Evaluation of current evidence and practice to inform development of a Standardised Neurological OBservation Schedule for Stroke (SNOBSS)

by

Alison McLoughlin

A thesis submitted in partial fulfilment for the requirements for the degree of Doctor of Philosophy at the University of Central Lancashire

June 2022

RESEARCH STUDENT DECLARATION FORM

Type of Award Doctor of Philosophy

School School of Nursing _

1. Concurrent registration for two or more academic awards

I declare that while registered as a candidate for the research degree, I have not been a registered candidate or enrolled student for another award of the University or other academic or professional institution

_____

2. Material submitted for another award

I declare that no material contained in the thesis has been used in any other submission for an academic award and is solely my own work

_____

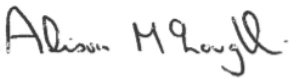
3. Collaboration

Where a candidate's research programme is part of a collaborative project, the thesis must indicate in addition clearly the candidate's individual contribution and the extent of the collaboration. Please state below:

4. Use of a Proof-reader

No proof-reading service was used in the compilation of this thesis.

Signature of Candidate



Print name: _____Alison McLoughlin_____

Abstract

Early neurological deterioration (END) is a poorly defined, but common complication significantly affecting outcome post-stroke. This thesis uses mixed methods in 3 phases to inform development of a consistent approach for neurological assessment and monitoring in acute stroke; allowing a range of staff to promptly identify changes, and take corrective action.

Phase 1

A scoping review identified 26 scales for neurological assessment and monitoring in acute stroke. Several reviews allowed comparison of the clinimetric properties of 20 scales where data utility was available. There was limited evidence to support the use of specific scale(s), and none had been fully tested across a whole stroke population. Yet, the review clarified the importance of assessments allowing early detection of change in individual items, rather than the total score, for END detection. The review clarified the key clinimetric properties to be established by future research.

Phase 2

A UK-wide survey of stroke units (n=125) demonstrated extensive variation in neurological assessment and monitoring practice. Most units use the Glasgow Coma Scale or AVPU (Alert, Voice, Pain, Unresponsive), for monitoring, which are not stroke-specific and only highlight late signs of deterioration (e.g. altered consciousness).

Phase 3

Semi-structured interviews (n=23) utilising Normalisation Process Theory explored current practice and barriers and facilitators for implementation of a new standardised assessment. Staff recognised the need for better guidance and practice change for this important element of care.

An expert group agreed on the Standardised Neurological OBservation Schedule for Stroke (SNOBSS). The SNOBSS and decision flowchart were presented to clinicians to consider acceptability and implementation concerns.

Recommendations which reduce variations in clinical practice and inform future research will progress SNOBSS development and implementation. SNOBSS development is a first step towards consistent stroke-specific monitoring to identify neurological changes, specifically END in acute stroke.

Table of contents

List of Appendices

List of Tables

List of Figures

Abbreviations

α	Cronbach's Alpha
AIS	Acute ischaemic stroke
ASU	Acute Stroke Unit
AVPU	Alert, Voice, Pain, Unresponsive scale
BI	Barthel Index
BIASP	British & Irish Association of Stroke Physicians
BSV	Between-Subject Variation
CAS	Complex Adaptive System
CFA	Confirmatory Factor Analysis
COSMIN	COnsensus-based Standards for the selection of health status Measurement INstruments
CNS	Canadian Neurological Scale (or CNSS Canadian Neurological Stroke Scale)
CRN	Clinical Research Network
CRSU	Complex Review Support Unit
CSS	Chinese Stroke Scale
CT	Computerised Tomography
DoH	Department of Health
EA	Expected Agreement
ECASS	European Cooperative Acute Stroke Study
ED	Emergency Department
EFA	Exploratory Factor Analysis
END	Early Neurological Deterioration
e-NIHSS	Expanded National Institutes for Health Stroke Scale
EHR	Electronic Health Record
ESS	European Stroke Scale
EWS	Early Warning Score
FA	Factor Analysis
FAST	Facial weakness, Arm weakness, Speech problems, Time to call 999 test
FOUR Score	Full Outline for UnResponsiveness Score
GCS	Glasgow Coma Scale
GDPR	General Data Protection Regulation
GRADE	Grading of Recommendations, Assessment, Development and Evaluation (GRADE) principles
HASU	Hyperacute Stroke Unit

HCA	Health Care Assistant
HCP	Health Care Professional
HRA	Health Research Authority
HRQL	Health-Related Quality of Life
HSS	Hemispheric Stroke Scale
IC	Internal Consistency
ICC	Intraclass Correlation Coefficient
ICH	Intracerebral Haemorrhage
ICP	Intracranial pressure
IRT	Item Response Theory
ISD	Information Services Division
IVBSS	Israeli Vertebrobasilar Stroke Scale
k	kappa statistics
LoA	Limits of Agreement
LOC	Level of Consciousness
LVO	Large Vessel Occlusion
MCA	Middle Cerebral Artery
MCANS	Middle Cerebral Artery Neurological Score
MDT	Multi-Disciplinary Team
MEND	Miami Emergency Neurological Deficit
MeSH	Medical Subject Headings
MESSS	Modified Edinburgh- Scandinavian Stroke Scale
MIC	Minimally Important Change
mK	Mean kappa
MK	Modified kappa
MMSE	Mini-Mental State Examination
mNIHSS	Modified National Institutes for Health Stroke Scale
MRC	Medical Research Council
mRS	Modified Rankin Score
n	Number
ND	Neurological Deterioration
NEWS	National Early Warning Score
NGT	Nominal Group Technique
NHS	National Health Service
NICE	National Institute for Clinical Excellence
NIHR	National Institute for Health and Care Research

NIHSS	National Institutes for Health Stroke Scale
NIRSA	Northern Ireland Statistics and Research Agency
NPT	Normalisation Process Theory
NPV	Negative Predictive Value
NSNF	National Stroke Nursing Forum
OA	Observed Agreement
OPTIMIST	Optimal Post-Tpa-Iv Monitoring in Ischemic Stroke
PA	Percentage Agreement
PCA	Principal Component Analysis
PIC	Participant Identification Centre
PPV	Positive Predictive Value
PSAT	Post Stroke Assessment Tool
PSS	Poststroke seizure
RCP	Royal College of Physicians
RA	Rasch Analysis
RCN	Royal College of Nursing
RCP	Royal College of Physicians
REC	Research Ethics Committee
SDC	Smallest Detectable Change
SEM	Standard Error of Measurement
SIAS	Stroke Impairment Assessment Scale
SIGN	Scottish Intercollegiate Guidelines Network
SIP	Stroke in Progression Scale a shortened version of the SSS
sNIHSS	Shortened or Simplified National Institutes for Health Stroke Scale
sNIHSS-EMS	Shortened National Institutes for Health Stroke Scale for Emergency Medical Services
SNOBS	Standardised Nursing OBservation for Stroke
SNOBSS	Standardised Neurological OBservation Schedule for Stroke
SOP	Standard Operating Procedure
SSEF	Stroke Specific Education Framework
SSNAP	Sentinel Stroke National Audit Programme
SSS	Scandinavian Stroke Scale (or SNSS Scandinavian Neurological Stroke Scale)
STEMH	Science, Technology, Engineering, Medicine, and Health
STOC	Stroke Thrombolysis Observation Chart
TBI	Traumatic Brain Injury
TM	Telemedicine

TSS	Toronto Stroke Scale
UCLan	University of Central Lancashire
UCV	Unwarranted Clinical Variation
UK	United Kingdom
WHO	World Health Organisation
Wk	Weighted kappa
WSV	Within-Subject Variation

Acknowledgements

I would like to start by thanking the NIHR for the funding and opportunity of the Fellowship that has allowed not only the development of this thesis but for me to undertake a robust research training programme. There are so many people I would like to thank both in my professional and private lives for supporting me before and during this Fellowship, however, I am unable to name them all here but will thank you personally at the first chance I get.

Extreme thanks must go to my supervisory team who have individually and collectively helped me in more than they will ever know and to whom I will always be grateful. Professor Liz Lightbody my Director of Studies has been a rock by my side for more years than both of us would probably like to admit. I will never be able to thank her enough for her patience and support through good times and bad and her unending help in me finally getting here. Professor Dame Caroline Watkins is an inspiration in terms of helping others develop their research careers and without her support and vision I would not be where I am today and for that I especially thank her. Dr Philippa Olive, you have been an integral part of the supervisory team and my personal and professional development for many years now and long may that continue. Thank you for helping me grapple with things in a friendly and safe environment, I have grown in so many ways with your support. Professor Chris Price, you have a special gift for keeping me grounded and focused on what matters and I am indebted to you.

I thank Professor Ruth Harris for providing me with mentorship and valuable advice since I met her at my Fellowship interview. I would like to thank her for helping me navigate the Fellowship process with a vision of the future. So many people supported me to take this opportunity, but special thanks must go to Gemma Whiteley, Kenny Finlayson (Research Design Service), the NIHR CLAHRC NWC, the Clinical Academic Faculty, the Research Development Group, and the Lay Research Group at Lancashire Teaching Hospitals NHS Trust as well as my supervisory team for supporting my NIHR Fellowship application.

This thesis would not have been possible without the support of the NIHR Complex Review Support Unit, the NIHR Clinical Research Networks, UK-wide Research and Development Departments, the Expert Group members, and the individual participants. Thank you all for finding the time to support this project, especially with everything else the world was throwing at you at the time. Special thanks must go to Professor David Barer as his previous influential work in this area helped shape several ideas for this thesis. He was also an extremely valuable

member of the expert group where he was able to impart his extensive knowledge and experience.

Thanks to all the Stroke Research Team but especially to Kulsum Patel, Steph Jones, Rachel Stockley, and the rest of the old BB445 crew for words of encouragement whenever they were able. I also need to thank Ellie Smith and Jane Gibbon for their help and positivity.

COVID had multiple impacts on the project but the most positive one has been the development of a virtual network that has kept me motivated and enabled me to succeed with a host of writing retreats. Special mentions to the Post Graduate Research Society, Tina Mckee, Dr Emma Jones, Nicola Gaskins, and Dr Dorothy Hardy. I look forward to continuing the motivation and who knows meeting in person one day.

I want to end by thanking all my family and friends for putting up with me and helping me to be able to pursue this process to the end. Even those no longer with us have driven my desire to succeed and I hope they would have been pleased. Hazel Dickinson, Susan Gallagher, Julie Grant and Cat LaMoon deserve a special mention for totally being there whenever I have needed you despite my absence. Ian my husband has been my guiding light through everything the world throws at us and again he has not let me down thank you for believing in me as always, and totally sustaining me throughout this process you are my hero. Finally, to Tegan and Niall, you are my reasons for striving and without you, there would be no purpose. You have been absolutely amazing throughout this time, and I am so proud of the people you are becoming. I hope this process will encourage you to follow your dreams, work hard, and aim high (attention to detail).

Chapter 1 Introduction

This chapter introduces the topic of stroke and the importance of organised stroke care. It will outline early neurological deterioration (END) as a potential complication after stroke and establish the importance of neurological assessment and monitoring. There will be a discussion of how and why variation in neurological assessment and monitoring may exist, and some of the factors that could influence this from individual to system levels. The chapter will highlight the current lack of evidence and thus guidance in how neurological assessment and monitoring should be completed and the potential benefits of addressing this. It will conclude with an overview of the PhD and the thesis structure by chapter.

1.1 Stroke and its impact

The World Health Organization's (WHO) definition of stroke is: "rapidly developing clinical signs of focal (or global) disturbance of cerebral function, with symptoms lasting 24 hours or longer or leading to death, with no apparent cause other than of vascular origin." (WHO MONICA Project Investigators, 1988 pg.108). Currently, there are more than 100,000 strokes per year across the British Isles (Royal College of Physicians Sentinel Stroke National Audit Programme [SSNAP]; 2017b; Information Services Division [ISD] Scotland, 2017). Stroke is the fourth single leading cause of death in the UK (National Records of Scotland, 2016; Northern Ireland Statistics and Research Agency [NIRSA], 2016; Office for National Statistics, 2016). Although stroke mortality has halved in the last two decades (NHS Digital, 2018) the disabling effects remain devastating for people's lives. Stroke is the leading cause of disability in the UK: almost two thirds of stroke survivors leave hospital with a disability (Stroke Unit Trialists' Collaboration, 2013). Stroke incidence is continuing to rise and unless further action is taken, the number of survivors left with disability after stroke is set to increase by a third (Patel et al., 2017). An established way to reduce death and dependency is through good quality organised acute stroke care (Stroke Unit Trialists' Collaboration, 2013).

1.2 Acute stroke care

The importance of specialist stroke care, and its benefits for all stroke patients, has been known for decades. The National Stroke Strategy (Department of Health, 2007) highlighted important components of an acute stroke unit which included rapid treatment of stroke and associated complications. Common codes of practice around assessment, management, and rehabilitation have been identified as characteristics of an effective package of stroke care (Langhorne & Pollock, 2002). Acute stroke care is particularly challenging because it encompasses multiple interventions within a complex system. This underlying complexity

means that many elements remain poorly defined and inconsistently implemented leading to imprecision in care practices (Langhorne et al., 2010). Many elements of care are interrelated making it difficult to ascertain which components make acute stroke care perform better than general wards. However, it is thought that improved patient outcomes can be achieved through comprehensive and standardised processes, especially in relation to the prevention and management of complications during the initial hours and days after the stroke (Govan et al., 2007). A key complication of acute stroke is early neurological deterioration (END), and this thesis focuses on its identification and management.

1.3 Early neurological deterioration (END)

END describes the worsening of symptoms in the hours and days following acute stroke. Reported occurrence of END varies across studies with between 5- 40% of stroke patients deteriorating in the first 24hrs (Seners et al., 2015). Incidence of deterioration is known to differ across stroke populations, for example, older patients and those with more severe strokes are at higher risk, however, it is a potential complication for all stroke patients. (Langhorne et al., 2010). There are multiple causes of END which are explained in greater depth in Chapter 2. Many are irreversible and reflect natural progression of the stroke however some are modifiable and treatable.

Prevention of END has been suggested as a key factor contributing to reduced mortality and better outcomes associated with care provision in a formal stroke unit setting (Govan et al., 2007; Roquer et al., 2008). Despite its clinical importance, the study of END and its prevalence whether resulting from stroke progression or other causes has been hindered by the lack of standardisation in definitions and assessment procedures (Birschel et al., 2004; Siegler & Martin-Schild, 2011). These inconsistencies in definition could explain some of the variation in END recognition. There is an extensive range of definitions within the literature in terms of the assessment scale employed, the threshold employed, and the timeframes under investigation (Appendix 1.1)

Disability, and in some cases death, after stroke could be reduced through timely recognition of and appropriate intervention to address modifiable causes of END (Helleberg et al., 2016; Kwan & Hand, 2006). Where END represents progression awareness of deterioration would support care planning and communication with families and carers. Effective neurological assessment and monitoring after stroke to identify END, with an appropriate response has the potential to improve outcomes for acute stroke patients.

1.4 Neurological assessment and monitoring

Neurological assessment and monitoring has for decades been recognised as an important element of acute stroke care in National and International policy and guideline documents. The National Stroke Strategy (Department of Health, 2007) stated, "intensive physiological and neurological monitoring in the early phase of a stroke supports early treatment that halts stroke progression and prevents more brain cells being damaged" (pg.23). The current National Clinical Guidelines for Stroke state that neurological assessment and monitoring is one of the fundamental components of acute stroke care (Royal College of Physicians, 2016a). Despite these clear statements of importance, there are no specific guidelines on the format, content, and frequency of how neurological assessment and monitoring should be achieved except for the small proportion of patients who receive thrombolysis and/or thrombectomy (Jauch et al., 2013; Royal College of Physicians, 2016a). However, only 12% of people who have a stroke will receive these treatments (SSNAP, 2019).

The thrombolysis/thrombectomy monitoring guidelines, which advocate intensive monitoring for the first 24 hours after treatment were based on the NINDS rt-PA trial (National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group, 1995). Until recently the role and required intensity of stroke neurological assessment and monitoring has never been robustly examined in research (Benedetti et al., 2021). Some work is underway with the ongoing Optimal Post-Tpa-Iv Monitoring in Ischemic Stroke (OPTIMIST main) trial which aims to establish whether less intense monitoring is at least as effective (non-inferior) as standard monitoring on functional outcomes for stable acute ischaemic stroke (AIS) patients (Faigle et al., 2020). However, this is only for patients after thrombolysis/thrombectomy treatment so is concerned with updating limited guidance rather than addressing monitoring requirements for all.

The frequency of monitoring assessments, for all stroke patients, is just one element that will impact success or failure (Tarassenko et al., 2006). The intermittent nature and user-dependent nature of neurological assessment and monitoring is also likely to limit its effectiveness. This has been identified with other intermittent monitoring regimens such as Early Warning Scores (EWS) (Downey et al., 2017). The purpose of ongoing monitoring is to identify deterioration in a patient in a timely manner because treatment of its cause(s) where possible could result in better patient outcomes (Time is Brain).

The probability of identifying END will increase with longer time windows between assessments (Martin & Price 2018). However, this could delay identification and it seems logical that the more we assess the more likely we are to detect deterioration. A single centre study found increased neurological monitoring to be beneficial. In this study, semi-intensive monitoring including a neurological component, was associated with reduced mortality at one year in patients with severe stroke, although it showed no influence on dependency (Silva et al., 2005). There is currently no evidence to support optimal time frames or frequencies across the stroke population.

Frequency must be balanced with other factors such as sensitivity to detect change, resources available, and acceptability to both patients and staff. Studies have shown that there are accuracy and compliance problems with the completion and documentation of physiological observations despite them being crucial to the prevention of deterioration (Le Lagadec & Dwyer, 2017; Tysinger, 2015). This could be intensified with neurological monitoring as it is often thought of as more time consuming and complicated for staff (Iacono, Wells, & Finnerty, 2014; Izumi-Richards, & Simon, 2016). If requested too often it may be hard to balance with other workload and may cause disturbance of rest for patients. If called for less frequently detection of deterioration could be delayed, and it could impact on patient experience and outcomes.

Guidance in terms of what needs to be done to ensure clinical effectiveness of neurological assessment and monitoring to detect END in all stroke patients is required. The current lack of guidance for best practice to detect END with neurological assessment and monitoring means different stroke units could have widely varying practices. There are multiple different scales available for assessment and monitoring which are introduced in Chapter 2. It is possible there is widespread variation in the assessment and monitoring regimes that stroke units employ in terms of what (i.e., content), when (i.e., format) and how often (i.e., frequency) neurological assessment and monitoring is completed. Such inconsistency could cause unwarranted clinical variation (UCV) in patient care and outcomes.

1.5 Unwarranted clinical variation (UCV)

This thesis has been inspired by the variation the author has observed in neurological assessment and monitoring in both clinical and research practice over multiple years. Variation in neurological assessment and monitoring could be preventing detection of END resulting in potentially modifiable causes (Chapter 2) going untreated which could be leading to worse

outcomes for patients. UCV is variation that can only be explained by differences in health system performance (Kennedy et al., 2010). Identification and eradication of practices where variation is common is critical to achieve more equitable and evidenced-based healthcare (Harrison et al., 2019).

Currently the level of variation in neurological assessment and monitoring practice is unknown. Potential cases of UCV should trigger an exploration of that aspect of the care pathway (Harrison et al., 2019). Many aspects of clinical practice, such as neurological assessment and monitoring, remain empirically based on previous observation or experience rather than having a basis in clinical science (Wennberg, 2002). This research aims to explore new possibilities for neurological assessment and monitoring practice through working at the edge of knowledge and experience (Plesk & Greenhalgh, 2001). To achieve an understanding of UCV in current practice of this complex element of care not only the level of variation but the interplay of factors that affect it need to be explored.

Information on variation allows examination of the systems affecting clinical decisions and raises important questions concerning the efficiency and effectiveness of healthcare (Wennberg, 2011).

One way of exploring this element of care in more depth is to adopt a systems view that would take into account the people involved and the wider systems in which they operate (Clarkson, 2018). There are many factors that may lead to variation in clinical practice such as an individual's knowledge, the working practices of the team, or hospital policies. A systems approach underpinned by complexity theory would emphasise the importance of understanding the underlying factors that interconnect and affect neurological assessment and monitoring practice.

Complexity theory acknowledges that the world, like healthcare, is complex, non-linear, dynamic, and unpredictable (Zimmerman et al., 1998). The overall function of any element of care is affected by components of the system in which it operates and their relationships and interactions with each other. (Braithwaite, 2018). Healthcare is itself a Complex Adaptive System (CAS) incorporating a dynamic network of different independent agents whose goals and behaviours can conflict with unpredictable results. Such systems cannot be reduced to individual components as it is the interactions that result in the overall behaviours (Thompson et al., 2016).

This study will explore the interrelationship and interactions of factors that impact on neurological assessment and monitoring practice within the CASs in which it is delivered (Chapter 3). However, to structure where these components fit within the healthcare setting it will adopt an individual, unit, and organisational level systems framework. Through examining the conditions (or contexts) of this element of care across these levels it may be possible to explain why variations might occur (Gill & Turbin, 1999; Greener & Mannion, 2009). This will help identify factors that could be targeted to help reduce variation and implement future changes in practice.

1.6 Factors linked to variation in performance of neurological assessment and monitoring

The previous section introduced why factors at individual, unit, and organisational levels should be explored. This section outlines some factors that have been shown to impact and others that could potentially influence neurological assessment and monitoring practice. Awareness of these factors has helped shape the design and delivery of the research. However, this research is open to the discovery of known and unknown factors and contexts that impact on current practice in neurological assessment and monitoring practice and will utilise a mixed-methods approach to explore them (Chapter 3).

1.6.1 Individual

At the individual level, performance of neurological assessment and monitoring is directly influenced by differences in the way it is undertaken, interpreted and rated (Birschel et al., 2004). Neurological assessment has been found to differ widely between colleagues on a given unit (Gocan & Fisher, 2008). This could occur for many reasons such as the different professional group of the individuals, their experience, the training they have received, the scale they use, or the way they assess neurological status. In the author's masters study inter-rater reliability utilising one stroke specific assessment scale was found to be poor with wide ranging variation in scores on individual items of assessment (Mcloughlin et al., 2022). Variation can be further amplified by differences in the way assessments are reported (Iacono et al., 2014). Collectively these factors could accumulate and further magnify UCV in busy acute stroke unit environments.

1.6.2 Unit

Neurological assessment and monitoring practice has been shown to vary within and between hospitals (Gocan & Fisher, 2008; Iacono et al., 2014; Wells-Pittman, 2020). Without guidance

on how, when, and how often to complete neurological assessment, stroke units are likely to be completing this element of care differently. Differences in outcome have been shown in other elements of stroke unit care through Sentinel Stroke National Audit Programme (SSNAP) data (e.g., mortality improves for patients treated in stroke units with higher nurse staffing levels) (Royal College of Physicians, 2016b). SSNAP has been collecting data since January 2013. SSNAP is undertaken by all stroke units in Great Britain and Northern Ireland and measures processes of care (clinical audit) and the structure of stroke services (organisational audit) against evidence-based standards with a view to monitor and improve standards of care. SSNAP data is reported for over 95% of all stroke admissions in England, Wales, and Northern Ireland.

Currently, neurological assessment and monitoring data collected in the SSNAP clinical dataset comprises of:

- a baseline National Institutes for Health Stroke Scale (NIHSS) for all stroke patients entered in the audit
- a follow-up NIHSS score at 24 hrs for patients who receive thrombolysis.

This limited dataset does not provide any detail on ongoing assessment and monitoring practices. Current practice and the extent of variation across stroke units within the UK, is unknown along with the factors that influence it.

1.6.3 Organisational

Wider organisational level factors that may contribute to variation in neurological assessment and monitoring include staffing shortages, service design, and delivery systems. Staffing shortages, within the NHS generally and specifically in stroke, are likely to have effects on service provision and how elements of care are delivered now and in the future. A report published by the British & Irish Association of Stroke Physicians (BIASP) reported a growing shortage of stroke consultants in the UK and that lack of specialists is limiting the ability of the NHS to deliver high-quality stroke care (Hart et al., 2019).

There is a growing need to look at staffing in more flexible ways which may mean that traditional job roles may merge across disciplines so that the highest level of stroke care can be delivered (SSNAP, 2019). Without adequate levels of skilled staff, there are limitations in what care can be delivered. It is important to understand what capacity there is in the system

for staff to undertake specific elements of care, such as neurological assessment and monitoring.

Total staffing numbers are not the only consideration in terms of staffing because to provide effective stroke care the multidisciplinary team must also be skilled (MDT) (Royal College of Physicians, 2016a). Ongoing education and training are considered an essential foundation for safe, effective care (Greatbatch, 2016). Lack of knowledge, skills, and competence represent barriers to optimal evidence-based practice in stroke care (Baatiema et al., 2017). High-quality stroke-specific education and training are supported through the Stroke-Specific Education Framework (SSEF) (Health Education England, 2022). However, it is not clear what if any specific training to support neurological assessment and monitoring is provided to staff. There are also reported barriers to healthcare staff receiving education and training including time, accessibility, and financial issues (Ward & Wood, 2000).

Resource issues in terms of stroke unit bed capacity is another organisational factor that impacts the delivery of quality stroke care. As well as total bed capacity the organisation and appropriate use of stroke beds impact on care quality. Hospital crowding has been shown to reduce the likelihood of patients being directly admitted to a stroke unit despite the evidence that acute stroke care is associated with better outcomes (Darehed et al., 2017; Moore, 2022).

Current improvement drives to re-organise stroke services to provide more centralised and specialist stroke services are underway due to clear evidence that these service designs that are better staffed, have the latest equipment, and are open 24 hours a day provide better clinical outcomes (Davie et al., 2013; Hunter, Davie et al., 2013; Hunter, Fulop et al., 2018). Different service designs and delivery systems have the potential to generate differences in the delivery of elements of care.

Another factor that has the potential to affect levels of variation in neurological assessment and monitoring is the increased use of electronic observation systems and electronic health records (EHRs) across the NHS (Hodgson et al., 2021). These systems have the potential to reduce variation if they are developed around guidelines with a view to improve care provision (Pelletier, 2010). However, there is potential that these systems can create variation, with multiple ways to perform a task or the design not being congruent with the work to be done (Thomas et al., 2020).

Neurological assessment and monitoring practice is an element of care with limited clinical guidance. Even where best practice is known, such as all patients should be admitted to a stroke unit within four hours extensive variation have been highlighted, (NHS RightCare, 2017; SSNAP, 2019). Exploration of current practice in neurological assessment and practice to explore levels and identify factors that cause variation is warranted.

1.7 PhD overview

1.7.1 Aim

This thesis aims to begin the development of a consistent plan of how the neurological effects of stroke should be assessed, recorded, and monitored over time through development of a Standardised Neurological Observation Schedule for Stroke (SNOBSS).

1.7.2 Objectives

This PhD programme of research will:

- Identify the scales used or available for neurological assessment and monitoring in the acute phase of stroke.
- Develop an overview of the clinimetric properties (e.g., reliability, responsiveness, measurement error, time to complete) of the scales identified to create a synopsis of the effectiveness of the scales to detect END.
- Establish current practice and briefly explore clinicians' experiences of neurological assessment and monitoring in the acute phase of stroke. This will clarify the current level of variation in neurological assessment and monitoring across the UK.
- Determine knowledge, understanding, and acceptability of neurological assessment and monitoring in acute stroke and explore the barriers and facilitators to its implementation in clinical practice.
- Agree the content and design of the Standardised Neurological Observation Schedule for Stroke (SNOBSS).
- Provide recommendations* for ways to improve standardisation and reduce variation in current clinical practice.
- Identify the next steps for development of a consistent plan of how the neurological effects of stroke should be assessed, recorded, and monitored over time through recommendations * for future research.

*all recommendations will be derived from key findings of all the work within the thesis

The aim is to develop a tool (SNOBSS) based on evidence collected through the PhD. It will incorporate the items that there is evidence to suggest are best to assess changes in neurological status, specifically END. An associated decision flowchart will also be developed to maximise consistency and the process of escalation based on local expert decisions about the patient condition and what could be done if deterioration is noted. The SNOBSS is not a scale to measure deterioration, but rather a tool to allow better identification of END, improve communication and understanding of patient status, and contribute toward optimum patient care and outcomes (Theofanidis et al., 2015).

If neurological assessment and monitoring after stroke was evidence-based and standardised it would have the capability to ensure effective stroke care provision, improve efficiency, prevent deaths and disability caused by stroke, and reduce UCV and health costs. Although not all END is treatable, early recognition of deterioration will allow better prognostication, communication, and end of life planning for patients and families. The following section describes the thesis structure before the chapter summary.

1.8 Thesis structure

This thesis has eight chapters:

Chapter One:

Presents the main aim and objectives of the programme of research and thesis. It will introduce neurological assessment and monitoring as an important element of organised stroke care, and how END, despite a lack of definition, is a potential complication for all stroke patients. The potential for variation in clinical practice of neurological assessment and monitoring is highlighted. It also how multiple factors need to be explored to understand current practice and the present level of variation that exists before a Standardised Neurological OBservation Schedule for Stroke (SNOBSS) can be developed.

Chapter Two:

Discusses the importance of neurological assessment and monitoring to identify changes in neurological status. It provides an overarching justification for the work, including:

- Discussion that improved consistency of assessment and monitoring, better recognition of neurological deterioration and timely instigation of intervention/s could improve outcomes for patients.

- Known mechanisms for Early Neurological Deterioration (END) focusing on treatment options to address deterioration in acute stroke patients.
- Commonly used scales in clinical practice for neurological assessment and monitoring in acute stroke patients.
- Awareness that there are multiple actual and perceived barriers that could affect successful implementation of the SNOBSS or other changes to neurological assessment and monitoring practice in the future.

Chapter Three:
The methodology chapter summarises the overall methodology and theoretical underpinnings of the project. It outlines the overall programme of research and describes the mixed methods approach taken. This chapter also includes issues encountered with ethical processes, an overview of alterations to study design and delivery and concludes with a brief outline of the impact COVID-19 had on project delivery.

Chapter Four:
Describes a series of reviews presenting the clinimetric properties of all the scales that are used or available for neurological assessment and monitoring in acute stroke. The results are presented to allow direct comparison within and between different scales across a range of properties. The discussion highlights implications to practice of the results.

Chapter Five:
The extent of variation in clinical practice of neurological assessment and monitoring across the UK was unknown. This chapter presents the Neurological Assessment Practices after Stroke Survey. It describes the development and delivery of a UK wide survey to explore current practice and understanding of neurological assessment. The results provide a synopsis highlighting variations across participating stroke units.

Chapter Six:
Describes the design and delivery of the semi-structured interviews which were completed with a range of staff from sites sampled from the responses received in the survey (Chapter 5). These interviews determined knowledge, understanding and acceptability of neurological observation in acute stroke. Through the application of Normalisation Process Theory (NPT) they identified barriers and facilitators to implementation in this element of care.

Chapter Seven:

Explains the development of the SNOBSS and associated decision flowchart by an expert stakeholder group using consensus techniques. It describes the iterative process of discussion and review of information through which the SNOBSS was created. This chapter also presents opinions provided by interested clinicians when the SNOBSS was presented to them to obtain feedback on face validity and applicability to their pathways.

Chapter Eight:

This chapter provides an overall discussion of the thesis. This includes the main findings of the study. It highlights the strengths and limitations as well as the implications and recommendations for clinical practice and future research. It articulates this study's original contribution to knowledge before providing a short conclusion of the whole project.

1.9 Chapter summary

This chapter has introduced the topic of stroke and the importance of organised stroke care. END has been identified as a potential risk for all stroke patients and neurological assessment and monitoring practice as an important, if complex, element of care. Due to the limited guidance that exists and a range of factors there is expected variation in current practice that needs to be explored. The chapter has outlined the objectives and structure of the thesis with the main aim being the development of a Standardised Neurological OBservation Schedule for Stroke (SNOBSS). The next chapter, chapter 2, provides the background and greater justification for the work completed throughout the thesis.

Chapter 2 Background

Chapter 1 introduced the topic of stroke and the importance of organised stroke care. It presented the significance of stroke neurological assessment and monitoring in relation to the detection of Early Neurological Deterioration (END) and the impact of END on patient outcomes. The potential causes of variation in neurological assessment and monitoring were discussed, including the current lack of guidance in how and when it should be completed.

This chapter provides justification for the work within the thesis by further explaining the need for, and benefits of, creating a Standardised Neurological OBservation Schedule for Stroke (SNOBSS) and associated decision flowchart. It begins by discussing definitions and prevalence of neurological deterioration (ND) generally, and END specifically, explaining why the identification of both is poorly understood. Secondly, it summarises the main potential causes of deterioration and treatments where appropriate. Commonly used scales available for neurological assessment and monitoring are then introduced along with the importance of understanding their measurement properties. Finally, the importance of understanding actual and potential barriers and facilitators, to the future successful implementation of the SNOBSS and associated decision flowchart, will be discussed.

2.1 Neurological deterioration (ND)-definition and prevalence

Multiple terms have been or continue to be used in the literature to describe a clinical worsening of symptom severity in stroke patients (Aslanyan et al., 2007; Birschel et al., 2004; Castillo, 1999; Helleberg et al., 2014 & 2016). Stroke or neurological progression has been used in instances where the worsening is permanent and relates to the progression of the ischaemia, haemorrhage, or tissue necrosis. Neurological deterioration (ND), the focus of this thesis is the more general terminology to describe a worsening in the patient's functional condition (regardless of underlying cause or reversibility) (Castillo, 1999).

ND has been the focus of many studies but there is a lack of standardisation in terminology and assessment procedures which can be clearly seen in the example studies in Appendix 1.1. Incidence of ND varies depending up on the definition used and the timing of assessments. (Britton and Roden 1985; Davalos, Cendra et al. 1990; Jorgensen, Nakayama et al. 1994; Toni, Fiorelli et al. 1995; Itoh, Shioi et al. 1996; Davalos, Toni et al. 1999; Tei, Uchiyama et al. 2000).

There are extensive ranges in the timing of assessments across previous studies including the first 24hr period (Davalos et al., 1999; Ovesen et al., 2015; Siegler et al., 2013), the first 72 hrs (Castillo, 1999; Birschel et al., 2004; Bugnicourt et al., 2011; Helleberg et al., 2016), within 1-7 days (Barber et al., 2004; Castillo, 1999; Davalos et al.,1999; Indredavik et al., 2008; Jorgenson et al., 1994; Kwan & Hand, 2006; Weimar et al., 2005) or even from recovery assessment to 90 days post stroke (Aslanyan et al., 2007). This thesis has focused on neurological assessment and monitoring and recognition of early neurological deterioration (END) within the first 72 hours of stroke. This pragmatic decision has been taken as it is understood most deterioration occurs within the first three days after stroke (Siegler et al., 2017).

Incidence rates of END specifically have also been shown to vary widely across studies, from 16% to 43%, depending upon the definition and how and when it is measured (Britton & Roden, 1985; Davalos et al., 1990 & 1999; Itoh et al., 1996; Jorgensen et al., 1994; Tei et al., 2000; Toni et al., 1995).

There are multiple stroke scales available to assess stroke severity and neurological deficits (formally identified in a scoping search in Chapter 4). In terms of END, multiple different scales have been used within studies that measure different neurological deficits (Appendix 1.1). The utilisation of different scales (measured over varying time intervals) in definitions of END causes confusion regarding identification and incidence in the medical literature. However, even when studies use the same scale, different criteria for defining END are applied leading to variation in what constitutes END. For example, the National Institutes for Health Stroke Scale (NIHSS) is a commonly used scale to identify END but the change in score chosen to signify a deterioration in the patient's condition has differed widely across studies. An NIHSS increase of four or more points has been accepted in many studies (Arenillas et al., 2002; Bugnicourt et al., 2011; Cui et al., 2022; Ovesen et al., 2015). Others have used an increase of 3 or more points (DeGraba et al, 1999), at least a two-point increase (Helleberg et al., 2014; Siegler & Martin-Schild, 2011; Wei et al., 2020), or even a single point increase (Aslanyan et al., 2007). Some of the study definition ranges give precedence to certain signs, such as conscious level in the Scandinavian Stroke Scale (SSS) (Barber et al., 2004; Birschel et al., 2004; Helleberg et al., 2016). Therefore, a drop in level of consciousness, which is linked to poor outcome (Oxbury et al., 1975) should warrant quicker identification and possible intervention regardless of the total score. Overall, it appears the degree of change chosen to represent significant neurological worsening in many studies is often arbitrary. Only one study included agreement on definitions by expert consensus (Birschel et al., 2004).

Studies that define END differ widely in their primary aims and design (Appendix 2.1). Many of the studies are concerned with identifying factors that predict or cause END (Arenillas et al., 2002; Bugnicourt et al., 2011; Davalos et al., 1999; Flemming et al., 1999; Leira et al., 2004; Mayer et al., 1994; Sorimachi & Fujii, 2010; Sun et al., 2012; Weimar et al., 2005). One study focused on the usefulness of a score to predict deterioration (Miyamoto et al., 2017). Other studies focus on outcomes after deterioration (Geng et al., 2017; Kwan & Hand, 2006; Maramattom et al., 2004) or both prediction of END and the subsequent outcomes (Helleberg et al., 2014; Ovesen et al., 2015). Two observe deterioration within evaluation of specific treatments (Roden-Jullig et al., 2003; Wahlgren et al., 2007).

The studies range massively in the numbers of participants included from 24 (Maramattom et al., 2004) to 6483 (Wahlgren et al., 2007) affecting reproducibility and statistical power. Various sampling criteria could cause issues with selection bias in some studies. All the studies expect two focus on specific stroke types (Birschel et al., 2004; Kwan &Hand, 2006). Although this is potentially justified as different stroke types have different risk factors for deterioration it means that the results are bot generalisable across the whole stroke population. Several studies are based purely on secondary or retrospective data analysis which provide inferior evidence to prospective studies (Cui et al., 2022; Davalos et al., 1999; Mayer et al., 1994; Ovesen et al., 2015; Sun et al., 2012).

Due to the differences in studies including across definitions and study design meta-analysis would be complex, if possible, across some studies further limiting generalisability. Given the heterogeneity of the populations studied, lack of robust methods in studies and the inconsistencies in definitions of END, it is not surprising that the incidence rates vary considerably and that there is no clear guidance in clinical practice in terms of how best to identify END through neurological assessment and monitoring. END will remain difficult to quantify whilst problems with definitions and lack of agreement on the assessment processes persist.

Despite varied definitions and frequency of END across studies END has been shown to be a fundamental risk to all stroke patients and consistently predicts poor outcomes (Seners et al., 2015; Siegler et al., 2013). Theoretically prevention or reversal of END could improve outcomes for patients but the use of a standardised framework for recognition and reporting is essential. Clinical practice needs clearer, more standardised guidance in terms of what assessment needs to be completed and when, to be able to accurately identify END (Birschel et al., 2004; Helleberg et al., 2014 & 2016; Roden-Jullig et al., 2003; Siegler et al., 2013). Standardisation of

the definition of END and the assessments used to identify it would allow more comparable results across organisations and countries and should improve quality by reducing variation and improving efficiency (Kwan & Hand, 2006).

2.2 Early neurological deterioration (END)- risk factors and underlying causes

Reasons for END after stroke can sometimes be difficult to separate from other factors, notably the underlying pathophysiology of the stroke itself. This section will summarise a wide variety of possible risk factors and mechanisms that have been associated with END in acute stroke. These will be presented under the headings of irreversible or modifiable. Irreversible factors cannot be addressed by treatment whereas modifiable ones often reflect secondary or non-neurological causes and are potentially treatable.

2.2.1 Irreversible risk factors associated with END

There are intrinsic irreversible risk factors for END linked to the individual or the stroke itself.

2.2.1.1 Individual factors

Risk factors linked to the individual include time since stroke onset, age, coronary heart disease, degree of carotid stenosis, triglyceride levels, biochemical factors, and history of diabetes.

Natural patterns of END have been suggested such that patients who present early are more likely to deteriorate than those who present with symptoms later. However, no obvious causes for this have yet been identified and it could be due to END having already occurred in patients who present later, and so can be considered an irreversible consequence of the underlying stroke (Weimar et al., 2005).

Increasing age has often been cited as a risk factor for both ND and END (Britton & Roden, 1985; Davalos et al., 1990; Jorgensen, Nakayma, Raaschou et al., 1994; Toni et al., 1999). Although increased age seems to indicate an increased risk of suffering from END it appears patients of any age are at risk (ibid).

Coronary Heart Disease was identified as a risk factor for END after stroke by The European Cooperative Acute Stroke Study (ECASS I) study. They concluded that poorer collateral blood supply when a higher prevalence of severe extracranial or intracranial atherosclerotic disease

is responsible (Davalos et al., 1999). Strokes caused by carotid large vessel disease have been associated with early onset and higher rates of END (Khatri et al., 2012; Kim et al., 2013). Systemic atherosclerosis may also affect the risk of END; however, there is still no certainty about the causes and prediction of the deterioration (Sümer & Özön, 2018).

High serum lipids including triglycerides are a well-known risk factor for stroke but levels at both ends of the normal range have been shown to impact after the stroke. High and low triglyceride values are associated with haemorrhagic transformation, and low triglyceride values are associated with an increased risk of deterioration (Choi et al., 2012).

A previous history of diabetes has been associated with poorer outcomes after all stroke types by most, but not all, studies (Lau et al., 2019). As well as increased incidence of END with previous diabetes it is also associated with increased severity (Davalos et al., 1990; Helleberg et al., 2014; Tang et al., 2016).

2.2.1.2 Stroke factors

Risk of END will vary depending upon the direct mechanism of the stroke and the subsequent pathophysiological chain of events. Stroke, whether ischaemic or haemorrhagic in origin is an abrupt neurological condition caused by impaired blood and oxygen supply to areas of the brain (Kuriakose & Xiao, 2020). If this lack of perfusion cannot be rectified cell death will occur which causes disruption of the plasma membrane leading to swelling and oedema which could lead to mass effect and herniation (Liang et al., 2015).

In terms of stroke type, both intracerebral haemorrhage (ICH) and ischaemic are prone to END but the timing has been shown to differ. ICH patients are prone to deterioration earlier, within the first 24 hrs (Ovensen et al., 2015; Yu et al., 2020). Whereas, in ischaemic strokes, the first 48 hours are associated with potential instability and worsening (DeGraba et al., 1999; Summers et al., 2009). Related pathological mechanisms that could impact on development of END include perilesional blood flow reduction, haematoma expansion, clot progression, inflammation, energy failure, loss of homeostasis, acidosis, increased intracellular calcium levels, excitotoxicity, free radical-mediated toxicity, cytokine-mediated cytotoxicity, complement activation, impairment of the blood–brain barrier, activation of glial cells, oxidative stress and infiltration of leukocytes (Brouwers & Goldstein, 2012; Gelderblom et al., 2009; Helleberg et al., 2014; Qureshi et al., 2003; Suh et al., 2008; Wang et al., 2007; Woodruff et al., 2011).

Various stroke related irreversible risk factors have been associated with END including stroke severity and type. There is some evidence to suggest END risk is mainly determined by the severity and extent of early injury (Arenillas et al., 2002; Cuadrado-Godia et al., 2013) however, deterioration can occur across all stroke severities (DeGraba et al., 1999). It is perhaps not surprising that END has been shown by some to be more common in patients with larger infarct volumes (Davalos et al., 1990, Jorgensen et al., 1994). In larger vessel occlusion (LVO) strokes a positive correlation has been shown between END and the presence of an internal carotid artery (ICA) or untreated middle cerebral artery (MCA) occlusion (Weimar et al., 2005). MCA flow velocity changes and poor cerebral hemodynamic reserve have also been linked to a more severe final outcome (Alexandrov et al., 2004; Baizabal-Carvallo et al., 2014). However, lacunar infarctions have been shown to have greater association with END when compared to other ischemic stroke subtypes with similar severity, possibly because symptoms can evolve over a longer time interval and it is easier to demonstrate a change (Steinke & Lay, 2002; Tei et al., 2000; Yamamoto et al., 1998).

Biomarkers are naturally occurring indicators that predict physiologic or disease states, or increased disease risk (Kim, Moon & Bang, 2013). Several proinflammatory cytokines are released early after the onset of brain ischemia, but it is unknown whether inflammation predisposes to neurological deterioration. The evidence concerning the association of excitotoxic amino acids and proinflammatory cytokines in CSF and blood to subclinical stroke course and prognosis is growing rapidly (Martin & Price, 2018). Different substances and stroke subtypes are being investigated but glutamate and glycine (Castillo et al., 1997), interleukin (IL)-6 and tumour necrosis factor (TNF)-a (Vila et al., 2000) have been identified as positively correlated with END after acute stroke. Despite these advances, around half of those patients that deteriorate are stated to have no clear mechanism and more scientific advances are required to understand whether there are indicators of reversible tissue injury (Seners et al., 2015).

2.2.2 Modifiable factors associated with END

Potentially modifiable factors associated with END include high body temperature, infection, electrolyte or glucose abnormalities, hypoxemia, extremes of or variation in blood pressure, seizure, and medication effects. It is important that END is identified in a timely manner so that if modifiable causes are found they can be treated appropriately, although it remains unclear from much of the current evidence whether these factors are causes and/or consequences of END.

2.2.2.1 Temperature and infection

The link between elevated body temperature and END is well established (Helleberg et al., 2014; Weimar et al., 2005,). Two separate meta-analyses have shown that high body temperature after stroke leads to significantly higher morbidity and mortality (Greer et al., 2008; Hajat et al., 2000). Between 40% and 61% of patients after stroke will develop a fever (Azzimondi et al., 1995; Castillo et al., 1998). Hyperthermia could increase the metabolic demands of the brain (Nemoto & Frankel, 1970), cause changes in the blood brain barrier, and promote acidosis and release of excitatory neurotransmitters (Busto et al., 1989). Experimental models have shown that hyperthermia increases cerebral lesions and the volume of infarcted tissue (Busto et al., 1987; Ginsberg et al., 1992). The most frequent cause of fever after stroke is infection but it can also be an expression of cell necrosis or change in thermoregulatory mechanisms (Powers & Sheld, 1996; Przelomski et al., 1986).

Thirty percent of stroke patients will develop infections in the first week post-stroke (Aslanyan et al., 2004). Infection post stroke increases the likelihood of death (Heikinheimo et al.,2013) and the extent of disability (Wartenberg et al., 2011). These data, combined with the growing body of evidence from other neurological disorders (Murta & Ferrari, 2013) indicate that infection has a detrimental effect on damaged brains and can cause rapid and severe deterioration. As well as bringing about fever, infection can also facilitate electrolyte imbalance and hypoxemia, other potential causes of END, which could theoretically cause further cell death within the ischemic penumbra (Ginsberg & Busto, 1998).

2.2.2.2 Electrolyte or glucose abnormalities

Severe electrolyte imbalance can affect neurological functioning and is a potential cause of END in stroke patients. For instance, hyponatremia or hypernatremia which are more common in neurological patients, cause changes in the brain cells due to changing plasma osmolarity and can have potentially devastating neurological effects (Tisdall et al., 2006).

High serum glucose levels on admission are also associated with END regardless of previous history of diabetes (Siegler et al., 2013). Stress hyperglycaemia is a well-recognized phenomenon that can occur within hours of stroke in people with and without pre-existing diabetes (Lau et al., 2019). Stress hyperglycaemia has been associated with poorer outcomes and END, particularly in those without known diabetes (Shimoyama et al., 2014), and left untreated can result in cell death (Cryer, 2007).

2.2.2.3 Hypoxaemia

Hypoxaemia is common after stroke and can be attributed to pneumonia, aspiration and respiratory muscle dysfunction, sleep apnoea, pulmonary embolism, and cardiac failure. Although the brain has several vascular adaptations to cope with a certain threshold of hypoxemia maintaining adequate saturation, and treating the underlying cause, is vital to prevent further damage to the brain (Ferdinand & Roffe, 2016). Routine supplementary oxygen therapy for all patients is not advocated but in those found to be hypoxic, oxygen could prevent or treat END by ensuring sufficient oxygen to prevent further brain cell death (Roffe et al., 2017).

