM-IV, ICD-10, NINDS-AIREN). Stroke. 1996;27(1):30-6.
306. Erkinjuntti T, Ostbye T, Steenhuis R, Hachinski V. The effect of different diagnostic criteria on the prevalence of dementia. N Engl J Med. 1997;337(23):1667-74.
307. Mansutti I, Saiani L, Palese A. Detecting delirium in patients with acute stroke: a systematic review of test accuracy. BMC Neurol. 2019;19(1):310.
308. Bauer RM, Iverson GL, Cernich AN, Binder LM, Ruff RM, Naugle RI. Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. Arch Clin Neuropsychol. 2012;27(3):362-73.
309. Geraghty J, Ablewhite J, das Nair R, Lincoln N, Kontou E, Roffe J, et al. Results of a UK-wide vignette study with occupational therapists to explore cognitive screening post stroke. International Journal of Therapy and Rehabilitation. 2020;27(7):1-12.
310. Newman CGJ, Bevins AD, Zajicek JP, Hodges JR, Vuillermoz E, Dickenson JM, et al. Improving the quality of cognitive screening assessments: ACEmobile, an iPad-based version of the Addenbrooke's Cognitive Examination-III. Alzheimers Dement (Amst). 2017;10:182-7.
311. Cannon P, Larner AJ. Errors in the scoring and reporting of cognitive screening instruments administered in primary care. Neurodegener Dis Manag. 2016;6(4):271-6.
312. Wong S, Jacova C. Older Adults' Attitudes towards Cognitive Testing: Moving towards Person-Centeredness. Dementia and geriatric cognitive disorders extra. 2018;8(3):348-59.