2.2.2.4 Blood pressure

Observational studies and post hoc analyses of clinical trials have demonstrated higher rates of END and worse outcomes in patients presenting with the extremes of blood pressure as well as with haemodynamic variability (Sare et al., 2009; Vitt et al.,2019). There are numerous potentially treatable causes for elevation in blood pressure after stroke including pre-existing hypertension infection, pain, stress, and raised intracranial pressure. Acute intervention to reduce elevated blood pressure is recommended in ICH within specific limits (NICE, 2019; NICE, 2021). However, in acute ischaemic stroke, there is no evidence to support lowering blood pressure unless the patient is receiving reperfusion therapy (Bath et al., 2018). Low blood pressure, although less common can also cause END due to impaired cardiac output and potential causes include sepsis, cardiac arrhythmias such as atrial fibrillation (AF), heart failure, and hypovolaemia (Appleton et al., 2016).

2.2.2.5 Seizure

Poststroke seizure (PSS) can impact neurological status both during the event and in the postictal phase, whilst it can occur early or late in the clinical course of stroke, this discussion is limited to early acute onset seizures that occur within 24 hrs of stroke onset. It is not always obvious that seizure activity is occurring due to their focal nature or stroke-related paralysis, and sudden changes in physiological observations may be the only indicator. ICH, more severe stroke (regardless of type), and those affecting the cerebral cortex are at increased risk of PSS. Seizures after ICH are attributed to irritation by products of blood metabolism. Several causes have been suggested for early PSS following ischaemic stroke including hypoxemia, metabolic dysfunction, hyper, and hypo perfusion, and disturbances in electrophysiology (Myint et al., 2006). PSS management is awaiting reliable clinical practice guidelines around monitoring and treatment (Xu, 2019).

2.2.2.6 Medication

Initiation and abrupt discontinuation of some medications have been suggested to acutely impair neurological function after brain injury (Goldstein, 1995), including centrally acting drugs such as benzodiazepines, opiates, and anticonvulsants (phenytoin and phenobarbital). A full and accurate drug history is important to identify anything that could impact on the patient's neurological function and be an explanation for END.

2.2.3 Summary

Multiple factors including age, stroke severity, type, and time since stroke, have been shown to put patients at a greater risk of END and it is appropriate to monitor these patients more closely. However, END is a potentially serious and unpredictable complication following a stroke and all patients should receive monitoring to identify change in their condition. Identification, and where possible treatment of the underlying cause of END should occur to improve outcomes.

There are potentially modifiable factors that if identified and treated quickly may prevent END from persisting and causing secondary brain injury. Recognition of END and the identification and treatment of remediable factors, where they exist, is vital to improving outcomes (Birschel et al., 2004). Even when nothing can be done to reverse these underlying causes of END it is important deterioration is still identified as some patients could be eligible for urgent transfer to a neuroscience centre for treatment such as craniectomy or hypothermia (Georgiadis et al., 2002; Schwab & Hacke, 2003).

The literature identifying risk factors and causes of END after stroke is often contradictory. This could be largely due to the differences in trial design including definitions and timeframes utilised. Many studies look retrospectively for predictors of deterioration with specific causes in mind and no way of accounting for confounding variables (Cui et al., 2022; Davalos et al., 1999; Mayer et al., 1994; Ovesen et al., 2015). Much of the END literature also pre-dates the introduction of treatments such as thrombolysis and thrombectomy. Patterns of END and identification of underlying causes reported here could be confounded by treatment effects (Alexandrov & Grotta, 2002).

Greater standardisation in assessment could lead to better recognition of END. Prospective monitoring might help predict groups that would benefit from advanced imaging or

measurement of biomarkers which in time might provide insight into the mechanisms involved in END. With increased recognition of the mechanisms, further treatment options may become available (Siegler et al 2016). There is potential that in the future it might be possible to develop and deliver therapies targeting these sequelae or even allow prophylactic treatment, such as neuroprotective agents, before END is even detected (DeGraba et al., 1999; Martin & Price, 2018; Weimar et al., 2005). Assessments that identify the important information about short-term changes in a patient's condition and aid communication are needed (Baron et al., 2013).

2.3 Scales available for neurological assessment and monitoring after stroke

There is a wide range of scales available to use within acute stroke practice for the assessment and monitoring of neurological status and these will be explored in more detail in Chapter 4. However, the three scales used most in clinical practice are:

- The National Early Warning Score (NEWS) 2 (Royal College of Physicians, 2017a)
- Glasgow Coma Scale (GCS) (Teasdale & Jennett, 1974)
- National Institutes of Health Stroke Scale (NIHSS) (Brott et al., 1989)

The NEWS 2 is an aggregate scoring system allocated to physiological measurements to guide the identification and management of parameters that vary from normal (Royal College of Physicians 2017a). It improves the detection and response to clinical deterioration in adult patients. NEWS 2 is considered a key element of patient safety and improving patient outcomes through a pragmatic approach that emphasises system-wide standardisation. It provides a surveillance system for any hospital patient for tracking their clinical condition, alerting the clinical team to any medical deterioration, and aiming to trigger a timely clinical response. (ibid). NEWS 2 was mandated in acute and ambulance trusts in 2018 by NHS England. However, it is not specific to a stroke population and does not assess or provide feedback on functional neurology.

The GCS is regarded as a quick and easy scale to administer for assessing level of consciousness (Teasdale & Jennett, 1974). Whilst it has been widely adopted in stroke practice, it was developed and validated to measure conscious levels in traumatic brain injury. It is insensitive to the detection of focal neurological deficit which is vital in stroke. The GCS has been shown to incorrectly identify 56% of stroke patients as having no neurological deficit when tested

against the NIHSS (Nye et al., 2012). As it is based on recording the patient's best response it is also criticised for failing to capture fluctuation (Lowry, 1999). As with all scales errors in technique affect the accuracy and inter-rater reliability (Basauhra Singh et al., 2016; de Souza & Woodward, 2016).

The NIHSS is an established measure to assess the severity of neurological deficits of acute stroke patients. It was developed in the early 1980s as a research tool to allow consistent reporting of neurological deficits in acute-stroke studies, particularly the early trials of thrombolysis and putative neuroprotectants (Brott et al., 1989). Based on these studies an NIHSS score >5 was previously used to be a prerequisite in the decision to treat with thrombolysis. It has limited validation in ICH, patients with co-morbidities, and very severe stroke patients as these groups were often excluded from such trials.

The NIHSS is criticised for its heavy weighting towards motor and language function and under representation of both posterior and right hemisphere lesions (Gottesman et al., 2010; Linfante et al., 2001). Ordinal scales such as the NIHSS are also open to other measurement limitations such as ceiling effect. Ceiling effect occurs when the highest possible score or close to the highest score on a test or measurement instrument is reached, thereby decreasing the likelihood that the testing instrument has accurately measured the intended domain (Salkind, 2010).

Despite its critics, the NIHSS has been adopted for a wide variety of uses, including acute assessment and decision-making in discharge planning. It is part of the clinical decision process to treat with thrombectomy (score requirement of ≥ 6) and decompressive hemicraniectomy (total score >15 with a decrease of one or more in level of consciousness score) (Royal College of Physicians, 2016a). In order to collaborate in the Sentinel Stroke National Audit Programme (SSNAP) stroke units must achieve an 80% compliance of NIHSS collection on arrival to hospital and 90% NIHSS at 24hrs for those patients who receive thrombolysis (SSNAP, 2021).

This section has discussed the three most common scales used in clinical practice. The NEWS and GCS are not stroke specific. The NIHSS whilst well-established has known limitations. A reproducible and valid method for measuring neurological deficit and detecting END is required to monitor patients after acute stroke. Clinical teams need to be able to assess the type and severity of neurological impairments accurately and effectively, to be able to monitor change in neurological status, and to examine responses to treatment. Different measures will have different advantages and disadvantages and it may be that one scale may not be suitable

for all stroke patients. Before making any recommendations on standardising neurological assessment and monitoring, the clinimetric properties of all available scales, need to be thoroughly understood and appraised (Chapter 4).

Understanding the properties of the current scales will inform the development of the SNOBSS. The introduction of SNOBSS should facilitate the systematic identification of meaningful neurological changes at the bedside when it is used by a range of clinicians, but this would require further testing. It could ultimately improve outcomes for patients when implemented into clinical practice. However, to successfully implement changes in clinical practice an awareness of the actual and potential barriers is required.

2.4 Barriers and facilitators

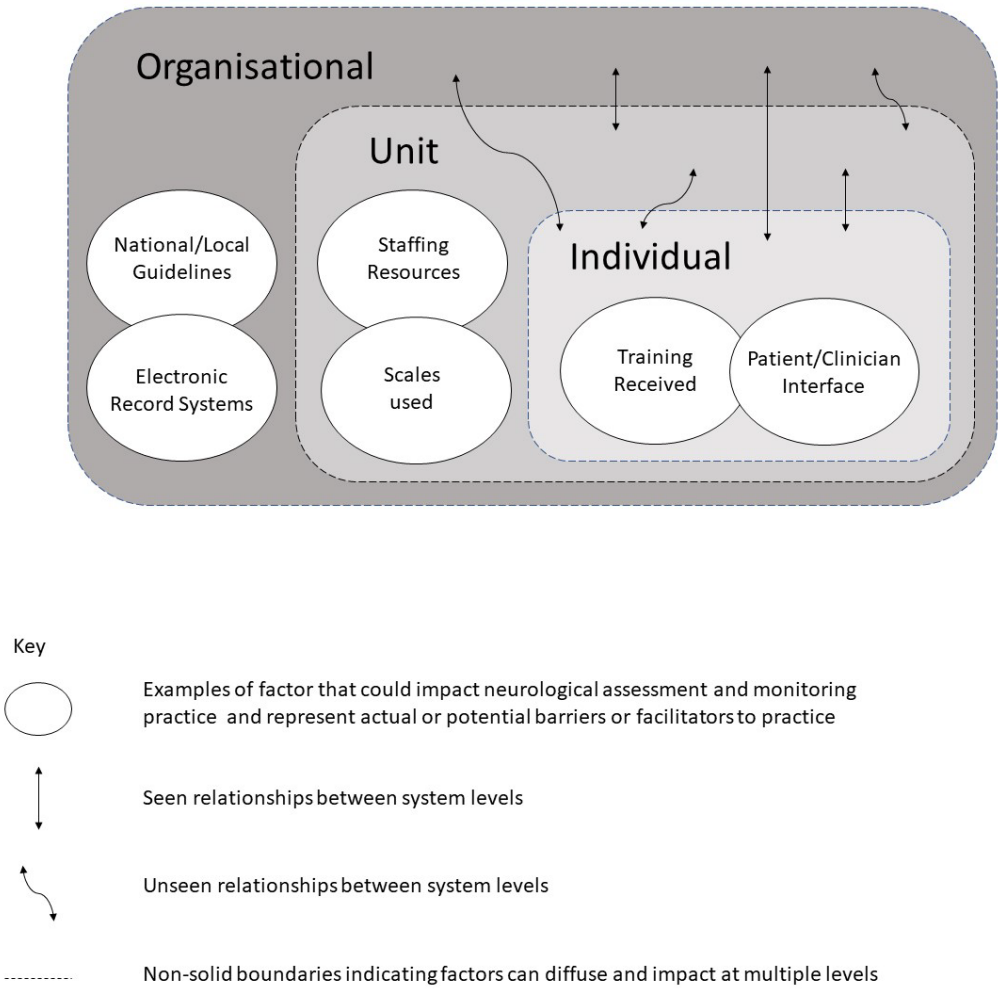
Barriers and facilitators to interventions need investigating to be able to develop guidelines, and recommendations for future implementation, that are truly applicable to practice. Even where evidence-based guidelines exist for acute stroke management their adoption is often delayed due to a range of barriers and facilitators (Baatiema et al., 2017). Multiple actual and perceived barriers and facilitators affect how changes are implemented into clinical practice. Within the stroke literature these include the lack of protocols and pathways (Williams et al., 2013), limited staff capacity (O'Rourke et al., 2013; Purvis et al., 2014), lack of skills or self-efficacy to apply the intervention (Stecksén et al., 2014, Van Der Weijden et al., 2004) or low motivation to implement an evidence-based therapy (Meurer et al., 2011).

Some specific barriers in relation to neurological assessment and monitoring are already known in the literature such as the time taken to complete assessments (Yanko & Lang, 2013). The need for education, especially with the NIHSS, has been established for a long time (Andre, 2002). Despite this, there remains a lack of adequate teaching resources for educating clinicians regarding how to use certain scales (Richardson et al., 2006; Yanko & Lang, 2013). Even when training is provided there can be inconsistent scoring among clinicians, especially with comatose or uncooperative patients as it can be difficult to master the performance and interpretation of components of the examination (Nye et al.,2012; Richardson et al., 2006; Yanko & Lang, 2013). Maintenance of knowledge and skills and sustained competence in assessment remains an established problem in practice (Gocan & Fisher, 2008). This can be complicated further if the scale, or language used within it, is complex and raters do not feel competent or comfortable with the assessment.

Although there has been an improvement over time, collection and reporting bias can still occur, with factors such as patient severity or treatment options affecting completion of neurological assessment (Reeves et al., 2013).

However well-constructed and accepted, the SNOBSS and decision flowchart will not be able to change and standardise practice unless it is adopted by all staff. As introduced in Chapter 1 neurological assessment and monitoring is a complex element of care within a Complex Adaptive System (CAS). Figure 2.1 has been created, by the author, to provide a visual representation of the CAS in which neurological assessment and monitoring sits in practice. It shows the factors that impact neurological assessment can occur at an individual, unit, or organisational level and that factors can be interconnected within and across levels. The complexity is represented in multiple visual ways including the overlapping of the factors within the systems. The dotted borders represent that system levels do not have solid boundaries and the impacts can diffuse into other levels, and the arrows signify that factors can interact in both seen and unseen ways between multiple levels in different ways.

Figure 2.1 Visual representation created to show that neurological assessment and monitoring is a complex element of care within a Complex Adaptive System (CAS) and examples of some of the factors that impact its delivery at an individual, unit, or organisational level.



The factors included within Figure 2.1. are examples of actual or potential barriers and facilitators to the implementation of the SNOBSS, although many more could be present. These chosen examples represent tangible factors such as the current rapid expansion of electronic observation systems which are geared to a general hospital population and not stroke specific observations. However, the research within this thesis is open to other less empirical impacts on neurological assessment and monitoring practice such as social elements around teamwork and organisational culture that can impact on all aspects of care delivery.

For successful implementation of the SNOBSS to occur in the future a broad assessment of barriers and facilitators within this complex element of care is needed. This thesis will also consider that actual and potential barriers may vary between different individuals, units, and

organisations. The survey (Chapter 5) and the interviews (Chapter 6) explore clinicians' views about their actual and perceived barriers and facilitators to neurological assessment and monitoring. The results from these will allow the development of an overview of the understanding and awareness that clinicians hold about the importance of, and any misconceptions around, neurological assessment and monitoring. This overview will help shape future targeted recommendations and suggested interventions to implement the SNOBSS and decision flowchart. These recommendations will increase the potential of the SNOBSS, and decision flowchart, being successfully adopted into clinical practice in the future.

2.5 Chapter summary

This chapter has provided further justification for the standardisation of assessment and monitoring, through the development of the SNOBSS, and the work within the thesis. It has summarised some of the difficulties in defining and identifying END and the multiple factors that cause or are associated with it. Although different stroke subtypes and characteristics have been shown to affect the prevalence of END, it remains a potential risk for all patients after stroke and appropriate neurological assessment and monitoring is needed for all stroke types and severities.

The chapter went on to highlight that a range of scales is currently available for neurological assessment and monitoring and emphasised that the measurement properties of these should be investigated for their suitability to detect neurological change and to assist in the development of the SNOBSS. The chapter finally discussed the complexity of neurological assessment and monitoring and the importance of exploring the factors that impact on it to identify actual and potential barriers and facilitators to successfully implement the SNOBBS within clinical practice in the future.

The next chapter, Chapter 3, is the methodology chapter which will outline the theoretical underpinnings that shape and guide the programme of research. It will outline the overarching questions the thesis aims to address, and the mixed methods approach designed to answer them. It will also discuss ethical issues, alterations made, and the impact of COVID-19 on project delivery.

Chapter 3 Methodology

This chapter will summarise the overall methodology and theoretical underpinnings of the project. It will discuss study design and justify why a mixed method approach was chosen. Details of specific methods employed to address specific questions will be found within the individual phase chapters (Chapters 4-7). This chapter will discuss issues encountered with ethical processes before providing an overview of alterations to study design and delivery that occurred during the PhD programme of research. It will conclude with a brief outline of the impact COVID-19 had on project delivery.

3.1. Context and theoretical underpinnings

In chapter 1 neurological assessment and monitoring was introduced as a complex element of care within acute stroke practice. Chapters 1 and 2 highlighted some of the multifaceted influences that impact its place within an acute stroke care pathway. In addition, myriad factors occurring at the individual, unit, and/or organisational levels that impact use and potentially cause variation in neurological assessment and monitoring were also highlighted.

As such, this thesis investigates a complex element of care within a non-linear and dynamic complex adaptive system (CAS) (Figure 2.1 pg.26). Complex implies diversity and a wide range of elements, adaptive means the capacity to change and suggests the ability to learn from experience. With the system being a set of connected or interdependent things (Begun et al., 2003). CASs are additionally complex due to the interactions between different agents affecting the overall behaviour of the system (Thompson et al., 2016). The interconnectedness between individuals and levels of the system means that one action or inaction can change the context for others so that CASs are neither stable nor predictable (Plesk & Greenhalgh, 2001). Consequently, the research design needed to be able to take account of this complexity and adaptive nature of stroke care systems.

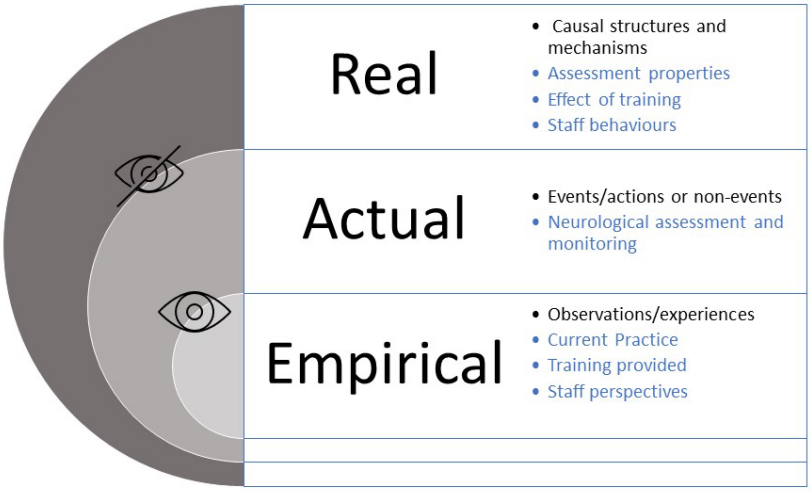
To develop the Standardised Neurological OBservation Schedule for Stroke (SNOBSS) and associated decision flowchart evidence needs to be gathered and combined from multiple sources (Pype et al., 2018). Multiple factors impact on neurological assessment and monitoring and therefore there are several potential solutions to how the SNOBSS, and decision flowchart could be developed. The research, therefore, needs to be exploratory, rather than seeking an ultimate truth.

Traditional philosophical approaches to the nature of reality do not fit well with this viewpoint. The positivist outlook is that there is only one reality, based on objectivity and truth, and that it is 'out there', waiting to be discovered (Oltmann & Boughey, 2011). A positivist approach was not appropriate here given the complexity of neurological assessment and monitoring intervention and the healthcare settings in which it takes place. Conversely, constructivism asserts that people construct their own understanding and knowledge of the world through experiencing things and reflecting on those experiences (Honebein, 1996) suggesting that there could be as many realities as there are people in the world (Coghlan & Brydon-Miller, 2014). The ontological position of constructivism also did not fit well with the aims of this research to inform the development of a Standardised Neurological Observation Schedule for Stroke (SNOBSS). This study, stands in between these ontological positions and therefore needed an alternative viewpoint. Critical realism offers that alternative position as it neither rejects nor endorses the different stances offered by the positivist and constructivist paradigms (Julnes et al, 1998; Pawson & Tilley, 1997).

Realism is a way of understanding and explaining the nature of being or existing (ontology) (Bhaskar 1989), and critical Realism evolved from the writings of the philosopher Roy Bhasker (Bhasker, 1978, 1979, 1989). It is critical in a Kantian sense as it accepts knowledge as local and historical (epistemic relativity) but not that all viewpoints must be equal (judgemental relativity) (Mingers et al., 2013). It is realist in that it does not believe that everything can be empirically observed and measured (positivism) but that there are properties that impact the world independent of our knowledge (Levers, 2013). Critical Realism, therefore, distinguishes between the 'real' world and the 'observable' world. According to Bhasker, it is more meaningful to search for, understand, and describe causative or generative mechanisms than to seek the "absolute truth" (Wilson & McCormack, 2006, p.46).

For Bhasker (Bhasker, 1978, 1979, 1989), reality exists at three levels – the empirical (experienced), the actual (every event whether experienced or not), and the real (where causative structures and mechanisms exist) (Houston, 2001) which are illustrated in Figure 3.1.

Figure 3.1 The three layers of reality in critical realism and what exists at those levels with examples based on neurological assessment and monitoring.



Key

	Represent that some things are observable
	Represents that some things are unobservable
Black Text	What exists at that level of reality
Blue Text	Examples of what exists or might exist at that level of reality

This research aimed to explore the factors that impact on neurological assessment and monitoring (as well as the evidence base underpinning measurement properties of neurological assessment tools used in practice) rather than only describe current practice. However, it is also recognised that some aspects of practice and factors that impact neurological assessment are observable and measurable but that there are likely to be others that will be unobservable elements and mechanisms that impact on practice.

The data that will be collected, the observations (survey of practice and training) and experiences (staff and expert panel perspectives), and what is currently known (reviews of measurement properties of neurological assessment tools used in practice), exist in the Bhaskar's (1978) empirical layer. The causal structures and mechanisms that generate neurological assessment and monitoring reside in the real layer. Mechanisms are the

"underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest" (Astbury & Leeuw, p. 368). They connect the inputs and outputs within the system, have causal powers, and generate the observed events (neurological assessment and monitoring) (Dalkin et al., 2015). These often-unobservable mechanisms may include the effect of training and education received or not, how staff as individuals or teams behave concerning neurological assessment and monitoring, and elements of the assessment process itself that impact on its measurement properties and the assessment produced, but may themselves not be measurable. The eye and crossed eye within Figure 3.1 represent the observable (seen) or less observable (unseen) and the move from the measurable empirical layer to the often-unmeasurable mechanisms that affect the delivery of the event or neurological assessment.

Critical realism has been adopted and endorsed in research across a range of disciplines. It provides a suitable underpinning for this project to illuminate and explore the complexity of health care (McEvoy & Richards, 2003; Pawson & Tilley, 1997) yet remain focused on real problems (Mingers, 2011; Syed et al, 2010). Research, underpinned by critical realism, has the scope to guide policy recommendations for change by providing scientific explanations of complex problems (Cruickshank, 2011).

Interactions of components within a system result in the overall behaviours of the system. In a CAS there is a collection of individuals (agents) whose actions are interconnected. However, this can be unpredictable as one individual's actions can change the context for others (Plesk & Greenhalgh, 2001). Individuals often have little control over system level changes. This research aimed to explore new possibilities of neurological assessment and monitoring practice by working at the edge of knowledge and experience (ibid). Critical realism provides a strong underpinning to explore such a complex element of care within a CAS.

Interrelationships and interactions will be explored at the individual, unit, and organisational levels as any CAS is not the sum of these parts. This will lead to multiple perspectives having to be considered at each stage but will help guide analysis and interpretation. Logically it may enable a clearer understanding of what may work best in relation to neurological assessment and monitoring across multiple situations (Gill & Turbin, 1999; Greener & Mannion, 2009).

A goal of the SNOBSS and decision flowchart is to reduce unwarranted clinical variation, but this cannot happen unless it is successfully implemented into practice. Implementation science is defined as the scientific study of methods to promote the systematic uptake of research

findings and other evidence-based practices into routine practice to improve the quality and effectiveness of health services (Eccles & Mittman 2006). One key aim of implementation science is to understand what influences outcomes and the mechanisms by which implementation is more likely to succeed (Nilsen, 2015).

It is important to think about implementation at an early stage in intervention development (Skivington et al., 2021). Therefore, this project will consider implementation throughout identifying what may work best in different settings over time to support future implementation of the SNOBSS (Westhrop et al., 2011). Normalisation Process Theory (NPT) (May & Finch, 2009) was utilised in the interview phase, this theory of implementation explains how practices become embedded into practice. CASs are dependent upon what has gone before, they can be highly creative, and transformational change can emerge at any given time, therefore, they can provide multiple paths for action. Application of NPT with a critical realism lens will added explanatory power to the exploration of implementation across multiple CASs.

Change within a CAS can be difficult to predict. This is especially true as the level of input does not necessarily correlate to the size of the change observed. Within CASs like healthcare, small changes can be more attractive than large-scale changes which can create resistance (Zimmerman et al., 1998). However, within a CAS the feedback that occurs in the system and between agents can generate change or stability dependent upon the relationships (Begun et al., 2003). This means that even within two systems that appear similar, significant differences can develop over time, and this is important as we are looking to change practice over multiple sites.

Critical realism contributes to the understanding of complex issues in healthcare practice adding ontological depth to the scope of inquiry. Critical realism accepts the existence of different types of knowledge and accommodates different research methods to source them (Bergin et al., 2008; Sayer, 2000). Applying a critical realism lens offers to enhance understanding. There are several research questions that need to be answered to gain a better understanding of the practice, meaning, and significance of the complex intervention of neurological assessment and monitoring after stroke. Critical realism supports the use of appropriate questions and the use of appropriate methods to answer them (Walsh and Evans, 2014). Including mixed methods research designs (Mingers, 2004; Venkatesh et al., 2013).

3.2. Medical Research Council (MRC) framework

This thesis is exploring an intervention that is complex in terms of both content and the context in which is conducted. The Medical Research Council (MRC) framework supports the development and evaluation of complex interventions. It can help prioritise research questions and the design and conduct of research with an appropriate choice of methods (Skivington et al.,2021a).

The MRC framework encourages consideration and use of diverse research perspectives to support the appropriate choice of research questions and methods to address them. It acknowledges that precise answers to narrowly defined questions are not always the most useful and often broader more complex questions are needed in order to generate the most useful evidence in complex situations (Deaton & Cartwright, 2018). Throughout the research has paid attention to future implementation of the intervention in the real world by exploration of barriers and facilitators. This is important as early consideration of implementation increases the potential of developing an intervention that can be widely adopted and maintained in real-world settings (Campbell et al., 2018)

The framework includes a checklist of six elements to be addressed throughout the research process (Skivington et al., 2021 pgs.131-132). These elements have been considered below in relation to the research to describe some of the methodological decisions:

Addressing Uncertainties

There are multiple uncertainties in relation to neurological assessment and monitoring as an intervention. The research team have identified and prioritised which are the most important to address. The broad formulated research questions (Table 3.1) aim to explore neurological assessment and monitoring at a systems level from different perspectives. Mixed methods, utilising multiple data sources were therefore chosen as the most suitable to address the research questions.

Engaging Stakeholders

Multiple perspectives were sought in designing the research from stakeholders who deliver or will benefit from the intervention. Their involvement in the development enhanced understanding of the uncertainties related to the intervention. It was intended to have greater PPI involvement to enhance the delivery, evaluation, and dissemination of the research (Section 3.5). An interactive expert group developed the SNOBSS (Chapter 7) and helped

define the next set of uncertainties that needed to be addressed in future research (Chapter 8).

Considering context

The MRC framework emphasises the importance of context and understanding interventions within the systems that they interact with in order to support implementation (Craig et al., 2018). This research was designed to explore the influence of context on delivery of the intervention. The MRC framework supports the exploration of neurological assessment and monitoring within a CAS as it appreciates that systems cannot be explained in terms of individual parts but that there needs to be an awareness and understanding of the whole system. This systems perspective encouraged consideration of how neurological assessment and monitoring may be influenced by many elements of the system. This was supported by the application of the critical realism lens to be open to the multiple, often indirect, or unintended routes through which the context may impact on the intervention. Through improving understanding important implications for decision making could be identified.

Developing and refining programme theory

The research has not focused on creating programme theory as there were too many uncertainties to consider initially. As current practice and therefore the intervention is currently not well designed it would be difficult to extrapolate into what conditions cause what effects. However, the research drew on appropriate existing theories, such as Normalisation process theory (NPT) as a framework to evaluate existing practice (Interviews- Chapter 6) and illicit knowledge for the development of the SNOBSS and its future implementation.

Refining the intervention

Once the preliminary version of the SNOBSS was developed through synthesis of the research findings it was refined through review from the expert group members and the evaluation work completed by the wider clinicians (Section 7.5). Useful information was also gathered to consider in the design of the future testing and evaluation of the SNOBSS. It is anticipated that there will be further refinement once further research is completed.

Economic considerations

It was not possible to explore economic considerations within this thesis. The MRC framework places emphasis on this work as a key element of all phases of a project rather than simply assessing cost-effectiveness. The future recommendations (Chapter 8) discuss the importance of its inclusion. However, this research aimed to identify what neurological assessment and

monitoring currently looked like across a range of systems and economic considerations were not feasible in these initial stages.

3.3. Mixed methods approach

This thesis uses an overall mixed method approach. Specific methods for each phase of the research within the thesis are fully described within each appropriate chapter. Methods describe the strategies, processes, or techniques used in research data collection and analysis (University of Newcastle Library, 2022). Research methods are often divided into quantitative which generates numerical data and qualitative which generates non-numerical data (Tariq & Woodman, 2013). A mixed methods research approach in its simplest terms is defined as research that integrates qualitative and quantitative methods (Tashakkori & Creswell, 2007). However, there are numerous classifications and applications of mixed method designs often representing different disciplines and using different terminology (Taskakkori & Teddlie, 2010).

The choice of design is based on the research purpose. A mixed methods approach suits this project due to its complex and exploratory nature. Table 3.1 outlines the broad research questions, methods to address them, and intended outputs. It shows that there is a mixture in data collection and analysis of both quantitative and qualitative methods. A combination is needed as the different questions required different approaches but also because the author is aware that there are multiple truths and multiple perspectives that need to be considered (Creswell, 2009; Johnson & Onwuegbuzie, 2004). This research aimed to tease out as many of the multiple potential solutions or conclusions as possible (Krathwohl & Smith, 2005).

Table 3.1 Overview of research questions - showing the broad questions, the methods adopted to address them, and the intended output from each question.

Question	Method/s	Output
What are the measurement properties of the tools available for neurological assessment and monitoring? (how good are they at identifying END?)	Scoping review to identify scales.	Identification of scales used or available for neurological assessment and monitoring in acute stroke
	Series of reviews looking at clinimetric properties of identified scales	Collation of the evidence on clinimetric properties of all scales identified.
What does current practice look like? (e.g., are there protocols, what tools are used, frequency of observations)	Survey of all UK stroke units that admit acute stroke	Overview of current practice across the UK
What are clinicians' experiences of neurological assessment and monitoring practice? (knowledge, understanding, acceptability)	Survey of all UK stroke units that admit acute stroke	Summary of clinician experiences
	Semi-structured interviews	
What are the barriers and facilitators in this element of care?	Survey of all UK stroke units that admit acute stroke.	Description of the barriers and facilitators
	Semi-structured interviews	
How could we standardise neurological assessment and monitoring?	Expert Group consensus based on knowledge, experience, and results of the reviews, survey, and interviews.	Standardised Neurological OBservation for Stroke Scale (SNOBSS) and decision flowchart
What are the best ways to make changes in this element of care?	Expert Group consensus based on knowledge, experience, and results of the reviews, survey, and interviews.	Potential implementation strategies for the SNOBSS

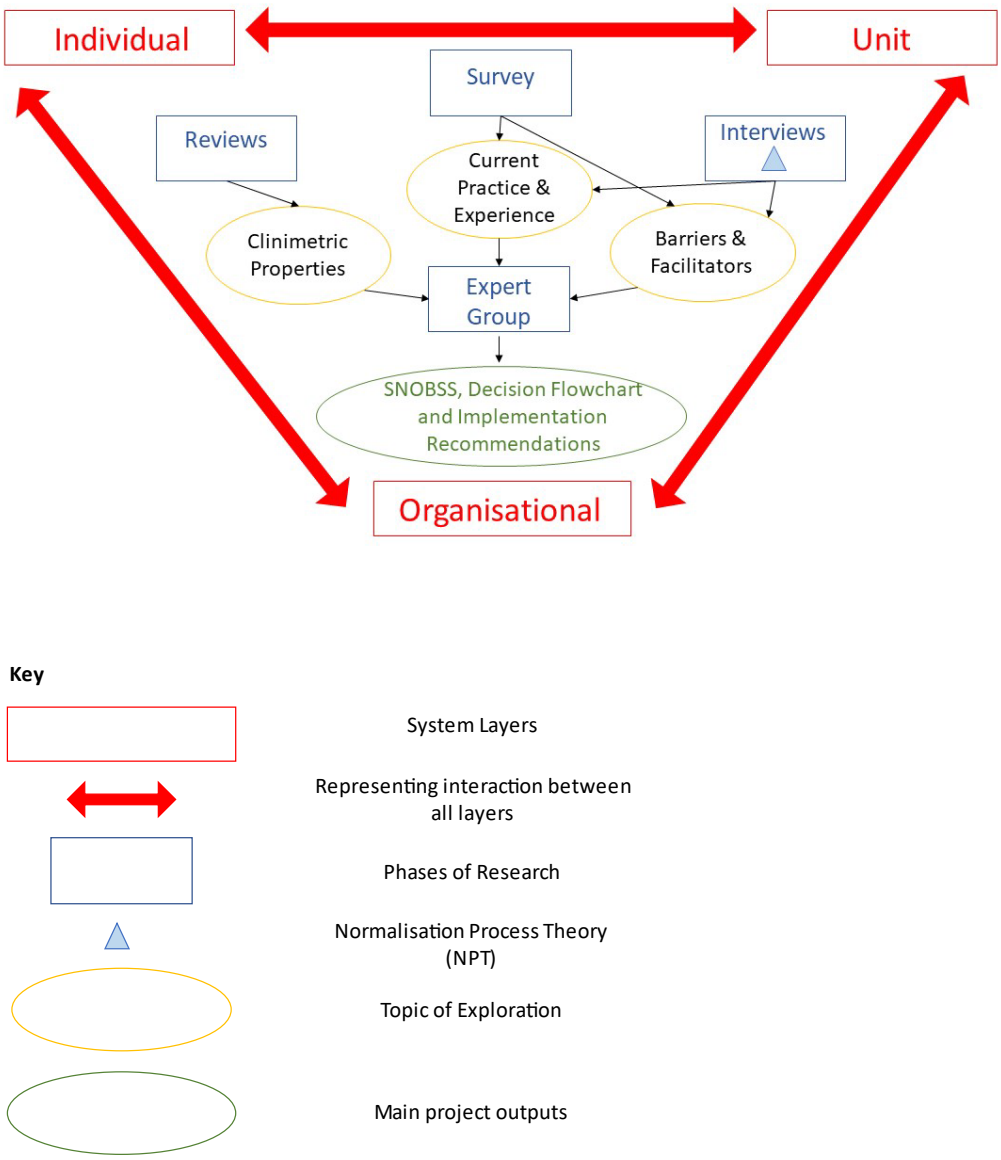
Key

Shows colour coding of methods and chapter location.

Colour Key	Method	Chapter
	Reviews	4
	Survey	5
	Interviews	6
	Expert Group	7

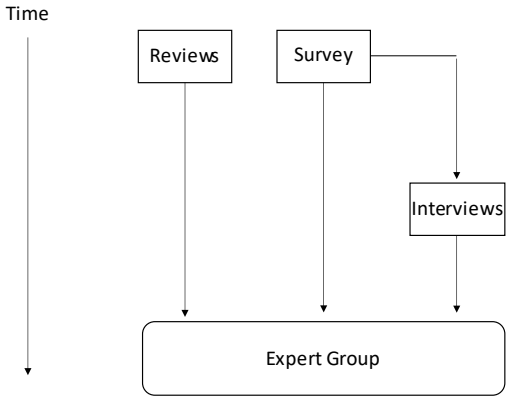
As already stated previously this thesis is exploring a complex element of care within a complex adaptive system (CAS). Figure 3.2 provides a visual overview of the project. The diagram clearly shows that there is the potential for interaction between all aspects of the system as would be expected in a CAS. Within the context of the CAS the reviews, survey, and interview phases of the research will be completed. The key results from these phases will then be presented to the expert group. The expert group with use this data alongside their knowledge and skills to develop the SNOBSS and decision flowchart as well as consider recommendations for future implementation.

Figure 3.2 Visual representation of the exploration of neurological assessment and monitoring practices after stroke



The separate research phases within this thesis i.e. the reviews, survey, and interviews were designed and delivered to address their specific questions (Table 3.1) before being converged and presented to the expert group. This overall exploratory mixed method design is based on the premise that a single data set is not sufficient and mixes quantitative and qualitative methods at the design level as different questions need to be answered that require different data types (Creswell, 2006). Figure 3.3 shows that the reviews and survey were run concurrently with the interviews and expert group following sequentially.

Figure 3.3 Visual representation of how each phase fed into the overall mixed methods approach utilised over time.



The reviews involved the concurrent collection and analysis of quantitative and qualitative data within the literature which was merged to better understand the clinimetric properties of scales. The survey questionnaires combined a mixture of quantitative and qualitative questions. Although some of the open questions increased understanding or developed a complementary picture of the quantitative data they did not allow for an in-depth understanding of the issues in question.

The interviews were developed to provide a more in-depth understanding. Data around site characteristics from the survey were used to identify and purposely select the hospitals that were approached for the in-depth interviews (explanatory participant selection model). Broad findings from the survey especially in relation to differences and unexpected results fed into the interview schedule to obtain further detail or explanation (explanatory follow-up explanations model) (Creswell, 2003).

Overall, this mixed methods thesis has addressed multiple questions but has completed convergence and collaboration where appropriate within and between separate data collection methods. The use of a mixed-methods approach has allowed findings to be built up and developed where complementary data was obtained.

3.4. Personal Stance and Reflexivity

The author decided to undertake the PhD because she had seen widespread variation in the ways that neurological status was assessed and monitored and was frustrated by the lack of guidance to support her own and others' practice. She had observed that variation could have an impact on both clinical and research outcomes. This work was a way that she could, as a nurse, make a difference through supporting the development of the evidence-base in this important element of care. She was keen to develop a SNOBSS that would be usable and meaningful in clinical practice.

The author acutely understood that repeated assessment of neurological status to monitor for deterioration is a complex element of care that can be affected by numerous factors at an individual, unit and organisational level. Patient factors, such as type of stroke, result in huge variation in the signs of symptoms of stroke dependent upon the location and severity of the damage within the brain. It was the author's view at the outset that a one size fits all approach in terms of identifying deterioration might be difficult and therefore it was important to factor different stroke types into the data collection and consideration in the design of the SNOBSS.

Being a nurse with clinical and research experience in stroke care had obvious advantages in terms of the subject area and context. However, having that experience and belief about this being an important aspect of care could influence the approach to analysis and interpretation of the data within the PhD (Creswell 2007; Pope et al. 2000). The author was aware that her experience around variation in practice might not be reflected in other areas and was keen to explore practice without an expectation that variation was inherent everywhere. Her research experience also made her aware of the potential void between what individuals and teams think they do and the reality of their practice. Having that awareness helped to ensure that the work followed insightful enquiry and was not biased based on previous experience. The research aimed to identify factors that would impact on current practice further testing and implementation of the SNOBSS (Walsh & Evans, 2004).

In order to challenge the author to collaborate beyond her personal and professional viewpoints, critical realism was chosen to underpin the research. This encouraged the holistic examination of the complex phenomena of neurological assessment and monitoring practice from multiple perspectives (Walsh and Evans, 2014). The research was therefore designed and conducted to explore this complex element of care within the complex systems it resides as credibly as possible without being influenced by the author's own values and beliefs. Every attempt was made to prevent personal biases affecting the research and this was supported by the oversight of an experienced supervision team.

The work within this PhD represents the first steps towards a consistent plan of how the neurological effects of stroke should be assessed, recorded, and monitored over time. It is hoped that the work within this thesis will initiate and drive forward the development of accurate and effective use of neurological monitoring in acute stroke and ultimately lead to changes in practice to improve patient care and outcomes.

3.5. Patient and Public Involvement

The work within this thesis is a continuation of ideas and work from the applicant's Masters, as a result there has been longstanding input and suggestions from stroke survivors, families, and carers, who have provided valuable feedback and input into this work. The patient and public involvement (PPI) groups involved before and during this work felt that the variation in practice (clinical and research) the author had witnessed was unacceptable and had the ability to add to the distress of patients, and their families, following a stroke. The PhD proposal was reviewed and commented on by several members of the Lay Research Group (LRG) at a local secondary care NHS Trust and a stroke survivor who has experience with Stroke Research at a national level. The LRG reviewers were crucial in terms of the project development and through several rounds of review and alteration ensured that the aims, methods and intended outcomes were clear and logical to a wider audience. They helped develop a proposal that aimed to provide an "informed intervention that will give stroke patients the best chance of receiving the right treatment and support at the right time".

It was intended to have a project specific PPI advisory group, who would have been involved throughout the PhD. They would have had an active role in the management of the research, helping with the interpretation of emergent findings, promotion of the study and dissemination of findings to a wider audience. It was also planned to have two PPI members in the expert group as it was felt that PPI views around importance, acceptability and tolerance of neurological monitoring would be crucial to the development of the SNOBSS.

Initial consultations with stroke groups in Preston, Warrington, and Aintree were undertaken to recruit PPI members. The opportunity to be involved was also advertised on the people@UCLan website.

In the introductory meetings and discussions with potential PPI members they were positive about the project and affirmed its importance. However, it was clear that patients did not remember interventions such as neurological monitoring in the acute phase. There was also an

assumption amongst stroke survivors and carers that if something was being done it was for a good reason.

Twenty-one individuals expressed an interest in being involved in the study moving forward. Their preference was that the project work was brought to their existing support groups for consideration rather than additional meetings organised at the University or centrally. Key contacts within the groups were established to co-ordinate regular attendance and to feed back to the groups about the project. Unfortunately, further PPI involvement was prevented once the COVID pandemic hit. It was hoped that involvement would be temporarily delayed, but stroke group meetings were not re-established until the study was in write up stage.

3.6 Ethical process issues

All research projects inherently have ethical considerations and should follow codes and guidelines to protect both the participants and the researcher (European Commission, 2013). The survey and the interview portions of the research required both Health Research Authority (HRA) and University approval as the research was being completed within NHS institutions for the purposes of an educational award. The project was considered low risk as the topic of neurological assessment and monitoring practice was not considered potentially sensitive to the staff participants. There was a slight risk that it could expose poor or bad practice within stroke units. This could have been at an organisational unit level (survey) or an individual level (interviews). Strategies were developed to deal with these situations should they arise.

Despite the project being considered low risk the project experienced unexpected delays in approval. The University ethics board requested changes to the consent process for the survey. Instead of consent being inferred by the return of the survey, they requested agreement by email needed to be in place before the survey was sent out. This ensured that General Data Protection Regulation (GDPR) requirements were fully adhered to if not exceeded. The HRA also introduced changes in their application process which caused further delays as paperwork had to be amended to meet the new requirements.

As the survey and interview phases involved NHS staff participants and not patients, they were eligible for proportionate review by the HRA. The HRA reviewer decided that the hospitals involved should be classified as Participant Identification Centres (PICs) rather than research sites as the study was being remotely managed. This initially meant that despite the study having been adopted onto the Clinical Research Network (CRN) portfolio it was not eligible for accruals. This had implications in terms of research departments being willing to support the

study. An appeal to HRA failed, however, a research manager with extensive CRN experience advised on how to get the study registered for accruals via the local CRN. Agreement took months (finally confirmed on 19/02/2020) but was essential to the success of the project, as without this many sites would not have been willing to participate.

Despite the streamlining of NHS approvals through HRA, there has not been a standard and linear process to follow. Nearly all research departments had different approaches to the way they managed the project approval process resulting in multiple different requirements. The change to PIC status should theoretically have reduced the administrative burden but most Trusts still requested all the usual research site requirements and documentation despite the hospitals being deemed PICs. Research approval administration was further protracted and complicated by requirements due to the COVID-19 pandemic.

3.7. COVID-19 impact

The COVID-19 pandemic caused a considerable impact on the delivery of the research causing significant delays across the timeline of the project. The reviews were affected during the initial phase of the pandemic as both the British and local libraries were unable to supply any papers for which they had to request print versions.

Research approvals through NHS sites were underway as the pandemic hit. Due to the National requirement to place non-essential research on hold the study was affected and delayed in multiple ways. Where approvals were in place the study was placed on hold and recruitment paused for several months. These study interruptions also increased the administrative burden in terms of on-hold and restart documentation. In departments where approvals had not yet been granted the study had to wait until the on hold was lifted and research departments were able to consider approving this type of research again. Even after non-COVID research was restarted some sites did not have the capacity or infrastructure to restart or approve the project.

Due to the ongoing issues with COVID, and risk assessments within the NHS sites, survey distribution had to move from postal to electronic distribution. The delays to approval and restart meant that a few potential participants that had been identified at the beginning of 2020 had retired or changed roles in the intervening period and new participants needed to be identified. Across Scotland, a nationwide change of NHS email addresses, during the on-hold period, caused administrative problems upon restarting the survey phase of the study.

Temporary closures and reconfigurations to accommodate stroke care safely through the pandemic also caused issues in engaging with the clinical stroke teams. Due to service reconfigurations, expedited through the pandemic, the number of eligible sites for survey completion reduced.

The ongoing effects of COVID both in the peak waves and on either side caused issues with both questionnaire and interview completion. Even where potential participants were keen to complete there were delays as they struggled to complete due to worsening staffing issues brought about by COVID absences, general sickness, and the backlog of holidays that needed to be taken within NHS sites.

The requirement to work from home for the duration of the pandemic has had implications for material resources, support, and opportunities for networking and informal development. The author had two children being home schooled for large portions of the project which had a detrimental effect on progress although work patterns were adjusted to try and accommodate this with the least disruption.

3.8. Alterations to project

Alterations to the design and delivery of specific methods will be described in the relevant chapters. This section outlines a major alteration to the study design that occurred due to the impact of COVID-19 and delays to the programme of work. The original project plan included the implementation of the SNOBSS and decision flowchart in one site. The aim was to gather fidelity data to see how well the schedules were adhered to as well as feedback from staff (surveys and focus groups) to check for acceptability and highlight any issues or additional training needed. The findings would have been used to improve and refine the implementation strategy developed by the expert group. Testing of the SNOBBS developed in this PhD will be a fundamental part of the post-doctoral project plan and will be discussed in more detail in Chapter 8 (Discussion and Conclusion).

3.9. Chapter summary

This chapter has described the overall methodology of the project. Neurological assessment and monitoring is a complex element of care delivered within a CAS. The context of the care can be impacted from an individual, unit, or organisational level. This project completed mixed methods research to explore multiple topics that impact the delivery of neurological

assessment and monitoring before sharing the data with an expert consensus group to consider in the design of the SNOBSS and its associated decision flowchart.

Critical realism provides a strong theoretical approach for the conceptual framework, design, and execution of this research. It recognises that healthcare systems are complex, non-linear, dynamic, and unpredictable. It will ensure investigation of the topic with consideration to the empirical, actual, and real levels of reality searching for both observable and unobservable factors that impact this element of care after stroke. Using a theory-driven approach, the relationships between mechanisms of action and the contexts (or conditions) in which they are triggered could be revealed. This could lead to an explanation of outcomes that will enable a clearer understanding of what interventions might work, which will help provide recommendations for the implementation of changes in this area of practice. It is hoped this project will assist in the development and delivery of more standardised neurological assessment and monitoring after acute stroke.

The next chapter, chapter 4, describes and presents the results from a series of reviews. Initially, a scoping review was undertaken to identify all the scales that are used or available for neurological assessment and monitoring in acute stroke. Further reviews were then undertaken to develop an overview of the clinimetric properties of the scales identified. The results are presented to allow direct comparison within and between different scales across a range of properties. Implications to practice of the clinimetric property data are highlighted and discussed throughout.

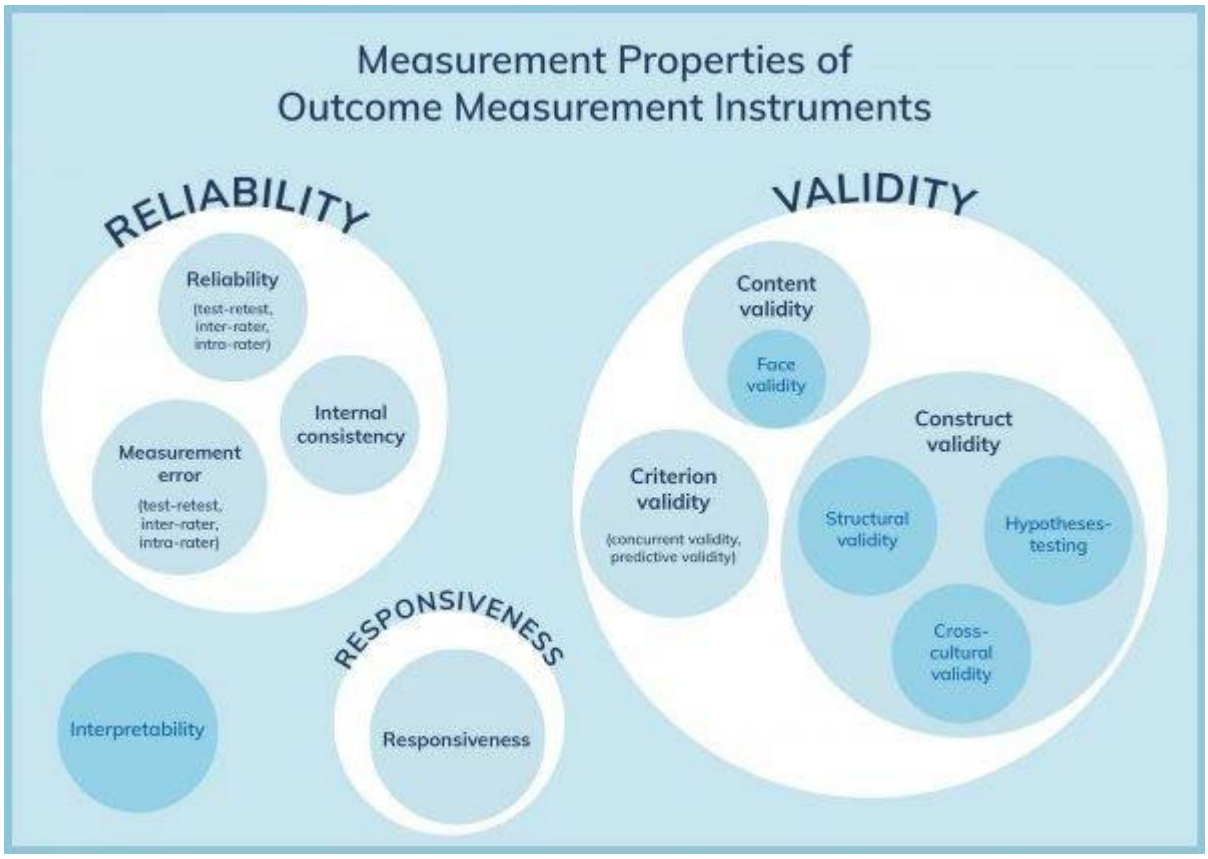
Chapter 4 Clinimetric Property Reviews

Assessments are central to clinical practice and health research. Scales used to undertake assessments need to measure the parameters of interest and be fit for purpose (Harrison, McArthur, & Quinn, 2013). Clinimetrics is 'a measurement of clinical phenomena' and promotes the use of clinical expertise opposed to pure statistical techniques to appraise and develop measurement instruments (Feinstein, 1987). Clinimetrics has been chosen as it integrates knowledge from a range of disciplines like psychometrics, epidemiology, and biostatistics to provide researchers and clinicians with the best methods and ways to assess, appraise and improve the methodological quality of their measurements.

This chapter details a series of reviews to describe the clinimetric property data of scales used, or available for, neurological assessment and monitoring in acute stroke. Firstly, a scoping review identified 26 scales. Secondly, searches on the scales' clinimetric properties were completed utilising the COSMIN (COnsensus-based Standards for the selection of health status Measurement INstruments) clinimetric property search design (Terwee et al., 2009). These searches are based around the COSMIN taxonomy (Figure 4.1.). The figure shows the relationships of measurement properties developed through an international delphi study to reach consensus on the terminology and definitions of measurement properties (Mokkink et al., 2010b). The taxonomy shows the grouping of the measurement properties, or aspects of measurement properties into three quality domains: reliability, validity, and responsiveness. Interpretability is included separately as although not a measurement property it was considered important for evaluating health measurements (ibid). After presenting the methods and results of these reviews the chapter concludes with a discussion of the overall findings including potential implications for clinical practice. The reviews were guided by input from the supervisory team, the stroke research team's information specialist (Janet Reed), and the NIHR Complex Review Support Unit.

Figure 4.1 COSMIN taxonomy of relationships of measurement properties.

COSMIN (COnsenus based Standards for the selection of health status Measurement Instruments)



Note. Adapted from (by original authors) The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes by Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. 2010, Journal of Clinical Epidemiology, 63 (7), p. 741. (doi.org/10.1016/j.jclinepi.2010.02.006). Reprinted with permission.

4.1 Methods

This section will describe the methods for both the scoping review and the subsequent separate clinimetric reviews. The scoping review identified all scales that are used, or available for, neurological assessment and monitoring in the acute phase (first 72 hours) after stroke. The subsequent clinimetric property reviews (26) obtained the clinimetric data for each of the scales identified from the scoping review.

4.1.1 Scoping Review

Chapter 2 introduced the scales that the author has seen used in clinical practice but there are other scales available. A scoping review approach was chosen to identify all scales used or available as this was a broad research question that wanted to explore all relevant literature

regardless of study design (Arksey & O'Malley, 2005). The review was designed to provide a comprehensive overview of all available scales and refine the subsequent clinimetric reviews through the development of robust search terms for each scale.

4.1.2 Scoping review- search strategy

Five databases (Medline (Ovid), Embase (Ovid), CINAHL (Ebsco), PsycINFO (Ebsco) and HMIC (Ovid)) were searched to identify scales in use. The search design was an iterative process of narrowing and broadening the terms to ensure comprehensive coverage. The final search terms were broad as more specific search terms limited useful results in that scales known to the author were not identified. The final design included all the Cochrane stroke strategy syntax and the terms neurologic examination, neurologic deterioration, and deterioration. As an example, the Medline search is presented in Appendix 4.1. A targeted grey literature search was undertaken between 23rd January 2019 and 22nd March 2019 and involved searches of the OpenGrey database, systematic review databases and research registries. The full list of grey literature sources is reported in Appendix 4.2.

4.1.3 Scoping review- inclusion and exclusion criteria

Inclusion and exclusion criteria for the scoping review were developed. Most scoping reviews aim to include all relevant papers on a topic. However, this review aimed to identify all scales used, or available, for neurological assessment and monitoring so these criteria mainly relate to the scales rather than the literature.

Scale inclusion

- Created for, or used in, neurological assessment and monitoring in the acute phase (first 72 hours) after stroke onset
- Only English language
- Any date range
- Any publication type.

Scale Exclusion:

- Created for identification of stroke or large vessel occlusion and not ongoing assessment
- Used solely for prognostication or prediction of outcome
- Used for interoperative or procedural monitoring
- Sedation scales
- modified Rankin Scale (mRS) (global disability score)
- Designed for use after 72 hrs (e.g., scales used in rehabilitation).

If the purpose of the scale was unclear on identification further information on scale use was sought prior to any inclusion decision.

4.1.4 Clinimetric property reviews search strategy

The search strategy for each of the clinimetric property reviews followed the principles and nomenclature of systematic reviews alongside considering the five elements (construct search, population search, instrument search, measurement properties filter, and exclusion filter) of COSMIN clinimetric property search design (Prinsen et al., 2018; Terwee et al., 2009). The elements and their application to the development of the search strategies are outlined below:

1. The construct search includes terms to search for the construct to be measured by the scales (e.g. assessment, monitoring or neurological deterioration). In the final searches theses search terms were omitted as it was apparent that inclusion severally limited results obtained and may have prevented appropriate data being identified.
2. The population search contains search terms for the population of interest. This was kept simple and consistent with exploded generic search terms for stroke.
3. The instrument search comprises the specific search terms for the scale of interest. Rapid scoping searches were completed for each scale in each of the databases to identify MESH terms and keywords to ensure all potential search terms and abbreviations were included.
4. The measurement properties filter ensures search terms related to the clinimetric properties are included. The COSMIN measurement properties search filters were created to address difficulties undertaking clinimetric reviews due to poor indexing, large variation in terminology and poor reporting of measurement properties (COSMIN, n.d.). The COSMIN search filters for use in Medline (Ovid), Embase (Ovid), and HMIC (Ovid) have been validated (Terwee et al., 2009). For the EBSCO databases (PsycInfo and CINAHL) two versions of non-validated translations of the Ovid search filter were available (ibid). As it was not clear which filter would be the most sensitive, both translations were run for the NIHSS. Due to the large number of papers returned, it was not feasible to complete a full side by side comparison of the search results and a random selection of papers (n=10) unique to each search were selected and reviewed for relevance. From this review of 20 papers, the larger search filter translated by Inger Abma, Radboud UMC was the only one to identify any papers of relevance and therefore this filter was used.
5. COSMIN recommend an exclusion filter to remove irrelevant records from searches such as case reports and animal studies. However, this was not utilised as the aim was

to capture all relevant data. Manual screening was undertaken to prevent any relevant papers being lost.

Searches using the chosen syntax were completed in Medline (Ovid), Embase (Ovid), CINAHL (Ebsco), PsycINFO (Ebsco) and HMIC (Ovid) for the 26 scales identified in the scoping review. Appendix 4.3 shows the Medline search terms for the NIHSS search as an example). Duplicates were removed and then titles, and abstracts were screened. Endnote software (Version Endnote X9) [64 bit], Clarivate Analytics, Philadelphia, PA, USA) was utilised throughout for management of references.

4.1.5 Clinimetric property reviews inclusion and exclusion criteria

Inclusion and exclusion criteria were developed that were broad to try and ensure data was not missed. However, they were also required to be pragmatic given the resources available to complete them.

4.1.5.1 Inclusion criteria

Papers were included if:

- they were written in English
- participants (or selection of) had a clinical diagnosis of stroke and were adults aged 18 or over
- data presented were relevant to neurological assessment and monitoring
- they included primary research methods and/or presented original data on one or more of the clinimetric properties of interest.

4.1.5.2 Exclusion criteria

Papers were excluded if:

- they presented data for other purposes such as prognostication or prediction of large vessel occlusion (LVO)
- they were audits, opinion or discussion papers, reviews, and editorials
- cross-cultural validity studies
- they involved retrospective data analysis (except in the papers focusing on scale construct through factor analysis).

If the clinimetric content was unclear from the abstract, the paper was put through for full text review to ensure papers were not excluded prematurely. If unclear at full text stage whether a paper was eligible for inclusion this was discussed with the Director of Studies and the final

decision was agreed and documented. Citation tracking was completed, and any papers meeting the inclusion and exclusion criteria not already identified were included. Details of numbers of papers for each scale by clinimetric property can be found in Appendix. 4.4.

4.1.6 Clinimetric property reviews data extraction

There were four key steps to data extraction. Firstly, data on the scale's characteristics were collected. Secondly, the scale's content in terms of items measured by the scales was assembled. Thirdly, data was extracted around the clinimetric properties before a final assessment of the methodological quality of the studies was undertaken in line with the COSMIN taxonomy. The COSMIN taxonomy was developed through an international Delphi study which decided which were the most important measurement properties and the most adequate terms and definitions (Mokkink et al., 2010a).

4.1.6.1 Scales' characteristics

The following characteristics for all scales, where available, were collected:

- Year of scale development
- Original purpose of the scale (identification, measurement, prognostication)
- Whether originally developed for research or clinical use
- Staff Group/s the scale was originally designed for
- Scale type (whether nominal, ordinal, interval or ratio)
- Possible scores- indicating the minimum and maximum scores obtainable within the scale
- Value indicating greater degree of severity- whether a higher or lower score indicates a greater degree of severity within the scale
- Stroke Type/ Circulations Assessed- if the scales were developed for specific stroke types or specific circulatory systems of the brain.

4.1.6.2 Scales' content

All items tested within the scales are derived from elements of the medical neurological examination and measure elements of neurophysiological function. However, the items assessed by the scales differs, so this data was extracted to allow comparison.

4.1.6.3 Clinimetric property data extraction

Data was extracted using pre-defined proformas for both original (first paper introducing scale and its development) and later papers (papers that report one or more clinimetric property of

the scale) These proformas (Appendix 4.5 & 4.6) were designed to ensure that the contextual factors (e.g., setting, health care professionals, stroke type) were collected alongside the clinimetric property data. Time to complete, although not a clinimetric property, was collected and presented. Data was grouped by property, summarised, and explained. Tables were used where possible to allow comparison between scales.

4.1.6.4 Assessment of methodological quality

Every study included was assessed for each clinimetric property it contained using the appropriate COSMIN box of standards (Mokkink et al., 2018). These checklists evaluate the methodological quality of studies on standards related to design and preferred statistical methods. Standards were assessed on a four-point rating system (very good, adequate, doubtful or inadequate). To determine the overall rating of quality the "worst score counts" principle was applied (Terwee et al., 2012).

4.2 Results

This section first provides results of the scoping review an overview of the literature obtained regarding the clinimetric properties of the 26 scales identified. It will then present information about scale content and characteristics before presenting the clinimetric properties data and assessment of methodological quality for each scale. The results will be presented in chronological order (oldest to newest scale) unless otherwise stated.

4.2.1 Scoping review

This broad search to identify scales resulted in 23,010 records being identified (22,968 from databases and 42 from other sources). Following removal of 4376 duplicates, title and abstract screening was undertaken on 18,634 records to identify stroke scales named or used within the literature. All publication types were included (abstracts, opinion or discussion papers, reviews, letters, and editorials). Full text was sought if the title or abstract did not identify the stroke scale used. Citation searching of relevant papers identified in the search, including forward tracking using the Science Citation Index, was completed in order to identify additional papers not captured by the search strategy that could include different scales. Grey literature searching did not identify any additional scales.

The scoping review identified 19 scales (List 1) (and 7 modified versions of the National Institutes for Health Stroke scale (List 2):

List 1 showing the 19 stroke scales identified:

- National Institutes for Health Stroke Scale (NIHSS) (Brott et al., 1989)
- Glasgow Coma Scale (GCS) (Teasdale & Jennett, 1974)
- Scandinavian Stroke Scale (SSS) or Scandinavian Neurological Stroke Scale (SNSS) (Scandinavian Stroke Study Group, 1985)
- Canadian Neurological Scale (CNS) or Canadian Neurological Stroke Scale (CNSS) (Cote et al., 1986)
- Full Outline for UnResponsiveness (FOUR) Score (Widjicks et al., 2005)
- Standardised Nursing OBservation for Stroke (SNOBS) (Birschel, 2005)
- Middle Cerebral Artery Neurological Score (MCANS) (Edwards et al., 1995)
- Miami Emergency Neurological Deficit (MEND) (Physio-pedia, 2023)
- European Stroke Scale (ESS) (Hantson et al., 1994)
- Unified Neurological Stroke Scale (Edwards et al., 1995)
- Japan Stroke Scale (Gotoh et al., 2001)
- The Hemispheric Stroke Scale (HSS) (Adams et al., 1987)
- The Mathew Stroke Scale (Mathew et al., 1972)
- The Orgogozo Scale (Orgogozo et al., 1983)
- Toronto Stroke Scale (TSS) (Brown et al., 1990)
- Chinese Stroke Scale (CSS) (Chen, 1995)
- Israeli Vertebrobasilar Stroke Scale (IVBSS) (Gur et al., 2007)
- Post Stroke Assessment Tool (PSAT) (*Stubits et al., 2015*)
- Modified Edinburgh- Scandinavian Stroke Scale (MESSS) (Criteria for the degree of clinical neurological impairment in stroke patients,1995).

List 2 showing the 7 modified versions of the National Institutes for Health (NIHSS):

- Expanded National Institutes of Health Stroke Scale (e-NIHSS) (Olivato et al., 2016)
- Modified National Institutes of Health Stroke Scale (mNIHSS or Modified NIH Stroke Scale (Lyden et al., 2001)
- Shortened/ Simplified versions of the National Institutes of Health Stroke Scale, Shortened/Simplified NIH Stroke Scale, Shortened/Simplified NIHSS (Lyden et al., 2001)
- NIHSS-11 (Lyden et al., 2001)
- NIHSS-8 (Demeestere et al., 2017; Garnett et al., 2010)
- NIHSS-5 (Lee et al., 2016)

- Shortened NIH Stroke Scale for Emergency Medical Services (sNIHSS-EMS) (Purrucker et al., 2017)

All eligible identified scales were included regardless of their age. However, in order to give an indication of which scales are the most recently used a table of their most recent occurrence with the literature indexed in PubMed was created (Appendix 4.7).

4.2.2 Clinimetric property reviews

4.2.2.1 Overview of the literature

Twenty-six searches were completed, one for each scale identified in the scoping review. The Chinese Stroke Scale (CSS) and the Modified Edinburgh Scandinavian Stroke Scale (MESSS) were acknowledged to assess the same items (Zhao et. al, 2018). Upon commencing data extraction, it became evident that the Orgogozo and the Middle Cerebral Artery Neurological Score (MCANS) and the modified NIHSS (mNIHSS) and NIHSS-11 also assessed the same items, so their searches were amalgamated. The NIHSS-8 and the Hunter NIHSS 8 were found to contain different items, so were treated as two separate scales. Scale characteristics data was obtained from scales and the literature and is presented in Table 4.3. Copies of all the scales that are available in Appendices 4.8- 4.20. The authors of the CSS/ MESSS and the PSAT were contacted to obtain copies of scales, but no response was received. Content data is therefore only presented for the 22 available scales (Table 4.4).

The number of papers with relevant clinimetric data across all scales was less than expected and there were notable differences in the number of papers identified for each scale. Two letters (*Berthier et al., 2013, Schmulling et al., 1998*) and eight abstracts (*Binz et al., 2013; Boutot et al., 2013; Brown et al., 1990; Cabal et al., 2018; Guterud et al., 2019; Isahaya, 2017; Peters et al., 2012*) were included as they contained primary data that was not reported elsewhere. The data presented in the Berthier letter was combined with data from another paper (Berthier et al., 2012) and treated as one reference throughout the review. Although limited data on design is routinely found in abstracts and letters, which can hamper quality assessment, the need to obtain all data outweighed this potential risk. Throughout this chapter the abstracts and the *Schmulling et al.* (1998) letter are italicised for easy of identification. Two theses were included, one was identified through the searches (Specogna, 2013) the other through citation (Birschel, 2005). No scale had data for all clinimetric properties (Table 4.1).

Table 4. 1 Visual indication of the numbers of papers by scale that include information on clinimetric properties of interest.

	Total No of Papers	Content Validity	Criterion Validity (concurrent)	Construct Validity	Internal Consistency	Inter-rater Reliability	Intra-rater Reliability	Measurement Error	Responsiveness	Time to complete
Mathew										
GCS									*	
Toronto			*							
SSS										
CNS				H						*
Hemispheric								*		*
NIHSS				*						
MCANS/ Orgogozo										
Unified										
ESS		*								
Chinese										
SNOBS										
MEND										
Japan										
mNIHSS / NIHSS-11				*		*			*	
NIHSS-8				*					*	
NIHSS-5				*					*	
FOUR Score									*	
IVBSS										
Hunter NIHSS-8										
sNIHSS										
e-NIHSS					*		*	*		
PSAT			*							
sNIHSS-EMS										

Key

Colour	Number of Papers
	0
	1-5
	6-10
	10+

*Includes paper that reported covered but either not or not applicable to review question.
[H] -only paper to present validity hypothesis

Overall, the quantity of data across all scales and properties was low except for inter-rater reliability and measurement error data for the NIHSS. It is worth noting that although some papers focused on a single property, many presented data across two or more properties.

Papers were initially identified for the CSS/ MESSS, the Miami Emergency Neurological Deficit (MEND) exam, the PSAT and the sNIHSS-EMS (the shortened NIH Stroke Scale for emergency medical services), however, upon commencing data extraction no usable data applicable to the review was presented so these papers were excluded. One abstract for the PSAT claimed to report on validity but actually reported significance levels rather than true results for its correlation with the NIHSS so also had to be excluded (*Stubits et al., 2015*). One paper stated it reported on construct validity for the NIHSS, mNIHSS, NIHSS-8 and NIHSS-5 but the online results were not available, and the author did not respond to a data request (Lee et al., 2016). However, the paper did contain data on other clinimetric properties so was not fully excluded. Following these limitations and exclusions, clinimetric data extraction was only possible for 20 scales (Table 4.2). However, all scales were included in the scale characteristics and content if there was information available to show trends and patterns.

Table 4.2 Scales for which clinimetric data extraction was possible.

Mathew
GCS
Toronto
SSS
CNS
Hemispheric
NIHSS
MCANS/ Orgogozo
Unified
ESS
SNOBS
Japan
mNIHSS / NIHSS-11
NIHSS-8
NIHSS-5
FOUR Score
IVBSS
Hunter NIHSS-8
sNIHSS
e-NIHSS

4.2.2.2 Scale Characteristics

Key characteristics of each scale are presented in Table 4.3. All but two scales, GCS and FOUR score, were created for stroke populations. Ten scales were developed for use in hospital clinical practice: GCS, SSS, CNS, SNOBS, MEND, Japan, FOUR Score, s-NIHSS, e-NIHSS (extended), and PSAT. Three were for pre-hospital use: NIHSS-5, Hunter NIHSS-8 and the sNIHSS-EMS. Eight originated from research settings: Mathew, Toronto, HSS, NIHSS, MCANS, Unified, ESS, and the mNIHSS, though some, such as the NIHSS, have been adopted into clinical practice. For one scale, the NIHSS-8 the original purpose and context of development is not clear.

Many scales were developed for specific staff groups, although this was not always stipulated. Earlier scales and those developed for research were more often created for use by medical staff. Modified or simplified versions of scales, generally with fewer but the more reliable items of the original scales were tailored for use by a multi-disciplinary team. Data is limited on the PSAT, but it appears to have been developed specifically for nurses (*Stubits et. al, 2015*).

Most of the scales are classified as ordinal. However, many, like the NIHSS, are not a true ordinal scale of impairment as some items conflict (e.g., there will be no evidence of Extinction in a comatose patients) and nonassessable items must score zero (Muir et al., 1994). The Japan stroke scale is a weighted parametric scale. The Toronto and the IVBSS are arbitrarily weighted scales. The basis of the weighting for the Toronto scale is not reported whereas the IVBSS is based on expert opinion or a given value (Gur et al., 2007). The SNOBS, a shorter scale derived from the SSS, is used as a nominal scale in practice but has been assessed as an ordinal scale (Birschel, 2005). The MEND is a purely nominal scale. As the Chinese / MESSS and the PSAT are not available their scale type is unknown.

The stroke type or circulation the stroke specific scales were created to assess varied (Table 4.3), though this was not always stated e.g the PSAT did not specify but was only reported in acute ischaemic stroke. Some were developed exclusively for ischaemic strokes e.g. HSS and some for both ischaemic and ICH strokes e.g. CNS, Unified, and Japan. However, the CNS was not designed to be used for individuals whose Level of Consciousness (LOC) is worse than drowsy. Some stipulated they were for specific arterial territories such as the MCANS and the ESS were created for middle cerebral artery strokes, the Mathew and the NIHSS for both anterior and posterior circulation, the SSS anterior territory, and the IVBSS the vertebrobasilar circulation. Although many scales reported they could be, or have been used for multiple or all

stroke types and territories there was evidence to suggest that they had not been developed for that e.g. the NIHSS was originally being developed for use in ischaemic stroke research trials and the SNOBS was derived from the SSS so really based on anterior circulation. It appears that scales have been adopted into use across stroke types without having been developed or tested within whole populations so they could potentially not be fit for purpose. There was limited data on the Toronto scale, so an author was contacted via e-mail, to confirm the data provided however it was not possible to confirm the intended purpose of the Chinese Stroke Scale/ MESSS.

Table 4.3. Key characteristics of stroke scales presented chronologically

Tool/ Scale	Year developed	Original Purpose	Developed for Clinical or Research	Staff Group	Scale Type	No of items	Possible Scores	Value indicating greater degree of severity	Stroke Type and / or Circulations Assessed
Mathew	1972	Measurement	Research	Neurologists	ordinal	10	100 to 0 (death)	Lower	Anterior and Posterior
GCS	1974	Measurement and prognostication	Clinical	Multidisciplinary team	ordinal	3	3-15	Lower	None (Designed for Traumatic Brain Injury)
Toronto	1976	Measurement	Research	Physicians	weighted arbitrarily	16	0-44	Higher	All stroke types
SSS	1985	Measurement	Clinical	Developed for non-neurologists.	ordinal	9	2-56	Lower	Anterior
CNS	1986	Measurement	Clinical	Generic staff	ordinal	8 or 6 if comprehension deficit	1.5-11.5	Lower	Designed to be used in any stroke type- not with LOC below drowsy
HSS	1987	Measurement	Research	Neurologist or neurosurgeon	ordinal	25	0- 100	Higher although uses GCS as LOC score	Infarction
NIHSS	1989	Measurement	Research	Initially physicians on trials but has been rolled out	ordinal	15	0-42 (actually 35)	Higher	Anterior and Posterior Circulation

Tool/ Scale	Year developed	Original Purpose	Developed for Clinical or Research	Staff Group	Scale Type	No of items	Possible Scores	Value indicating greater degree of severity	Stroke Type and / or Circulations Assessed
				to other professionals					
MCANS/ Orgogozo	1991	Identification	Research	Unclear but trial investigators	ordinal	10	0-100	Lower	Middle cerebral artery stroke
Unified (MCANS & SSS)	1992	Measurement	Research	Unclear	ordinal	12	0-158	Lower	Ischaemic and ICH
ESS	1994	Measurement	Research	Neurologists in seminal paper- Trial investigators	ordinal	14	0-100	Lower	Middle cerebral artery stroke
Chinese Stroke Scale/ MESSS	1995	Measurement	Unclear	Unclear	unclear	8	0-45	Higher	Unclear
SNOBS	1998	Measurement	Both	Any staff completing bedside assessment	ordinal or nominal	5 (from SSS)	n/a	Lower	All stroke types
MEND	2000	Measurement	Clinical	All clinical staff on stroke pathway including paramedics	nominal	11	n/a	Presence of abnormal finding	All stroke types
Japan	2001	Measurement	Clinical	Physicians (neurologists, internists and neurosurgeons)	weighted parametric	10 (17 questions)	-0.38-27.86	Higher	All stroke types
mNIHSS/	2001	Measurement	Research	Unclear	ordinal	11	0-32	Higher	All stroke types

Tool/ Scale	Year developed	Original Purpose	Developed for Clinical or Research	Staff Group	Scale Type	No of items	Possible Scores	Value indicating greater degree of severity	Stroke Type and / or Circulations Assessed
NIHSS-11									
NIHSS-8	2002	Identification & Measurement	Not documented	Unclear	ordinal	8	0-24	Higher	All stroke types
NIHSS-5	2002	Measurement	Clinical (pre-hospital)	Paramedics and stroke team members	ordinal	5	0-16	Higher	All stroke types
FOUR Score	2005	Measurement	Clinical	Intensive care unit staff. (neuro nurses, neurology residents or fellows, and neuro-intensivists.	ordinal	4	0-16	Lower	Not purely stroke. Used with or without endotracheal intubation
IVBSS	2007	Measurement	Unclear	Unclear	weighted arbitrarily	11	0-44	Higher	Vertebrobasilar stroke
Hunter NIHSS-8	2010	Identification & Measurement	Clinical (pre-hospital)	Paramedics and stroke team members	ordinal	8	0-20	Higher	All stroke types
sNIHSS	2011	Measurement	Clinical	Unclear	ordinal	12	0-35	Higher	All stroke types
e-NIHSS	2015	Measurement	Clinical	Unclear	ordinal	15	0-42 (3 items changed)	Higher	All stroke types. Extended to include more posterior signs
PSAT	2015*	Measurement	Clinical	Nurses	Unknown	4 (unknown total number of questions)	0-63	Higher	Acute ischaemic stroke

Tool/ Scale	Year developed	Original Purpose	Developed for Clinical or Research	Staff Group	Scale Type	No of items	Possible Scores	Value indicating greater degree of severity	Stroke Type and / or Circulations Assessed
sNIHSS-EMS	2017	Identification & Measurement	Clinical (pre-hospital)	Paramedics and stroke team members	ordinal	9	0-29	Higher	All stroke types

* indicates year of publication of abstract reporting scale

4.2.2.3. Scale Content

Stroke signs and symptoms are diverse due to the heterogeneity of the disease. As scales were created for different purposes the components of the nervous system that they assess varies (Table 4.4). Each item of assessment within the scales assesses a different component. The total numbers of items by scales ranges from 3 (GCS) to 25 (HSS). Although many of the scales assess common components, there is widespread variation across scales in what items they include and how they measure the component, making comparisons challenging.

To aid evaluation, scale components have been organised into four domains (Table 4.4):

- Alertness: items that evaluate Level of Consciousness (LOC)
- Vision or Sensory: items that consider aspects of ability to see or feel
- Involuntary: items that assess autonomic actions such as Pupillary Response and Reflexes
- Voluntary: items assessing somatic actions such as voluntary movement or reports on ability of these functions.

All scales except the NIHSS-5 include at least one item that assesses LOC in some way. Eight scales assess both visual and sensory function (Toronto, HSS, NIHSS, MEND, Japan, mNIHSS, e-NIHSS and the IVBSS), two only vision (ESS and NIHSS-5) and two only sensory (Mathew, sNIHSS-EMS). The IVBSS is the only scale to formally assess Diplopia.

The most common items for assessment of involuntary features were Gaze, Conjugate Eye Deviation or Extraocular Eye Movements (19 scales) and Extinction/Neglect (7 scales). Less common items were Tone, Reflexes, and Respiration. There were multiple items (n=22) to assess voluntary movements, most frequent were Motor Power (19 scales), Speech and Language Function (18 scales), and Facial Palsy (16 scales). Of the six scales that don't assess Best Language, four assess Dysarthria (NIHSS-8, Hunter NIHSS-8, PSAT, and sNIHSS-EMS). The FOUR Score and the PSAT do not to assess Speech in any capacity.

Table 4.4 Components of the nervous system assessed by items within stroke scales

	Mathew	GCS	Toronto	SSS	CNS	Hemispheric (HSS)	NIHSS	MCANS/Orgogozo	Unified	ESS	SNOBS	MEND	Japan	mNIHSS/NIHSS-11	NIHSS-8	NIHSS-5	FOUR Score	IVBSS	Hunter NIHSS-8	sNIHSS	e-NIHSS	sNIHSS-EMS
Alertness																						
Level of Consciousness (LOC)	✓	✓	✓	✓$_p$	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	-	✓	✓	✓	✓	✓	✓
LOC Questions	-	-	-	-	-	-	✓	-	-	-	-	✓	-	✓	✓	-	-	-	✓	✓	✓	-
LOC Commands	-	✓	-	-	-	✓	✓	-	-	✓	-	✓	✓	✓	✓	-	✓	-	✓	✓	✓	-
Orientation	✓	-	-	✓$_l$	✓	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-
Vision or Sensory																						
Visual Fields	✓	-	✓	-	-	✓	✓	-	-	✓	-	✓	✓	✓	-	✓	-	✓	-	-	✓	-
Diplopia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-
Sensation	✓	-	✓	-	-	✓	✓	-	-	-	-	✓	✓	✓*	-	-	-	✓	-	-	✓	✓
Involuntary																						
Gaze/ Conjugate Eye Deviation Extraocular Eye Movements	✓	-	✓	✓$_p$	-	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	✓*	-
Pupillary Abnormality	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-
Extinction/Neglect	-	-	-	-	-	✓	✓	-	-	-	-	-	✓	✓	✓	-	-	-	✓	-	✓	-
Muscle Tone	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Upper Limb Tone	-	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-
Lower Limb Tone	-	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-
Upper and lower limb asymmetry	-	-	-	-	✓$_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Plantar Reflexes	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-
Deep Tendon Reflex	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

	Mathew	GCS	Toronto	SSS	CNS	Hemispheric (HSS)	NIHSS	MCANS/Orgogozo	Unified	ESS	SNOBS	MEND	Japan	mNIHSS/NIHSS-11	NIHSS-8	NIHSS-5	FOUR Score	IVBSS	Hunter NIHSS-8	sNIHSS	e-NIHSS	sNIHSS-EMS
Pathologic Reflexes	✓	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Brainstem Reflexes	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-
Respiration	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-
Voluntary																						
Facial Palsy	✓	-	✓	✓$_{I}$	✓$_{\$}$	✓	✓	✓	-	✓	-	✓	✓	-	✓	-	-	-	✓	✓	✓*	✓
Motor Power- Affected Arm	✓	-	✓	✓ p&I	-	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	-	-	✓	✓	✓	✓	✓
Motor Power- Unaffected Arm	✓	-	-	-	-	-	✓	-	-	-	-	✓	-	✓	-	-	-	✓	✓	✓	✓	✓
Proximal Arm	-	-	-	-	✓$_{1}$	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distal Arm	-	-	-	-	✓$_{1}$	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Motor Power-Affected Leg	✓	-	✓	✓ p&I	-	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	-	✓	✓	✓
Motor Power- Unaffected Leg	✓	-	-	-	-	-	✓	-	-	-	-	✓	-	✓	✓	✓	-	✓	-	✓	✓	✓
Proximal Leg	-	-	-	-	✓$_{1}$	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distal Leg	-	-	-	-	✓$_{1}$	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Foot Dorsiflexion	-	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-
Shoulder function	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Hand (movement, power)	-	-	-	✓$_{I}$	-	-	-	✓	✓	-	-	-	✓	-	-	-	-	-	-	-	-	-
Wrist Extension	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-
Finger Strength	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-
Ataxia	-	-	✓	-	-	-	✓	-	-	-	-	✓	-	-	-	-	-	✓	-	✓	✓*	-
Gait/ Walking	-	-	-	✓$_{I}$	-	✓	-	-	✓	✓	-	-	-	-	-	-	-	✓	-	-	-	-

	Mathew	GCS	Toronto	SSS	CNS	Hemispheric (HSS)	NIHSS	MCANS/Orgogozo	Unified	ESS	SNOBS	MEND	Japan	mNIHSS/NIHSS-11	NIHSS-8	NIHSS-5	FOUR Score	IVBSS	Hunter NIHSS-8	sNIHSS	e-NIHSS	sNIHSS-EMS
Performance/Disability Status	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Speech/ Best Language	✓	✓	✓	✓$_{l}$	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	-	-	-	✓	✓	-
Dysarthria	-	-	✓	-	-	✓	✓	-	-	-	-	✓	-	-	✓	-	-	✓	✓	✓	✓	✓
Dysphagia	-	-	✓	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-
Dementia			✓																			
Higher cortical function (frontal, parietal			✓																			

Key:

*signifies different way of assessing to original NIHSS- only marked for modified NIHSS scales

$_{\S}$ Assessed in both parts of the CNS

$_{1}$ Section A1 of CNS used where no comprehension deficit present

$_{2}$ Section A2 of CNS where there is comprehension deficit present

$_{p}$ prognostic score elements of SSS

$_{l}$ long-term score elements of SSS

^ described as assessing limb coordination

4.2.2.4. Scale Language

Language and communication barriers have been shown to affect quality of care (Flores et al., 2003; Flores et al., 2005; Scheppers et al., 2006). To ensure effective communication people need to interpret language in the same way. However, the way we interpret language is affected by multiple factors such as cultural background, educational background as well as mood and personality. The effectiveness will vary depending upon those involved in any communication.

Acute stroke patients are at increased risk of language difficulties influencing assessments and care outcomes compared with other conditions (Rohde et al., 2018). Dependent upon the area of brain affected by the stroke the ability to comprehend and/or produce speech may be affected (Price et al., 2010). Studies estimate that language impairments occur anywhere between 15% to 42% of acute stroke patients Ineffective communication whether through misunderstanding of meaning or impairment can have a direct influence on neurological assessment and monitoring especially if eliciting subjective patient symptoms (Rohde et al., 2018). Patients may not understand what is asked of them especially if the test is complex or the instructions difficult to understand which can be the case when medical terminology, or other elements of the assessments are complicated. Even amongst cognitively intact individuals with no speech impairment elements of language assessments, such as in the NIHSS, have been shown to be inconsistent amongst different individuals (Burns et al., 2014). It will therefore be important in development and testing of the SNOBSS to consider the language used to ensure as much consistency of application as possible across multiple individuals when applied by different staff members.

4.3 Clinimetric Properties Data and Assessment of Methodological Quality

This section presents the clinimetric properties data in the following order: content validity, criterion (concurrent) validity, construct validity, internal consistency, inter-rater reliability, intra-rater reliability, measurement error and responsiveness. This is followed by time to complete data. Data for each property is presented and summarised in turn before synthesising the results in a final discussion around implications.

4.3.1 Content validity

Content validity is the extent to which the assessment adequately measures the construct under investigation (de Vet et al., 2011). In clinimetrics a construct is a well-defined and

precisely demarcated subject of measurement. Unless scales provide a clear description of the construct that is being measured content validity cannot be properly assessed.

Ten papers included information on content validity each on a different scale (Adams et al., 1987; Birschel, 2005; Brott et al., 1989; Cote et al., 1986; Garnett et al., 2010; Gotoh et al., 2001; Hantson et al., 1994; Lyden et al., 2001; Scandinavian Stroke Study Group, 1985; Wijdicks et al., 2005).

A five-step process of content validation (de Vet et al., 2011 pgs. 156-159) was used to present and appraise papers' content validation data. These five steps are discussed in turn before an overview of which steps were considered or completed for each scale is presented (Table 4.6).

Step 1: Consider information about construct and situation

Only one paper claimed to report underlying constructs of a scale, although, the information presented was about the underlying internal structure of the mNIHSS scale compared to the full NIHSS and did not therefore actually provide construct details (Lyden et al., 2001). No other papers reported information on construct. All papers provided information on the situation, that is the purpose of measurement for the target population, and these are presented in Table 4.5. below. No papers reported clear information on the theoretical background underpinning the scales. All scales (n=20) that were included in these reviews are multi-item reflective models (measuring items that are manifestations of the extent of the stroke).

Table 4.5 Situation (purpose of measurement for the target population) reported by scale and paper.

Scale	Paper	Situation (purpose of measurement for the target population) reported
SSS	Scandinavian Stroke Study Group 1985	Developed for a haemodilution study. For use by non-neurologists.
CNS	Cote 1986	Developed as a scale that could assess both conscious and aphasic patients.
Hemispheric	Adams 1987	Developed as a standardised neurological assessment scoring instrument for use in a multicentre trial of hypervolemic haemodilution in acute hemispheric stroke.
NIHSS	Brott 1989	Developed as a stroke neurological examination scale for use in acute stroke therapy trials.
ESS	Hantson 1994	Developed to detect therapeutic effect and match treatment groups in stroke trials.
SNOBS	Birschel 2005	Developed to detect significant changes in clinical neurological status.
Japan	Gotoh 2001	Developed as a quantifiable scale for estimation of severity that provides information on the relative weights of the items included.
mNIHSS/ NIHSS-11	Lyden 2001	Developed to improve the NIHSS scale for use in clinical research.
FOUR Score	Wijdicks 2005	Developed as a new coma scale to address the shortcomings of the GCS. Allows assessment of verbal ability in intubated patients and brainstem reflexes.
Hunter NIHSS-8	Garnett 2010	Developed as an eight-point version of the NIHSS for use by paramedics in the prehospital setting to assess patient's potential eligibility for stroke thrombolysis.

Step 2: Consider information about content of the measurement instrument

All papers either described all items or provided a copy of the scale. The level of detail about the focus and development of the scales' content provided in the literature varied considerably. For the ESS, despite claims there was no detail on how content validity was achieved (Hantson et al., 1994). The SSS focused on including items easy to assess and of functional significance to patients and the decisions were based on previous expert opinion (Scandinavian Stroke Study Group, 1985). The CNS also focused on including items associated

with functional status. It removed Gaze Paresis as it was previously stated that it associated other items assessed (LOC or Motor Power) (Oxbury et al., 1975) so was deemed redundant and because from their own experience, they felt nurses found it difficult to evaluate (Cote et al., 1986).

The HSS, NIHSS and Japan used variable studies to guide their content development. The HSS authors applied a graded neurological examination to 25 hemispheric stroke patients and selected the most pertinent deficits that could be reliably and rapidly assessed in patients. They felt that functional measures were impractical to assess in the acute period and removed them (Adams et al., 1987). The NIHSS authors conducted a pilot study of 10 ischaemic stroke patients (within 3 weeks of onset) and completed qualitive comparison with other scales. The composite included items to test for mental status and the presence of neurological signs in the distribution of each of the major arteries of the brain (Brott et al., 1989). The Japan authors completed multivariate analysis of the data of 1274 stroke patients (within 72 hrs of onset) admitted to Keio University Hospital in a 5-year period. Items were selected as variables that could predict functional dependence or death. These items were then re-evaluated and modified to improve distribution and sensitivity over a 4-8 week period with 65 new patients (48 ischaemic & 17 ICH) to create a temporary unweighted stroke scale. Relative weights were then calculated using conjoint analysis to measure the relative importance of the items within the scale (Gotoh et al., 2001).

Content development for some scales were based on previously assessed properties of previous or earlier versions of scales. The SNOBS authors selected items based on their prognostic relevance after stroke that they thought would be the most useful to detect change. They chose items from the SSS that had previously shown good reliability (Birschel, 2005). The mNIHSS creators used former clinimetric analysis of the NIHSS and factor analysis (technique to reduce large numbers of items into related factors) to delete poorly reproducible or redundant items (Lyden et al., 2001). The FOUR Score developers provided justification for the addition of items and the provision of extensive detailed instructions and pictures to follow to overcome shortcomings of the GCS, such as the inability to reflect the severity of coma (Wijdicks et al., 2005). During Hunter NIHSS-8 development items were chosen based on their published discriminatory values and practicality for paramedics in the prehospital setting (Garnett et al., 2010).

Step 3: Select an expert panel

The Hunter-8 NIHSS was the only scale that reported independent expert assessment on content validity. A workshop was held including senior ambulance personnel (including clinical, operational, and training personnel) and members of the John Hunter Hospital Acute Stroke Team. This group felt none of the available tools were suitable so decided to select items from the widely used NIHSS as described above (Garnett et al., 2010). All the other scales chosen content was overseen by the expert researchers/clinicians that developed them.

Step 4: Assess whether the content of the measurement instrument corresponds with the construct

During scale development consideration should be given to whether the scale is relevant and comprehensive for what it is measuring. The NIHSS validity was tested prospectively in 65 acute stroke patients by comparing scale scores with measured infarct size on CT (computerised tomography) scan at 1 week and clinical outcome at 3 weeks (Brott et al., 1989). To confirm content validation of the weighted Japan scale it was tested in 133 acute stroke patients (96 ischaemic and 37 ICH) and the distribution of the categorized variables was evaluated by the authors (Gotoh et al., 2001). The mNIHSS was reported as appearing to be clinimetrically identical to the original NIHSS and the same data was used for checking validation and reliability (Lyden et al., 2001).

The FOUR score was the only scale that got users to assess face validity, an aspect of content validity. Nine examiners were asked to grade on a five-point Likert scale for the following statements: (1) The FOUR score is clinically relevant and easy to use; (2) The FOUR score is obtained in a matter of minutes; (3) The FOUR score is a good alternative to GCS; (4) The FOUR score is a better score than GCS when looking for depth of coma or patient deterioration; and (5) The FOUR score is a coma assessment scale I would use if it becomes generally accepted. All raters agreed or strongly agreed (4 or 5) with all five statements (Wijdicks et al., 2005).

Step 5: Use a strategy or framework to assess the correspondence between the instrument and construct

None of the papers describe a strategy or framework through which they addressed relevance and comprehensiveness during the development process.

No papers described all of the five steps of content validation (Table 4.6). Some steps were completed and described more than others. Consideration of information about content of the measurement instruments (step 2) was the most completed and no papers described using a

strategy or framework to assess the correspondence between the instrument and construct (step 5).

Table 4.6 Which of the five steps of content validation were completed per scale

Tool	Paper	Considered information about construct and situation	Considered information about content of the measurement instrument	Selected an expert panel	Assessed whether the content of the measurement instrument corresponds with the construct	Used a strategy or framework to assess the correspondence between the instrument and construct
SSS	Scandinavian Stroke Study Group 1985					
CNS	Cote 1986					
Hemispheric	Adams 1987					
NIHSS	Brott 1989					
ESS	Hantson 1994					
SNOBS	Birschel 2005					
Japan	Gotoh 2001					
mNIHSS/ NIHSS-11	Lyden 2001					
FOUR Score	Wijdicks 2005					
Hunter NIHSS-8	Garnett 2010					

Key

Yes	
Partially	
No	

Summary

Content validity should be an integral part of scale development. All papers presented or described the scale which would be expected but the reporting of content validity was limited and inconsistent. The ESS paper provided no detail on content validity despite discussing its importance in the development of a good stroke scale (Hantson et al., 1994). Although more information was available for the SSS it was not presented or reported as content validity data (Scandinavian Stroke Study Group, 1985). The papers generally focused on face validity and the addition or removal of items rather than full content validity. The variability in content validation reported could be because the criteria for content validity assessment is mainly subjective and unquantifiable, so the importance has been overlooked.

Choice of items, and subsequent removal of others needs further exploration. Items need to be chosen based on whether they are clinically valuable to the phenomenon of interest (identification of change and END). COSMIN state that content validity is the most important measurement property and if overlooked the risk of bias is both subjective and extensive (Terwee et al., 2018). None of the scales provided a clear description of the construct to be measured or their theoretical underpinning. In sum, all the papers rate inadequate on the risk of bias checklist.

4.3.2 Criterion Validity

Criterion validity is defined as 'the degree to which the scores of a measurement instrument are an adequate reflection of a gold standard' (Mokkink et al., 2010a). The level of agreement required should be pre specified but this can be difficult to do as it depends upon the situation in which it is being used. Multiple factors such as costs, burden, false positives, and false negatives will impact on the criterion validity of scales. Criterion validity may therefore be reported as sufficient rather than optimal. Criterion validity can be separated into concurrent validity and predictive validity (de Vet et al., 2011). This thesis only presents concurrent validity data as it is the evaluative properties of scales and not their predictive applications that are of interest.

4.3.2.1 Concurrent Validity

There is no gold standard available for neurological assessment and monitoring (Specogna, 2013). Therefore, what is presented are comparisons with other well-established tests within stroke. For any scales that are ordinal, which the majority were, the spearman rank correlation coefficient should be used (Ramzai, 2020) because it determines the strength and nature of

the monotonic relationship between two variables rather than the strength and direction of the linear relationship. The Spearman correlation coefficient, *r*s, can take values from +1 to -1. The closer the value is to zero the weaker the association. Negative correlations are obtained when comparing scales, where greater severity is indicated by higher scores in one and lower scores in the other.

Thirteen papers presented data on concurrent validity across twelve scales (Adams et al., 1987; Bessenyei et al., 2001; Brott et al., 1989; Cote et al., 1989; Gur et al., 2007; Hantson et al., 1994; Lee et al., 2017; Lyden et al., 2001; Meyer et al., 2002, 2005, 2008; Olivato et al., 2016; Wijdicks et al., 2005). One abstract compared the Mathew Scale, the HSS, the Toronto stroke scale and the Barthel Index but used a Pearson correlation coefficient which is only suitable for continuous data, so this had to be discounted (*Brown et al., 1990*). Table 4.7. shows all the data and includes the methodological quality assessment of each paper. Papers are presented alphabetically as opposed to chronologically by scales as many of the papers present on more than one scale or involve comparisons with other identified scales. Statistical significance was provided in a few papers, but these are not presented as they do not provide information about the strength of the relationship.

Table 4.7 Concurrent validity data available for identified scales

(Data only shown for different sub-groups or examiners when combined data not available)

Paper	Scale of Interest	Scale compared with	Spearman rank correlation coefficient	
Adams 1987 (VG)	HSS	Inverted BI	0.87	
Bessenyei 2001(VG)	NIHSS	Mathew	-.903	
	NIHSS	SSS	-.896	
	NIHSS	Orgogozo	-.858	
	Mathew	SSS	.899	
	Mathew	Orgogozo	.826	
	SSS	Orgogozo	.91	
Brott 1989 (VG)	NIHSS- Baseline	Baseline Lesion Volume (NCCT)	0.39	
	NIHSS- Baseline	7-day lesion volume	0.78	
	NIHSS- 7 day	7-day lesion volume	0.74	
Cote 1989 (VG)	CNS- LOC Item	NE- LOC Item	0.574 (0.384-0.764)	
	CNS- Orientation	NE- Orientation	0.716 (0.583-0.849)	
	CNS- Speech	NE- Speech	0.691 (0.581-0.801)	
	CNS- Weakness	NE- Weakness	0.767 (0.695-0.839)	
	CNS- Global	NE- Global	0.755 (0.720-0.830)	
Gur 2007 (VG)	IVBSS	NIHSS	0.767	
	IVBSS	mRS	0.726	
	NIHSS	mRS	0.585	
Hantson 1994 (VG)	ESS	MCANS	0.95	
	ESS	CNS	0.93	
	ESS	MCANS	0.95	
	ESS	SSS	0.94	
	ESS	BI	0.84	
	ESS	mRS	0.86	
	ESS- Motor score	Brunnstrom Fugl-Meyer score	0.92	
Lee 2017 (VG)			Doctors	Nurses
	FOUR	GCS	0.871	0.914
Lyden 2001 (I)			PT	rtPA
	mNIHSS at baseline	mNIHSS at 90d	0.50	0.47
	mNIHSS at 2h	mNIHSS at 90d	0.56	0.65
	mNIHSS at 24h	mNIHSS at 90d	0.74	0.76
	mNIHSS at 7-10d	mNIHSS at 90d	0.77	0.80
	mNIHSS at baseline	Barthel at 90d	-0.46	-0.52
	mNIHSS at 2h	Barthel at 90d	-0.57	-0.64
	mNIHSS at 24h	Barthel at 90d	-0.72	-0.76
	mNIHSS at 7-10d	Barthel at 90d	-0.75	-0.76
	mNIHSS at 90d	Barthel at 90d	-0.82	-0.79
	mNIHSS at baseline	mRS at 90d	0.50	0.56
	mNIHSS at 2h	mRS at 90d	0.60	0.69
	mNIHSS at 24h	mRS at 90d	0.73	0.81
	mNIHSS at 7-10d	mRS at 90d	0.76	0.82
	mNIHSS at 90d	mRS at 90d	0.83	0.86
	mNIHSS at baseline	GOS at 90d	0.48	0.57

Paper	Scale of Interest	Scale compared with	Spearman rank correlation coefficient	
	mNIHSS at 2h	GOS at 90d	0.59	0.67
	mNIHSS at 24h	GOS at 90d	0.71	0.78
	mNIHSS at 7-10d	GOS at 90d	0.75	0.79
	mNIHSS at 90d	GOS at 90d	0.82	0.85
	mNIHSS at baseline	LV at 90d	0.47	0.51
	mNIHSS at 2h	LV at 90d	0.53	0.62
	mNIHSS at 24h	LV at 90d	0.61	0.65
	mNIHSS at 7-10d	LV at 90d	0.59	0.63
	mNIHSS at 90d	LV at 90d	0.61	0.65
Meyer 2002 (I)	NIHSS	BI	-0.165	
	NIHSS	mRS	0.219	
	NIHSS	mNIHSS	0.944	
	mNIHSS	BI	-0.238	
	mNIHSS	mRS	0.296	
	mRS	BI	0.819	
Meyer 2005 (I)			Remote	Bedside
	NIHSS	BI	-0.22	-0.19
	NIHSS	mRS	0.25	0.26
	mNIHSS	BI	-0.21	-0.19
	mNIHSS	mRS	0.24	0.26
	NIHSS	mNIHSS	0.93	0.95
Meyer 2008 (I)	NIHSS	mRS	0.71	
	mNIHSS	mRS	0.71	
			Bedside	Remote
	NIHSS	mNIHSS	0.98	0.97
Olivato 2016 (VG)			Observer 1	Observer 2
	eNIHSS	NIHSS	0.933	0.930
Wijdicks 2005 (VG)	GCS	FOUR	0.92	

Key
Methodological quality assessment: Very Good (VG) Adequate (A) Doubtful (D) Inadequate (I)
NE= Neurologic Examination
NCCT= Non- contrast computer tomography (CT) scan
BI= Barthel Index
mRS= Modified Rankin Score

d= days
GOS= Glasgow Outcome Scale
PT= Placebo-treated patients
rtPA= Thrombolysis- treated patients

Summary

The concurrent validities reported generally show strong relationships in the comparisons made. However, correlation is not truly indicative that scales measure the same constructs. Statistically concurrent validity can be calculated even if variables measure completely different constructs.

Many of the comparisons are with the Barthel Index (BI) or Modified Rankin Score (mRS). These scales measure functional independence or disability and are not necessarily suitable for the acute timeframe. The data utility is further diminished by the testing of concurrent validity in the sub-acute phase and the use of 90-day comparators. Due to the range of comparisons, there is insufficient data to show superiority of any scale. The data compares scales but does not indicate which would be suitable for identification of END.

Methodological assessment of quality across all papers was initially deemed as very good as they all reported correlations. However, all the papers that compared the NIHSS to the mNIHSS had to be deemed inadequate as they compared a long and shortened version of a scale with the same responses so potentially introduced bias.

4.3.3 Construct Validity

Construct validity is about determining how well the measurement instrument measures what it is supposed to. This can include evidence about internal relationships, relationships with scores of other scales or differences between relevant groups (Mokkink et al., 2010a). This thesis focuses on identification and measurement of END in acute stroke and the scales featured attempt this through assessing items that represent functional ability of the patients. The theoretical constructs are based around the standard neurological examination. Constructs are therefore abstract representations of the theoretical structure and not directly observable but assumed manifestations of the underlying pathology (Portnoy & Watkins, 1993). The scales can be treated as reflective models, and it would be expected that items will correlate. Scales can represent several constructs (e.g., left and right hemisphere strokes) as long as the items are clearly associated with them and the score reflects the dimensional structure of the scale (Nunnally, 1979).

There are three aspects of construct validity: cross-cultural validity, hypotheses testing, and structural validity. Cross-cultural validity is concerned with culturally adapted or translated versions of scales. This is omitted from the thesis as translated versions of scales were excluded from these reviews.

Hypotheses testing involves testing relationships of scores on one scale with scores on another. The scales can measure similar constructs (convergent validity) or dissimilar constructs (discriminant validity) and differences between subgroups of patients. There is always an assumption that the instrument validly measures the construct of interest (de Vet et al., 2011).

Within the literature several papers purported to have measured construct validity but without generation of hypotheses based around the construct to be measured. This meant that they were actually reporting was concurrent validity, so the data was presented in that section. Only one paper presented and tested a discriminative validity hypothesis that the GCS would not correlate as well with the standard neurologic exam as the CNS (Cote et al., 1989). The results in Table 4.8. show that if adopting the generic hypotheses that correlations with instruments measuring similar constructs should be ≥0.50 the CNS correlates with the standard neurologic exam across all items whereas for two of its three items the GCS does not. Although the GCS can crudely pick up neurological deficit the CNS achieves better discrimination (ibid). Both subsections of the study were assessed for methodological quality and deemed adequate.

Table 4.8 Comparison of the abilities of the GCS and the CNS to measure neurological status (taken from Cote et al., 1989)

	Initial neuro exam item	**Scale Item**	**Correlation**	**95% Confidence Intervals**	**No of patients**
GCS	LOC	Eye opening	0.277	-0.072, 0.626	74
	Orientation & Speech	Best verbal response	0.643	0.475, 0.811	72
	Weakness	Best Motor Response	0.363	0.198, 0.528	77
	Total Score	Total Score	0.563	0.418, 0.708	77
CNS	LOC	LOC	0.702	0.457, 0.947	79
	Orientation & Speech	Orientation & Speech	0.749	0.612, 0.886	74
	Weakness	Weakness	0.664	0.513, 0.815	66
	Total Score	Total Score	0.769	0.675, 0.863	79

Structural validity is defined as the degree to which scores of a scale are an adequate reflection of the dimensionality of the construct to be measured (Mokkink et al., 2010a). Dimensionality of scales, whether uni or multi, is important as some clinimetric properties are assessed differently between the two. Item test statistics including factor analysis are statistical processes that can establish how individual items cluster around a dimension (Boone, 2016; DeCoster, 1998; Joliffe & Cadima, 2016). They can describe and explain how a large set of independent items correlate to underlying factors in terms of loading. By studying the factor loadings, interpreted as correlation coefficients, it determines how well the factors explain the

data. A group of items in a scale may represent any number of underlying factors, from a single factor to the total number of items.

Nine papers stated they presented data related to structural validity by item fit statistics. However, two papers had to be excluded. The first an unapplicable theoretical paper that described statistical models for the NIHSS based on item response theory (IRT) (Iramaneerat et al., 2009). The second reported completing Confirmatory Factor Analysis (CFA) for mNIHSS, NIHSS-11, sNIHSS-8, and sNIHSS-5. However, the data file was unavailable, and the author did not respond to requests (Lee et al., 2016). Therefore, results from seven papers (Bessenyei et al., 2001; Edwards et al., 1995; Lyden et al., 1999, 2001, 2004; Millis et al., 2007; Zandieh et al., 2012) are presented for six scales (Mathew, SSS, NIHSS, MCANS, Unified and mNIHSS).

Three types of item fit statistics were used: principal component analysis (PCA), factor analysis (FA), and Rasch analysis (RA). Results are described by item test statistic rather than by scale. Appendix 4.21. show the structural validity data by scale showing which papers completed item test statistics, and in what populations. It also shows the method used, the number and descriptors of factors obtained, as well as the broad purpose for completion.

PCA was used in one paper across four scales and identified the factors within the Mathew scale, the SSS, the NIHSS, and the MCANS/Orgogozo (Bessenyei et al., 2001). Three papers completed exploratory factor analysis (EFA) on the NIHSS, two found two factors (Lyden et al., 1999, 2004) and the other found four factors (Zandieh et al., 2012). Two of these papers then completed confirmatory factor analysis (CFA) on the NIHSS with both identifying four factors (Lyden et al., 1999, 2004). Another paper completed both EFA and CFA as part of the development of the mNIHSS and identified four factors (Lyden et al., 2001).

CFA was completed on the SSS and two factors were found and the MCANS/ Orgogozo where three factors were identified (Edwards et al., 1995). Only one paper used RA which identifies two factor for the NIHSS (Millis et al., 2007).

Summary

Although different techniques are used, they all identify patterns in the correlations between variables. They identify and name underlying concepts in the scales providing some insight into the pathophysiological mechanisms that combine the items. The different methods of item test statistics identified different numbers of factors. PCA identified more factors in both the SSS and the NIHSS than FA. In PCA a larger number of factors is associated with higher

sensitivity of the scale. Therefore, the NIHSS would be considered more sensitive than the other scales. However, this aligns with common sense at it contains more items that cover more neurological components. The mNIHSS performed similarly to the NIHSS but this would be expected as is it a modified version with redundant items removed.

The results initially appear consistent when the same method is used on the same scale e.g., all three papers that report CFA on the NIHSS found four factors. However, different studies exclude different items and the items do not always load on the same factors. In some cases, factors split load where they load onto more than one factor or items correlate with each other to produce a factor despite having little underlying meaning for the factor (Tabachnick & Fidell, 2007). This measurement invariance shows that the underlying structure can change based on the sample used. Most of the studies tested scale structure in a purely ischaemic population. It is advised to use a heterogeneous sample rather than a homogeneous sample as similar populations lower the variance and factor loadings (Kline, 1994). The methodological quality of the studies ranged from inadequate to very good (Appendix 4.21). However, this is based on statistical methods and does not factor in population heterogeneity. It would therefore be useful to test scales with a wider stroke population, including ICH patients, to check that the structure and factors hold in these populations.

4.3.4. Reliability

Overall reliability refers to the extent to which scores for patients who have not changed are the same for repeated measurements under several conditions (Streiner & Norman, 2008). Measurements are seldom perfect especially if they involve subjective measurement of symptoms. Other influences such as the way instructions are given, or encouragement provided will add further subjectivity (de Vet et al., 2011). It is important to note that reliability is a characteristic of an instrument used in a population, and not just the instrument. The conditions of testing define the type of reliability established: internal consistency, inter-rater reliability, and intra-rater reliability (ibid).

4.3.4.1. Internal Consistency

Internal consistency is the degree of the interrelatedness among the items when utilising different sets of items from the same multi-item measurement instrument (Tavakol & Dennick, 2011). Simply it is a measure of the extent to which items assess the same construct. If there is one item that measures something different, this item will have a lower item total correlation. As internal consistency is a function of the mean correlation between items, and the number

of items on a scale, longer scales with fewer choices will generally report higher coefficients (Cote et. al, 1988).

Nine papers presented data on internal consistency (IC) (Table 4.9). The higher the value of Cronbach's alpha (α) the greater the correlation. A α value of 0.70 or more is generally reported to show good IC of a scale (Taber, 2018). Good IC was reported in all scales except the Mathew and the SSS. This could be due to the fact α tends to underestimate the IC of scales with fewer than 10 items (Herman, 2015). However, the results for the CGS, CNS, and the FOUR Score have not been reduced by having fewer items. The quality of all papers was rated as doubtful or inadequate.

Table 4.9 Results by scale and paper for internal consistency calculated using Cronbach's Alpha (α)

Scale	Paper	α
Mathew	*Brown 1990* (I)	0.54
GCS	Wijdicks 2005 (D)	0.88 1st rater
		0.84 2nd rater
Toronto	*Brown 1990* (I)	0.72
SSS	Edwards 1995 (I)	0.59 whole sample
		0.53 CVA
		0.56 ICH
		0.59 SAH
		0.58 TBI
CNS	Cote 1986 (D)	0.896
	Cote 1989 (D)	0.792
HSS	*Brown 1990* (I)	0.88
MCANS	Edwards 1995 (I)	0.65 whole sample
		0.79 CVA
		0.76 ICH
		0.75 SAH
		0.76 TBI
ESS	Hantson 1994 (D)	0.92
Japan	Gotoh 2001 (D)	0.998
FOUR Score	Lee 2017 (D)	0.843 doctors 0.868 nurses
	Wijdicks 2005 (D)	0.86 1st rater
		0.87 2nd rater

Key
Methodological quality assessment: Very Good (VG) Adequate (A) Doubtful (D) Inadequate (I)

Summary

Although the high α coefficients initially indicate a high degree of IC it could be misleading. What the data tells us is that every item is measuring something similar to some of the other items (Tavakol & Dennick, 2011). Calculating α to give an overall measure of IC of a scale is not

appropriate if the scale is not unidimensional and around the dimensionality of the scales presented is lacking. From the structural validity data, it was clear that the Mathew Scale, the SSS and the MCANS were not unidimensional. All papers were assessed as either doubtful or inadequate because where IC is reported for a multidimensional total scale, it should be rated 'inadequate' and if there is no information on the structural validity or dimensionality, this standard can be rated with 'doubtful' (Mokkink et al., 2017 pg. 49).

4.3.4.2 Inter-rater Reliability

Inter-rater reliability is defined as the extent to which the measurement records the same values in the same nonchanging patient, at the same point in time, by different examiners (Streiner & Norman, 2008). It is important that examiners are consistent in their scoring otherwise variation is introduced. Fifty-one studies presented data on inter-rater reliability across 17 scales. Several papers report on more than one scale. The NIHSS was the most investigated scale with 35 papers presenting data. Across the other scales, two had four papers (CNS and mNIHSS), three had two papers (GCS, SSS/SNOBS, and FOUR Score), and ten had only one paper reporting some form of reliability (Mathew, HSS, MCANS, Unified, ESS, Japan, NIHSS-8, IVBSS, sNIHSS and e-NIHSS).

Several different statistical methods were utilised across the 51 studies (Appendix 4.22). To aid comparison across and between data will not be presented chronologically. A broad overview of the studies will be presented followed by data collated by statistical method:

- percentage agreement (4 papers)
- kappa statistics
    - simple kappa (k) (24 papers)
    - weighted kappa (Wk) (26 papers)
    - mean kappa (mk) (2 papers)
    - modified kappa (Mk) (1 paper),
- observed and/ or expected agreement (3 papers)
- intraclass correlation coefficients (ICC) (25 papers)
- other reliability co-efficients (10 papers).

Where 95% confidence intervals were provided in the papers they will be shown in brackets. Sub-group analyses are presented only if overall results were not provided. Further detail illustrating specifics of both inter-rater and intra-rater reliability studies by scale, along with COSMIN methodological assessment of quality is presented in Appendix 4.23. Forty-one of the

studies reporting inter-rater reliability were conducted in a pure stroke population. Of these, six had purposively selected patients who were more alert and able to co-operate with the examination. Seven studies were completed in a suspected stroke population and three in a population that included some stroke patients. Five studies used actors to simulate patients who had had a stroke patients. Six studies used patient video recordings with no live patient assessment being completed.

In terms of stroke types, 37 studies used a mixture although often it was not stated what types were included. Where stroke types were reported most were ischaemic. Eleven papers reported purely on ischaemic stroke patient populations and three on ICH populations.

Timing since onset of stroke was poorly reported. Thirty-five studies did not specify time since onset, and only eight studies included patients in the period of interest (the first 72 hours from onset). The numbers of examiners involved in the studies ranged from one to 8214. In 13 studies the numbers of examiners were unknown. Thirty-one studies reported 20 or less examiners and seven 20 or more examiners. Examiners professional groupings were reported in all but two studies, 30 studies were conducted with medical examiners, 17 had a mixture of health care professionals (HCPs), and two were only nurses. Settings ranged across pre-hospital, hospital, and community. Fifteen studies involved telemedicine. Four studies involved data from the development of, or results of, NIHSS certification process. Most studies were completed in a controlled research context and only three reported being completed in uncontrolled clinical practice (Demeestere et al., 2017; Specogna, 2013; Specogna et al., 2013).

Thirty papers reported training and 16 did not. However, from those that mentioned training only 15 papers specifically reported training in the scale or scales of interest. Within three of those fifteen training was only for certain professional groups or those not previously trained in the scale (Demeestere et al., 2017; Dewey et al., 1999; Gonzalez et al., 2011). Though training and certification in the NIHSS is considered a requirement for its reliable and valid use (Andre 2002), less than half (n = 15 /36) reported whether all examiners were certified.

Four papers presented percentage agreement (PA) data related to the NIHSS. Three of these calculated PA for total NIHSS scores, which ranged from 92% (Peters et al., 2012) to 95% (*Boutot et al., 2013*; Shafqat et al., 1999). The fourth paper calculated PA by item twice using both television and videorecorder playback and a telemedicine system up to one year apart (LaMonte et al., 2004). Five items, level of consciousness (LOC), LOC commands, Gaze, Right Arm, and Ataxia had 100 % agreement at both time points. All other items reported less than

100% agreement at one or both rating time-points with Facial Palsy recording the worst percentage agreement with 80% in one rating but only 20% in the other (LaMonte et al., 2004).

Kappa (k) statistics were the most frequently used method to assess reliability. Kappa quantifies agreement between examiners above that which would be expected by chance (Harrison et al., 2013). Values for kappa can range from -1 (agreement less than chance) through 0 (expected agreement by chance) to 1 (total agreement). There are a variety of techniques for calculating kappa statistics depending upon study methods (Sim & Wright, 2005), for example, numbers of examiners (Cohen's for 2 and Fleiss for more than 2) and balancing of groups (e.g., jackknife). The kappa techniques used within all the included literature are not routinely reported. Comparisons have occurred despite it not being technically correct to compare variable methods, this appears been accepted in the literature to date (McHugh, 2015). For this study, the Landis and Koch (1977) classification system has been applied to all k and weighted kappa results tables in this section (Table 4.10). This allows visual representation of the variation in inter-rater reliability scores.

Table 4.10 Colour key for all kappa (k) statistics presented based on Landis and Koch (1977). These classifications and this colour coding is used throughout the presentation of the reliability k statistics in this chapter.

kappa range	0.81-1	0.61-0.80	0.41-0.60	0.21-0.40	<0.20
Definition of agreement	Very good	Good	Moderate	Fair	Poor
Key					

Data for the NIHSS and modified versions is presented separately to other scales. Twenty-five studies reported k statistics for the NIHSS, four for the mNIHSS and one each for the NIHSS-8, sNIHSS and e-NIHSS. Reliability was calculated for either total score or by individual item assessed. The total score reliability, calculated by any form of k statistic, for the NIHSS and its modified versions is illustrated in Table 4.11.

Table 4.11 Kappa (k) statistics calculated for the total score of the NIHSS and its modified versions.

Paper	Scale	Test	Reported Result (confidence intervals in brackets)
Alasheev 2017	NIHSS	k	0.17 (.11, .27)
		Wk	0.91 (.87, .95)
Gur 2007		Wk	0.87 (.79, .96)
Liman 2012		Wk	0.69 (.51, .87) (hospital vs. ambulance)
			0.79 (.59, .98) (ambulance vs. video)
Meyer 2002		Wk	0.969 (.678, **1.261**)
Schmulling 1998		mk	0.61 (SD=0.17) (trained)
			0.33 (SD=0.22) (untrained)
			0.45 (SD= 0.2) (combined)
Wu 2017		Wk	0.71 (.62, .79)
Meyer 2002	mNIHSS	Wk	0.988 (.696, **1.280**)
Demeestere 2017	NIHSS-8	Wk	0.69
Gonzalez 2011	sNIHSS	Wk	0.73 (.43, 1)
Olivato 2016	e-NIHSS	k	0.968

Key
k- kappa
mk- mean kappa
Wk- weighted kappa
Numbers- confidence intervals as reported (should not be greater than 1)

Agreement on the NIHSS total score was good or very good when Wk was used but poor with simple k. Training may have a beneficial effect of inter-rater reliability (*Schmulling et al., 1998*). However, *Schmulling et al.,* (1998) provided no justification for the use of mk. One paper calculated grouped NIHSS (score=0– 5, 6–12, and >13) rather than total score which could have inflated the reliability (Wu et al., 2017). All the modified versions of the NIHSS had either good or very good total score reliability.

Results by individual items from the studies that calculated k values for the NIHSS and e-NIHSS are presented in Table 4.12, and from studies that calculated Wk values for the NIHSS, mNIHSS, NIHSS-8 and sNIHSS in Table 4.13.

Table 4.12 Inter-rater reliability by item and study for the NIHSS and e-NIHSS when calculated using kappa (k) values.

Study	LOC	Questions	Commands	Gaze	Visual Fields	Facial Palsy	Right Arm	Left Arm	Right Leg	Left Leg	Ataxia	Sensory	Best Language	Dysarthria	Extinction (Neglect)
NIHSS															
Alasheev 2017	.71 (.4, 1)	.53 (.4, .7)	.36 (.15, .62)	.51 (.3, .68)	.44 (.26, .67)	.16 (.01, .35)	.66 (.48, .81)	.59 (.40, .72)	.53 (.33, .66)	.64 (.50, .79)	.32 (.14, .53)	.51 (.33, .66)	.57 (.41, .76)	.55 (.40, .68)	.47 (.30, .69)
Albanese 1994 1st	.52	.69	.77	.74	.80	.30	.96	.77	.81	.54	.48	.58	.66	.52	.69
Albanese 1994 2nd	.44	.65	.71	.74	.80	.28	.95	.82	.79	.49	.47	.59	.65	.51	.66
Anderson 2011	1.0 (1, 1)	1.0 (1, 1)	*	.48 (-.02, .98)	.85 (.60, 1)	.66 (.36, .96)	1.0 (1, 1)	1.0 (1, 1)	1.0 (1, 1)	.92 (.76, 1)	.46 (-.16, 1.)	1.0 (1, 1)	.79 (.51, 1)	.73 (.45, 1.)	1.0 (1, 1)
Birschel 2005	.31			.69	.58	.32	.59	$	.64	$	-0.1	.52	.72	.73	.45
Brott et al 1989	.49	.80	.58	.82	.81	.57	.85	$	.83	$	.57	.60	.64	.55	.58
Dewey 1999 (1)	.83	.71	.95	.81	.74	.74	.79	.88	.84	.63	.60	.71	.56	.69	.81
Dewey 1999 (2)	.83	.56	.85	.81	.91	.53	.76	.80	.52	.36	.55	.47	.56	.69	.76
Dewey 1999 (3)	.83	.61	.81	.81	.74	.70	.71	.76	.56	.48	.42	.37	.71	.66	.66
Goldstein 1989	.50	.64	.41	.33	.57	.22	.77	$	.78	$	-0.16	.50	.79	.32	.61
Josephson 2006	.572	.814	.855	.350	.807	.440	.883	.774	.771	.831	.025	.907	.396	.725	.746
LaMonte 2004 (4)	*	.58	*	*	*	-.011	*	.74	.44	.72	*	.58	.58	.58	.62
LaMonte 2004 (5)	*	.58	*	*	.44	.69	*	*	.58	.44	*	*	.67	.38	.58
Lyden 1994a (6)	.62 (.63, .69)	.68 (.65, .70)	.00 (-.63, .64)	.02 (-.45, .49)	.94 (.83, 1)	.38 (.34, .40)	.79 (.69, .87)	.79 (.76, .81)	.71 (.68, .74)	.80 (.77, .82)	.23 (.06, .38)	.94 (.84, 1.)	.39 (.35, .42)	.72 (.67, .75)	.54 (.48, .60)
Lyden 1994a (7)	.42 (.12, .70)	.90 (.85, .93)	0.93 (.75, 1)	0.51 (.42, .59)	.81 (.77, .84)	.20 (.17, .23)	.94 (.84, 1)	.92 (.82. 1)	.66 (.58, .74)	.95 (.87, 1)	.56 (.66, .85)	.81 (.76, .84)	.57 (.53, .60)	.42 (.32, .50)	.53 (.47, .57)
Lyden 2005	.46 (.39, .53)	.77 (.64, .90)	.92 (.75, 1)	.70 (.39, 1)	.72 (.57, .87)	.38 (.27. .49)	.72 (.54, .90)	.65 (.51, .79)	.64 (.53, .75)	.64 (.51, .77)	.21 (.12, .30)	.73 (.53, .93)	.64 (.53, .75)	.56 (.39, .73)	.57 (.40, .74)

Study	LOC	Questions	Commands	Gaze	Visual Fields	Facial Palsy	Right Arm	Left Arm	Right Leg	Left Leg	Ataxia	Sensory	Best Language	Dysarthria	Extinction (Neglect)
Lyden 2009	.43 (.01, .51)	.77 (.66, .84)	.81 (0, 0.86)	.45 (.03, .63)	.57 (.27, .62)	.25 (.14, .32)	.51 (.28, .65)	.52 (.21, .62)	.59 (.49, .64)	.66 (.64, .73)	.15 (.6, .22)	.54 (.17, .68)	.58 (.37, .71)	.46 (.28, .58)	.60 (.49, .64)
Nanri 2013	.64	.64-.82	1.0	.68	.75	.03-.21	.79-1.0	.40-.62	.35-.48	.36-.63	-.10	.56-.89	**.73-.82**	.71-.90	**.41-.73**
Schmulling 1998 (8)											.34				.32
Schmulling 1998 (9)				.06	-.02						-.03			.18	
Specogna et al., 2013 (10)	.26 (0, .54)	.54 (.28, .79)	.75 (.47, 1)	.08 (0, .36)	.43 (.24, .63)	.24 (.02, .46)	.39 (.16, .62)	.53 (.30, .76)	.41 (.16, .65)	.29 (.07, .50)	0 (unreliable)	.17 (0, .43)	.68 (.40, .96)	.35 (.06, .64)	.17 (0, .46)
Specogna et al., 2013 (11)	.54 (.26, .83)	.32 (.04, .59)	.65 (.33, .98)	.45 (.18, .72)	.39 (.12, .65)	.52 (.22, .83)	.62 (.40, .84)	.80 (.58, 1)	.67 (.43, .91)	.72 (.51, 0.93)	.34 (.08, .60)	0.35 (.06, .64)	.78 (.50, 1)	.43 (.17, .70	.26 (0, .55)
Specogna 2013	.345 (.043, .646)	.5 (.239, .761)	.687 (.416, .958)	.177 (-.111, .464)	.489 (.238, .740)	.242 (.011, .472)	.468 (.236, .700)	.734 (.476, .992)	.404 (.111, .698)	0.454 (.238, .670)	-.047 (-.175, .082)	.237 (-.016, .491)	.640 (.371, .909)	.417 (.125, .709)	.157 (-.130, .445)
Wu et al 2014 (12)	.703	1	.84	.862	.605	.709	1	.77	1	.845	1	.606	1	1	.494
Wu et al 2014 (13)	.46	.832	.467	.689	.36	.37	.616	.715	.616	.704	1	1	1	.819	.348
e-NIHSS															
Olivato 2016	1	1	1	.967	1	.949	.983	.966	1	.973	1	1	0.953	.946	1

Key

* All patients able to perform the task and given same score so unable to calculate k statistic

[$] Reliability was only carried out on affected limb score and therefore only one kappa calculated (score reported under Right Limb).

Numbers- assumed kappa values reported were over 1 so assumed 1 had been transposed and removed

Neuro = neurologist

(1) Neuro 1 vs. 2
(2) Neuro 1 vs. nurse
(3) Neuro 2 vs, nurse
(4) Neuro 1 Telebat vs. 2 videotape
(5) Neuro 2 Telebat vs. 1 videotape
(6) Tape 1
(7) Tape 2
(8) Trained
(9) Untrained
(10) physicians v. nurses
(11) nurses v. nurses
(12) Live
(13) Recorded

Table 4.13 Inter-rater reliability by item and study for the NIHSS and modified versions when calculated using weighted kappa (Wk).

Study	LOC	Questions	Commands	Gaze	Visual Fields	Facial Palsy	Right Arm	Left Arm	Right Leg	Left Leg	Ataxia	Sensory	Best Language	Dysarthria	Extinction (Neglect)
NIHSS															
Alasheev 2017	.71 (.4, 1)	.74 (.6, .86	.46 (.2, .53)	.76 (.61, .86)	.57 (.40, .77)	.43 (.26, .55)	.90 (.79, .96)	.86 (.72, .93)	.77 (.64, .86)	.90 (.81, .95)	.29 (.09, .55)	.51 (.25, .68)	.60 (.27, .81)	.68 (.58, .78)	.50 (.31, .76)
Anderson 2011	#	1.0 (1, 1)	*	#	.92 (.77, 1)	.72 (.47, .98)	#	1.0 (1, 1)	#	.97 (.90, 1)	.35 (-.25, .95)	1.0 (1, 1)	.85 (.65, 1.)	.75 (.409, 1)	#
Berthier 2012 & 2013	1.00	1.00	1.00	.84	.96	.75	.88		.82		.61	.95	.94	.67	.87
Demaerschalk 2012	.67 (0.48, 0.86)	.94 (0.87, 1.0)	.89 (0.74, 1.0)	.72 (0.53, 0.90	.91 (0.82, 1.0)	.59 (0.46, 0.72)	.79 (0.65, 0.93)	.83 (0.75, 0.91)	.79 (0.68, 0.91)	.79 (0.71, 0.87)	0.03 (-0.12 to 0.17)	.64 (0.46, 0.83)	.75 (0.63, 0.88)	.68 (0.55, 0.82)	.61 (0.44, 0.79)
Dewey 1999 (1)	.58 (0.29, 0.87)	.57 (0.24, 0.90)	.86 (0.66, 1.0)	.46 (0.21, 0.71)	.62 (0.35, 0.89)	.79 (0.93, 0.95)	.85 (0.73, 0.97)	.82 (0.57, 1.0)	.84 (0.68, 1.0)	.53 (0.26, 0.80)	.25 (-0.18, 0.68)	.73 (0.53, 0.93)	.56 (0.29, 0.83)	.60 (0.29, 0.91)	.77 (0.53, 1.0)
Dewey 1999 (2)	.68 (0.52, 0.84)	.44 (0.15, 0.73)	.68 (0.35, 1.0)	.50 (0.21, 0.79)	.90 (0.76, 1.0)	.59 (0.41, 0.77)	.81 (0.65, 0.97)	.78 (0.54, 1.0)	.54 (0.36, 0.72)	.39 (0.08, 0.70)	.16 (-0.15, 0.47)	.52 (0.28, 0.76)	.59 (0.34, 0.84)	.62 (0.31, 0.93)	.65 (0.34, 0.96)
Dewey 1999 (3)	.50 (0.07, 0.93)	.67 (0.47, 0.87)	.58 (0.25, 0.91)	.38 (-0.07, 0.83)	.63 (0.36, 0.90)	.73 (0.57, 0.89)	.77 (0.61, 0.93)	.84 (0.70, 0.98)	.62 (0.46, 0.78)	.62 (0.38, 0.86)	.11 (-0.14, 0.36)	.39 0.12, 0.66)	.72 (0.52, 0.92)	.67 (0.47, 0.87)	.53 (0.20, 0.86)
Liman 2012 (4)	1	.64	1	.63	.57	1	.83	1	1	.86	0	.59	.92	1	0
Liman et al 2012 (5)	1	.77	1	.63	.57	1	.83	1	1	.86	0	.77	.92	1	0
Meyer 2002	.457 (.164, 0.749)	.937 (.645, **1.229**)	.943 (.652, **1.235**)	.662 (.369, .954)	.876 (.589, **1.163**)	.742 (.460, **1.024**)	.959 (.669, **1.249**)	.971 (.678, **1.263**)	.975 (.684, **1.267**)	.947 (.654, **1.239**)	.690 (.407, .973)	.892 (.601, **1.183**)	.841 (.555, **1.127**)	.289 (-.005, .583)	.891 (.599, **1.183**)
Meyer 2005	1	.93 (.79, 1)	1	1	.93 (.93, 1)	.22 (.00, .45)	.82 (.57, 1)	.88 (.71, 1)	.80 (.56, 1.)	.74 (.47, 1.	.34 (.00, .68)	.80 (.59, 1)	.73 (.42, 1)	.61 (.30, .91)	.80 (.51, 1)

Study	LOC	Questions	Commands	Gaze	Visual Fields	Facial Palsy	Right Arm	Left Arm	Right Leg	Left Leg	Ataxia	Sensory	Best Language	Dysarthria	Extinction (Neglect)
Meyer 2008	.87 (.62, 1)	.96 (.87, 1)	1	.60 (.22, .99)	.78 (.61, .96)	.62 (.42, .82)	.97 (.92, 1)	.95 (.89, 1)	.89 (.79, .98)	.95 (.88, 1)	.65 (.11, 1)	1	.89 (.77, 1)	.75 (.46, 1)	.72 (.47, .97)
Shafqat 1999	*	.75	.29	.41	.60	.40	.82		.83		-.07	.48	.65	.55	.77
mNIHSS															
Meyer 2002	-	.937 (.645, 1.229)	.943 (.652, 1.235)	.661 (.369, .954)	.876 (.589, 1.163)	-	.959 (.669, 1.249)	.971 (.678, 1.263)	.975 (.684, 1.267)	.947 (.654, 1.239)	-	.910 (.618, 1.202)	.841 (.555, 1.127)	-	.891 (.599, 1.183)
Meyer 2005	-	.92 (.79, 1)	1	1	.86 (.65, 1)	-	.82 (.57, 1)	.84 (.64, 1)	.80 (.56, 1)	.74 (.47, 1)	-	.83 (.60, 1)	.69 (.33, 1)	-	.80 (.51, 1)
Meyer 2008	-	.96 (.97, 1)	1	.60 (.22, .99)	.78 (.61, .96)	-	.97 (.92, 1)	.95 (.89, 1)	.89 (.79, .99)	.95 (.88, 1)	-	1	.89 (.76. 1)	-	.72 (.47, .97)
NIHSS-8															
Demeestere 2017	0.54	0.68	0.34	0.43	-	0.54	0.79	0.66	-	-	-	-	-	0.53	0.26
sNIHSS															
Gonzalez 2011	.99 (.98, 1)	1	.63 (.32, .95)	1	-	59 (.27, .91)	-	.74 (.44, 1)	-	.62 (.30, .94)	98 (.74, 1)	-	.99 (.75, 1)	.66 (.36, .96)	-

Key

* All patients able to perform the task and given same score so unable to calculate k statistic

Examiners assigned only two values so unable to calculate Weighted k statistic

Numbers- confidence intervals as reported (should not be greater than 1)

Neuro = neurologist

(1) Neuro 1 vs. 2

(2) Neuro 1 vs. nurse

(3) Neuro 2 vs, nurse

(4) hospital vs. ambulance

(5) ambulance vs. video

Generally, reliability was better when calculated by Wk as opposed to k, with more items achieving above fair reliability. However, the only scale to achieve very good reliability across all items was the e-NIHSS when calculated by simple k (Olivato et al., 2016). For the NIHSS the two papers that showed the best reliability over most items when calculated by k were small studies involving telemedicine (Anderson et al., 2011; Wu et al. 2014). Eight items across the NIHSS when calculated by k achieved poor reliability across at least one study (LOC Commands, Gaze, Visual Fields, Facial Palsy, Ataxia, Sensory, Dysarthria and Extinction). When utilising Wk, two items showed poor reliability, but this was repeated over two or more studies (Ataxia and Extinction).

In the k calculations Left Arm is the only item that scores greater than fair. For the Wk calculations seven items achieve higher than fair across (LOC, LOC-Questions, Visual Fields, Right Arm, Left Arm, Right Leg and Best Language). Although technically possible, you would not expect to see a negative kappa value as this indicated that agreement has been less than by chance. Negative kappa values were reported for Ataxia (Birschel, 2005; Goldstein et al., 1989; Nanri et al., 2013; Shafqat et al., 1999; Specogna, 2013), and Visual Fields (*Schmulling et al., 1998*) indicating these are the most unreliable items.

All versions of the NIHSS showed varying reliability for items when calculated by Wk. The results overall show that the reliability of items can be affected by the populations studied, the examiners and the context in which the examinations take place but that there are no items that repeatedly and consistently have very good reliability across all studies.

One study calculated a reported modified kappa (Mk) (Josephson et al., 2006) (Table 4.14). This data could not be combined with the data in Tables 4.12 & 4.13 as a Mk is strongly correlated with percentage agreement (PA) and generally higher than traditional k so the Landis and Koch classifications for strength of agreement could not be applied. However, the Mk was chosen to statistically allow for the fact that the patients included were not a sample that covered all possible scores for all items. Best Language and Facial Palsy were the least reliable items in this study (ibid).

Table 4.14 Inter-rater reliability by item for the NIHSS calculated using modified kappa (Mk) in Josephson et al., 2006.

Item	Mk
LOC	0.886
Questions	0.849
Commands	0.993
Gaze	0.936
Visual Fields	0.925
Facial Palsy	0.652
Right Arm	0.845
Left Arm	0.966
Right Leg	0.891
Left Leg	0.871
Ataxia	0.803
Sensory	0.962
Best Language	0.596
Dysarthria	0.848
Extinction	0.843

For none-NIHSS based scales, nine studies report reliability of items across nine scales. Results are presented separately for items that assess alertness, vision and sensory and involuntary function (Table 4.15) and items that assess aspects of voluntary function (Table 4.16). Data for the SNOBS is presented with the SSS as although it uses less items they are assessed in the same way.

Across both tables for none-NIHSS scales there was widespread variability in reliability across some items. For example, LOC the mostly commonly included items across all scales had values ranging from 0.22 to 0.91. Although this will be affected by statistical methods and study characteristics it illustrated that reliability, even in commonly assessed items varies widely. For items that encompass alertness, vision and sensory, and involuntary function (Table 4.15) the items that score poor are orientation and visual fields and for voluntary function it is Facial Palsy all from within the Mathew Scale (Gelmers et al., 1988). This could indicate that the way this scale assesses items is flawed but it could be a result of the small study size with only 12 non acute ischaemic patients included.

Table 4.15 Inter-rater reliability for items that encompass alertness, vision and sensory, and involuntary function for none-NIHSS stroke scales.

		LOC	LOC Questions	LOC Commands	Orientation	Visual Fields	Diplopia	Sensation	Gaze/ Deviation	Pupillary Abnormality	Extinction/ Neglect	Upper Limb Tone	Lower Limb Tone	Limb asymmetry	Plantar Reflexes	Pathologic Reflexes	Brainstem Reflexes	Respiration	Total Score
Mathew																			
Gelmers 1988	k	0*	-	-	.189	.159	-	.265	0*	-	-	-	-	-	-	-**	-	-	-
GCS																			
Lee 2017 (SS)	Wk	.867 (.755, .979)	-	.894 (.796, .992)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.796 (.694, .898)
Lee 2017 (CS)	Wk	.857 (.724, .990)	-	.846 (.707, .985)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.750 (.623, .877)
Wijdicks 2005	Wk	.77 (.69, .85)	-	.88 (.81, .96)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.82 (.76, .87)
SSS/ SNOBS																			
Birschel 2005	k	.22	-	-	-	-	-	-	.73	-	-	-	-	-	-	-	-	-	-
Lindenstrom 1991	M	.738 (.558, .918)	-	-	.741 (.584, .898)	-	-	-	.727 (.525, .929)	-	-	-	-	-	-	-	-	-	-
CNS																			
Birschel 2004	k	.46	-	-	-	-	-	-	-	-	-	-	-	.45	-	-	-	-	-
Cote 1986	M	-	-	-	.744, 1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cote 1989	M	.535 (.372-.698)	-	-	.835 (.672-.998)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ESS																			

		LOC	LOC Questions	LOC Commands	Orientation	Visual Fields	Diplopia	Sensation	Gaze/ Deviation	Pupillary Abnormality	Extinction/ Neglect	Upper Limb Tone	Lower Limb Tone	Limb asymmetry	Plantar Reflexes	Pathologic Reflexes	Brainstem Reflexes	Respiration	Total Score
Birschel 2005	k	.3	-	-	-	.78	-	-	.46	-	-	-	-	-	-	-	-	-	-
Hantson 1994	Wk	0.69	-	0.72	-	0.85	-	-	0.81	-	-	-	-	-	-	-	-	-	-
Japan																			
Gotoh 2001	Wk	0.83	-	-	-	0.91	-	0.78	0.81	0.81	0.91	-	-	-	0.78	-	-	-	-
FOUR																			
Lee 2017 (SS)	Wk	.911 (.825, .997)	-	.818 (.706, .930)	-	-	-	-	-	-	-	-	-	-	-	-	.832 (.665, .999)	.624 (.314, .934)	.742 (.626-.858)
Lee 2017 (CS)	Wk	.912 (.814, .999)	-	.780 (.639, .921)	-	-	-	-	-	-	-	-	-	-	-	-	.847 (.671, .999)	.659 (.220-.999)	.733 (.600, .866)
Wijdicks 2005	Wk	.78 (.70, .87)	-	.80 (.72, .88)	-	-	-	-	-	-	-	-	-	-	-	-	.81 (.70, .91)	.78 (.68, .88)	.82 (.77, .88)
IVBSS																			
Gur 2007 (1)	Wk	0.86 (.694, **1.182**)	-	-	-	0.93 (.820-**1.622**)	0.76 (.612, **1.132**)	0.69 (.622, **1.229**)	0.90 (.713, **1.262**)	-	-	-	-	-	-	-	-	-	0.78 (.595, **1.132**)
Gur 2007 (2)	Wk	0.82 (.694, **1.162**)	-	-	-	0.78 (.589, **1.652**)	0.76 (.652, **1.163**)	0.65 (.654, **1.297**)	0.75 (.634, **1.343**)	-	-	-	-	-	-	-	-	-	0.90 (.592, **1.027**)

Numbers- confidence intervals as reported (should not be greater than 1)

* All investigators except one score all patients equally

** All patients are scored identically by all investigators

k= simple kappa Wk =weighted kappa M= mixed kappa

SS= Suspected Stroke CS= Confirmed Stroke

(1)- first hospital setting

(2)- second hospital setting

Cote 1986 did not report the calculated kappa just the confidence intervals.

Cote 1989 saw the inclusion of more items into the scale.

Table 4.16 Inter-rater reliability for items that assess voluntary function for none-NIHSS stroke scales.

		Facial Palsy	MP- Affected Arm	MP- Unaffected Arm	Proximal Arm	Distal Arm	MP- Affected Leg	MP- Unaffected Leg	Proximal Leg	Distal Leg	Foot Dorsiflexion	Hand	Wrist Extension	Finger Strength	Ataxia	Gait	Performance/ Disability Status	Speech/ Best Language	Dysarthria	Dysphagia
Mathew																				
Gelmers 1988	k	.126	.909	.637	-	-	.455	.399	-	-	-	-	-	-	-	-	.563	.758	-	-
GCS																				
Lee 2017 (SS)	Wk	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.828 (.726, .930)	-	-
Lee 2017 (CS)	Wk	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.769 (.637, .901)	-	-
Wijdicks 2005	Wk	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.77 (.68, .85)	-	-
SSS/ SNOBS																				
Birschel 2005	k	.57	.55	-	-	-	.61	-	-	-	-	.76	-	-	-	.65	-	.78	-	-
Lindenstrom 1991	M	.608 (.390, .825)	.752 (.654, .850)	-	-	-	.688 (.578, .798)	-	-	-	-	.763 (.657, .869)	-	-	-	.912 (.859, .965)	-	.860 (.784, .936)	-	-
CNS																				
Birschel 2005	k	.47	.61	-	-	-	.56	-	-	-	.7	.78	-	-	-	-	-	.62	-	-
Cote 1986	M	.535, 1	-	-	.788, 1	.785, .974	-	-	.722, .842	-	-	-	-		-	-	.934, 1		-	-
Cote 1989	M	.771 (.598,	1 (.737,**1.294**)	-	.801 (.697,	.750 (.646,	.722 (.445	-	.753 (.643,	.798 (.685,	-	-	-	-	-	-	-	.682 (.554,.810)	-	-

		Facial Palsy	MP- Affected Arm	MP- Unaffected Arm	Proximal Arm	Distal Arm	MP- Affected Leg	MP- Unaffected Leg	Proximal Leg	Distal Leg	Foot Dorsiflexion	Hand	Wrist Extension	Finger Strength	Ataxia	Gait	Performance/ Disability Status	Speech/ Best Language	Dysarthria	Dysphagia
		.944)			.905)	.890)	, .999)		.863)	.911)										
ESS																				
Birschel 2005	k	.29	.64	-	-	-	.53	-	-	-	.6	-	.5	.6 7	-	.7	-	.76	-	-
Hantson 1994	W k	.62	P .72 R .65	-	-	-	P .71 F .69	-	-	-	.6 4	-	.7 7	.7 8	-	.78	-	.79	-	-
Japan																				
Gotoh 2001	W k	.67	.86	-	-	-	.84	-	-	-	-	.85	-	-	-	-	-	.78	-	-
IVBSS																				
Gur 2007 (1)	W k	-	.82 (.544, **1.025**)	-	-	-	.80 (.618 , **1.28 6**)	-	-	-	-	-	-	-	.88 (.588 , **1.13 3**)	.80 (.578 , **1.09 6**)	-	-	.68 (.584 , **1.18 3**)	.86 (.684 , **1.25 6**)
Gur 2007 (2)	W k	-	.86 (.576, **1.043**)	-	-	-	.85 (.696 , **1.24 3**	-	-	-	-	-	-	-	.79 (.577 , **1.13 4**)	.84 (.657 , **1.29 6**)	-	-	.68 (.567 , **1.18 8**)	.81 (.645 , **1.29 6**,

Numbers- confidence intervals as reported (should not be greater than 1)
No items on the FOUR Score within this table
Cote 1986 did not report the calculated kappa just the confidence intervals.
Cote 1989 saw the inclusion of more items into the scale.
MP=Motor Power

Wk =weighted kappa M= mixed kappa
SS= Suspected Stroke CS= Confirmed Stroke
(1)- first hospital setting
(2)- second hospital setting

Inter-rater reliability data across all the other studies and items was variable. Unlike the NIHSS scales, which have standardised assessment procedures except for the e-NIHSS, this data represents items assessed in different ways which could account for some of the variation seen. However again other study characteristics have the potential to affect the reported reliabilities of items. Reliability of all items and the total scores of the GCS and the FOUR score items appeared good or very good, across the two reporting papers. However, the studies were of poor methodological quality and not in completed in a pure stroke population. One study did not blind the GCS ratings and 62.5% of patients had no detectable deficits which could have increased the inter-rater agreement (Lee et al., 2017). The second study found an indication of reduced reliability with more alert patients, however, only 24% of the patients were stroke and there may be other condition-based confounding factors (Wijdicks et al., 2005).

For items in the SSS the reduced reliability of Facial Palsy in comparison to other items in the Lindenstrom et al., (1991) paper could have been affected by the statistical limitations of not being able to calculate with WK as the assessment is dichotomous or because the item is difficult to assess. A further study calculating the SSS items showed worsening classifications for Motor Power Arm and Leg, and LOC. However, this study excluded those who would have scored zero on LOC (not able to react to verbal command) which may have impacted reliability (Birschel, 2005) and further illustrates the effect of population on reliability.

In the CNS two items LOC and Facial Palsy had moderate agreement, whilst all other items were either good or very good (Cote et al., 1986, 1989). All items in the ESS showed good reliability (Hantson et al., 1994). In the Japan stroke scale, all items were classified as good or very good (Gotoh et al., 2001) as they were in the IVBSS (Gur et al., 2007).

Observed agreement (OA) is the proportion of cases for which all the examiners agree. Expected agreement (EA) is the proportion of agreements that are expected to occur by chance because of the examiners scoring in a random manner. They differ from PA as they can be calculated across multiple examiners and were reported in two papers, one for the Mathew scale (Gelmers et al., 1998), the other the NIHSS (Lyden et al., 1994a). Across both, the OA values are higher than the EA values which would indicate the examiners agree beyond what could occur by chance alone. However, OA is still less than 0.5 for two items in the Mathew score, Homonymous Hemianopia (0.39) and Facial Palsy (0.47) and one item Facial Palsy (0.47) for the NIHSS.

Intraclass correlation coefficients (ICCs) are statistics that describes the ratio of the between-subject variation (BSV) to the within-subject variation (WSV) (Gwet, 2008). They provide an approximation of consistency across ratings and therefore only provide an estimate of inter-rater reliability but are useful as they can be adjusted to calculate for multiple examiners and ratings. ICCs are measured on a scale of 0 to 1; where 1 represents perfect reliability with no measurement error, and 0 indicates no reliability.

Twenty-two separate papers calculated intraclass correlation coefficients (ICC) for total scale scores (Table 4.17). There are several formulas for calculating an ICC but very few papers reported which method they used. Even within named methods, such as Fleiss, there are different approaches dependent upon the selection and comparison of the examiners. Four papers reported the calculation without referencing the method (Albanese et al., 1994; Dewey et al., 1999; Shafqat et al., 1999). All the ICCs reported are higher than the 0.70 cut-off generally considered acceptable (Nunally & Bernstein, 1994). The paper that used a one-way random effects model reported the lowest ICCs, these were completed in an uncontrolled clinical environment and performed by multiple different examiners (Specogna et al., 2013).

Table 4.17 Intraclass correlation coefficients (ICCs) with 95% confidence intervals for total scale scores.

(ICC is calculated as an estimate of inter-rater reliability. Method of calculation is included where provided. Subgroups are only reported if overall ICC not provided.)

Scale	**Paper**	**Intraclass Correlation 95% CI in brackets**	**Method if reported**
SSS	Birschel 2005	0.97 (0.94 to 0.99)	
SSS^	Birschel 2005	0.96 (0.93 to 0.98)	-
CNS	Birschel 2005	0.92 (0.86-0.96)	-
	Cote 1989	0.924 (0.896-0.951)	-
	Nilanot 2010	0.87 (0.81 to 0.92)	-
	Specogna 2013 Thesis	0.797 (0.507 to 0.926)	-
HSS	Adams 1987	0.95	-
NIHSS	Albanese et al 1994	0.96	-
	Anderson et al 2011	0.98 (0.96 to 0.99)	McGraw and Wong
	Birschel 2005	0.93 (0.87 to 0.97)	-
	Chapman 2016	0.96 (0.92 to 0.98)	-
	Dewey 1999	0.95 neurologists 0.92 neurologist 1 & nurse 0.96 neurologist 2 & nurse	-
	Geisler 2019	0.87	-
	Goldstein 1997	0.94 initial 4 cases	-

Scale	Paper	Intraclass Correlation 95% CI in brackets	Method if reported
		0.92 4 new cases at 3 months 0.95 ratings of all 8 cases	
	Govindarajan 2015	0.96 (0.87 to 0.99)	-
	Lyden 2005	0.94 (0.84 to 1)	-
	Lyden 2009	0.85 (0.72 to 0.90)	-
	Meyer 2005	0.94 (0.89 to 0.99)	Fischer
	Meyer 2008	0.97 (0.95 to 0.99)	Fischer
	Nilanot 2010	0.90 (0.85 to 0.94)	-
	Shafqat 1999	0.97	-
	Singer 2005	0.953	-
	Specogna 2013	0.78 (0.58 to 0.89) Physicians vs. Nurses 0.75 (0.55 to 0.87) Nurses vs. Nurses	Fleiss
	Specogna 2013 Thesis	0.934 (0.865 to 0.969)	-
	Wu 2014	0.997 (0.992 to 0.999) real-time 0.993 (0.975 to 0.999) recorded sessions.	-
	Wu 2017	0.88 (0.84 to 0.91)	-
ESS	Birschel 2005	0.95 (0.91 to 0.98)	-
SNOBS	Birschel 2005	0.93 (0.88 to 0.97)	-
SNOBS~	Birschel 2005	0.94 (0.88 to 0.97)	-
mNIHSS	Birschel 2005	0.94 (0.89 to 0.97)	-
	Meyer 2005	0.95 (0.91 to 0.99)	Fischer
	Meyer 2008	0.98 (0.97 to 0.99)	Fischer
sNIHSS	Birschel 2005	0.94 (0.89 to 0.97)	-
	Gonzalez 2011	0.97	-

Key

^ excluding gait

~ excluding eye movements

Nine papers reported reliability co-efficients other than the ICC across six scales. Five papers calculated coefficients for total score (Table 4.18) and four by items (Appendix 4.24). Unlike an ICC, none of these methods consider the magnitude of the differences between scores. There was a mixture of coefficients of concordance and agreement. Measures of association are less dependable as reliability parameters. Although unable to directly compare, especially across different methods the total score coefficients indicated that the SSS was the most reliable and the IVBSS the least in the populations tested. The Berthier papers indicate that reliability is affected by professional group in which neurologists had better reliability in total score than non-neurologists.

Table 4.18 Correlation coefficients, other than intraclass correlation coefficients, and method, for total scores by scale and paper.

Scale	Paper	Correlation Coefficient	Method
SSS	Lindenstrom 1991	r= 0.963 R(S)= 0.954	Pearson Spearman rank
NIHSS	Berthier et al., 2013	0.96 neurologists 0.85 non-neurologists	Regression line model validated by the Kendall coefficient
	Demaerschalk 2012	r=0.949	Pearson
	Wang 2003	0.9552	Pearson
	Gur 2007	0.942	Spearman rank
IVBSS		0.864	Spearman rank

The coefficients presented by item (Appendix 4.24) are also not directly comparable due to the different methods used but they illustrate again that reliability will be different by item across scales. Most coefficients were calculated at 0.7 or above indicating good reliability but several scored lower. Eye Movements (SSS & MCANS), Inattention (NIHSS), Facial Palsy (NIHSS, MCANS & ESS), Motor Power Arm (SSS & NIHSS), Motor Power Leg (CNS & ESS), Wrist Extension (ESS) and Ataxia (NIHSS) were <0.7 (Birschel, 2005; Edwards et al., 1995). There is not enough data to conclude whether this has been affected by the way items are assessed or by population differences in the included studies, but these items have been shown to have reduced reliability across several methods and studies.

Untestable items for patients will impact the reliability calculated though this was poorly reported across the papers. Across 54 assessments of the HSS, 110 of 1350 items (8%) were not testable. The authors Adams et al., (1987) found that 90% of these were items for Neglect, Visual Construction or Stereognosis. However, there was good agreement on what could be accurately assessed with only 8 examples of disagreement and most of these were in patients with Dysphasia (ibid). Three papers reported information on untestable items for the NIHSS. In one seminal paper (Brott et al., 1989), it was reported that within a population of 65 patients a mean of 1.3 items per patient would be untestable, they found Ataxia was untestable in 46% of patients, mainly due to total hemiplegia, and that Visual Fields were untestable in 17% of patients. Another paper found 33 untestable items across 20 patient assessments and the items not reported most were Extinction and Dysarthria (Goldstein et al., 1989). Aspects of clinical practice e.g., being off the unit for tests, staff shortages, other clinical priorities, or reluctance to disturb sleeping patients could interrupt assessment. However only one paper

reported that access to scans, treatment and other clinical priorities meant they had to exclude 54 (4.5%) evaluations (Nanri et al., 2013).

Summary

There are a range of different statistical methods used within the literature to calculate inter-rater reliability. Percentage agreement (PA) can only be calculated between two examiners and is unable to account for chance agreement or that examiners may have guessed the score which could account for its limited reporting. For the papers that have reported PAs, for total NIHSS scores, the results appear good. However, the studies were small and one paper classed scores as agreeing unless they differed by more than three points, so did not report absolute agreement (*Boutot et al., 2013)*.

Kappa (k) statistics were the most utilised methods. Standard kappa statistics assume that examiners have knowledge of the distribution of the given characteristic or some tendency that would allow them to reproduce these probabilities. However, this assumption is inaccurate in a testing environment where the selection of patients included in the test does not constitute a random sample of all possible patients, and the distribution of characteristics is unknown to examiners (Josephson et al., 2006). Most of the studies were not completed in a heterogenous stroke population and none with a full range of potential deficits so we cannot conclude how well these scales perform across the entire stroke population. Some studies selected less severe and more co-operative patients making them easier to assess and which would affect the reported reliability. As k is affected by prevalence, reliability may be underestimated in rarer items, and low k values may not necessarily reflect low rates of overall agreement.

Weighted Kappa (Wk) can be used when there is a meaningful difference in scores, such as within ordinal scales. It assigns less weight to agreement as categories are further apart. However, the determination of weights for a Wk is a subjective issue on which even experts sometimes disagree (Viera & Garrett, 2005). It is therefore unclear if weighting is comparable across studies. Wk, despite its potential merits, cannot be used in unbalanced group sizes. Specific methods used to calculate k and Wk across studies were often not reported. There is potential that incorrect methods have been utilised in studies or methods were chosen based on data obtained rather than planned and calculated utilising the most suitable method. This has affected the methodological quality of studies as it is not always possible to ascertain they have utilised the correct methods.

Results can be directly influenced by the statistical method used to calculate k. One paper clearly demonstrated this by using the same data to calculate total score reliability of the NIHSS. The authors reported a poor classification when calculated using k and very good using Wk (Alasheev et al., 2017). Overall inter-rater reliability is generally lower when calculated by k (Table. 4.12) rather than Wk (Table.4.13). This would be expected as the weighted version considers the degree of disagreement rather than total disagreement. The Landis and Koch classifications were cautiously applied to allow comparisons between scales. Although the classifications provide an indication of the level of reliability there is no formal agreement on what is acceptable for a scale used in clinical practice. Many of the studies had reliability levels calculated that if they were replicated in practice could impact clinical decision making and patient care. Observed agreement (OA) and/or expected agreement (EA) were rarely reported despite it being a fundamental part of calculating kappa statistics. Other reliability coefficients are interpreted as suggestive but not definitive estimates of reliability because the scales are ordinal.

There is limited data presented for reliability of total score of scales but because most scales are ordinal in nature, it is not strictly valid to compute a total score through the addition of the sub-items. Total scores could agree even with fundamental disagreement in the underlying items as lack of agreement in items is cancelled out through totalling (Demeestere et al., 2017). Untestable items will influence the population sample and also affect calculated reliability of total scale scores. Assessing the reliability of individual items is more useful and appropriate for clinical practice.

Across the reliability data for all scales, there are no items that consistently classify as very good. Certain items such as Ataxia, Facial Palsy, Visual Fields, Gaze, and Extinction/Neglect seem to be less reliable across all statistical methods and scales than items such as Motor Power and LOC- Commands. Removal of items with poorer reliability has been employed by some scales such as the mNIHSS. The scale development paper reported the number of items showing poor values decreased from 20% to 14% and that for the mNIHSS, 55% of the items showed excellent agreement, compared with 40% for the NIHSS (Lyden et al., 2001). Although the removal of unreliable items can be a consideration in scale development it is not the only concern. Items need to be chosen to reflect specific deficits caused by different stroke types and severities and their ability to detect change in neurological status. Thought must be given the purpose of the assessment as if items were chosen purely based on reliability, they may not wholly reflect what is needed from the assessment. Consideration should be given to other

ways to make items more reliable such as better descriptors of assessment, training, and competency assessments.

All reliability parameters are dependent upon the sample in which they are tested and provide a snapshot of the reliability of that instrument in that population. The range of variation in values seen where multiple same method calculations have occurred, such as in the ICC data, is indicative of population variance. Most studies are completed either in a purely ischaemic stroke population or have very small numbers of ICH stroke included. As populations do not represent all potential deficits, the scales have not been fully tested. Most examiners across all the studies are medical staff and do not reflect the professional groups undertaking these assessments in clinical practice. Furthermore, in some studies the examiners are very experienced practitioners, often involved in the scale development, which may have a positive effect on reliability. These experienced examiners do not reflect the examiners in clinical practice who will have different levels of knowledge, skills, and experience that could all impact on assessment (Albanese et al., 1994; *Schmulling et al., 1989*).

Most studies were completed in controlled environments not representative of real clinical practice. Eleven of the NIHSS studies (Albanese et al.,1994; Chapman Smith et al., 2016; Goldstein et al., 1997; *Govindarajan et al., 2015*; LaMonte et al., 2004; Liman et al., 2012; Lyden et al., 1994a; Lyden et al., 2005; Nanri et al., 2013; Wu et al., 2014) and one for the sNIHSS (Gonzalez et al., 2011) involved actors, simulated patients or video tapes. Although this removes patient variation across multiple examiners and assessments, and makes it easier to ensure independent ratings, these studies may possibly overstate the agreement.

Independent examinations imply that the first administration has not influenced the subsequent administrations. For a high quality, inter-rater reliability study, administrations should be independent. There is an assumption across most of the studies that there is independence of examinations with the scales, but this is not assured. Only sixteen of the studies stipulated that the examiners assessments were carried out independently (Appendix 4.23- blue text within time interval column). By not ensuring independence of scores, the inter-rater reliability could have been overinflated, an example of this was where GCS were not blinded (Lee et al., 2017).

4.3.4.3 Intra-rater Reliability

Intra-rater reliability defines the extent to which the scale records the same values in the non-changing patient, at the same point in time, by the same examiner (Steiner and Norman,

2008). It is a measure of the consistency of an individual scoring under the same set of circumstances. Eight papers calculated intra-rater reliability and core study details, including method utilised are presented by scale in Appendix 4.25. As some studies included both inter-rater and intra-rater reliability further study details and methodological quality assessments are available in Appendix 4.23. Three papers reported intraclass correlation coefficients (ICC) (Table 4.19). An ICC of 1 would indicate that any variation in the data is caused by differences in the subjects rather than examiner inconsistency (Gwet, 2008). The results indicate good intra-rater reliability in both the CNS and NIHSS within time intervals of 1 week to 3 months. However, ICC is generally used to assess continuous data, and as both of these scales are ordinal the method is inappropriate.

Table 4.19 Intraclass correlation coefficients (ICCs) for intra-rater reliability by scale

Scale	**Intraclass correlation coefficient (ICC)**	**Paper**
CNS	0.99	Nilanot 2010*
	0.97	
	0.98	
	0.96	
NIHSS	0.97	Albanese 1994
	0.93	Goldstein 1997
	0.97	Nilanot 2010*
	0.96	
	0.97	
	0.96	

* 4 ICCs calculated as reported for 2 stroke fellows and 2 internal medicine residents respectively

Intra-rater reliability of the NIHSS was also calculated by PA, mk, and k (Table 4.20). PA is the simplest method to assess intra-rater reliability and can be used with any scale type. PA has been calculated for total score (*Binz et al., 2013*), and individual items (La Monte et al., 2004). Intra-rater reliability varied both across individual items and total score. Interestingly, inconsistency between examiners' ratings was still seen despite the use of videos to remove patient variation (ibid). Only one paper reported a time interval (Brott et al., 1989).

Table 4.20 Intra-rater reliability for the NIHSS calculated for items and total score by percentage agreement (PA), mean kappa (Mean k) or kappa (k) statistics.

Study	Test	LOC	Questions	Commands	Gaze	Visual Fields	Facial Palsy	Right Arm	Left Arm	Right Leg	Left Leg	Ataxia	Sensory	Best Language	Dysarthria	Extinction (Neglect)	Total Score
Binz 2013	PA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	90
LaMonte 2004 (1)	PA	100	100	100	100	100	60	100	100	80	80	100	80	60	100	100	-
LaMonte 2004 (2)	PA	100	100	100	100	80	80	100	80	80	100	80	80	80	80	80	-
Brott 1989	Mean k	-	-	-	-	-	-	-	-	-	-	-	-	neurologist			0.77
Brott 1989	Mean k	-	-	-	-	-	-	-	-	-	-	-	-	neurology house officer			0.70
Brott 1989	Mean k	-	-	-	-	-	-	-	-	-	-	-	-	neurology nurse clinician			0.68
Brott 1989	Mean k	-	-	-	-	-	-	-	-	-	-	-	-	ED nurse clinician			0.66
LaMonte 2004 (1)	k	*	*	*	*	*	0.4118	*	*	0.5833	0.7059	*	0.5833	0.3750	*	*	-
LaMonte 2004 (2)	k	*	*	*	*	0.4444	0.6429	*	0.7368	0.4444	*	0.0000	0.5833	0.3333	0.6875	0.6154	-

Key
(1)- neurologist 1
(2)- neurologist 2
* total agreement achieved

Wk was used to assess the reliability of items in the ESS (Table 4.21) Reliability was seen to vary however the number of examiners is unknown so it difficult to draw any real conclusions about the intra-rater reliability. The time interval was reported as between one and two hours (Hantson et al., 1994) which could have inflated results due to recall.

Table 4.21 Intra-rater reliability by item of the ESS calculated by weighted kappa (Wk).

Study	LOC	Comprehension	Speech	Visual field	Gaze	Facial Movement	Arm Position	Arm Raising	Wrist Extension	Finger Strength	Leg Position	Leg Flexing	Foot Dorsiflexion	Gait
Hantson 1994	*	*	0.82	1.00	0.65	0.94	0.86	0.90	0.82	0.69	0.67	0.70	0.73	0.78

*No Wk value possible as all patients were scored identically.

One paper reported reliability coefficients for all items of the Japan stroke scale of >0.996 when utilising α (Gotoh et al., 2001). Full results are not presented as although suggestive of intra-rater reliability they are an estimate and are not comparable with any other data.

Summary

Only a small number of studies reported on intra-rater reliability. A range of methods and approaches were used making comparison difficult. Variation in repeated measurement can come from multiple sources not just the scale in use. In studies of intra-rater reliability there is an assumption of clinical stability, but performance can change based on patient fatigue, and patient and examiner experience with and recall of assessments. Patient variation was addressed in four studies by using patient videos for all or repeated ratings (Albanese et al., 1994; Goldstein et al., 1997; LaMonte et al., 2004; Nilanot et al., 2010) but the effect on results is unclear due to limited comparative studies. Clinician recall could introduce bias and falsely inflate the results of studies. For several studies the time interval is not documented (*Binz et al., 2013*; Gotoh et al., 2001; LaMonte et al., 2004) or was over a short period of time such as 1-2 hrs (Hantson et al., 1994) and up to 24hrs (Brott et al., 1989). The time interval in intra-rater reliability studies is important as there is a need to balance between repeated assessments being close enough to prevent a change in the patient condition affecting results but long enough to prevent examiner recall inflating results.

4.3.4.4 Measurement Error

Measurement error is the difference between a measured score (what is recorded) and a true score (correct reflection of the item) (de Vet et al., 2011). There are two types of errors random and systemic. Random errors are naturally occurring and occur randomly across measurements. Systemic errors are where a consistent error or bias is introduced across a series of measurements for example an examiner who always scores certain items high or low. There are two formal parameters of measurement error: Standard Error of Measurement (SEM) and the Bland and Altman method (Bland & Altman, 1986).

SEM reflects the magnitude of measurement error, the amount of variation that could be expected in a single patient's score on repeated assessments. Two papers reported SEM for the NIHSS, the first 1.81 (Albanese et al., 1994) and the second 2.02 (Specogna, 2013). Bland-Altman plots allow identification of any systemic differences between measurements by displaying the difference between scores and the average difference between scores. The limits of agreement (LoA) are calculated from the SEM and describe the scores between which 95% of the difference between repeated measurements will fall. LoA provide an impression of the size of the measurement error. Changes in scores within the LoA are likely due to measurement error. Those that fall outside provide information on change beyond measurement error or the smallest detectable change (SDC).

Five studies plotted a Bland- Altman and/or or reported the calculated LoA for the NIHSS. In one paper the LoA on the plot visually ranged from -5 to 3 but this was not reported in the paper. Instead, they reported the average difference between bedside (face-to-face) and remote (video) NIHSS scores (0.25, 95% CI 1.00 to -0.50). Despite some outliers of -4 which they reported was due to confusion on the scoring sheet for one assessment, there were wider differences evident than the reported averages seem to indicate (Chapman Smith et al., 2016).

The other four studies reported LoA of -4.44 to 5.61, -3.46 to 2.42, -4 to 3.6, and -4.14 to 7.26 respectively (Demaerschalk et al., 2012; Geisler et al., 2019; *Govindarajan et al., 2015*; *Guterud et al., 2019*). Three of the studies found that variation increased as total score increased indicating that the more severe the stroke the greater the degree of measurement error (Demaerschalk et al., 2012; Geisler et al., 2019; *Guterud et al., 2019*). Several other papers also showed a tendency for greater score variation in more severe strokes (Josephson et al., 2006; Meyer et al., 2002, 2005 & 2008). However, variation could be high even in less severe stroke populations (Alasheev et al., 2017; Anderson et al., 2013).

Two studies used the SEM and the Minimal Detectable Difference (MDD) to estimate the sensitivity of the NIHSS to detect change (Specogna, 2013; Specogna et al., 2013). The MDD identifies the smallest amount of change that is required to detect any improvement or deterioration. It is a statistical measure involving logistic and linear regression about how likely it has captured 'true' improvement or deterioration (Specogna et al., 2013). Although it does not describe clinically meaningful score changes both studies calculated an $MDD_{95.}$ This indicates the change in score that would be needed to conclude with 95% certainty that it reflected real neurological changes and not natural errors in measurement or the degree of error that affects individual NIHSS measurements. The reported MDD_{95} values were 10 or more points (Specogna et al., 2013) and 6 or more points (Specogna, 2013).

During the review, it became evident that many studies showed or allowed variation in scoring without it being formally classed as a form of error. Allowing inherent error has important implications for the reported data in these reviews and clinical practice so it has been collated and reported. The Mathew scale showed a difference of between three and 14 points despite a small sample size of only 12 patients (Gelmers et al., 1988). Fifteen papers reported differences of two or more points in patient scores for the NIHSS and/or the mNIHSS (Appendix 4.26). Differences reported were broad, with up to 10 points reported for the NIHSS and 4 points for the mNIHSS. One paper reported on the items that were involved in the disagreements showing that in 10 incidences of disagreement all involved Arm or Leg Paresis and four cases involved differences in the assessment of Facial Paresis (Geisler et al., 2019). For the e-NIHSS the examiners disagreed on total score by 1 point in 4 out of 47 patients (Olivato et al., 2016). Training either in neurology or the NIHSS scale seemed to reduce differences in scores (Berthier et al., 2012,2013; *Schmulling et al., 1998*).

Some studies included in the review utilised data from the widely used NIHSS certification process (Josephson et al., 2006; Lyden et al., 1994a). The certification tests are identified as pass or fail using an outlier method, where an outlier is a score given by 12% or less by the original NINDS rt-PA investigators (Lyden et al., 1994a). This results in nearly 25% of questions allowing for more than one 'correct' answer. Individuals could technically have outlier scores across all patients in test 1 or 2 and still pass. Across the 11 patients in these tests differences in score ranged from 6-21 points. Even when the data is reanalysed using only examiners who passed 6 of the 11 patients, 55% had a four or more-point difference in NIHSS score from the 5th to 95th percentile (Josephson et al., 2006). As reanalysis had little effect on the results, it suggests that there is a problem with the NIHSS itself rather than poorly performing examiners being responsible for the variation (ibid). However, multiple individuals are certified in the

NIHSS even though they scored over a clinically significant difference in the assessments. Potentially there could be widespread variation of how clinicians assess and score the NIHSS despite receiving training.

Summary

Formal measurement error is not well reported with only minimal data for the NIHSS available. The studies' methodological quality varies in terms of design requirements; however, they are all technically inadequate because they have utilised inappropriate methods calculating a SEM or LoA which is suitable for continuous scales and the NIHSS is ordinal.

Despite the ability to calculate a SEM from an ICC formula many studies with an ICC did not present a SEM, this may have been because they used homogenous samples in which it can be difficult to detect change rather than a withholding of data. Both studies that calculated a SEM were completed in an uncontrolled clinical setting with ICH patients. Levels of error in clinical practice need more robust testing.

Measurement errors are considered acceptable if the SDC is smaller than the values for minimally important change (MIC). However, the LoAs reported indicate the SDC is larger than quoted values for the MIC if we use a 2-point change on the NIHSS as a sensitive measure of deterioration (Siegler et al., 2013). Although the MDD does not assess clinically meaningful changes the values calculated indicate real errors associated with use of the NIHSS.

Studies allowing differences in scores to be acceptable undermines the fundamental principles of measurement. This is of particular concern when variation is accepted above the 2-point difference acknowledged as a marker for change in the NIHSS (Kasner, 2006). The style of reporting across all papers played down differences that could affect clinical decision making and outcomes for patients. The lack of standardisation within NIHSS certification may also have clinical implications. Multiple individuals are certified in the NIHSS even though they scored over a clinically significant difference in the assessments. There could be widespread variation of how clinicians assess and score the NIHSS despite receiving training. There is also the potential that END could be being missed in clinical practice.

4.3.5. Responsiveness

Responsiveness is the ability of a measurement to detect change over time in the construct to be measured (Mokkink et al., 2010a). This review is specifically concerned with responsiveness

of the scale or items within it changing when the patient's condition changes. There are two approaches to assessing responsiveness: a criterion approach and a construct approach.

In a criterion approach you identify a suitable measure, such as a gold standard, and determine the strength of the relationship between changes in scores on the instrument and the chosen standard. Using an appropriate population sample and a pre-defined required level of agreement, changes in scores are obtained independently but over the same period. A construct approach involves comparing different instruments and their measurement properties based on hypotheses about expected relationships, differences, and changes. This requires a detailed description of the construct including the conceptual model. It is important to gather empirical data and assess the consistency of results and hypotheses prior to discussion. Ideally evidence would be combined from both approaches to draw conclusions. Both approaches assume that the comparison scales are responsive.

Eight papers stated they reported responsiveness data across seven scales. However, none of the papers measured responsiveness as stated by the COSMIN standards as the validity of a change score. Two papers, one for the mNIHSS (Lyden et al., 2001) and one for the GCS and FOUR score (Wijdicks et al., 2005) reported on responsiveness of the scale in prediction of outcome. Another reported on the effect size of four NIHSS scales (NIHSS, mNIHSS, NIHSS-8 & NIHSS-5) for detecting a difference in mRS, mini–mental state examination (MMSE), and BI at 3 months (Lee et al., 2016). Effect size, although widely used in measurement literature calculates magnitude of change rather than the ability to detect change over time.

One abstract concluded that although the FOUR score performed similarly to GCS for the early detection of neurological deterioration (ND) both are less sensitive than subjective assessment by trained nurses (*Zink et al., 2012*). Another calculated the cut point for change on the CNS as 1 point as it had the best results across sensitivity (0.933), specificity (0.508), positive predictive value (PPV) (0.318) and negative predictive value (NPV) (0.969) (Cote et al., 1989). Another compared the CNS to the NIHSS and the Stroke Impairment Assessment Set (SIAS) at assessing changes in motor function. The SIAS was deemed superior, however the CNS was better at detecting change in motor items over a 12-week period than the NIHSS (Seki et al., 2014) which indicated that the NIHSS motor items are not able to respond to change as well as other scales.

Two papers provided more information on the responsiveness of items within the NIHSS. In one, Facial Palsy, Dysarthria, and Best Language were shown to change minimally despite

improvement in patients (Brott et al., 1989). The other, identified the items of the NIHSS that were associated with patients who suffered ND after thrombolysis at three time points in the first 24 hours. LOC, LOC Commands, LOC total, Facial Palsy, all the Motor items and their total, and Sensation were significantly correlated with ND at one or more time points. LOC, LOC total, Left Arm Motor Drift, Right Leg Motor Drift, and Motor Limbs total, were related to ND on logistic regression analysis at two or more time points. Seven items, including Questions, Gaze, Visual Fields, Ataxia, Language, Dysarthria, and Extinction/ Inattention, were not significantly related to ND at any time (Nanri et al., 2013). These data are potentially useful in showing which items are responsive to change.

Summary

The level of data around responsiveness of scales to detect change in stroke patients was limited. No papers presented data that met the COSMIN requirements for assessment of responsiveness methodology. Nevertheless, some of the data could be clinically relevant and more data on responsiveness is needed to be able to ascertain which items can be used effectively to detect change. It is reasonable that the GCS and FOUR score would be less responsive to change than other scales as they are LOC scales and reduced LOC is a late sign of ND. Two studies explored change over longer periods (Brott et al., 1989; Seki et al., 2014). Responsiveness of scales to detect ND may vary over time so it is important that responsiveness of items is assessed across the first hours and days of stroke if they are to be used to identify END.

The items chosen and the way they are assessed will affect the responsiveness of the scale. Within the CNS the PPV, which represents detection of change, is low indicating the scale may not be suitable to detect change that would be clinically important (Cote et al., 1989). The sensitivity of some scale items to detect change is inherently inadequate. For example, the CNS rates LOC as present or absent and prevents deterioration from being registered (Hantson et al., 1994). In order to be able to identify change scale items need spread within responses to represent the distribution of the population over the item (de Vet et al., 2011). The upper and lower ends of a scale can exhibit sparseness of items suggesting that they have floor and ceiling effects (Bruce et al., 2013; Terwee et al., 2007). Both floor and ceiling effects can present problems in identifying change in a patient's condition in the extremities of scales. Patients with the highest or lowest scores cannot be distinguished from each other and change might not be detected resulting in limited responsiveness.

4.3.6 Time to Complete

Time to complete is not a measurement property but was included to explore an important aspect of feasibility and acceptability. Time is an important resource issue in clinical practice and therefore time to complete will directly impact the use of any assessment. Thirteen papers (five scales CNS, NIHSS, ESS, IVBSS, and sNIHSS) provided data on the time taken to complete an assessment (Appendix 4.26). All except the CNS reported precise timings (Cote et al., 1986). Across the scales the mean time to complete ranged from 2.9 mins to 15.09 mins.

Several studies for the NIHSS had to be excluded as the time reported, included other activities or measures, and not just time to apply the scale (Geisler et al., 2019, LaMonte et al., 2004, Meyer et al., 2008, Wu et al., 2014, & Wu et al., 2017). The time taken to complete the NIHSS ranged from 5 minutes 45 seconds to 15.09 minutes. The original NIHSS paper reported a mean assessment time of 6.6 minutes (Brott et al., 1989). The four examiners in this study (ibid) were individuals involved in the scale development. Assessment involved one examiner administering and scoring the scale and the other three scoring so you would expect this to be faster than other assessments. Still, these mean assessment times have been matched in other studies (Alasheev et al., 2017; Isahaya, 2017; *Peters, 2012*; Shafqat et al., 1999; Wang et al., 2003).

The ESS took on average 8.2 minutes to complete, but the setting, like in many of the other papers, is not indicative of acute clinical practice and only involved neurologists in the assessment (Hantson et al., 1994). The IVBSS reported a mean time of 5.5 minutes, however, this was calculated over two assessments of 18 patients in one setting so would need further verification (Gur et al., 2007). The fastest scale to complete was the sNIHSS with a mean assessment time of 2.9 minutes when completed at the bedside (Gonzalez et al., 2011), as a simplified version of the NIHSS with fewer items this is to be expected.

Multiple factors including setting and patient severity will impact on the time to complete an assessment. When compared in telemedicine studies, examination at the bedside was quicker than the remote examination (Alasheev et al., 2017, Shafqat et al., 1999 & Wang et al., 2003). One paper indicated that training may also reduce time taken to assess (*Isahaya, 2017*).

Summary

There is debate over what constitutes a suitable amount of time for a stroke scale assessment (Adams et al., 1987; Cote et al., 1986; Hantson et al., 1994). Time taken to complete repeated

assessments, has been cited as a barrier to routine use in clinical practice (Yanko & Lang, 2013). It is an important aspect to consider in terms of clinical resource especially as repeated assessments would be needed to identify END in a timely manner in multiple patients. Acceptability of and completion of assessments in clinical practice will be affected by the time taken to complete. Experience and training may impact on the time taken to complete however timings deemed acceptable in papers may be considered excessive in busy clinical environments.

4.4 Discussion

Despite many scales being developed that could be used for the clinical assessment of patients after stroke, none have been tested across all clinimetric properties. In view of the initial volume of literature, the limited data describing the performance of scales was an unexpected finding. Despite the lack of evidence on their clinimetric properties, several of these scales are routinely used in clinical and research practice. This discussion will be organised under four subheadings:

- Scale items and relevance to Early Neurological Deterioration (END)
- Study relevance and quality
- Clinimetric properties
- Clinical considerations.

4.4.1. Scale items and relevance to END

In order to recognise END, it is vital to be able to identify a change in a patient's condition. The evaluation of acute stroke is difficult because of heterogeneous clinical presentation, affecting factors such as speech, sensation, power, and equilibrium (Treves et al., 1994). The review results have shown that reliance on total scale scores to identify change in a patient's condition can be misleading and could potentially lead to changes being overlooked. Items can change in different directions as a patient's condition alters, whilst the total score remains unchanged. Changes in individual items are key to identifying change in condition as they are more responsive (Nanri et al., 2013). Assessment focusing on change within items themselves, rather than whole scales, would potentially be more useful within clinical practice. Utilising items to direct and clarify observation rather than focus on quantification of deficits could potentially ensure that change is noted more quickly, by a range of staff. However, these items need to be relevant, representative, comprehensive, and comprehensible. They need to be useful for the detection of END without being too complex.

Multiple items, mainly originating from the medical neurological examination, have been selected and used across multiple stroke scales. Some items are commonly included in scales (e.g., LOC, language) and others specific to certain scales (e.g brainstem reflexes) (Table 4.3). Performance of neurological assessment and monitoring will depend upon the choice of neurological signs to be used as indicators of change (Birschel et al., 2004). Item selection should be guided by consideration of validity (that it measures what it is supposed to), reliability, and responsiveness to detect END. It is difficult to decide, from the evidence in this review, which items are the most suitable to assess and monitor for END in acute practice.

Poorer reliability was clearly evident in certain items (Visual Fields, Gaze, and Ataxia), although variation was present across all the scales and items. Items that rely on subjective information, tend to have lower reliability (Gelmers et al., 1988). It is also clear that certain items are harder to assess and agree on than others. However, even within items with better classifications of reliability such as Motor Power, disagreements were common (Geisler et al., 2019). Subjective measurements can be as reliable as objective ones (Hahn et al., 2007). The aim is to limit error and variation as much as possible and obtain consistency of scores over infinite times (Streiner et al., 2015). Strategies to achieve this could involve having clear and easy to follow categories and methods of assessment, training and competency assessments.

Items can be reliable without being responsive to change. Pupillary Response was removed from the NIHSS as it was not deemed to change enough over the course of time to warrant inclusion (Brott et al., 1989). Seven items of the NIHSS (Questions, Gaze, Visual Fields, Ataxia, Language, Dysarthria, and Extinction/ Inattention) were shown to not be significantly related to ND at any time (Nanri et al., 2013). However, it was only completed in 43 patients after thrombolysis. Further evidence is needed about which items are most responsive to changes related to END. It is important that this is completed across a whole acute stroke population as changes associated with END will vary based on stroke type and severity.

Assessments need to be practical and understandable to those who use them (Wade, 2004). Once the items of importance are chosen it would be sensible to decide on the best way to assess and score them. Across scales there is extensive variation in the way items are assessed. Binary rating judgements tend to have higher agreement amongst examiners than multiple subjective options such as type of aphasia (Shinar et al., 1985). Limiting the number of variables would theortically reduce error as well as contribute to simplicity and utility of an assessment. However, extreme limiting of the options could result in a loss of sensitivity to change (Feinstein et al., 1986; Feinstein, 1987). Increasing options may prevent floor and

ceiling effects of assessments and allow better responsiveness. However, sensitivity to change has to be balanced with reproducibility as more options would bring a greater potential for variability in examiners choices which could reduce the possibility of detecting a meaningful change (Lyden & Lau, 1991). Inefficient items that lengthen the scale but have little discriminative value should be avoided (Prescott et. al, 1982) but these decisions should be based on more than reliability as was seen within the content validity data.

Ideally a few meaningful items could be chosen, that accurately reflect the patient's condition and allow identification and communication of END. The challenge is finding items that balance clinical utility and appropriate clinimetric properties. Certain items display better clinimetric properties dependent on the way they are assessed, such as motor items in the CNS compared with the NIHSS (Seki et al., 2014). The e-NIHSS was the only scale to achieve very good classifications of reliability across all items (Olivato et al., 2016). It has scoring criteria extended from the original NIHSS for Gaze, Facial Palsy, and Ataxia. Reliability amongst different examiners could increase if the assessment criteria of the items were clearer to understand and score. However, this would require further testing comparing methods of assessment for specific items.

4.4.2. Study relevance and quality

Multiple factors impact on the relevance of these reviews and the quality of the data within them. Many of the scales have been adopted into use across all strokes without further validation despite being developed for, or tested in, specific stroke types, mainly ischeamic. The NIHSS, for example, was originally developed for use with ischaemic stroke patients in The National Institute of Neurological Disorders and Stroke (NINDS) tPA trial (National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group, 1995) but is now used across all stroke types in practice. Two studies specifically considered the use of the NIHSS in ICH patients and showed poor reliability and responsiveness (Specogna, 2013; Specogna et al., 2013). This could indicate that the assumption that the NIHSS can be used to assess ICH patients may be wrong and this warrants further testing.

Any study of clinimetric properties is based on a snapshot of results and the results are dependent upon the populations studied. Different stroke subtypes have different clinical patterns of deterioration and may require different thresholds for intervention. Even within specific stroke subtypes none of the scales or items were tested across the full range of scoring criteria. Two papers for the NIHSS reported that all possible item score were represented (Lyden et al., 2005, Lyden et al., 2009). However, although the whole range of potential deficits

were included over the 18 patients included the examiners only scored six at a time and not all examiners completed every patient. Only one paper (NIHSS), allowed for the fact that their patients were not a random sample of all possible patients (Josephson et al., 2006). Further testing covering the full range of deficits in items would be needed to ensure assessments are suitable for the whole stroke population.

Variations in the populations studied could have impacted on the inconsistent results obtained across multiple scales and papers. Patient characteristics such as stroke severity, reduced conscious level, and communication ability affect the application and therefore the properties of the scales. There is a general trend that validation studies in more severe patients with marked decline in consciousness perform worse (Lee et al., 2017). The effects particularly occur in relation to reliability and measurement error. Greater score variation and degree of measurement error are seen across the studies in more severe strokes (Demaerschalk et al., 2012; Geisler et al., 2019; *Guterud et al., 2019;* Josephson et al., 2006; *Schmulling et al., 1998*). Six studies actively selected more alert and co-operative patients in which to test the scales (Adams et al., 1987; Alasheev et al., 2017; Birschel, 2005; Cote et al., 1989; Hantson et al., 1994; Nilanont et al., 2010). This selective approach would improve patient compliance and directly impact on the reliability reported, however, it reduces the quality and usefulness of the studies as they are not generalisable to the whole stroke population.

Study characteristics that may have affected results include study setting, choice of examiners, and training received. Very few of the studies represent true clinical practice. Only three were completed in uncontrolled clinical environments and their results indicate that this could have a profound effect on reliability (Demeestere et al., 2017; Specogna, 2013; Specogna et al. 2013). Numerous studies were not completed with acute stroke patients and some did not state the acuity of the population. Only seven were reported within the time period of interest (up to 72 hours post stroke). This is potentially important as alteration in condition may be different across different timeframes of stroke evolution. Detection of change over longer time periods (Brott et al., 1989, Seki et al., 2014) may not be relevant to the changes associated with early neurological deterioration (END).

Choice and number of examiners also impacts on study relevance and quality (de Vet et al., 2011). Most studies used a small number of set examiners and not a random selection of multiple examiners as would be seen in clinical practice. Only one described using a one-way random effects model to account for ratings being performed by different examiners (Specogna et al., 2013). Examiners were often doctors and in many studies were the

individuals developing the scale. Generally, examiners did not represent the multi-professional, and multiple grade workforce we would expect to be completing assessments and monitoring, so are not representative of practice.

The range of different statistical methods and techniques used across studies made direct comparisons difficult. Choice of reliability statistics is driven by multiple things such as number of scoring responses per item, numbers of examiners, patient numbers (too few and cannot calculate), and whether groups are balanced. Generally, the methods chosen, the indications for being chosen, and their potential limitations were poorly reported in the literature. This could be an indication that methods are chosen to suit data collected rather than an integral part of study design. Kappa was the most utilised method for assessing reliability, and which maybe because they are applicable to ordinal and interval data. Kappa can also analyse partial agreement for multivariate data of subjects rated on multiple characteristics accounts in a more effective way (Gwet, 2008). Although PA is not statistically sound, because it can only be used across two examiners, it could quickly show whether there is a high level of disagreement. This might be useful in clinical practice, especially across items, to check and train agreement within teams. Studies should be designed to ensure they test the scale for its intended purpose and not use the most convenient statistical method.

4.4.3. Clinimetric properties

Clinimetrics integrates knowledge from a range of disciplines like psychometrics, epidemiology, and biostatistics to provide researchers and clinicians with the best methods to assess, appraise and improve the methodological quality of their measurements. It promotes the use of clinical expertise rather than statistical techniques to develop and appraise measurement instruments (Feinstein, 1987). It was decided to focus on clinimetrics using the COSMIN taxonomy as it can help both evaluate and design health related measurements suitable for clinical practice and would therefore help evaluate current scales to provide important information for the development of the SNOBSS.

Clinimetrics is fundamentally important for healthcare measurement and assessment. For a clinical stroke assessment to be useful it should perform well across a range of properties. Whatever scale or assessment is used, it is important to know if it is adequate for its purpose. It is crucial to have practical and reliable instruments for accurate clinical monitoring (Finocchi et al., 2018). Assessments should be credible and have face validity, they should be reliable when repeated by single or multiple examiners and be easy to administer and score (Baron et al., 2013). The importance of properties depends upon the purpose of measurement. For

instance, if the aim is to accurately score a patient's functional neurological condition, then reliability (how consistent the scale is) and validity (whether the scale measures what it intends to) are crucial. Although, if we want to evaluate change, then responsiveness/change over time could be considered more important.

The theoretical underpinnings and internal structures of the scales included in the reviews were not well described or explored. In many studies there is an underlying assumption that the scales are unidimensional (despite an obvious hemispheric split in item assessment). Although all scales assess symptoms of stroke, they are created with different items and tested in different populations so therefore are composed of different theoretical constructs (Edwards et al., 1995). The heterogeneity of stroke means that items measure function related to specific lesion locations and therefore not all are relevant for all patients. Theoretically then stroke scales are a collection of separate unidimensional scales. Despite this inherent multidimensionality, the clinimetric properties are assessed by methods suited to unidimensional scales with unknown consequences on data quality.

Overall, the methodological quality across the clinimetric properties of scales was lower than expected. There were repeated inaccuracies in methods including ordinal scale scores being incorrectly totalled and treated as continuous parametric data. In some papers there was uncertainty in the clinimetric terminology itself. Some evidence was stated to represent a certain property but was actually another (e.g., concurrent not criterion validity) or was not present despite the paper stipulating it was. The ambiguity in the statistical tests added to difficulties in comparing property data within and between scales.

The range of scales developed illustrates that there is currently no consensus on what should be included in a meaningful stroke scale. All scales had content based on the neurological exam to identify a range of potential pathology. However, details on the processes and justifications for decisions about what the scales include is poorly reported. Only the Hunter NIHSS-8 used an expert panel, with clinicians independent of scale development, to create the scale. Moving forward clinical opinion should be sought on what items need to be monitored to identify END to ensure appropriate content validity for practice.

There was a lack of data on construct validity of the scales. Existing stroke scales were treated as unidimensional despite evidence to show their multidimensionality. It may not be possible to develop a single assessment that can effectively identify change that can be used across all stroke types as they will have different neurological presentations. It seems sensible to assume

that all stroke assessments should have an element of multidimensionality as they are going to be used to measure different hemisphere and arterial territory strokes. Either a multidimensional assessment or a series of assessments may be required to identify END in acute stroke. Assessments based on stroke location or different baseline severities might be warranted, making assessments more targeted and efficient (Millis et al., 2007).

Currently, no gold standard exists for neurological assessment and monitoring of stroke. Therefore, the concurrent validity data is varied in terms of comparison scales. Although the data indicates there is correlation between many measurements used within stroke populations it is meaningless without clarification of the constructs the scales measure. Scales may show correlation even if they measure different constructs. The construct validity data showed variability in methods and results. This can be the most difficult validity to establish, especially if, as in this case, the assessments are based on constructs that are not "real" (that is they are not directly observable). Assessments should ideally represent a small number of underlying factors (Lyden et al., 1999). This would indicate that items correlate well with each other which is preferable to each item representing a single factor.

Reliability is easier to describe and measure than validity (Lyden & Lau, 1991). This might account for why inter-rater reliability is the most tested clinimetric property across all the scales. Most scales reported high internal consistency. However, these results cannot be assumed to be accurate as Cronbach's alpha (α) should not be calculated unless the scales are unidimensional. The high correlation reported could instead indicate redundancy amongst items and that items are measuring similar concepts (lack of efficiency). In the future it would be helpful to assess for redundancy of items as it would falsely elevate reliability estimates. The kappa statistics for reliability allowed for the most comparison between data by utilising the Landis and Koch (1977) classifications. Across the literature the NIHSS, and its modified versions were generally less reliable than the other scales. However, as the results for the other scales are from a lesser number of small-scale studies this could be due to other study characteristics and not be truly indicative of improved reliability.

Independence in examinations is important when assessing inter-rater reliability. Independence was not clearly reported in many of the studies. Several reported simultaneous assessments, to ensure patient stability, but if the scoring was not truly independent it could have overinflated the reliability calculations. There is less data on intra-rater than inter-rater reliability across the scales. This is common and partially explained by the tendency to underestimate the importance of data reproducibility (Gwet, 2008). However, intra-rater

reliability is especially important in terms of repeated assessment to monitor for END. Frequent assessment might be completed by the same clinician, meaning the results might technically be affected by recall. Although logically if regular assessment is completed by the same individual, they may be better placed to detect subtle changes.

Level of error across the literature was higher than anticipated even before the additional variation in scoring not classified as error was extracted. Sources of error can be attributable to the incorrect use of the scale but there are signs to suggest that some of the scales have inherent error within them. Extensive variation in allowed scores was seen across several papers meaning that the error with the use of scales is not fully reported. In future testing, level of error should be an integral part of studies. As the level of error is likely to be influenced by variations in factors such as study population and training it is vital it is tested in large scale studies.

Sensitivity to detect change was poorly reported across the scales with only two papers reporting the minimal detectable difference (MDD) and this was across the whole score of the NIHSS. The reported MDD_{95} values were 10 or more points (Specogna et al., 2013) and six or more points (Specogna, 2013). These studies did not assess clinically meaningful changes on the NIHSS but evaluated the errors associated with rating the NIHSS using a statistical distribution-based method. However, these poor results indicate the potential for far higher error in the NIHSS than has been previously considered. The studies were performed in uncontrolled clinical environments only in ICH, elements have not been tested elsewhere. The results, indicate that further evidence is needed about measurement error in the NIHSS and other scales.

Responsiveness of a scale, although crucial, Is challenging to test. Scales need to measure a specific item accurately and not under or overestimate otherwise responsiveness is lost. Data on responsiveness was only available across four scales (GCS, CNS, NIHSS, and FOUR score). The GCS and FOUR score were shown to be less sensitive to change than subjective assessment by trained nurses (Zink et al., 2012) which calls into question their usefulness to clinical practice.

4.4.4. Clinical considerations

The GCS and FOUR score are not stroke specific scales but level of consciousness assessments and their use in stroke practice has not been robustly tested. Interestingly, for the GCS, inter-rater scores dropped when tested in the confirmed stroke patients particularly in verbal

assessment. This suggests that the GCS is not suitable to be used in a population where speech deficits are linked to the condition and not just the level of consciousness (Lee et al., 2017). When 15-42% of stroke patients suffer from some form of speech deficit (Inatomi et al., 2008; Kadojić et al., 2012; Ryglewicz et al., 2000) the widespread use of the GCS after acute stroke may not be warranted. Change in level of consciousness is a late indicator of deterioration, for all but the most severe strokes, therefore these scales are potentially not suitable for widescale use to detect END.

Inter-rater reliability across all items was found to be variable and no items show repeated 'very good' according to the Landis and Koch (1977) classifications. These classifications did aid broad comparison and showed the extreme ranges of reliabilities across studies, scales, and items. Nevertheless, there is no agreement on what the required level of classification should be for scales or items used in clinical practice. No subjective scale will ever perform perfectly but clinicians need to be more aware of scale limitations and be involved in deciding what is an acceptable level on which to base clinical decision making.

Reproducibility, within and between examiners is a key concern in repeated assessments in clinical practice. All these assessments rely on complex subjective decisions. Rather than just assigning a score it is important that they notice a change in a patient's condition over time. Assessments in clinical practice are not independent or blinded and cannot be as clinicians need to see and share changes and trends in condition. Added to this different staff might complete the assessments based on factors such as work allocation and staffing levels. Therefore, there could be a tendency towards agreement with previous findings, be that with yourself or others. Although this would improve reliability of scales it could have a detrimental effect on responsiveness with change not being picked up.

There is a need to find the balance between clinically useful and adequate clinimetric measurement properties. If the only objective was excellent clinimetric data, you could use multiple scales for patients with different characteristics however this might not be practical in practice. As already stated, different stroke types and severities exhibit different neurological presentations. Identification of whether an item is present or absent is easier to score than change within a specific item. Therefore, for clinicians it is harder to notice change in a deficit than a deficit itself. To further compound this difficulty scales also have a potential ceiling effect beyond which change cannot be recorded. This is especially of concern in severe strokes where within scale items it is unable to spot when someone has deteriorated.

Stroke severity can affect the application and clinimetric properties of assessments in practice. In some patients, items may become untestable e.g., unable to assess Speech in a comatose patient. Greater levels of measurement error were also associated with more severe stroke presentations. Formal assessment of measurement error was limited across the studies, but sixteen papers allowed variation in scoring without it being formally classed as error. Allowed variation in scoring means that different people are assessing items differently and this has potentially serious consequences especially if care decisions are affected by the results. It is of vital importance that all clinicians who assess a patient are using the same criteria otherwise measurement error is intrinsic in clinical practice.

Performance of neurological assessment and monitoring is affected by minor differences in the way items are rated, such as visual field assessment through counting or finger wriggling, (Birschel et al., 2004) as well as communication inconsistencies in reporting (Iacono et al., 2014). Reliability of assessment is modifiable through training, certification and use of standardised protocols (Harrison et al., 2013). Differences in training, experience and professional background could easily produce different results especially in terms of clinimetric properties and effectiveness in clinical practice. Training and its effect on clinimetric properties were not a direct focus of this review but the range of training made available across the studies is recorded in Appendix 4.23. A couple of the NIHSS studies suggested that training improved reliability (Berthier et al., 2012, 2013, *Schmulling et al., 1998*).

It has been stated that multiple professional groups can use the NIHSS with limited training (Lyden, 2017, Spilker et al., 1997) However, the study widely cited for nurses routinely using the NIHSS involved intensive and robust training over a month (Dewey et al., 1999). The NIHSS certification process was initially aimed at highly skilled and trained stroke neurologists with prior experience in using the scale (Hills et al., 2009). This review highlighted that NIHSS certification awards a pass despite allowing a range of scores across patients (Josephson et al., 2006). The NIHSS certification process does not provide any feedback on scoring, so people cannot learn from mistakes or feel confident in their choices. The assessment involves the clinician scoring a video of an expert applying the NIHSS and therefore does not prepare clinicians to competently apply the scale themselves. International surveys have also shown there is a lack of standardisation in the way GCS assessment is taught (Reith et al., 2016). Whatever scale or assessment is chosen in the future multiple levels of training and certification processes need to be evaluated (Berthier et al., 2012, 2013; Jones et al., 2018). Adequate training is needed to achieve and maintain competence of the clinical workforce to perform effective serial bedside monitoring to detect END.

Serial assessments require time to complete so the length of assessments is important in clincal practice with many other demands on clinican time (Dall'Ora et al., 2021). Regular assessment of multiple patients has major resource implications (Benedetti et al., 2021). Removal or reduction of items would reduce the time taken but this must be weighed up against requirements of the scales (Lyden, 1999; Lyden et al., 2001). If a scale is too short, it will ultimately reduce its ability to identify END. Training could reduce the time taken to assess (*Isahaya, 2017*) but there will be a limit to how much this can be reduced no matter how much training or experience is provided. Although generally successive use of the scale will reduce time taken to complete there will always be variation in how long patient assessments last (Meyer et al., 2008). Implications for time to complete assessment is an important consideration when deciding what is feasible and achieveable.

4.5. Strengths and Limitations

The reviews were completed in a systematic manner by the author but there was no consistent second reviewer involved in the process. Where clarification and support were needed this was sought from the highly experienced supervision team or the NIHR Complex Review Support Unit (CRSU).

There is the potential that some data was not captured by the clinimetric search strategy. However, search strategies were planned and completed as robustly as possible with the aim of including all available data. Extensive work was completed to ensure the correct medical subject headings (MeSH), keywords, and general search terms were used (Salvador-Olivan, 2019). Sources that are not indexed in the databases such as Birschel (2005) are even more challenging to locate. The addition of nontraditional literature e.g., abstracts risked limiting methodological quality. However, collation of maximum data for comparison outweighed any risk and overall, the methodological quality was poor across all literature sources. Due to limited and inconsistent data it was not possible to quantitatively pool results, but they have been presented to allow comparison across scales within clinimetric properties where possible. There is potential publication bias in the literature in terms of what has been published.

4.6. Chapter summary

This review has outlined some of the multiple factors that could influence the effectiveness of assessments to identify END in clinical practice. For a scale to be suitable for universal use it

must be able to maintain validity and reliability amongst multiple examiners in the full range of settings used and for all the purposes it is employed for. Fundamentally we need a practical and acceptable selection of items, or series of items, to use at the bedside that has adequate responsiveness to change seen in END. It must be able to maintain reliability, with limited error, amongst multiple examiners for a full range of patients, in a range of clinical settings.

Clinimetric properties remain an important aspect of any measurement tool and should help guide future development and testing of the SNOBSS. Although future decisions might not be led by the clinimetric data presented here the principles remain important. Clinicians are often pragmatic in their decisions and use what is available and fits best in their current practice and choice of an assessment scale is dependent upon multiple factors not just clinimetric properties. Stroke assessment must be simple, have clear definitions and be quick to score if it is to be adopted into regular routine clinical practice. Large scale prospective studies are needed with heterogenous populations of patients and examiners in true clinical practice to test items that are most important and relevant for identifying change associated with END across the whole stroke population.

The findings from the reviews indicate that the scales identified are not suitable to be used for routine neurological monitoring to detect END. In fact, the findings indicate that rather than a scale what is needed is an assessment containing a series of items to direct rather than quantify identification of change. These findings support the development of a new tool within the Standardised Neurological OBservation Schedule (SNOBSS) to enable detection of END in an acute stroke population. Key findings from these reviews will be fed into the development of the SNOBSS in chapter 7 and the recommendations for future research in chapter 8.

The next chapter, chapter 5, presents the Neurological Assessment Practices after Stroke Survey. It describes the development and delivery of a UK wide survey to explore current practice and understanding of neurological assessment and monitoring practice.

Chapter 5- UK wide survey of current practice and understanding of neurological assessment and monitoring practices after acute stroke

The previous chapter provided an overview of the clinimetric properties of the multiple scales available for neurological assessment and monitoring. This chapter presents the findings of a United Kingdom (UK)-wide survey to explore current practice and understanding of neurological assessment and monitoring practices after acute stroke. The chapter begins by describing the development of the questionnaire and the distribution of the survey before presenting the quantitative and qualitative results. The discussion focuses on the variation in practice and the desire for change in this element of care. Key practices and opinions that might impact on future practice are identified as well as recommendations for future research, some of which have been addressed within the interviews in Chapter 6.

5.1. Aims and Objectives

The aims of the survey were to:

- establish current practice and briefly explore clinicians' experiences of neurological assessment and monitoring in the acute phase of stroke.
- determine knowledge, understanding, and acceptability of neurological assessment and monitoring in acute stroke and explore the barriers and facilitators to its implementation in clinical practice.

The objectives of the survey were to:

- outline current practice and level of variation in relation to neurological assessment and monitoring across UK stroke units.
- ascertain how neurological deterioration is currently identified and managed.
- check clinicians' understanding of the importance of neurological assessment and monitoring and whether they feel change is warranted.
- identify barriers and facilitators to completing neurological assessment and monitoring for patients after acute stroke.
- establish the current level of training provided in this element of care.

5.2. Methods

5.2.1 Study design

Study design was a cross-sectional survey using a self-administered questionnaire. The survey was designed to provide a snapshot of current practice and ascertain the level of variation in neurological assessment and monitoring practices across the UK.

5.2.2 Setting and Recruitment

All UK hospitals identified through audit data as admitting patients immediately after stroke were invited to participate in the survey (n=168). Hospitals in England, Wales, and Northern Ireland were identified via the Royal College of Physicians' Sentinel Stroke National Audit Programme (SSNAP) audit data. Hospitals in Scotland were identified via the Scottish Stroke Care Audit. Each hospital was asked to complete one questionnaire and the aim was to achieve a minimum response rate of 60%.

5.2.3 Questionnaire development and content

Questions were formulated based on the author's knowledge and experience of neurological assessment and monitoring (Section 3.3) and the literature (Chapter 2). This included the range of scales available, whether specific patient characteristics, co-morbidities or stroke types warrant different monitoring regimens, and known barriers and facilitators to practice such as time and training available.

A questionnaire was designed for the study. Questions were developed to illicit:

- demographic information of the unit and respondent
- current practice in relation to neurological assessment and monitoring
- how deterioration is currently noted and acted upon
- clinicians' experience of using neurological assessment and monitoring
- understanding of importance of neurological assessment and monitoring
- training staff receive in relation to neurological assessment and monitoring

Care was taken to make the questions clear, concise, and without bias. Most questions were closed to ensure consistency in responses and allow easier comparison (Jones et al., 2013). Free text options were utilised in some questions, to allow the provision of an alternative answer, or elicit a more detailed response. Likert- type scales were employed to measure attitudes or how much the respondent agreed or disagreed with statements. In sections, where participants were asked about multiple level of agreement statements, they were presented in a random order with some alteration of the direction of questioning. This was to

prevent respondent apathy and allowed responses to be verified against other sections within the questionnaire. Particular attention was paid to question order and aesthetics to make visual aspects such as symmetry appealing to maximise response rates (Mahon-Haft & Dillman, 2010).

Development was iterative with multiple drafts. The questionnaire was reviewed by the student's supervisory team. Initial piloting applied a cognitive interviewing style during completion by an experienced nurse with both clinical and research experience. This involved the nurse completing the questionnaire in the presence of the author, discussing the interpretation of the questions, questionnaire design, and other factors that could impact completion (Willis, 2005). After changes, the questionnaire was then piloted by stroke nurse consultants and stroke specialist nurses at three different NHS Trusts. They reviewed the questions, and response options, and determined the length of time to complete the questionnaire (approximately 30 minutes). Changes were made following feedback.

The questionnaire (Appendix 5.1) was divided into 5 sections:

1. Unit Demographics
2. Neurological Assessment/Monitoring Practices
3. Neurological Deterioration
4. Experience
5. Training

Participants were also asked, on behalf of Trusts, if they had and would be willing to provide a copy of the following:

- policies/protocols or documents related to neurological assessment /monitoring
- dedicated neurological assessment/monitoring documentation
- policies/protocols or documents related to response if neurological deterioration is noted
- tools or documentation to support neurological assessment not listed in the questionnaire
- neurological assessment and monitoring competency documents

Extensive time and effort went into the design and execution of the questionnaire. The questionnaire needed to be well designed as it could not just be changed if problems in completion were identified after ethics had been approved. Standardisation of the questions should have led to greater precision in answers but there are always risks that participants will understand questions differently (Sapsford, 2007).

5.2.4. Study participants

A nonprobability convenience sampling technique based on ease of access was adopted to obtain a single participant to complete the survey at each site (Etiken et al., 2016). Clinicians were initially invited to participate in the survey via email alerts through professional networks (e.g., FutureNHS Collaboration Platform, National Stroke Nurses Forum). Where individuals did not volunteer each stroke unit was contacted to identify a potential participant. Participants were required to have a working knowledge of neurological assessment and monitoring practices within the unit there were no other stipulations. A conscious decision was made not to limit participants by their professional group as they can have different philosophies and norms which may have influenced some of their experience responses (Baxter & Brumfitt, 2008).

Potential participants responding to the invitation were emailed the participant information sheet. By return of email, they had to confirm they were willing to receive the questionnaire and if they were prepared to receive follow up reminders. They were then sent the questionnaire for their unit with a unique identifier code. This individual was responsible for ensuring the questionnaire was completed and returned, although other team members could be involved in the completion. Participation was voluntary and informed consent was implied on return of the completed questionnaire.

5.2.5. Survey Delivery

The questionnaire could be supplied either by post or by e-mail. Postal versions were supplied with a reply-paid envelope. E-mails included an attached word document, which could be printed, completed by hand, and scanned to return or completed electronically and emailed back. Potential participants had up to three months to complete the questionnaire. Multiple strategies were employed to improve return rates. Until returned, follow-up was by email three times at three weekly intervals and then by telephone twice at two weekly intervals. Entry to a voucher prize draw was also offered as motivation as monetary incentives have previously been shown to improve response rates (Jobber et al., 2004).

5.2.6. Data management and analysis

All data were managed following university procedures adhering to data protection and general data protection regulation (GDPR) legislation. Questionnaire data were entered into a customised Microsoft Excel (Version 2108, Microsoft Corporation, Redmond, WA, USA)

spreadsheet. All inputted data was validated against the original questionnaires to minimise data errors and missing data. If responses were missing but data was provided elsewhere in a free-text format it was inputted. Data was then uploaded to Statistical Package for the Social Sciences (SPSS, Version 28, IBM Corporation, Armonk, NY, USA) for analysis.

To ascertain the frequency of neurological assessment and monitoring the questionnaire asked respondents to circle the most common frequency of assessment, for different patient groups, across different time periods in the first 72 hours. For post thrombolysis and thrombectomy assessment, the time periods asked about were 0-8 hours, 8-16 hours, 16-24 hours, 24-48 hours, 48-72 hours, and beyond 72 hours. For ischaemic stroke (without thrombolysis or thrombectomy), ICH (with blood pressure alteration), ICH (without blood pressure alteration), potential hemicraniectomy, and other patient groups time periods divided into 0-24 hours, 24-48 hours, 48-72 hours, and beyond 72 hours. The questionnaire did allow participants to state if a particular patient group was not seen in their unit at all, (n/a patient group, e.g., some units might not deal with thrombectomy patients) or in a particular time period (n/a time period e.g. patients receiving thrombolysis in other higher acuity units), and it also asked after what time period neurological monitoring was discontinued.

In terms of frequency of neurological monitoring, extensive data was obtained but some respondents provided a free-text response rather than circling the most common frequency for each time point. Where enough detail was provided in free text to know the most common frequency for a time period these were added to the data set. However, other participants provided a range of frequencies for a time period that did not stipulate the length of time each frequency was used for (e.g., for the 0-8 hour period they might state 15 minutes, 30 minutes, and hourly). In these cases, the most frequent responses were recorded as the most common was unknown. Initially, in analysis, it was attempted to manage the most common and the most frequent responses separately. However, the extensive range of answers and options made it difficult to differentiate and see variation. Therefore, it was decided to amalgamate and present the most common frequency and, where that was not available, the most frequent responses together. This was justified as more frequent monitoring has potential impact in relation to staff time and resources and it was deemed better to overestimate frequency of monitoring than underestimate it. Also, all the most frequent responses reported were similar to the most common frequencies reported and represented small numbers (four or less), except for one time period in the thrombolysis patient group. Further explanation of that data is provided with the results. However, for all other patient groups the data is amalgamated and reported as the most common frequency to prevent misunderstanding.

Overall analysis involved descriptive statistics with results mainly reported as counts and percentages. Qualitative data from open-ended questions was collated to report trends and themes where appropriate using content analysis. Content analysis is defined as a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns (Heish & Shannon, 2005). The basic coding process in content analysis is to organize large quantities of text into much fewer content categories (Weber, 1990). Conventional content analysis was used to derive the coding categories from the text data and avoid using preconceived categories (Kondracki & Wellman, 2002). The author undertook the analysis and immersed themselves in the data to allow new insights to emerge based on participants' responses.

A potential limitation of content analysis is in terms of credibility in that it can fail to identify key categories and a complete understanding of the context and therefore findings do not accurately represent the data (Lincoln & Guba, 1985). However, this was addressed through prolonged engagement with the data, discussion of findings and process of analysis with supervisors and triangulation with other data where possible and appropriate (Lincoln & Guba, 1985; Manning, 1997). Missing data were monitored in terms of overall response rate and whether specific questions were poorly answered or regularly missed from completion.

5.3. Ethical and Local Approvals

Ethical approval was obtained from the Science, Technology, Engineering, Medicine, and Health (STEMH) ethics committee at the University of Central Lancashire (UCLan) (reference STEMH 1018) (Appendix 5.2) and under proportionate review from the Health Research Authority (HRA) (project ID 261850, REC reference 19/HRA/4113) (Appendix 5.3). The HRA proportionate review decided that the project should be completed under a participant identification centre (PIC) agreement. Local approval was obtained from each Trust's research department before questionnaires were sent to the stroke units.

5.4. Results

Information on survey response rates is presented first followed by data from the five sections of the questionnaire.

- Unit Demographics
- Neurological Assessment and Monitoring Practices
- Neurological Deterioration

- Experience
- Training

5.4.1. Survey Response Rate

The process of site recruitment is illustrated in Figure 5.1, of the 168 units originally identified, only 156 were eligible due to service reconfiguration where 12 units no longer admitted acute stroke. From the 156 eligible 125 (80%) returned the questionnaire (December 2019 - September 2021). Follow up procedures appeared effective as from 138 sent out only 11 questionnaires were not returned. Figure 5.2 is a map showing information on response rates by geographical regions.

Figure 5.1 Flow diagram of survey recruitment

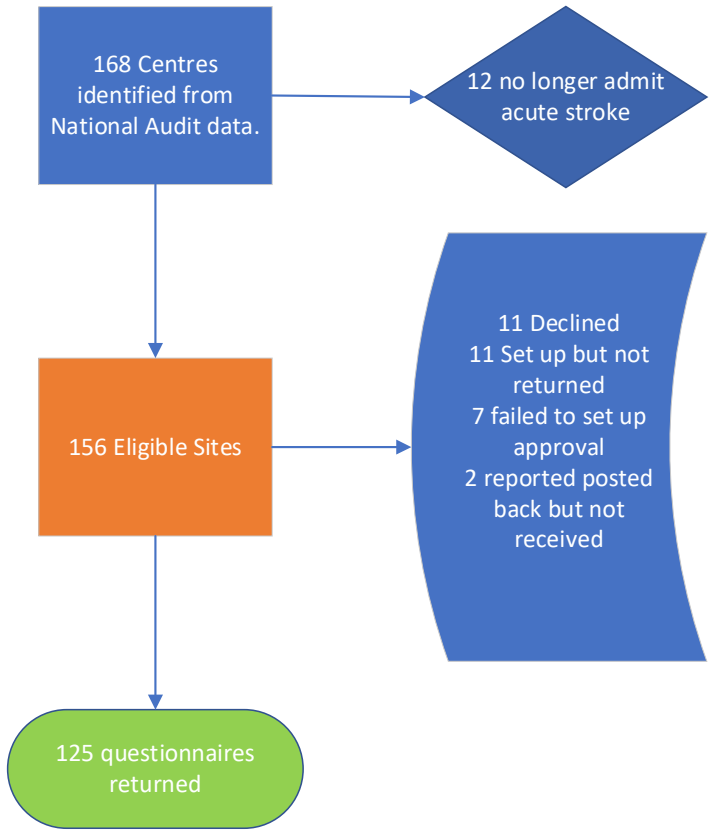
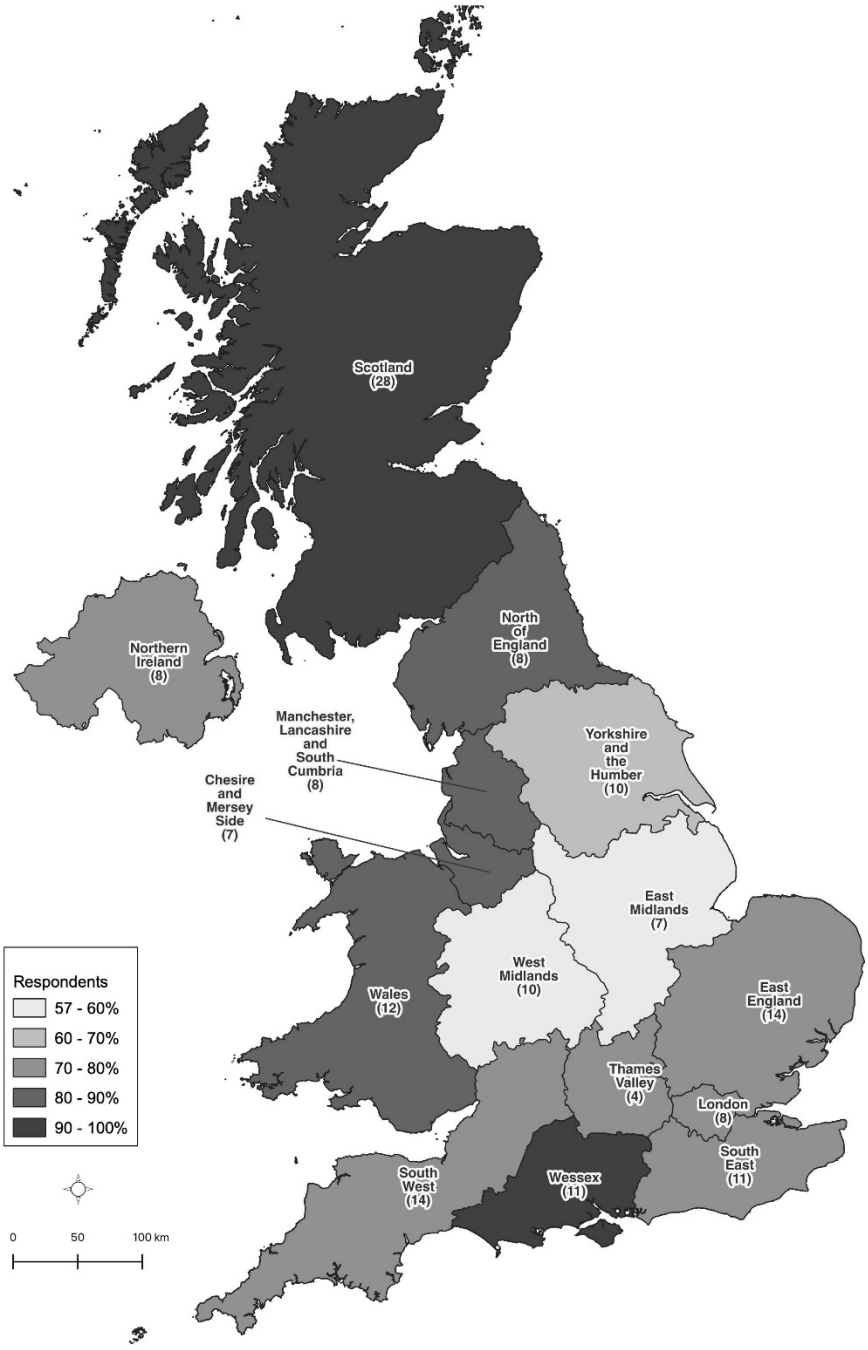


Figure 5.2 Map showing range of percentage response rates to questionnaire by geographical region

(numbers in brackets indicate total number of acute stroke units in geographical region).



Overall response rates were good and there was no distinct patterns around responders and non-responders to the questionnaire. The response rate was greatest across Scotland, but this could have been because recruitment commenced here before the COVID pandemic. The lowest levels of response were from Midlands regions. This geographical area undertook a lot of reconfigurations over the course of the study which may have accounted for reduced participation.

5.4.2. Unit Demographics

Participants were asked to enter information about their role, hospital setting, unit classification, total bed capacity, allocation of beds to stroke, and specialist stroke services provided.

Nurses were the largest respondent group (82%), followed by doctors (11%) and therapists (<2%). Most respondents held a clinical role (94%). The majority 54% (n=68) were from non-tertiary, general, district, or community hospitals with an Emergency Department (ED). Responses from large tertiary hospitals (take referrals from other hospitals) made up 41% (n=51) of responses and 5% (n=6), either didn't answer this question or were from hospitals without an ED.

Regarding unit classification, 3% (n=4) were reported as hyperacute stroke units (HASU), 30% (n=38) as acute stroke units (ASU), 62% (n=78) as both, 2% (n=3) reported to be neither and 1% (n=2) failed to respond. In terms of bed capacity and allocation to stroke, the number of beds in units ranged from 11-80 (median 28, IQR 14), and the number of dedicated stroke beds ranged from 0-67 (median 26, IQR 15). Seventy- two percent of units (n=90) units were dedicated to stroke. Respondents were also asked about bed types as per SSNAP data. There are three-bed types allocated dependent upon the acuity level they are used for, but this was poorly completed and cannot be reported.

Ninety-four percent of hospitals (n=117) reported providing thrombolysis treatment and 17% (n=21) thrombectomy. Less hospitals provide telemedicine services than receive it (Table 5.1) although 27% (n=32) both provide and receive telemedicine. Generally, use of telemedicine is higher in non-tertiary then tertiary hospitals indicating that they might need more collaboration to provide stroke specific decision making both in and out of hours. However, most telemedicine use was for out-of-hours services.

Table 5.1 Frequency and percentage of hospitals that provide and/or receive telemedicine and times of operation.

	Number of units (%)	**Number 9am – 5pm service (%)**	**Number out of hours (OOH) (%)**	**Number both (9 am-5pm and OOH) (%)**
Provide Telemedicine	39* (31)	8 (6)	24 (19)	7 (6)
Receive Telemedicine	60^ (48)	2 (2)	46(37)	12 (10)

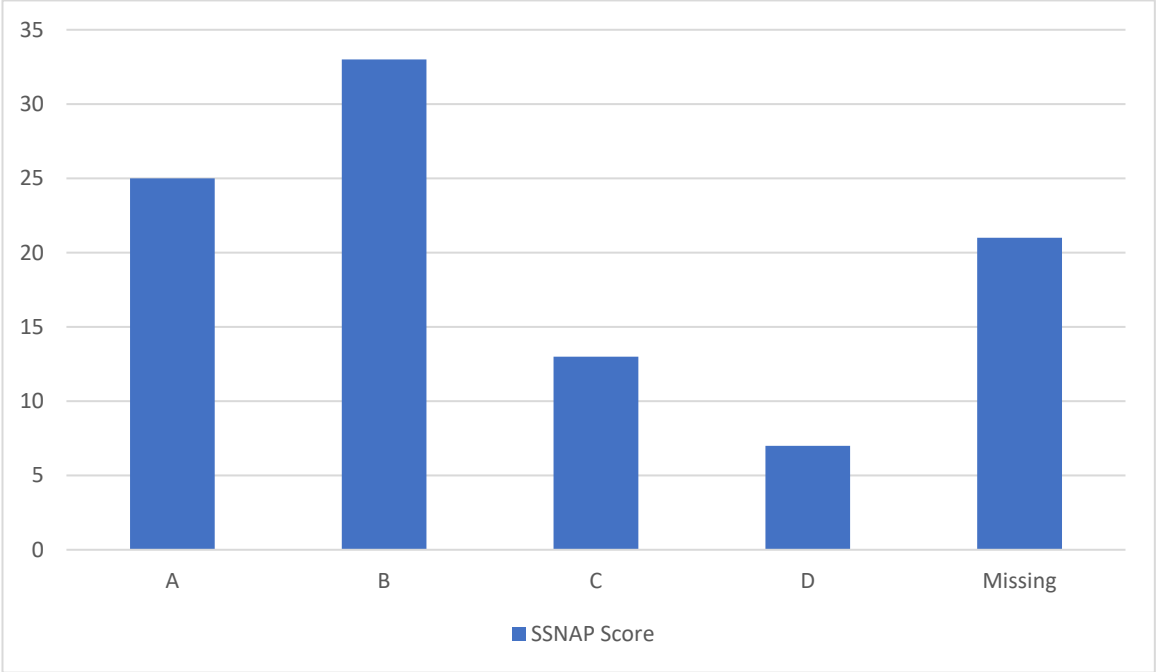
*Indicates 16 tertiary, 23 non-tertiary

^ Indicates 17 tertiary, 43 non-tertiary

Rehabilitation services were provided by 78% (n=98) of hospitals. Thirty-eight hospitals reported purely stroke-specific rehabilitation, the others had combined wards (stroke, neurology, or general).

In terms of overall SSNAP clinical audit score, most units scored a B (Figure 5.3 below). The 26 Scottish units are not represented as there is not a single score overview available within the Scottish Stroke Care Audit.

Figure 5.3 Frequency of the most recent overall SSNAP (Sentinel Stroke National Audit Programme) clinical audit score for UK units at the time of survey completion (excluding Scotland).



5.4.3 Neurological Assessment and Monitoring Practices

In this section participant responses about the practical aspects of completing neurological assessment and monitoring are reported.

5.4.3.1. Completion of Assessments

There was a multi-disciplinary approach to the completion reported. Doctors, physician assistants, nurses (including specialist nurses/nurse practitioners), therapists, healthcare assistants) HCAs, and students were all reported to be involved. In fact, in 38% (n=48) of units, no one professional group took overall responsibility, and it was described as a team approach. However, free-text comments suggested that doctors and nurses are the professionals most involved with nurses mostly responsible for completing assessments or monitoring in 54% (n=68) of units. Ninety-eight % of units (n=123) reported completing physiological observations at the same time as neurological monitoring. In 81% (n=101) units, both types of monitoring are completed by the same person. Where this did not occur (n=24 units), it was most common for healthcare assistants (HCAs) to complete the physiological observations and a trained member of staff to complete the neurological monitoring (n=14).

5.4.3.2. Documentation and Communication of Assessments

Documentation of the results of neurological assessment and monitoring was often recorded in multiple places. Most common were observation charts in 56% (n=70) of units and dedicated neurological assessment forms in 51% (n=61) followed by clinical pathway and care plans. There was a relatively even split between units that used paper-based patient notes (43%) and units that used electronic notes systems (42%). This was in keeping with the recording of physiological observations, where across the parameters of blood pressure, oxygen saturation, heart rate, and temperature, respondents reported that 53-54% of hospitals record electronically, 42-45% were paper-based, and 2-4% reported recording both electronically and on paper.

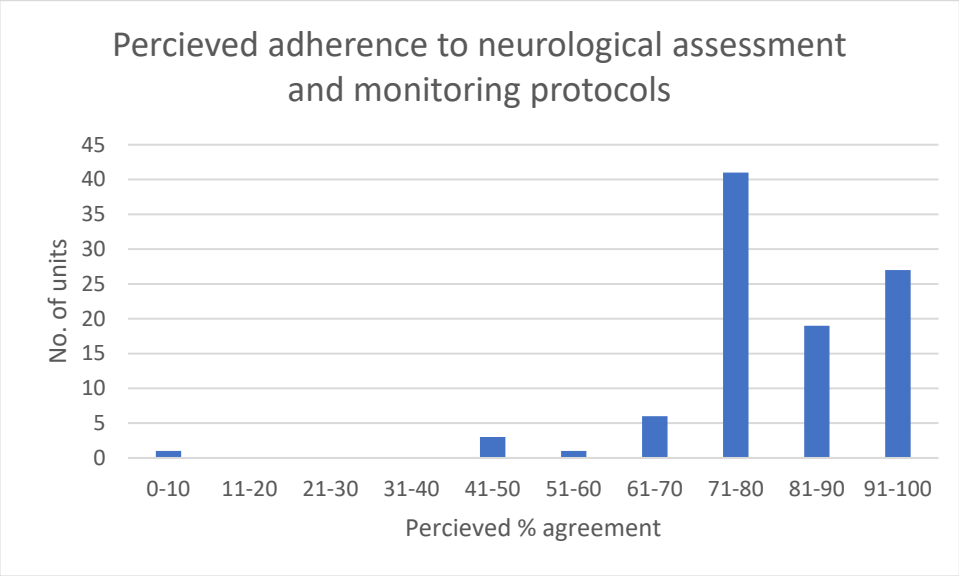
Inclusion of neurological assessments and monitoring in patient handovers varied, 56% (n=70) of respondents reported that this occurred regularly, 36 % (n=45) sometimes, and 8% (n=10) not regularly. Of those that reported regular inclusion in handover, approximately half (52 %, n=37) stated this was in twice daily shift handovers. The means of how this was handed over was less robustly answered, with 38% (n=27) of respondents having reported face-to-face verbal handover away from the patient, 16% (n=11) printed, written, or electronic

communication, and 9% (n=6) verbal handover at the patient's bedside. Thirty-seven percent (n=26) did not provide a response. Reasons for inclusion in handover were only reported by 23 respondents, and reasons covered four categories: to report abnormal results or any change/ deterioration (n=9), to provide/ ensure continuity (n=6), to record baseline or current status (n=5), and to report the frequency of monitoring (n=2).

5.4.3.3. Protocol for Assessments

Eighty percent (n=100) units reported having neurological assessment and monitoring protocols in place, 17% (n=21) reported no protocol, and 3% (n=4) did not know if the unit had a protocol. In the 100 units with protocols, 78% (n=97) reported they were stroke specific. However, when asked to what extent do you think your ward/unit adheres to the neurological assessment/monitoring protocol only 27 of the 100 units with protocols report >90% adherence (Figure 5.4.)

Figure 5.4 Perceived percentage adherence to neurological assessment and monitoring protocols.



5.4.3.4. Assessment Scales

All units reported using more than one scale for completion of neurological assessment and monitoring (Table 5.2). The most used scales are the Glasgow Coma Scale (GCS), AVPU (Alert, Voice, Pain, Unresponsive), and National Institutes for Health Stroke Scale (NIHSS). The general pattern across all responses was that the GCS and AVPU was used for regular and ongoing monitoring. The NIHSS was mainly used on admission and at key time points including at 2 and 24 hours post thrombolysis and/or thrombectomy which is in keeping with the requirements

for SSNAP. Thirty responses described occasional use of the NIHSS mainly instigated if the patient deteriorates or their condition changes. Only thirteen responses outlined more routine use of the NIHSS. Four reported daily use and the rest reported frequencies varied from hourly to weekly.

Table 5.2 Scales used in stroke units by numbers and percentages and an overview of the collated responses around what, when, or how often they use the scales.

Scale	**No. of units**	**% of units**	**Main reported use/s**
Glasgow Coma Scale (GCS)	120	96	Routine monitoring
AVPU (Alert, Voice, Pain, Unresponsive)	117	94	Routine monitoring
National Institutes for Health Stroke Scale (NIHSS)	117	94	On admission At set times for specific patient groups Patient deterioration or condition change Limited routine repeated use
Modified National Institutes for Health Stroke Scale (mNIHSS)	11	9	Similar pattern to NIHSS use
Standardised Nursing OBservations for Stroke (SNOBS)	11	9	Routine monitoring
Stroke Thrombolysis Observation Complication (STOC) Chart	4	3	Post thrombolysis monitoring
Neuro Observation Charts (unspecified)	2	2	Routine monitoring
Canadian Neurological Scale (CNS)	1	1	2 and 24 hours after thrombolysis Actioned if deterioration was noted within first 24 hours
FAST score (Face, Arm, Speech, Time)	1	1	Monitoring of Transient Ischaemic Attack (TIA) patients

5.4.3.5. Frequency of Assessments

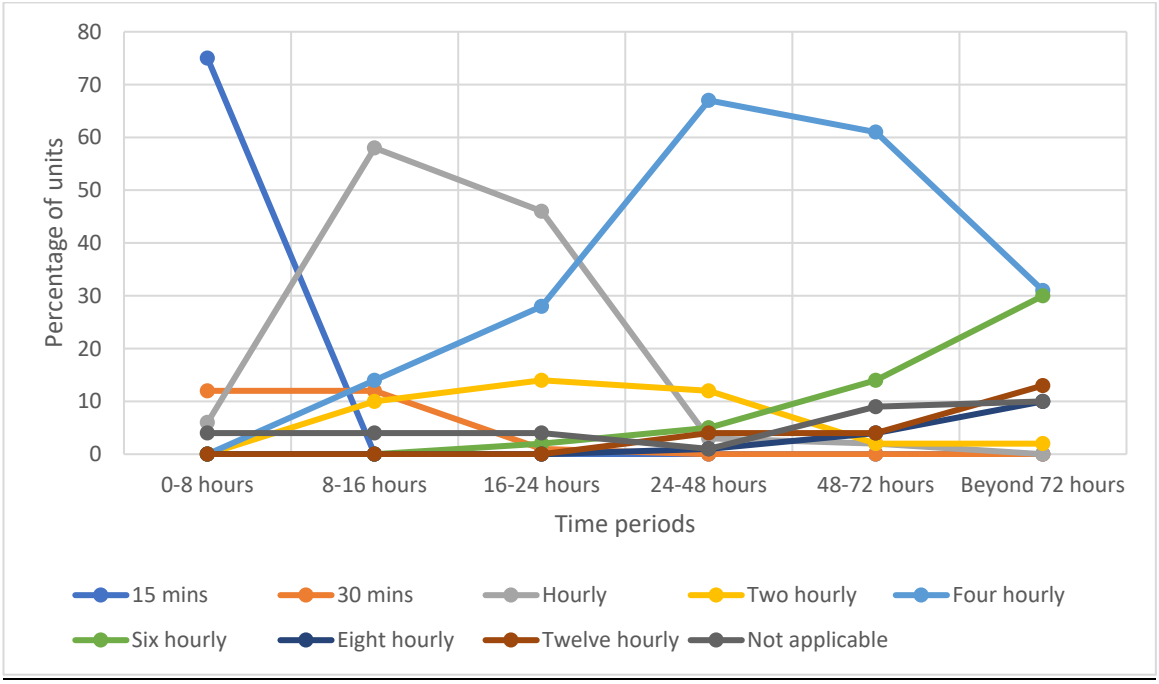
The most common or lowest reported frequencies of neurological assessment and monitoring for different patient groups across different time periods are presented visually using line graphs in Figures 5.4. to 5.8. The most common frequency for a time period, the number of units that reported for a specified time period, the number and percentage of units that reported the most common frequency by time period, and the range of frequencies used within that time period by patient group are presented in Tables 5.3 to 5.8. In order to be more representative of the data provided, the percentages of units that report the most common frequency within the tables are calculated from the number of questionnaires that provided responses for that time period, rather than the total number of questionnaires

returned. The frequency of assessment and monitoring was sometimes reported as being dependent on patient condition, but no detail on what that represented or justification about how those assessment frequency decisions were reached was provided.

Thrombolysis and Thrombectomy

The most common frequencies of neurological monitoring reported for thrombolysis patients are shown in Figure 5.5. Due to the amalgamation of the most common and the most frequent reported intervals of monitoring 75% of units reported 15-minute intervals. Despite guidelines available to support monitoring following thrombolysis this group and time period showed the greatest variability in reported frequencies across any group or time period. Twenty-two percent (n=28) responses actually reported 15 minutes as the most common frequency. However, within this group and period 68 free-text responses were reported representing 22 different frequency schedules with 66 of them starting at 15-minute intervals. This represents an overestimation of the most common frequency that is not seen in any other of the patient groups or time periods.

Figure 5.5 Most common reported frequencies of neurological assessment and monitoring by percentage of units for thrombolysis patients over different time periods.



Variation is visible in both the level of agreement on the most common frequency and the range of frequencies. For example, in Table 5.3 more units agree on the common frequency for 24 to 48 hours (74%) than they do in the beyond 72 hours period (36%).

Table 5.3 Most common frequency of neurological assessment, the number and percentage of units that reported the most common frequency based on the total number of units that provided responses, and the range of frequencies reported by time period after thrombolysis.

Time Period	Most common frequency	Total No. of units reported (time period)	No. (%) that reported the most common frequency	Range
0-8 hours	15 minutes	116	96 (83%)	15minutes to hourly
8-16 hours	Hourly	116	72 (62%)	30 minutes to 4 hourly
16-24 hours	Hourly	115	58 (50%)	30 minutes to 6 hourly
24-48 hours	4 Hourly	114	84 (73%)	Hourly to 12 hourly
48-72 hours	4 Hourly	109	77 (71%)	Hourly to 12 hourly
Beyond 72 hours	4 Hourly	107	39 (36%)	Two hourly to 12 hourly

The thrombectomy group (Figure 5.6) contains less data as 60% (n=75) of units did not manage patients after thrombectomy. The numbers of units reporting frequency of assessment and monitoring after thrombectomy increased over time as some patients would be repatriated to local units following treatment at a specialist centre (Table 5.4).

Figure 5.6 Most common reported frequencies of neurological assessment and monitoring by percentage of units for thrombectomy patients over different time periods.

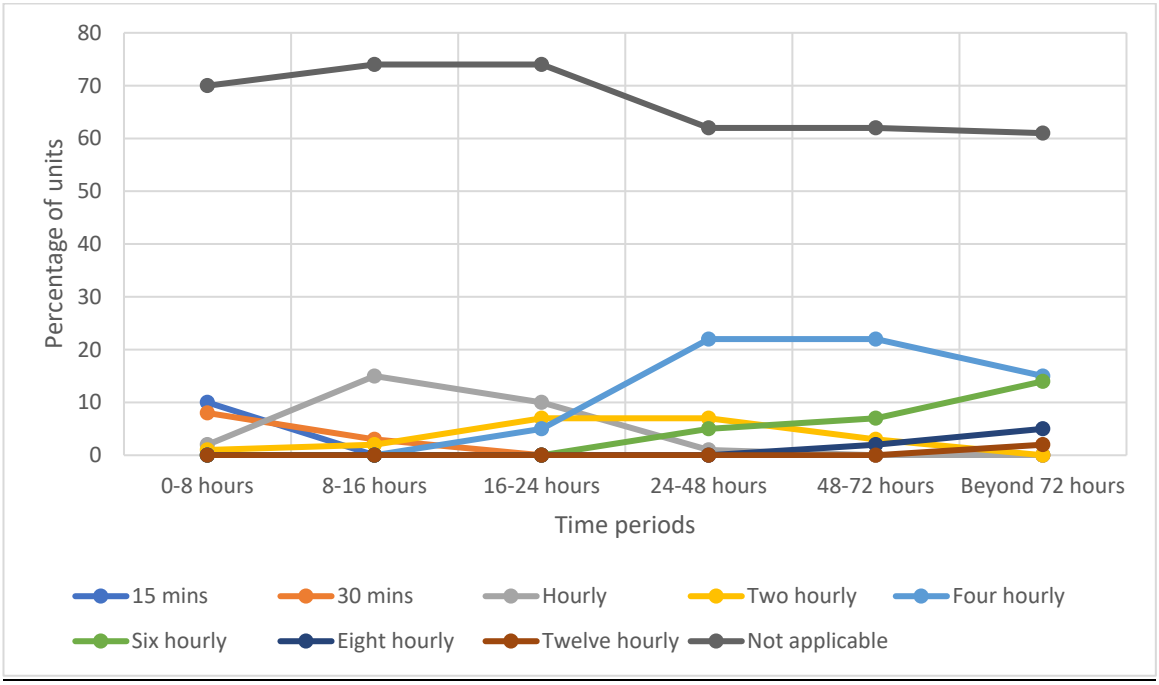


Table 5.4 Most common frequency of neurological assessment, the number and percentage of units that reported the most common frequency based on the total number of units that provided responses, and the range of frequencies reported by time period after thrombectomy.

Time Period	**Most common frequency**	**Total No. of units reported (time period)***	**No. (%) that report most common frequency**	**Range**
0-8 hours	15 minutes	25	11 (44%)	15minutes to two hourly
8-16 hours	Hourly	28	19 (68%)	30 minutes to 4 hourly
16-24 hours	Hourly	29	13 (45%)	30 minutes to 4 hourly
24-48 hours	4 Hourly	43	27 (63%)	Hourly to 6 hourly
48-72 hours	4 Hourly	43	27 (63%)	Two hourly to 8 hourly
Beyond 72 hours	4 Hourly	45	19 (42%)	Four hourly to 12 hourly

*Numbers exceed the current 24 thrombectomy units in the UK as data has also been supplied by the non-thrombectomy centres where patients await transfer or are repatriated post procedure.

There was a similar pattern of the most common frequency reported across the time periods for thrombolysis and thrombectomy (Figures 5.4 and 5.5). However, comparing Tables 5.3 and 5.4 a greater variation in assessment frequency range occurs after thrombolysis compared to thrombectomy, most notably from the 16-24 hour period onwards.

Ischaemic Stroke (without Thrombolysis or Thrombectomy)

Data for ischaemic stroke patients who do not receive thrombolysis or thrombectomy are presented in Figure 5.7 and Table 5.5. Four hourly assessment was the most common frequency across all time periods. However, the data indicates that the range of assessment frequency is greater across all time periods for this group than any other. This is evidenced by the extensive range of frequencies employed across all time periods (Table 5.5).

Figure 5.7 Most common reported frequencies of neurological assessment and monitoring by percentage of units for ischaemic stroke without thrombolysis or thrombectomy over different time periods.

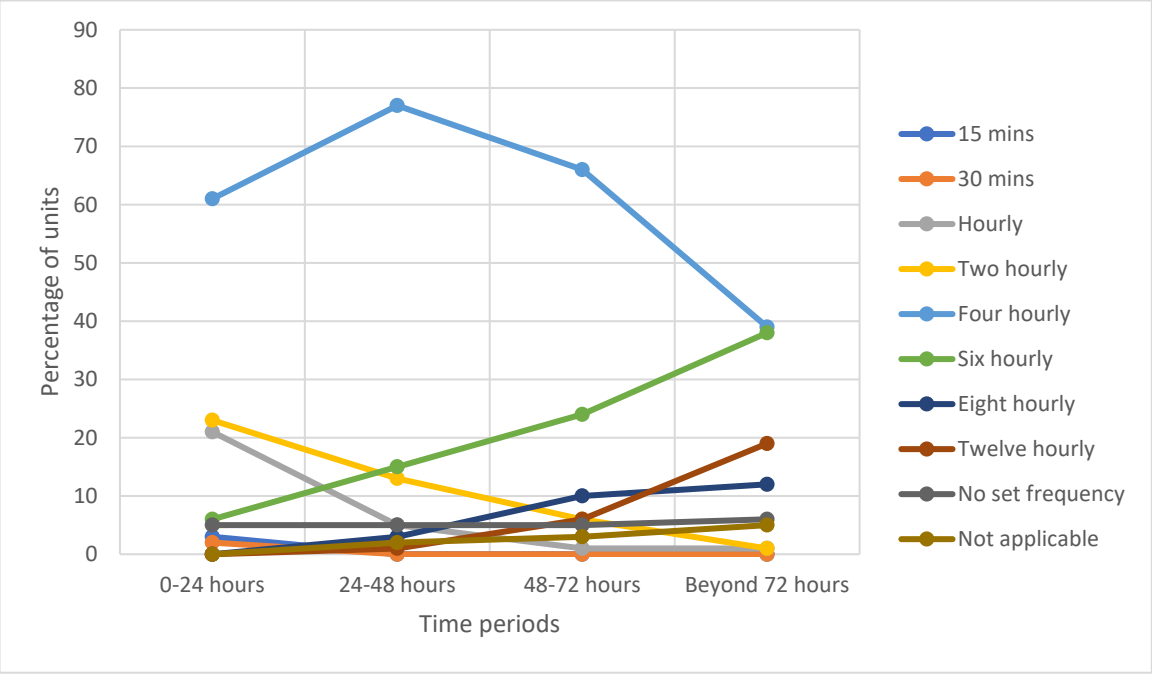


Table 5.5 Most common frequency of neurological assessment, the number and percentage of units that reported the most common frequency based on the total number of units that provided responses, and the range of frequencies reported by time period for ischaemic stroke without thrombolysis or thrombectomy.

Time Period	Most common frequency	Total No. of units reported (time period)	No. (%) that report most common frequency	Range
0-24 hours	4 Hourly	116	61 (53%)	15 minutes to 6 hourly
24-48 hours	4 Hourly	114	77 (68%)	Hourly to 12 hourly
48-72 hours	4 Hourly	113	66 (58%)	Hourly to 12 hourly
Beyond 72 hours	4 Hourly	110	39 (35%)	Hourly to 12 hourly

ICH (Intracerebral haemorrhage)

The frequencies of assessment and monitoring for ICH patients with blood pressure alteration are illustrated in Figure 5.8 and without blood pressure alteration in Figure 5.9. Initially, there appears to be some similarity in the ranges of frequencies across both groups with hourly being the most common frequency in the 0-24 hour period and four hourly beyond that for both ICH groups. However, the ranges in Tables 5.6 and 5.7 indicate that patients requiring

blood pressure alteration have greater frequency of neurological assessment and monitoring frequency and this was sustained over time. In both these patient groups, but particularly in those requiring blood pressure alteration, there were a larger number of units that neurological assessment frequency was dependent upon NEWS, BP protocol, or VitalPAC (clinical monitoring system). This indicates that the frequency of neurological monitoring was driven by physiological observations, depending upon blood pressure levels or the drug being used to manage it.

Figure 5.8 Most common reported frequencies of neurological assessment and monitoring by percentage of units for ICH with blood pressure alteration over different time periods.

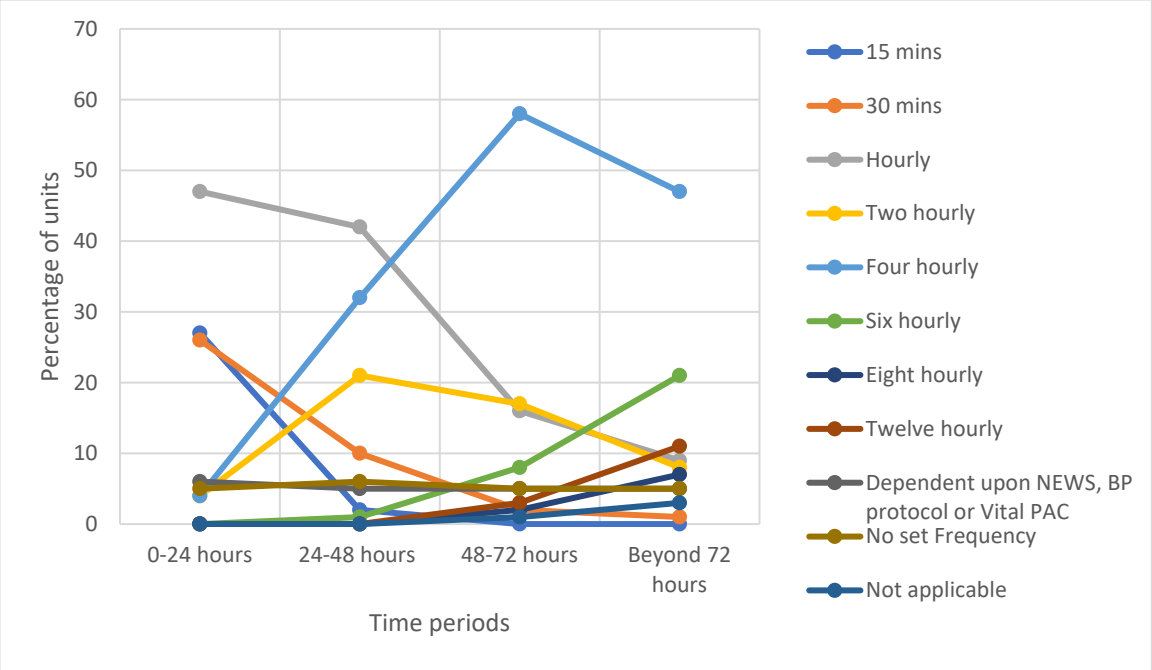


Figure 5.9 Most common reported frequencies of neurological assessment and monitoring by percentage of units for ICH without blood pressure alteration over different time periods

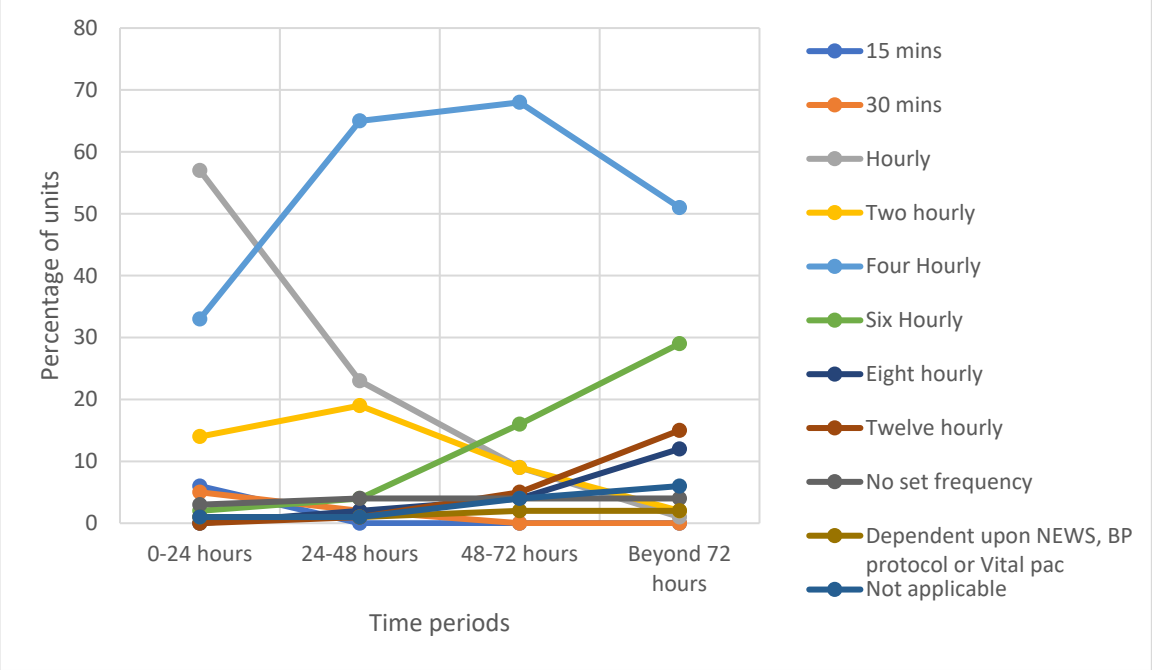


Table 5.6 Most common frequency of neurological assessment, the number and percentage of units that reported the most common frequency based on the total number of units that provided responses, and the range of frequencies reported by time period for ICH with blood pressure alteration.

Time Period	**Most common frequency**	**Total No. of units reported (time period)**	**No. (%) that report most common frequency**	**Range**
0-24 hours	Hourly	108	47 (44%)	15minutes to 4 hourly
24-48 hours	Hourly	108	42 (39%)	15minutes to 6 hourly
48-72 hours	4 Hourly	106	58 (55%)	30 minutes to 12 hourly
Beyond 72 hours	4 Hourly	104	47 (45%)	30 minutes to 12 hourly

Table 5.7 Most common frequency of neurological assessment, the number and percentage of units that reported the most common frequency based on the total number of units that provided responses, and the range of frequencies reported by time period for ICH without blood pressure alteration.

Time Period	**Most common frequency**	**Total No. of units reported (time period)**	**No. (%) that report most common frequency**	**Range**
0-24 hours	Hourly	117	57 (49%)	15 minutes to 6 hourly
24-48 hours	4 Hourly	116	65 (56%)	30 minutes to 12 hourly
48-72 hours	4 Hourly	111	68 (61%)	Hourly to 12 hourly
Beyond 72 hours	4 Hourly	110	51 (46%)	Hourly to 12 hourly

Potential hemicraniectomy

For patients potentially eligible for hemicraniectomy the data is shown in Figure 5.10. and Table 5.8. This group of patients, like thrombectomy, are not managed at all stroke units and may be sent to specialised or higher acuity units within the same or different hospital. Although there is variation in the range of frequencies there was more consistency for this patient group, the most common frequency remained regular and higher over a longer period than for any other group.

Figure 5.10 Most common reported frequencies of neurological assessment and monitoring by percentage of units for patients potentially eligible for hemicraniectomy over different time periods.

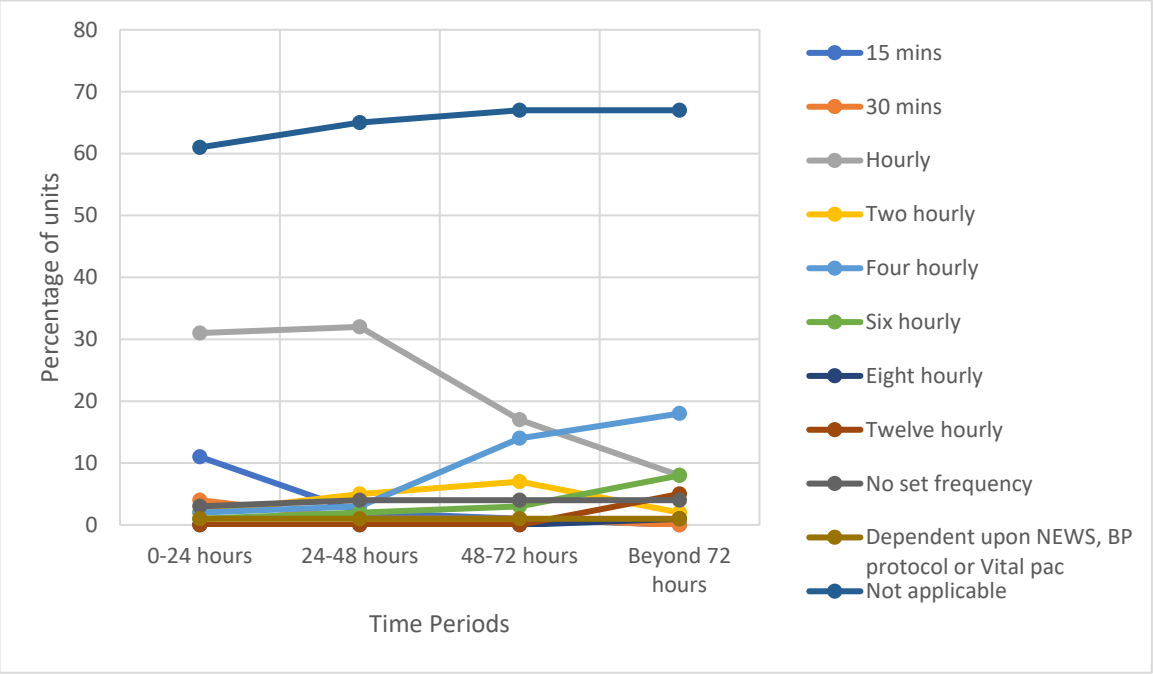


Table 5.8. Most common frequency of neurological assessment, the number and percentage of units that reported the most common frequency based on the total number of units that provided responses, and the range of frequencies reported by time period for patients potentially eligible for hemicraniectomy.

Time Period	Most common frequency	Total No. of units reported (time period)	No. (%) that report most common frequency	Range
0-24 hours	Hourly	51	31 (61%)	15 minutes to 6 hourly
24-48 hours	Hourly	45	32 (71%)	15 minutes to 6 hourly
48-72 hours	Hourly	48	17 (35%)	15 minutes to 6 hourly
Beyond 72 hours	4 Hourly	48	18 (38%)	Hourly to 12 hourly

Other patient groups

For the other patient groups, the data reported was limited and inconsistent. Ranges in frequency varied from hourly to 12 hourly with the most common frequency being 4 hourly up to 72 hours and 6 hourly after that.

5.4.4. Discontinuation of neurological assessment and monitoring

The number of units reporting frequencies generally reduces over time except following thrombectomy. This reduction likely represents monitoring being stopped at certain time periods. Overall, discontinuation data were poorly reported, only a small number of questionnaires provided information (Table 5.9). Most units continue with neurological monitoring beyond 72 hours with no set endpoint reported.

Table 5.9 Discontinuation of neurological assessment and monitoring presented chronologically within patient groups.

Patient Group	Discontinuation of Neurological Assessment and Monitoring	Number of Responses
Thrombolysis	Discontinued at 24 hours unless clinically indicated	1
	No set frequency after 24 hours, depends on patient's condition	2
	Discontinued at 48 hours	3
	Only completed after 48 hours if requested by a doctor	1
Thrombectomy	Discontinued at 24 hours unless clinically indicated	1
Ischaemic stroke (without thrombolysis or thrombectomy	Discontinued at 24 hours unless clinically indicated, depends on patient's condition	1
	No protocol beyond 48 hours	1
	Discontinued at 72 hours	1
	Reduce to twice a day (BD) if stable after 10 days.	1
ICH (with blood pressure alteration)	Discontinued after 48hrs unless clinically indicated	1
ICH (without blood pressure alteration)	No protocol beyond 48 hours	1
	Discontinued after 48 hours unless clinically indicated	1
Potential hemicraniectomy	Discontinued after 48 hours	1

ICH= Intracerebral haemorrhage

5.4.5. Missed neurological assessment and monitoring

Participants were asked about points of the patient pathway or times when neurological assessment or monitoring was more likely to be missed. Of the 110 respondents that completed this free-text entry question, 29 % (n=32) stated that neurological assessment and monitoring was not missed, the remaining 71% (n=78) questionnaires provided 148 separate responses which were grouped. The most common was classified as busy periods (n=53) and included specific times such as ward rounds (n=7), patient acuity (n=7), medicine rounds (n=6), handover period (n=6), emergency situations (n=4), mealtimes (n-=4), visiting times (n=3), and other multiple admissions (n=2).

The next most common were overnight or out of hours (n=29), when the unit is short-staffed or does not have stroke-trained staff (n=27), and when the patient is off the unit for investigations or therapy (n=18). Other reasons included patients not being within the stroke unit, either whilst they were in the emergency department, missed the pathway or were transferred to other hospitals (n=7). Monitoring could be missed due to frequency escalation or de-escalation caused by a change in patient condition (n=5), or where individualised frequencies and durations depend upon more than the presenting complaint (n=2). Scale choice, reason for use, and timing impact completion, AVPU or GCS were less likely to be missed, as completed regularly, as opposed to standalone post thrombolysis NIHSS assessments (n=1). Neurological assessment and monitoring may also not be completed on electronic systems (n=1), be missed due to lack of equipment (n=1), or not completed in patients identified as end of life (n=1). Three questionnaires mentioned using audits and senior staff overview to try and increase adherence to protocols and ensure completion.

5.4.6. Neurological Deterioration

Respondents were first asked about what it is that they would observe in a patient that would make them aware they had deteriorated. Multiple free text answers were possible and a total of 410 responses were received from 121 questionnaires. Responses were grouped into five categories of deterioration recognition:

- Changes in relation to assessments or scales (160 responses)

This was the most common theme and included: change in neurological assessment scores, specific scales, or across neurological examination or clinical observation, as well as new or worsening symptoms. The most common response was a change or reduction, in the GCS (n=71) indicating 57% of units use this as a marker the patient has deteriorated. Changes in neurological assessments generally (n=30) and specific scales: NIHSS (n=31) SNOBS (n=2) FAST (n=2) and CNS (n=1) were reported. Two respondents stated they did not use an exact score as a marker with one stating '*close observation relies upon the skill of knowing when a neurological change has occurred even without any change in score'* (Nurse Unit Manager, Large Tertiary Hospital).

- Changes noted in patient condition (106 responses)

The most common change in condition reported was linked to level of consciousness (LOC): altered or reduced LOC, increased drowsiness, and loss of alertness or responsiveness (n=41).

Others included unspecified change in presentation or patient condition (n=32) or change in function or level of dependence/physical condition (n=19) which included: changes in Speech/ Communication, Pupil Response, Muscle Tone, and Swallow. The final change reported was in behaviour (n=14) with eight specifically reporting patients becoming irritable or agitated as a sign of deterioration.

- Alterations in physiological observations (90 responses)

Change in the NEWS or physiological readings generally were cited by 67% of respondents (n=84) which included: change in blood pressure, pulse or heart rate, altered breathing or respiration rate, and oxygen saturation. Other responses included Airway, Breathing, Circulation, and Disability (ABCD) assessment (n=2), low urine output (n=2), and high or low BM (n=2).

- Specific symptoms (36 responses)

The three most common specific symptoms reported as indicating deterioration were confusion (n=11), headache (n=10), and vomiting or nausea (n=10). Others were seizures (n=2), lethargy (n=1), sweating (n=1), and reporting feeling different (n=1).

- Miscellaneous (18 responses)

The most common responses were concerned with having a gut feeling or intuition or knowing from experience that something is wrong with the patient (n=11). Changes being noted by others, either other members of the multi-disciplinary team (MDT) (n=3) or another observer/family member (n=2) were also mentioned.

Respondents were then asked about the actions that would be taken if deterioration was noted. Reports of actions are presented in descending order in Table 5.10. Escalation for medical review is the primary action in response to deterioration, reported by 99% of questionnaires. However, it is not possible to know from the data the order of actions and whether some actions are completed before the medical review or whether the review instigates the other actions.

Table 5.10 Actions reported completed by number and percentages of units when neurological deterioration is noted in an acute stroke patient.

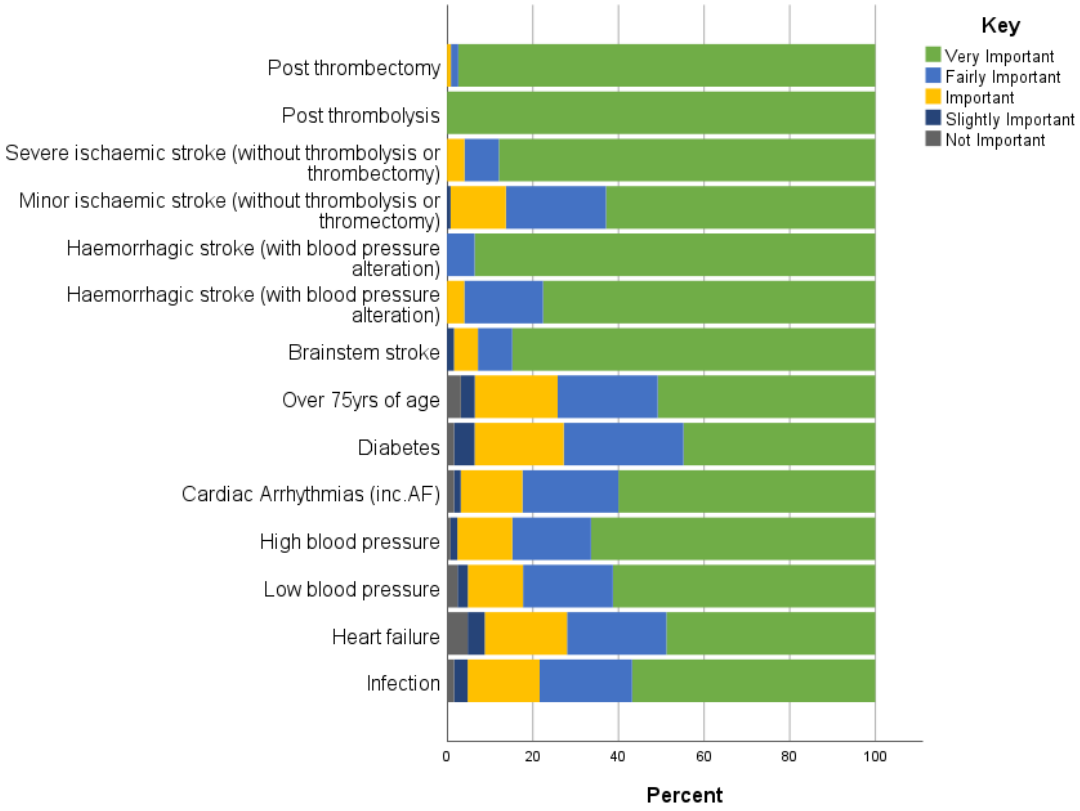
Action	**No. (%) of units who report completing the action**
Medical review	124 (99%)
Additional scan	119 (95%)
Additional observations	113 (90%)
Inform senior nurse	111 (89%)
Treatment to alter blood pressure	104 (83%)
Glycaemic control	92 (74%)
Neuro-surgical review	86 (69%)
Other	16 (13%)

5.4.7 Experience

This section first reports data on the perceived importance of, agreement with, confidence in, and general satisfaction with neurological assessment and monitoring in clinical practice. Second, perceived barriers and facilitators to providing this element of care are presented before concluding with respondents' opinions of whether change is needed and if so potential barriers to achieving it.

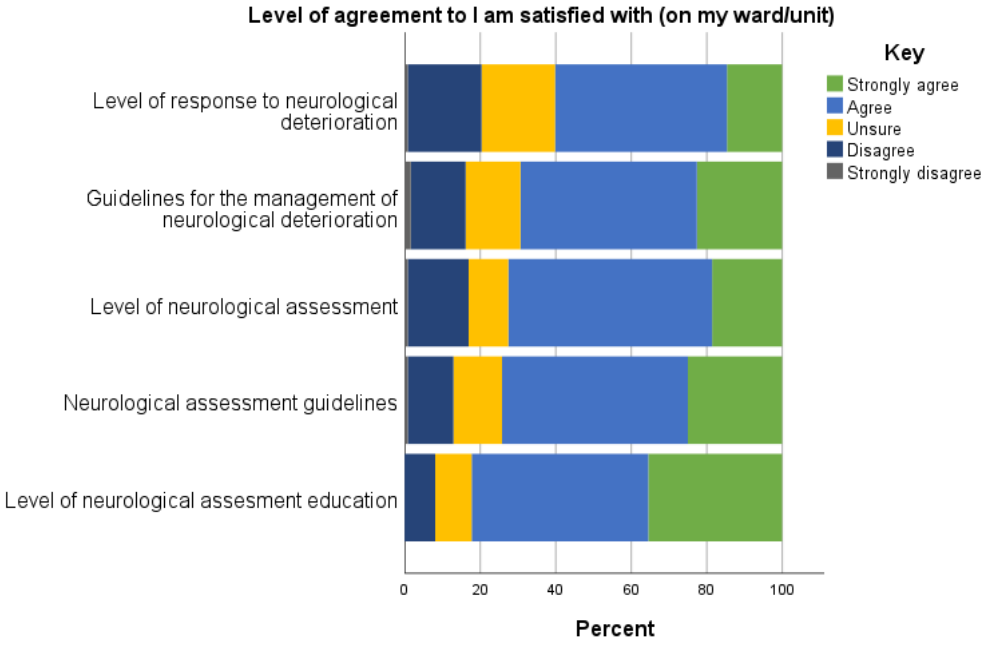
The perceived importance of monitoring based on certain patient characteristics or co-morbidities is presented in Figure 5.11. Although neurological monitoring is generally perceived as important across all groups it is influenced by treatments received, stroke type, and severity. Monitoring in patients who receive thrombolysis, thrombectomy, and blood pressure alteration after ICH is perceived to be very important by more respondents than any other group. More respondents agreed that it is very important to monitor brainstem stroke than ICH (without blood pressure alteration) and ischemic stroke (without thrombolysis or thrombectomy). Severity seems to affect the perceived importance of neurological monitoring with respondents reporting more importance for severe ischemic stroke than minor ischaemic stroke (without thrombolysis or thrombectomy). Only very small numbers reported any of the characteristics or co-morbidities as not being important to monitor neurologically. However, having heart failure and being over 75 years of age seemed least important.

Figure 5.11 Perceived importance of neurological monitoring for stroke patients based on characteristics or co-morbidities.



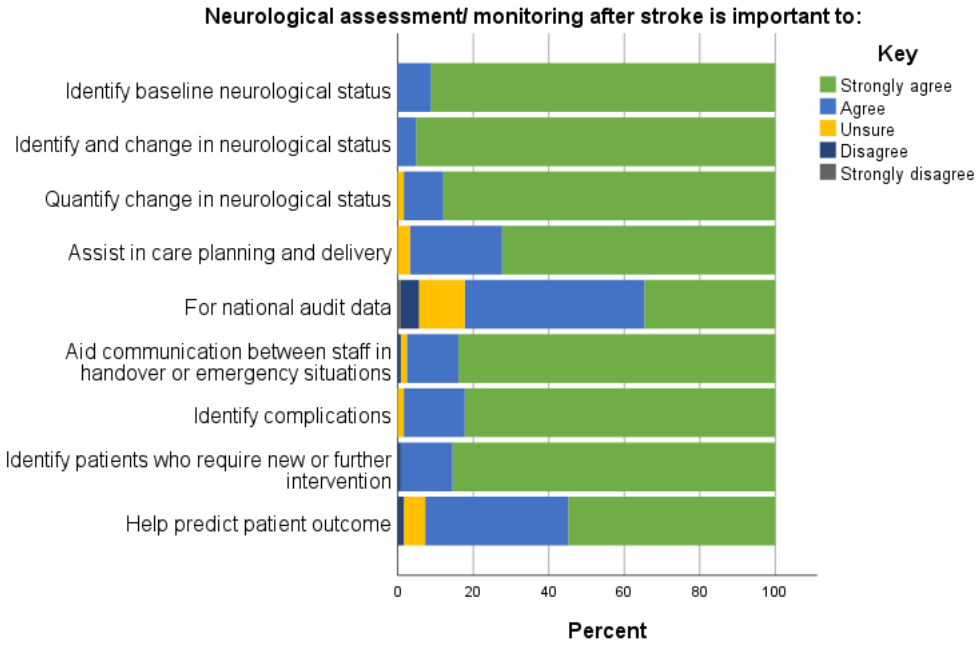
General satisfaction with education, guidelines, and levels of neurological assessment and response to deterioration is shown in Figure 5.12. Thirty-nine percent of respondents (n=49) disagree or are unsure they are satisfied with the level of neurological assessment education provided in their unit. More respondents were satisfied with the levels of neurological assessment and response to deterioration than they were with the guidelines to support these elements. Twenty-seven percent of respondents (n=34) agreed or strongly agreed that neurological assessment and monitoring is a neglected area of practice.

Figure 5.12 Respondents' satisfaction with aspects of neurological assessment and monitoring in their units.



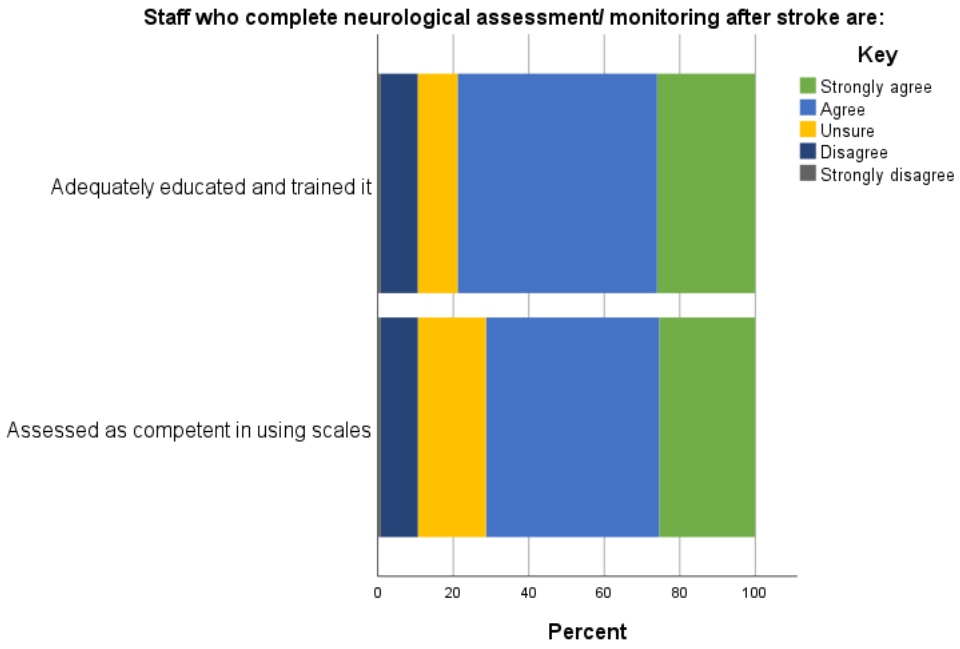
The responses to statements about why neurological assessment and monitoring is important are shown in Figure 5.13. Neurological assessment and monitoring is seen as valuable to many aspects of patient care but the aspect identified as most important was for the identification of change in neurological status. National audit data was deemed the least important reason.

Figure 5.13 Responses to statements around what neurological assessment and monitoring is important for.



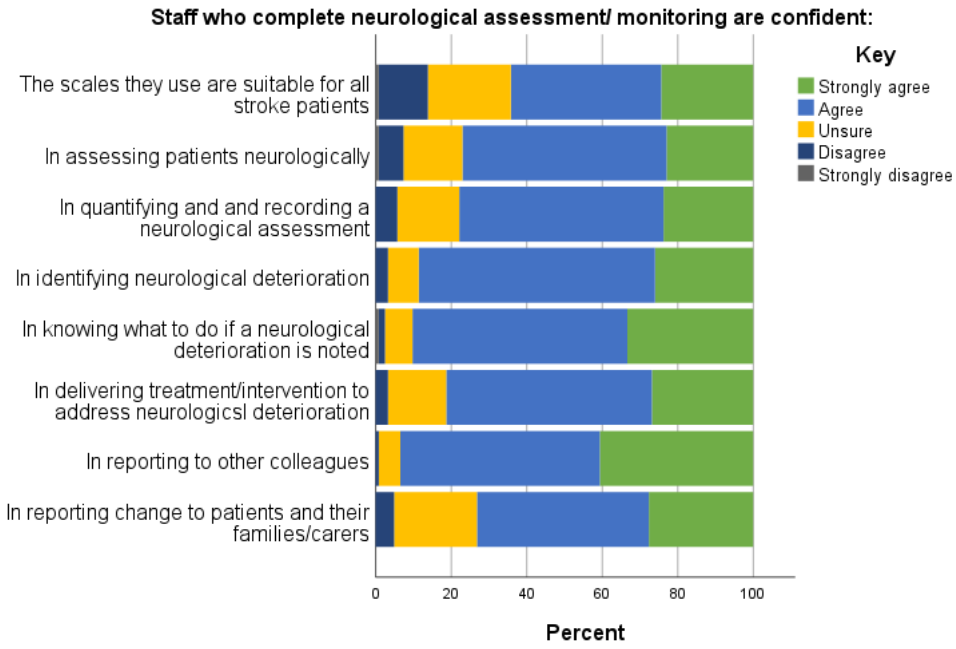
Levels of agreement about whether staff are adequately trained and competent in neurological assessment and monitoring are reported in Figure 5.14. The majority of respondents felt that staff were competent in using scales and adequately educated and trained in neurological assessment and monitoring. However, the responses showed a slight tendency to more certainty that staff were adequately educated than competent in the use of scales.

Figure 5.14 Respondents' views on staff education and competency in the completion of neurological assessment and monitoring.



Responses around statements of staff confidence in elements of completing neurological assessment and monitoring are reported in Figure 5.15. Disagreement with the statements was uncommon. However, the respondents reported being most confident in staff reporting change to other colleagues. Least confidence was conveyed in the suitability of scales for use with all stroke patients. Generally, staff are thought to be more confident in identifying deterioration than in performing neurological assessments.

Figure 5.15 Responses to statements around staff confidence in elements of completing neurological assessment and monitoring.



5.4.8. Barriers and Facilitators

Respondents were provided with 22 statements representing potential barriers and facilitators and were asked about their level of agreement with them. To aid interpretation the results are presented by theme:

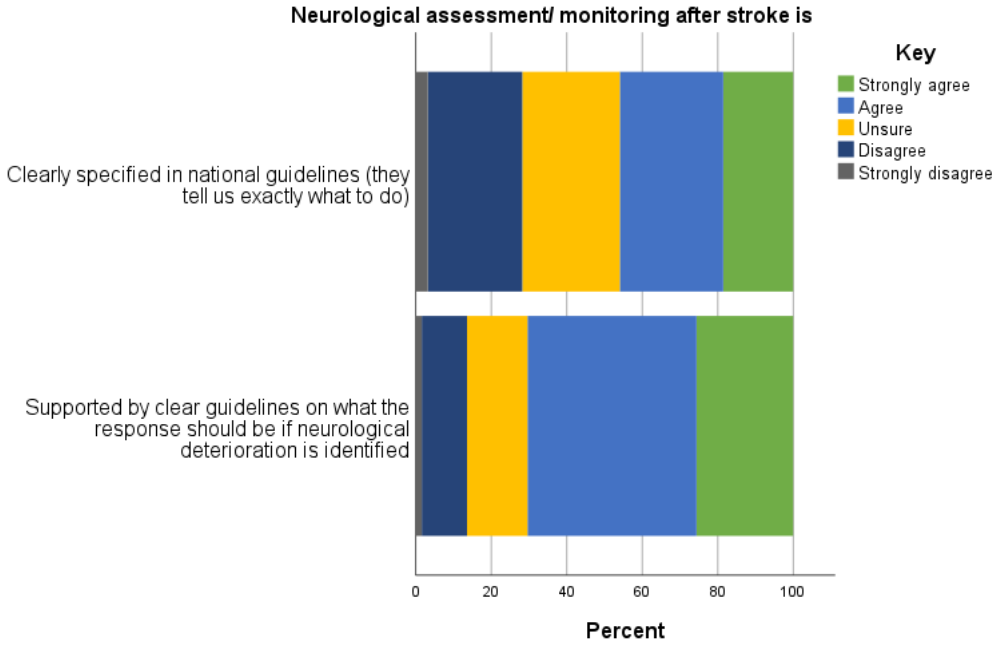
- Clear guidelines
- Importance for all stroke patients (should it be tailored to certain groups or individual patients)
- Importance in relation to change and deterioration
- Clinician and team factors
- Time
- Patient and family impact

- Completion and documentation factors.

5.4.8.1 Clear guidelines

Responses related to guidelines are shown in Figure 5.16. Fifty- four percent of respondents (n=68) disagreed or were unsure that national guidelines clearly specify what to do in terms of neurological assessment and monitoring. More respondents agreed or strongly agreed that there was better guidance on how to respond to deterioration that what to do in terms of monitoring to identify the deterioration potentially indicating clearer guidelines are warranted.

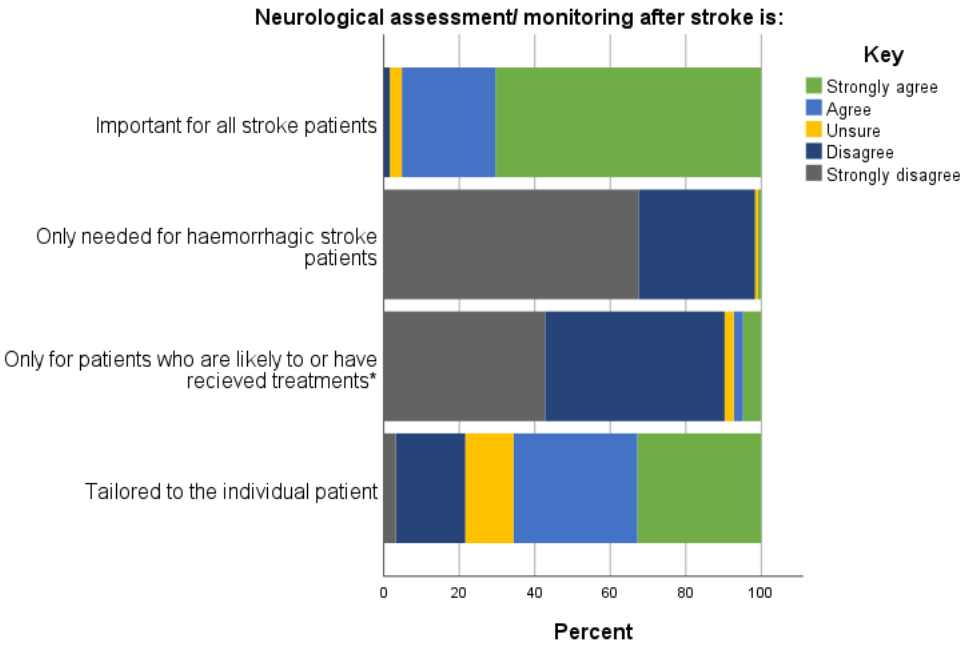
Figure 5.16 Responses to statements around whether clear guidelines exist in relation to neurological assessment/monitoring and response to neurological deterioration



5.4.8.2. Importance of neurological assessment and monitoring for all patients.

The responses around the importance of neurological monitoring for all stroke patients or whether it should be tailored to the individual, or groups of patients are shown in Figure 5.17. Most respondents, 95% (n=119), feel that neurological assessment and monitoring is important for all stroke patients. This was supported by large numbers disagreeing that it should only be for patients who are likely to or have received treatments (90%, n=112), or in ICH (98%, n=121). Tailoring to individual patients seems generally supported with 66% (n=82) agreeing and a further 13 % (n=16) unsure.

Figure 5.17 Importance of neurological assessment and monitoring for all stroke patients and whether it should be tailored to the individual patient.

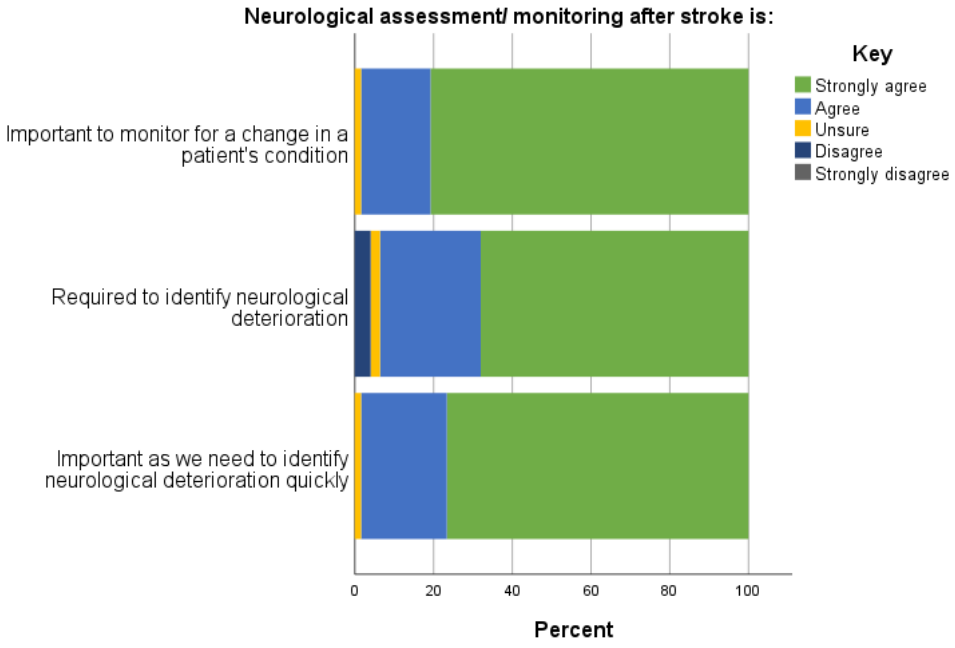


*e.g. thrombolysis, thrombectomy, or neurosurgical intervention

5.4.8.2. Importance of neurological assessment and monitoring to detect change and deterioration.

The responses about the importance of neurological assessment and monitoring to detect change, identify deterioration, and speed of identification are presented in Figure 5.18. Nearly all respondents agreed, with only less than 2% unsure (n=2), that it is important to monitor for a change in a patient's condition and identify neurological deterioration quickly.

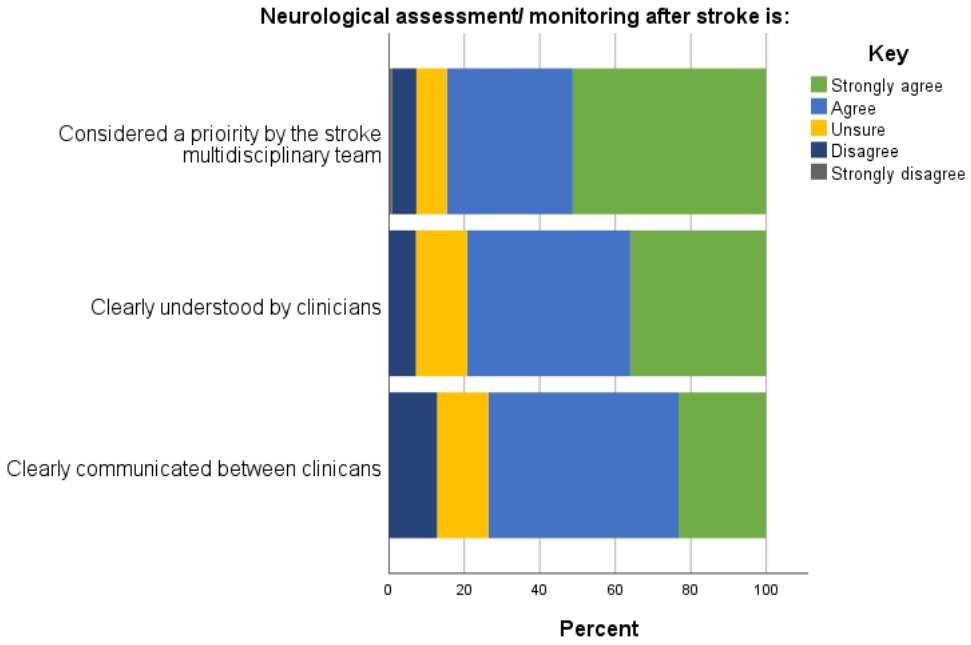
Figure 5.18 Responses about whether neurological assessment and monitoring is important to detect change and deterioration.



5.4.8.3. Clinician and team factors

Clinician and team factors included whether neurological assessment and monitoring is considered a priority by the stroke MDT and whether clinicians clearly understand and communicate it (Figure 5.19). Fifteen percent (n=19) were unsure or disagreed that neurological assessment and monitoring is considered a priority by the MDT despite the previously reported importance. Seventy-nine percent (n=99) agreed that it is clearly understood and 74 % (n=92) (agreed it is clearly communicated by clinicians.

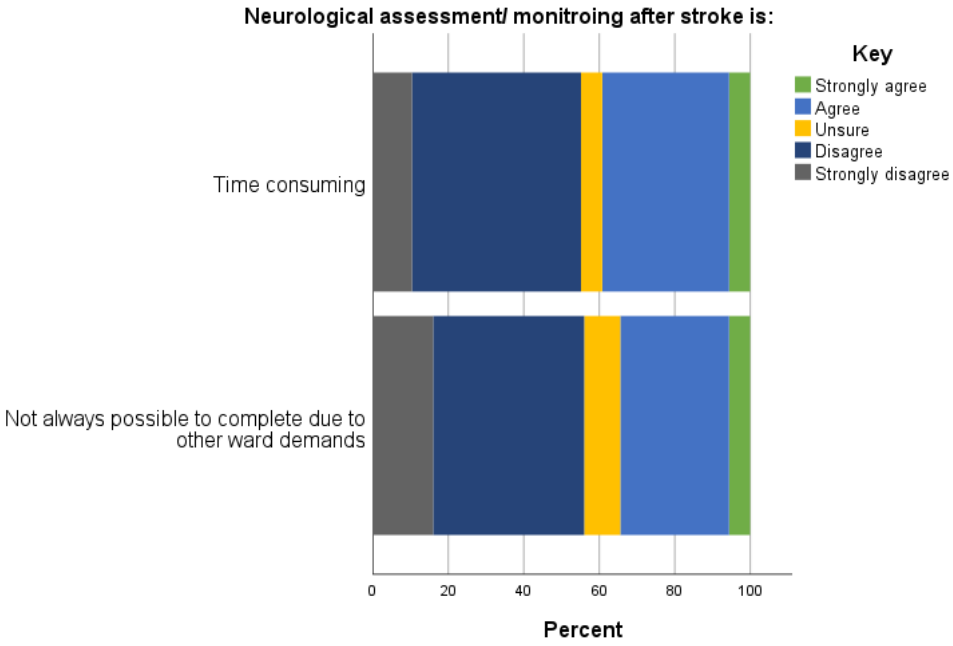
Figure 5.19 Respondents' views on whether clinician and team factors impact on neurological assessment and monitoring.



5.4.8.4. Time to complete neurological assessment and monitoring

Time factors responses (Figure 5.20) showed that 61% (n=76) disagreed or were unsure that neurological assessment and monitoring is time-consuming and 66% (n=82) disagreed or were unsure that completion was not always possible due to other ward demands.

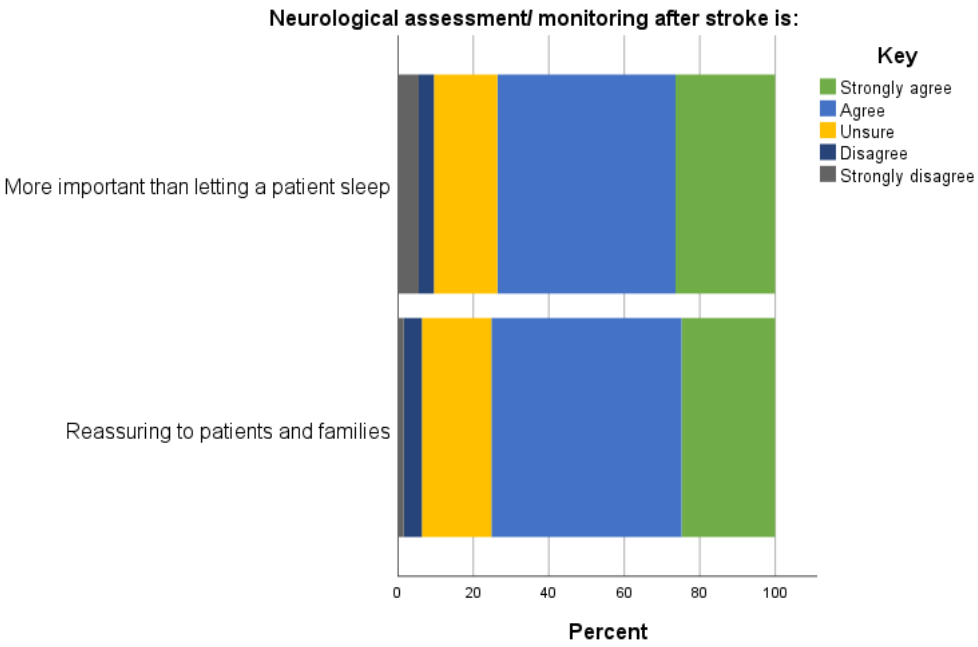
Figure 5.20 Respondents' responses to time impacting on the completion of neurological assessment and monitoring.



5.4.8.6. Patient and family impact

Patient and family impact results are shown in Figure 5.21. Ten percent (n=12) disagreed and 17 % (n=21) were unsure that neurological assessment and monitoring is more important than letting a patient sleep. Seventy five percent (n=94) agree that assessment and monitoring is reassuring to patients and families.

Figure 5.21 Responses to whether patient and family factors impact on neurological assessment and monitoring.

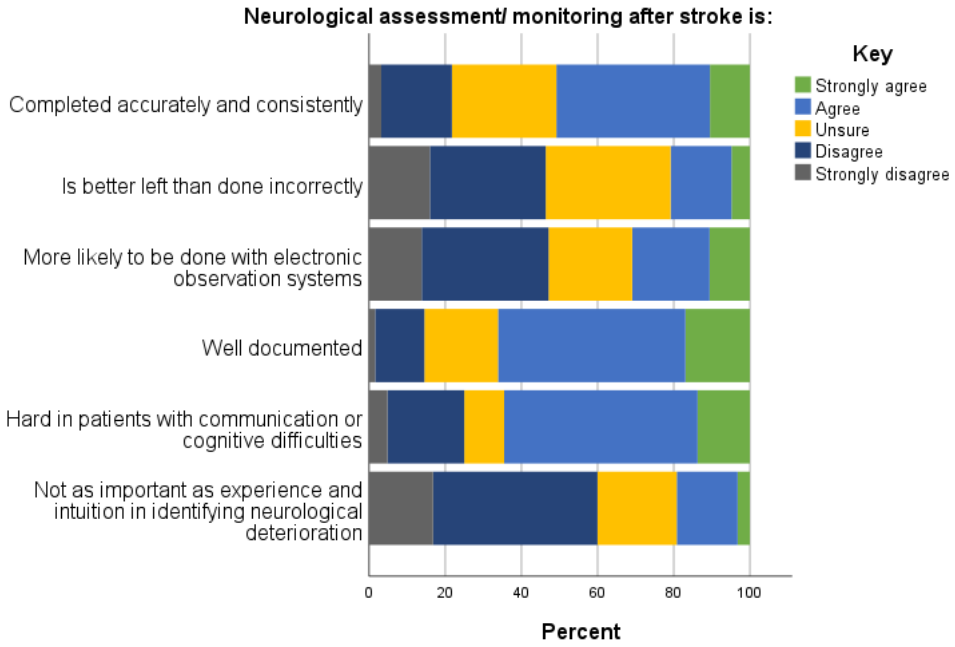


5.4.8.7. Completion and documentation factors.

The completion and documentation theme investigated practical issues that would need to be addressed if neurological assessment and monitoring were to change. Results are displayed in Figure 5.22. In terms of completion and documentation of neurological assessment and monitoring, half of the respondents (n=63) agreed that it is completed accurately and consistently, whereas 66% (n=83) agreed it is well documented. Just under half, 46% (n=58), disagreed that it is better to leave neurological assessment and monitoring than to do it incorrectly.

There was also disagreement that the use of electronic observation systems would improve compliance, with more disagreeing, 47% (n= 58), than agreeing, 31% (n=38), leaving 22% (n=27) unsure. Previous responses in the survey were validated when 64% (n=80) agreed that neurological assessment and monitoring is difficult to complete in patients with communication or cognitive difficulties. Nineteen percent (n=24) felt that experience and intuition are more important than formal neurological assessment and monitoring in identifying deterioration with 22% (n=28) unsure.

Figure 5.22 Respondents' views on completion and documentation factors around neurological assessment and monitoring.



5.4.8.8. Free-Text Responses

Thirty-one questionnaires provided multiple free-text responses around barriers around facilitators which were analysed and grouped into themes. More barriers (Table 5.11) were identified than facilitators (Table 5.12). The overall response rate to this section was low, however many responses reiterated other aspects that had been suggested or explored elsewhere in the survey, providing additional validation and in some cases more detail.

Table 5.11 Free-text responses for barriers to providing neurological assessment and monitoring

Barrier Theme (no. of comments)	**Detail within theme (no. of comments)**
Staffing (25)	Use of bank or agency staff and locum doctors (9)
	Inexperienced or new staff (e.g. newly qualified, new to stroke, overseas nurses) (8)
	Poor staffing levels (7)
	Limited nursing time prevents detailed and timely assessment (1)
Training (18)	Lack of training (8) especially around understanding how, when and why changes might occur in the patient 's condition
	Time for training due to staff turnover or being unable to get released from the ward (3)
	Staff not properly trained in the scales (GCS & NIHSS) (4)
	Assessment is subjective (1)
	Completion by less trained staff e.g., healthcare assistants (1)
	Lack of confidence (1)
Patient Factors (11)	Language (4)
	Cognitive issues (3)
	Deafness (2)
	No glasses (1)
	Ptosis (1)
Documentation (6)	Emergency department staff not documenting assessments (2) so baseline not available
	Use of electronic versus paper-based systems (2)
	Limb weakness was poorly understood or documented (1)
	Nurses follow previous documentation rather than completing a thorough assessment (1)
Ward Issues (5)	Aspects of ward dynamics such as patient numbers and acuity, turnover, delayed discharge, and visiting time (3)
	Lack of resources (1)
	Winter or all year pressures (1)
Scales (4)	Multitude of scales available (2)
	Limitations of the GCS and whether it is appropriate in stroke (2)
Lack of guidance (4)	The lack of clear and complete guidelines especially across all stroke types (3)
	Lack of evidence base to guide test-treatment pathways where in the event of deterioration there is little that can be done (1)
External Teams (2)	The lack of understanding of importance by other teams
	Outliers in other wards do not get neurological assessment and monitoring (only arrival to stroke unit)

There were multiple barriers identified within several themes despite the lack of completion by many respondents. The complexity of both neurological assessment and monitoring, the

systems in which it is delivered, and the factors that impact on it can be seen throughout these responses. The interrelatedness of barriers should also be acknowledged. Although staffing is the theme with the most responses it interrelates with many of the other themes including training and ward issues.

In Table 5.12 the suggested facilitators to neurological assessment and monitoring practice are presented. These facilitators are often answers to the barriers above in Table 5.11. Although very small numbers of responses there appears to be a desire for the SNOBSS to allow some tailoring, whether that is to patient groups or individual circumstances however is not clear.

Table 5.12 Free-text responses for facilitators to providing neurological assessment and monitoring

Facilitator Theme (no. of comments)	**Detail within theme (no. of comments)**
Education and competency (7)	Agreed National training (2)
	Use of competency assessments (2)
	Consistently trained staff with confidence (1)
	Practice as a team (so the process becomes commonplace) (1)
	Ability to contextualise the findings (1)
Tailored Approach (2)	Standardised evidence-based approach which allows for individualisation (1)
	Tailored approach rather than prescriptive monitoring (1)
Other (5)	Using shortened versions of a medical neurological examination (1)
	Experience/intuition (1)
	Knowing the patient so aware of trends (1)
	Support (1)
	Therapy intervention (assumed to mean included in therapy sessions) (1)

5.4.8.9 Change to neurological assessment and monitoring

The majority of respondents, 71% (n=89), thought that changes were needed in neurological assessment and monitoring. In terms of what they would change 86 questionnaires provided 152 free-text comments. The results organised under the five themes identified: Assessments, Guidelines, Training, Documentation, and Staffing (Table 5.13).

Table 5.13 Themes, sub-themes, justification of, and specific suggestions of changes from respondents that felt change was warranted in neurological assessment and monitoring.

Theme (n, %)	Sub- Theme (n, %)	Justification/explanations (n)	Specific suggestions for change (n)
Assessments (66, 74%)	GCS not appropriate or NIHSS more suitable (23, 26%)	Aphasia is lost in the assessment (4)	Adaptation of the GCS for aphasic and dysarthric patients (4)
		GCS was designed for traumatic brain injury (2)	Adaptation of the FAST tool using elements of the NIHSS to detect changes quicker (1)
		Communication difficulties (1) and confusion (1) can cause incorrect assessment	Using something like the STOC chart (1)
	Stroke specific assessment (not GCS or AVPU) that is up to date and validated in stroke (21, 24%)	Important to have a neurological baseline (1)	Monitoring needs to be considered in terms of - what do the team need to know and how this will change treatment (1)
	Scale Use (14, 16%)	-	Individualised (8)
			Easier to use (2)
			NIHSS too difficult (2)
			Increased frequency post thrombolysis is warranted (1)
			Less reliance on scoring and more on clinical judgement in conjunction with assessment (1)
	Specific Patient Groups (3, 3%)	NIHSS not suitable for posterior circulation (1)	TIA patients should receive neurological monitoring (1)
Guidelines (45, 51%)	Clearer guidelines and standards (27, 30%)	Ensure standard practice and equity of care (17)	What we should be doing for which patients including detail on frequency and duration (15)
		"the motivation for testing at the moment seems to be historical rather than based on any evidence" (1)	Across local, regional, and national arenas (7)
Training (29, 32%)	Need to have clear and standardised training (18, 20%)	-	Improved training in specific scales: NIHSS (6) GCS (3)
			Further training in language and cognition as they are hard to assess (2)
			All staff (2)
			Competency assessments (1)
			Multi-disciplinary team approach with shared learning (1)

Theme (n, %)	Sub- Theme (n, %)	Justification/explanations (n)	Specific suggestions for change (n)
			Addressing confidence and skills (1)
			Delivered Regularly (1)
Documentation (8, 9%)	Improvements in electronic observation and record keeping systems (5, 4%)	-	Improvements in electronic patient observation systems to provide a more in-depth neurological assessment (4)
			Clear documentation of neurological care plans by doctors (3)
			System that would immediately let clinicians know when there was a change (1)
Staffing (4, 4%)	Need for additional staff and resources 3 (3%)	-	Allowing continuity of care so subtle changes can be identified (1)

5.4.9. Training

The questionnaire aimed to explore the stroke-specific neurological assessment and management training that staff undertake. The aim was to provide an overview of courses provided including which staff groups received them, how it is delivered, and whether it is registered on the Stroke Specific Education Framework (SSEF). However, there was little consistency in the responses with some extensively completed with details of all training on the units, and others very little. A wide variation of courses were reported. After condensing similar courses together 113 separate course titles remained. Many were single centre locally delivered courses that without content detail were not comparable (e.g., short-stroke course, or in-house stroke induction training), and some were unrelated to neurological assessment and monitoring (e.g. Bobath, Advanced Life Support).

Details around specific scale training for the NIHSS, GCS, the National Early Warning Score (NEWS), AVPU (where this is reported separately to the NEWS), and SNOBS were extracted. Other courses may have included aspects of scale training, but this could not be ascertained from the data. Table 5.14 reports on the units that provide scale training and the professional groups included. Nurses are the group most likely to be trained in all scales, except the NIHSS, in which doctors and specialist nurses were more likely to be trained. It appears that training in

the SNOBS is only provided to nurses, but this may be misleading given the small amount of data.

Table 5.14 Numbers of units that reported receiving scale-specific training and percentages to show which professional staff groups receive that training.

Training in scale	**No of units**	**Percentage of units that report training for each professional group**								
		D	**PA**	**Nur**	**Spec Nur**	**SALT**	**OT**	**PT**	**HCA**	**Other**
NIHSS	69	78	36	65	75	7	6	10	4	4
GCS	8	63	50	100	63	25	13	25	25	13
NEWS*	6	40	40	100	100	20	20	20	80	20
AVPU	2	50	0	100	0	50	50	50	50	0
SNOBS	2	0	0	100	0	0	0	0	0	0

*Single response to professional groups trained was missing so percentages based on data from 5 units

D- Doctors

PA- Physician Assistants

Nur- Nurses

Spec Nur-Specialist Nurses

SALT- Speech and Language Therapists

OT-Occupational Therapists

PT- Physiotherapists

HCA- Healthcare Assistants

Information provided on course type, course format, and whether the courses are registered on the SSEF is displayed in Table 5.15. Where training was provided externally, who provided it was poorly reported e.g., for the NIHSS less than half provided responses (24/58). From the data, it appears most NIHSS training is online hosted via a variety of websites although some training is reported to be delivered internally face-to-face. For all other scales, there is greater tendency for internal courses. However, online delivery was prominent for the GCS and NEWS and in some cases, a mixture of formats is used. There is discrepancy in respondents knowing whether courses are registered on the SSEF. However, other than for the NIHSS there is no indication that we are comparing information on the same course content which could account for some of this variation.

Table 5.15 Numbers of units that reported receiving scale-specific training with information on course type, mode of delivery, and whether the courses were registered on the Stroke Specific Education Framework (SSEF).

Training in scale	**No of units**	**Course type**		**Mode of Delivery**		**Stroke Specific Education Framework (SSEF) registered**	
NIHSS	69	Internal	4	Face to Face	2	Yes	17
		External	58	Online	58	No	5
		Missing	4	Both	4	Don't know	36
GCS	8	Internal	6	Face to Face	4	Yes	0
		External	2	Online	4	No	2
						Don't know	6
NEWS	6	Internal	5	Face to Face	0	Yes	2
		External	0	Online	5	No	1
		Missing	1	Both	1	Don't know	3
AVPU	2	Internal	2	Face to Face	2	Yes	0
		External	0	Online	0	No	1
						Don't know	1
SNOBS	2	Internal	2	Face to Face	2	Yes	0
		External	0	Online	0	No	1
				Both		Don't know	1

Participants were also asked about whether informal training or mentorship opportunities existed for neurological assessment, and if so, what they are, as well as whether there are competency assessments in place. Fifty-four percent (n=68) reported informal local training or mentorship opportunities with 46% (n=31) specifically mentioning on-the-job, local, in-house, or informal training opportunities. Competency assessments were reported by 48% (n=61) of the questionnaires. However, further examination identified that 30% (n=18) classed the NIHSS online training as a competency assessment and were not referring to practice-based competency assessments.

5.4.10. Clinical Documentation Supplied

Twenty-three separate Trusts supplied clinical documentation. Only nine were protocols, guidelines, or standard operating procedures (SOPs) concerned with monitoring of which six were stroke specific. Five used GCS plus Limb Movements and Pupillary Response, two did not specify which scale, if any, was used, one used the Canadian Neurological Scale (CNS) and one used its own tailored observation tick box chart. Six provided specifics on the frequency expected three did not.

5.5. Issues with completion

No major issues were identified with completion of the questionnaire however certain sections were less well completed. As expected, open questions were the ones most regularly missed as these require more time and effort to complete (Holland & Christian, 2009). However, it was important to include some to yield descriptive information. Across most closed questions only small numbers were not answered. Where this occurred, it was single answers missing across questions rather than entire sections not completed which were likely due to human error. There were a couple of questionnaires where entire sections were not completed such as frequency of monitoring and training provided. This potentially indicated more reluctance to complete these sections as they were time consuming and possibly less accessible to busy clinicians. The most poorly completed section was the bed type data in unit demographics although this was felt to be due to the respondents not being aware of the classifications rather than issues with the question design.

5.6. Discussion

To the author's knowledge, this is the first UK-wide survey that has explored the practice and experiences of neurological assessment and monitoring after stroke. The survey is indicative of acute stroke services as they were selected through national audit data and supported by a high survey return rate (80%). There was excellent engagement with the research signifying this is an important

topic area for clinicians. Variation was identified in multiple aspects of neurological assessment and monitoring practice. A quarter of respondents felt that neurological assessment and monitoring is a neglected area of practice and crucially 89 (71%) of respondents feel that changes are needed in this element of care across stroke services which supports the work of the thesis.
The survey collected extensive data on clinicians' opinions and experience with neurological assessment and monitoring. Responses around satisfaction with and staff confidence in various aspects of practice showed that there is room for improvement in many areas. Although these elements are hard to measure and are based on perception, they add important justification for the need for change in this area of practice. The survey also broadly evaluated actual and potential barriers for both current neurological assessment and monitoring practice and future changes. These data have crucially provided indications of what needs to be considered in the future implementation of change in this area.

Nearly two-thirds of units identified as both a hyperacute and acute stroke unit. During the survey, it became clear that these categories do not have clear defining characteristics and can be selected based on opinion. This meant that meaningful comparisons based on unit type were not possible. However, new categories suggested for stroke units as part of the National Stroke Service Model (NHS England and NHS Improvement 2020) and the continued drive to reconfigure acute stroke services into specialist centres (Imison et al., 2014) could allow meaningful comparison in the future and be harnessed as drivers to reduce variation in neurological assessment and monitoring practices within and across services.

Across units, there was variation in which staff, in terms of both job role and grade, completed neurological assessment and monitoring. Although HCAs were often reported as responsible for the completion of physiological monitoring, neurological assessment and monitoring was deemed a trained professionals' role at many units. Nurses and doctors were reported as the professional groups most likely to be involved in neurological assessment and monitoring. In hindsight, it might have been beneficial to split this question between assessment and monitoring as it would have clarified the information from other survey responses where it seems that doctors and specialist nurses tend to complete more formal standalone assessments and nurses the regular ongoing monitoring.

Nurses being responsible for the regular ongoing assessments to monitor for change is consistent with the author's experience and the wider literature (De Leon Bendetti et al., 2021). However, the

survey data did suggest wider MDT involvement in many units, which was a positive finding as research shows that patients who have care provided by a specialised stroke MDT have improved outcomes (Clarke & Forster, 2015). Healthcare policy is driving the need for greater MDT working to bridge the workforce gap and improve quality by drawing on a broader range of skills and competencies. There is a desire for new ways of working with teams making effective use of different disciplines with a range of skills and experience (NHS, 2020). The 'one workforce' approach to MDT working aims to draw individuals together to achieve common goals (Health Education England, 2021). Involving all members of the MDT where possible in neurological assessment and monitoring could be beneficial and ensure better understanding and communication within teams. It could allow the maintenance of high standards despite staff shortages within the NHS generally and stroke services specifically as outlined in chapter 2. There is a growing necessity for the blurring of boundaries between professionals' roles to meet the needs of patients (CordisBright, 2018).

There is strong agreement throughout the survey data that neurological assessment and monitoring is important for all stroke patients to detect change, specifically deterioration. Prioritisation is given for certain characteristics such as receiving thrombolysis but overall agreement on what needs to be done, when, and for which patients is lacking. There were indications from the data that adherence to neurological monitoring is better when it is completed alongside physiological observations. As physiological parameters can impact neurological function and be indicators of deterioration they should be completed, understood, documented, and managed together to prevent complications, such as END, through the maintenance of homeostasis both physiologically and neurologically (Middleton, McEldruff, and Ward, 2011).

Documentation of physiological observations and neurological assessment and monitoring and associated decisions varies across units. Two-thirds of respondents agreed that documentation of neurological monitoring was well done, but there is room for improvement. Some of the discrepancies could be due to different systems of documentation being used. There was a relatively even divide in terms of both electronic health records (EHRs) (43% paper versus 42% electronic) and systems to record physiological observations (42 to 45% on paper and 53 to 54% electronically). Data from medical devices such as observation systems can be interfaced with EHRs but this can require custom interface applications (Evans et al., 2010). From the data, it is not clear if these have been routinely developed or implemented for neurological assessments in many units. Whether this is because electronic systems are not available, inflexible, or because the MDT view them as time-consuming or as increased workload needs investigating (Doods et al., 2014; Evans, 2016). Despite

the move to electronic records having been advocated for decades (Ornstein et al., 1992), there remains a hybrid collection of electronic and paper-based data (Institute of Medicine, 1997). This thesis does not provide enough data to advocate a particular documentation method, in terms of electronic versus paper. However, it does recommend that there should be continuity and consistency in how and where the results from physiological observations and neurological monitoring are recorded. It is important for continuity of care that assessments are accessible and comparable to ensure links and trends can be detected and communicated.

Communication beyond documentation of results and decisions in the form of patient handovers between staff was briefly explored in the survey. The majority of these handovers were reported as occurring orally. Just over half of respondents reported that neurological assessment and monitoring is regularly included in handovers. Further information should be gathered to explore whether there is a common language and understanding around neurological assessment and monitoring across the whole MDT or whether handovers are completed in a profession-specific format. Clear and concise transfer of information is known to enable continuity of and safety in patient care (Shahid and Thomas 2018). Therefore, effective communication of change, specifically deterioration, is essential, especially as communication has been shown to improve the performance of stroke teams (Cramm & Nieboer, 2011).

It is thought that well-developed and implemented clinical guidelines have the potential to reduce unwarranted variation and improve healthcare quality and safety (Langhorne et al., 2020; Panteli et al., 2019). Despite the national guidelines (Royal College of Physicians, 2016) lacking any specifications in terms of what we should be doing, when, and for which patients just under half of the participants agreed that guidelines clearly specified what to do. Approximately a quarter of the respondents were not satisfied with the guidelines in place.

One hundred units (80%) reported having neurological assessment and monitoring protocols of which 97% were reported to be stroke specific. However, the limited clinical documentation received from participants mainly covered patients receiving specific treatments, in particular thrombolysis. Only a small proportion provided any detail on what neurological assessment and monitoring should be completed in other patient groups indicating that further and more robust guidance is needed. This was supported by just under a third of those who felt change is needed calling for the introduction of standardised guidelines.

There was a consensus that guidelines on response to deterioration are better than on the monitoring itself. However, there is evidence that even where clear evidence-based guidelines exist they are not always well implemented and executed (Baatiema et al., 2017). This is mirrored in the survey where perceived adherence to protocols, where they existed, was reported as variable. This indicates that even where guidelines are in place additional variation could exist due to what is completed in practice not being what is advised. Guideline development for neurological assessment and monitoring will be particularly challenging not only because of the complexity of the element of care and its context of delivery but also due to the range of stroke types and severities.

Two-thirds of respondents agreed that assessment and monitoring should be tailored with a further eighth unsure. There is a historic assumption that individualised patient care benefits both patients and clinicians (Redfern, 1996). It does seem logical that tailoring of monitoring would be beneficial to patients and be a better use of healthcare resources including staffing. However, more research is needed to know across which factor(s) (e.g. patient groups, severity, co-morbidities) alterations can be made whilst retaining effective and efficient monitoring. In several questionnaires, the choice of frequency was reported as being dependent on patient condition, but no justification or reasoning was provided. Further work is needed to try and understand the decision-making processes currently employed in practice. There were concerns raised that when frequencies change monitoring is more likely to be missed. This would potentially have implications if different patients were on individualised frequencies. However, whether frequencies are standardised or tailored, systems would need to be in place to ensure that monitoring is not missed. Supporting protocols or guidelines could be a driver for future changes as they are associated with a positive impact on outcome (Jones et al., 2018) however other strategies will be needed to ensure guidelines are adhered to.

The survey showed that within each stroke unit there is a range of scales used, some stroke specific and some generic. The choice of scale used is dependent upon the time of and purpose of the assessment. The NIHSS is advocated for routine monitoring in international guidelines (Ashcraft et al., 2021; Boulanger et al.,2018). However, the survey clearly shows it is not used routinely for monitoring in the UK. This could be because the NIHSS is seen as too complicated and time consuming for this purpose (Richardson et al., 2006; Yanko & Lang, 2013). The GCS and AVPU are the scales most used for regular monitoring across the UK despite the GCS showing poor sensitivity to detect change after acute stroke (chapter 4). Only a very small number of sites use a stroke specific assessment, such as the SNOBS for routine monitoring. There was an awareness that scales might not be appropriate for purpose and many respondents reported a lack of confidence in their use.

Specifically, a quarter of those who felt change was warranted expressed that the GCS was not appropriate in a stroke population, but why it is continued to be used regardless needs exploring and addressing.

As well as variation in scale use, the survey illustrated that there is extensive variation in monitoring frequencies across all patient groups and time periods within stroke units across the UK. It also indicated that there is uncertainty in the frequencies that should be used across different patient groups and time periods. Variation was found to occur even where guidelines exist, such as in the 0-24 hour period after thrombolysis (Powers et al., 2019). The range of frequencies data added greater depth to the most common frequency data for instance in the ischaemic stroke (without thrombolysis or thrombectomy) group. In this group, four-hourly monitoring was commonly reported indicating this frequency is widely adopted by many units, but this group also had the broadest range of monitoring frequencies indicating a high level of uncertainty about what is the required frequency. As this represents the largest proportion of the stroke population, such variation has implications for care provision across the UK.

General trends were noted which showed patients who receive treatments (thrombolysis and thrombectomy) tend to have more frequent monitoring as do ICH patients compared to ischaemic stroke patients. Also, many units reduce the frequency of monitoring as time progresses with most patient groups reducing the frequency at 24-48 hours. However, in the potential hemicraniectomy patient group frequencies are maintained more frequently than four hourly for up to 72 hours potentially because this group may deteriorate over a longer period. There appeared to be more consistency in doing routine monitoring for typical and stable patients, but more uncertainty with complex patients and where there were significant changes in condition. The range of frequencies reported increases over the time periods for all patient groups which further indicates variation in practice and uncertainty across the UK about what frequency should be used and when neurological assessment and monitoring should be discontinued.

Information on discontinuation practice was requested however, it was poorly reported. Most units carry on neurological assessment and monitoring beyond 72 hours as standard. There is often a reluctance to reduce or discontinue any intervention within healthcare (DuBose & Mayo, 2020; Tappen et al., 2017). There is a risk that habitual practice is overriding need, but more research is needed into the best frequencies to use for all patients and time periods as well as about when to discontinue monitoring.

Regular neurological monitoring not being performed has been previously reported but little research has been done (De Leon Bendetti et al., 2021). This survey explored reasons why neurological assessment and monitoring may not be completed. The primary reasons cited were around the units being busy and competing priorities, or issues with staffing out of hours (overnight and weekends) which could result in missed assessments. Inadequate staffing and skill mix have previously been shown to cause poor recognition and management of deteriorating patients (Johnston et al., 2015; McGaughey et al., 2017). This was replicated in the data where inadequate staffing at any time was reported as impacting the completion of neurological assessment and monitoring. Reducing missed monitoring in clinical practice should in theory increase its ability to detect deterioration. Some of these factors identified would not be addressed by making changes to neurological assessment and monitoring practices but developing a system that is more time-efficient and can be completed by a greater number of staff could help improve completion.

Improvement in the detection of neurological deterioration, not only relies on monitoring being completed but the assessments themselves being completed correctly. The survey results show that only just over half of respondents feel that neurological assessment and monitoring is currently completed accurately and consistently. Incorrect application and completion of monitoring will limit its reliability and effectiveness to detect change (chapter 3). Multiple factors influence the effective completion of assessments including the patient being assessed.

Patient characteristics, specifically in terms of communication and cognition, were strongly identified as making assessment harder to complete. Approximately one in three patients after stroke will have some form of cognitive impairment or communication difficulty (Engelter et al., 2006; Patel et al., 2003). These problems can occur in the full severity range of stroke (Fens et al., 2013). Assessment skills with patients with cognition difficulties can be especially difficult for new staff (Tang et al., 2017). Clearer assessment criteria and better training could help address complications that arise in assessments due to different patient characteristics to ensure that deterioration is correctly identified.

Prompt identification and treatment where possible of END is crucial to prevent secondary brain injury and potentially improve outcomes for patients. The data on how deterioration is recognised was extensive and from the myriad of answers, it was unclear whether staff were considering neurological deterioration or more general deterioration. Initially, it was felt that the questionnaire should have more clearly specified neurological deterioration. However, on reflection, the data

reiterated the intrinsic link between neurological monitoring and physiological observations and highlighted that multiple factors affect and could be important to the recognition of deterioration.

More confidence was expressed in the detection of deterioration than the ability to use the assessment scales which supports the theory that clinicians currently use multiple indicators beyond neurological assessment scales to identify deterioration. The findings indicate that neurological deterioration is not always being identified independently and that staff rely on other physiological signs of deterioration to highlight change. This could be because the scales being used are not suitable to pick up on specific signs of neurological deterioration in patients.

The survey highlighted reliance on a scale score reduction, particularly in the GCS, and other indicators of changes in conscious levels to be aware that a patient has deteriorated. However, as reported in Chapter 3 changes in items within scores can occur but go unnoticed if the total score is relied upon as an indicator of change and individual items should be used to ensure change is identified. Also, alteration in level of consciousness is a late sign of deterioration. If stroke teams could pick up on early neurological changes potential actions might be more effective and potentially improve outcomes. There is a potential opportunity to improve the speed of recognition of deterioration by improving the identification of more subtle signs of stroke specific neurological change.

Experience and intuition have been highlighted in non-stroke-specific studies as important contributing factors to effective recognition of deterioration and referral (McGaughey et al., 2017, Massey, Aitken & Chaboyer 2015). Only a small number of participants (n=9) reported that their awareness of deterioration was due to gut feeling or intuition. Conversely, a fifth of respondents reported that experience and intuition were more important than neurological assessment and monitoring in identifying neurological deterioration, and about the same number were unsure. Experience is recognised as important in clinical practice (Bartel et al., 2014; Choudry et al., 2005) and staff who are experienced in stroke care may be able to pick up on subtle changes quickly without guidance. In order to ensure consistent care across stroke care in all settings and with a range of staff formal assessments of change are required however the process of identifying meaningful change should be standardised with assessments having easy to follow indicators.

When change is noted, it is important that there is an appropriate response (Siegler & Martin-Schild, 2011). The most common response reported is medical escalation. Whilst multiple other actions

were reported, there were no details on the order or hierarchy of these actions. It is unclear if other actions are instigated before the medical review or whether the review is pivotal to further actions. Standardisation in neurological assessment would hopefully provide a framework of communication between staff and strengthen articulation of concerns to medical teams.

The survey aimed to establish the current level of training provided in this element of care. However, there was little consistency in the completion of the training questions. From the data, only scale-specific training data could be collated, and this showed low levels of training not encompassing the whole MDT. Levels of training were generally low and not all members of the MDT were trained. Doctors and specialist nurses tend to be trained in the NIHSS whereas training in the scales used for routine neurological monitoring, such as the GCS, tends to be mainly for more general unit-based nurses. This variation in training across scales accords with the findings that different professional groups use different scales for different purposes as mentioned earlier. Several internal courses appeared to be delivered online, whereas traditionally internal courses are often face-to-face. It is unclear if this is because respondents did not realise training was delivered by an external provider or whether it could be an implication of data collection during a pandemic when face-to-face teaching was not allowed. Further work should identify current gaps in training and how this can be addressed within the context of staffing and their frequency of exposure to different patient scenarios.

Although staff expressed confidence in the training they received, they reported having less confidence in the competency of staff. This suggests that there is more to the development of competency and trust in others' skills than training alone. There was little indication of competency assessments currently in neurological assessment and monitoring practice and where they were reported it was mainly linked to specific scale training. NIHSS certification, despite being reported as a competency does not ensure proficiency and expertise in completion as outlined in Chapter 3. The development of meaningful competencies should be a key consideration in future neurological assessment and monitoring practice because limited knowledge and awareness and lack of skills and competence represent barriers to optimal stroke care (Baatiema et al., 2017).

5.7. Strengths and Limitations

Given the high response rate a representative sample of stroke units were included. However, using a convenience sampling method to enrol participants according to their availability and accessibility could have introduced bias, and affected generalisability. Many participants indicated how

important they felt research into this topic was and the response rate also indicated the engagement of clinicians despite the issues caused by the pandemic. The questionnaire was always completed by a professional who had a working knowledge of neurological assessment and monitoring in the acute stroke unit and focused on obtaining information about the service rather than personal opinion, so in theory this pragmatic sampling technique should not have impacted highly on data quality. Participants could involve other team members in completion, but it is unknown whether this occurred or whether the responses represented single opinions. In theory team collection may have been more robust rather than individual opinion but then this may have led to what was reported being led by agreement rather than reality

The questionnaires were mainly returned by senior clinicians. This was not considered a real limitation as senior clinicians may have a better understanding of practice and therefore the responses were more accurate. However, there is a risk that they might not have reported what they thought was or should be happening because they are not actively involved in neurological assessment and monitoring practice. It is also possible that they might not be aware of the difficulties more junior staff face. Future research should include all professional groups and grades involved.

This survey was intended to provide a 'snapshot' of current practice in relation to neurological assessment and monitoring but due to delays in completion including the global pandemic, it took one year and nine months to complete. However, no major changes to care provision or factors that influence it were identified during this period, so it is believed that the data is indicative of current practice across the UK. It was planned to compare units based on their national audit data in the form of the latest scores for the units in the Sentinel Stroke National Audit Programme (SSNAP) or Scottish Stroke Care Audit as these impact care provision in other areas. As this data was not well reported generally, or available for Scottish units, and because the data collection period was expanded it was decided inappropriate to make comparisons across data.

There was no way to externally validate the responses from sites, but the high volume of responses means that any issues in completion are unlikely to have affected the overall interpretation of the survey results. Some of the questions sought opinions, which could have many influences dependent upon individual, unit or organisational level factors. However, they all provide important context for understanding how to make improvements in patient monitoring and response.

5.8 Chapter Summary

This chapter outlined the findings of a UK-wide survey that explored variation in practice and clinicians' experiences of neurological assessment and monitoring practice. Data demonstrated that there was obvious variation both in terms of current practice amongst stroke units and clinicians' understanding of neurological assessment and monitoring across the UK which potentially could be leading to differences in outcomes for some stroke patients. The results showed there is a clear need and desire to make changes in neurological assessment and monitoring to reduce current uncertainty and variation in practice.

There was overall agreement that neurological assessment and monitoring is important for all stroke patients to detect change and instigate appropriate action. There is a need for more stroke specific neurological assessment and monitoring practices. Guidelines and protocols are needed that are specific to stroke, achievable in busy clinical environments, and result in appropriate action if deterioration is noted. Guidance on what we should be doing, when, how often, and for which patients could be tailored based on patient characteristics.

Further exploration across several factors is called for such as communication within teams, documentation, action on deterioration, and training. Although numerous barriers and facilitators could impact changes in this element of care the overall perception of importance and the desire for change are key drivers that can be harnessed to ensure meaningful change occurs.

The next chapter, Chapter 6, describes the design and delivery of the semi-structured interviews phase. These interviews allowed further and deeper exploration of neurological assessment and monitoring practices. Through the application of Normalisation Process Theory (NPT), they further identified and clarified barriers and facilitators to the implementation of changes in this element of care.

Chapter 6 -Interviews

The previous chapter reported the results of a UK-wide survey of current practice and experience of neurological assessment and monitoring practice. This chapter presents data on semi-structured interviews completed with a range of clinicians. These allowed deeper exploration of the use, understanding, and acceptability of neurological assessment and monitoring. They also provided insight into the barriers and facilitators around implementing and integrating changes in this element of care. Normalisation Process Theory (NPT) was utilised as a framework to underpin the interview development and analysis to identify factors that could impact on the implementation of future change in practice.

6.1 Normalisation Process Theory (NPT)

NPT was selected because it has been extensively used to support the work of implementing and embedding (i.e. normalising) complex interventions into practice (Bagot et al., 2017; Clarke et al, 2013; Gillespie et al., 2018). It has four constructs that represent different kinds of work that people do around implementing innovation or change to practice: Coherence, Cognitive Participation, Collective Action, and Reflexive Monitoring. NPT is a middle-range implementation theory that can underpin process evaluation of complex interventions in healthcare (May et al., 2020, 2021). NPT compliments the theoretical underpinnings of critical realism as it allows focus on accounts of the individual, unit, and system-level processes, practices, and ways of reasoning. NPT also recognises that context is active and dynamic and greatly impacts implementation processes and outcomes.

6.2 Aim

The aim of this component of the programme of research was to determine knowledge, understanding, and acceptability of neurological assessment and monitoring after stroke and to explore the barriers and facilitators to implementing a change to practice.

6.3 Objectives

- To ascertain from a range of staff their knowledge, professional experience, assumed importance, and acceptability of using neurological assessment and monitoring in patients after stroke.

- To identify and/or explore barriers and facilitators to the use and implementation of neurological monitoring in clinical practice (e.g., highlighting local standards and current systems of practice that assist or hinder neurological monitoring).

6.4 Methods

6.4.1 Study Design

Study design was qualitative using semi-structured interviews.

6.4.1.1 Justification of study method

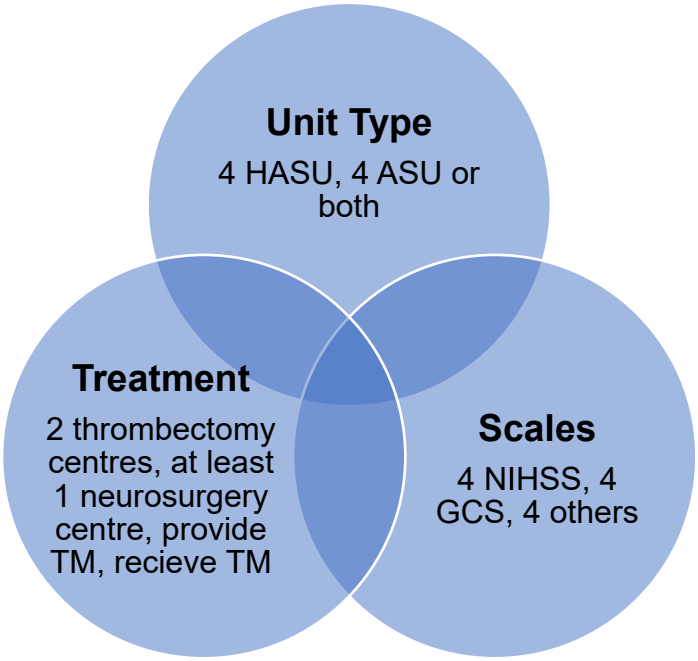
The choice of data collection method was important, because how the information collected is used, and what explanations it can generate are determined by the methodology and the analytical approach applied by the researcher (Teherani et al., 2015). The exploration required from this chapter was around capturing in-depth experience of neurological assessment and monitoring and how different factors within the system affect its delivery in clinical practice. Method options for eliciting experience include focus groups and interviews. Interviews were primarily chosen as it was felt that that they provided the ability to probe and prompt to ensure rich and in-depth data (Gallagher et al., 1993; Paradis et al., 2016) and that it was necessary to understand a range of different individuals experiences and understanding around neurological assessment and monitoring rather than a group agreement (Paradis et al., 2016). Although both methods have been shown to produce similar numbers of overall items, individual interviews have been shown to be more effective at generating a broad range of items on a per-person basis (Guest et al., 2017). Although in some situations the interpersonal and interactive nature of focus groups can allow production of a wider range of views and ideas than from a single respondent (Greenbaum, 2003; Kidd & Parshall, 2000) there is a risk that dominant personalities could have influenced the discussion and data collection (Gallagher et al., 1993; Paradis et al., 2016). Interviews were therefore felt to fit better with the critical realism explorative nature of the research as they allowed more detail and insight into the individual's personal thoughts, feelings, and world view (Morgan, 1998). Interviews also offered more flexibility in terms of scheduling with busy professionals and made them potentially more accessible for participants.

6.4.2 Site Selection

Sampling was designed to be purposive to reduce potential bias and maintain validity. Four sampling criteria were chosen: geographical location, type of unit, treatment provided, and scales used (Figure 6.1). Data about unit characteristics and service provision from the survey (Chapter 5) were

used to identify and purposely select sites to approach for the interviews (explanatory participant selection model [Cresswell, 2003]). It was planned that 15-25 staff would be recruited (approximately two interviews per site in 12 sites, across the UK). Sample size was based on a pragmatic decision of what was feasible within the time and resources available. This was an estimate of what would be needed to potentially generate adequate data to meaningfully explore the complex and multi-faceted nature of neurological assessment and monitoring practice (O'Reilly & Parker, 2013; Sim et al., 2018). The study was concerned with achieving sampling adequacy i.e. collecting quality data in terms of richness, depth, diversity and complexity rather than achieving a pre-determined quantity of data (Fusch & Ness, 2015). The supervisory team applied the sampling criteria and selected units for approach from anonymised spreadsheets. Reserve units were identified in case of no engagement from those sites selected.

Figure 6.1 Venn Diagram illustrating sampling criteria



TM= Telemedicine

6.4.2.1 Geographical location

The interviews were allocated across the four nations of the UK to achieve geographical spread (Table 6.1). The numbers of units needed across the four nations was calculated proportionally from the total number of units per nation (Table 6.1).

Table 6.1 Geographical spread allocation for interviews across the UK

	Total number of sites	**% of total sites (n=159)**	**Suggested interview numbers (proportional)**
NI	8	5	1-2 (1 site)
Scotland	28	17.6	4 (2 sites)
Wales	12	7.5	2 (1 site)
England	111	69.8	17 (8 sites)

6.4.2.2 Type of Unit

The aim was to sample from four sites that had a hyper acute stroke unit (HASU), four that had an acute stroke unit (ASU), or both. Sites that did not identify as either were not chosen, as it was felt they were not representative of most stroke services across the UK, generally receiving small numbers of stroke patients, which could potentially skew data.

6.4.2.3 Treatment offered

Knowledge, understanding, and acceptability of neurological monitoring might differ depending on services routinely available at sites. Therefore, it was decided to include two sites that provide thrombectomy service and a minimum of one site each that provide neurosurgical intervention such as hemicraniectomy, telemedicine services to others, or receives telemedicine services from others.

6.4.2.4 Scales used for assessment and monitoring

To incorporate sites that used different scales for assessing and monitoring neurological status, it was decided to sample four sites that used the National Institutes of Health Stroke Scale (NIHSS), four that used the Glasgow Coma Scale (GCS), and four that used other scales.

6.4.3 Study participants

The University ethics committee stipulated that a key contact at each site had to be used to advertise the opportunity to take part in the interviews to other staff members. A contact was identified at each of the chosen sites who was willing to circulate information about the study to the stroke team including participant information sheets and consent forms. After receiving the information these contacts were followed up three times at two weekly intervals. Potential participants were any member of staff who completed or had influence on neurological assessment

and monitoring in the stroke unit (e.g., nurses, nursing assistants, stroke physicians, therapists, and managers). They had up to 8 weeks from the time they received the information to decide whether to take part and contact the author. This meant that participants self-selected to take part in the interviews and there was no ability to purposively sample staff across any criteria including professional grouping or staff grades.

6.4.4 Interview content and development

Qualitative semi-structured interviews were chosen as they lend themselves to multi-method research. They are flexible and allow exploration without the risk of losing focus of the research questions, which might happen in unstructured interviews (Low, 2013). The initial interview schedule versions were devised utilising NPT constructs and the author's experience and knowledge in neurological assessment and monitoring (Appendix 6.1). The interview schedule underwent multiple changes and adaptions based on findings from the survey data, supervisors' opinions and advice, and external reviewers (both patient and expert representatives). Open-ended questions were used to explore views and experiences. To help ensure in-depth data was collected, some probes were pre-determined to encourage participants to expand on responses if required. The interview schedule was piloted to check that the questions were understandable to participants and assess potential interview duration (approx. 40 mins). The questions initially focused on individual practice and opinion. However, piloting identified that if questions were asked from an organisational perspective, the participant provided more open and useful feedback. This change in focus was adopted to allow participants to think broadly about the topic and context rather than interrogating and questioning their specific practice. The ordering of questions was altered to ensure questions flowed rather than followed the initial NPT order of questions. However, NPT remained the underlying theory in interview design and delivery. The final interview schedule (Appendix 6.2) was agreed upon by the study team.

6.4.5 Data collection

Full written consent was obtained before participation. Basic participant demographic data were collected as part of the interview (i.e., job title and length of time working in stroke services). All interviews were completed remotely by the author. Interviews were audio-recorded with the researcher taking field notes. These allowed items to be clarified later if needed without breaking the flow of participants' narrative and ensured some data collection in case of recording failure. Extensive notes were taken when consent to audio recorded was not provided to ensure maximum data collection.

The recordings were securely handled, stored, and transcribed (in line with GDPR and research governance guidelines [HRA, 2022a; International Council for Harmonisation, 2016]). Transcription was completed by an external company on behalf of the author. Transcriptions were anonymised, so it is not possible to identify individuals from them.

6.4.6 Data analysis

Thematic analysis (TA) was chosen as an accessible and useful method for identifying, analysing and reporting patterns within the data (Braun & Clarke, 2006). TA as a method fits with critical realism, the theoretical underpinning of the overall thesis, as it allows acknowledgment of the broader contexts and allows exploration of both the observable and unseen factors that impact on individuals' experiences (Bhasker, 1978, 1979, 1989). It allows for interpretation and not just description of the data allowing key concepts to be highlighted, followed by a process where patterns of association and possible reasons for these could be explored.

TA is flexible and allows themes to be developed in different ways. Primarily codebook TA was undertaken with NPT as the structured framework for developing and documenting the analysis (May & Finch, 2009; Braun & Clarke, 2021a). However, coding continued to be developed, defined, and refined throughout the entire analysis and new themes and sub-themes were developed through inductive data engagement until the data fitted well (Braun & Clarke, 2006; ibid). This blended approach to coding supported rich and detailed accounts of data whilst acknowledging complexity and allowed additional themes to be added to, or within, the NPT framework to facilitate analysis and collation (Graebner et al., 2012). Table 6.2 shows the final coding manual developed which includes social context, behaviour, and technical aspects of neurological assessment and monitoring practice. The themes house the participants multi-faceted accounts or stories.

The transcripts were checked against the audio recordings for accuracy and context before coding began using NVivo (Release 1.61) (Braun & Clarke, 2006). Each data item (interview transcription) was given equal attention in the active coding process. Portions of data were individually coded to specific themes (extract). "Trustworthiness" of coding was assessed by sample secondary coding by the supervision team. Multiple coding of extracts was undertaken when they fitted within more than one theme. All relevant extracts were collated under themes before data was interpreted by theme. This condensing and interpretation of data was agreed by the author and another member of the supervisory team not to measure reliability but to ensure coherence in the narratives written. There was focus on the tension and inconsistencies within themes as these provided better understanding of different viewpoints and experiences within the data. Prevalence data, the number of different

speakers who articulated the theme, was collected (ibid). This was not designed to show dominance or importance as the team were interested in all viewpoints and stories. Prevalence data is reported in more depth where it captures something important to the overall research questions such as in the deterioration and patient groups themes.

Table 6.2 Coding manual for the interviews based on Normalisation Process Theory (NPT) four core constructs.

Text in blue indicates themes identified during data analysis that are presented separately.

COHERENCE: the **sense making** work that people do when they are faced with using a new set of practices (chapter section 6.5.1)	
Differentiation	Perceived differences between old and new systems of work, that have consequences for how people operate in practice
Communal specification (Collective agreement about the purpose and function of the neurological assessment and monitoring and how it should/will work)	Current practice- how neurological assessment and monitoring occurs
	Deterioration- what would need to be seen in a patient that would indicate they had deteriorated
	Explanation- about why neurological assessment and monitoring is important and should be part of care
Individual specification	Individuals understand what the new practice requires of them
Internalisation	Perceptions of the value, benefits, and importance of neurological assessment and monitoring
COGNITIVE PARTICIPATION: the relational work that people do to build and sustain a new practice (chapter section 6.5.2)	
Initiation	Identification of key individuals who drive neurological assessment and monitoring forward.
Enrolment	People agree that neurological assessment and monitoring should be part of their work (included both engagement and reluctance with the practice
Legitimation	Governance procedures around neurological assessment and monitoring
Activation	Whether people work together, or not. Highlights pathway issues related to neurological assessment and practice
COLLECTIVE ACTION: the operational work that people do to enact a new practice (highlights the practical barriers and facilitators within clinical practice) (chapter section 6.5.3)	
Interactional workability	Staff and patients can perform the tasks required by neurological assessment and monitoring
Relational integration	Staff trust each other's work and expertise in neurological assessment and monitoring

Skill set workability	Who does the work involved in neurological assessment and monitoring (how it is allocated)
Contextual integration	The organisation adequately supports neurological assessment and monitoring
Patient Groups	Stroke types or patient characteristics that impact on neurological assessment and monitoring practice
REFLEXIVE MONITORING: the appraisal work that people do to assess and understand how a new practice affects them and others (chapter section 6.5.4)	
Systematization	People collect information about the impact of neurological assessment and monitoring
Communal appraisal	People collectively evaluate neurological assessment and monitoring as worthwhile
	Opinions on whether patients and carers evaluate neurological assessment and monitoring as worthwhile
Individual appraisal	Individuals evaluate neurological assessment and monitoring as worthwhile
Reconfiguration	Suggestions for modifications/ improvement in neurological assessment and monitoring practice based on current evaluation

6.4.7. Ethical and Local Approvals

Ethical approval was obtained from the Science, Technology, Engineering, Medicine, and Health (STEMH) ethics committee at the University of Central Lancashire (UCLan) (reference STEMH 1018) (Appendix 5.2) and under proportionate review from the Health Research Authority (HRA) (project ID 261850, REC reference 19/HRA/4113) (Appendix 5.3). The HRA proportionate review decided that the project should be completed under a participant identification centre (PIC) agreement. Local approval was obtained from each Trust's research department before a key contact was approached to distribute the invitations to interview within their stroke units.

6.5 Results

Interviews were planned to be completed after all the surveys were returned. However, due to delays caused by the COVID-19 pandemic, sites were selected for interviews after >40% of surveys had been returned. This allowed the survey to stay open to maximise return rates whilst allowing interviews to be completed within the PhD timeframe. Site selection for interviews occurred in March 2021 for Scotland and June 2021 for the other nations. Geographical representation was achieved, and all other key sampling criteria were covered except only two sites were pure HASUs.

Twenty-three interviews were completed, with 22 audio-recorded, between April and December 2021. Table 6.3 outlines the characteristics of the participants and the units they represented.

Table 6.3 Characteristics of interview participants and units by geographical location.

Geographical Location (GL)	Site No	Unit Type	Scale/s used	Thrombectomy	Telemedicine	Hemi-craniectomy	No of interviews	Participant Job Title/s	Time in stroke services
Scotland	1	Both	NIHSS, GCS, AVPU, SNOBS	No	No	No	2	Consultant Physician	>10
								Stroke Research Nurse	>10
	2	ASU	mNIHSS, GCS, AVPU	No	Provides	Yes	2	Deputy Charge Nurse	5-10
								Deputy Charge Nurse	5-10
Wales	3	ASU	NIHSS, GCS, AVPU	Yes	No	No	2	Physio Team Leader/ Stroke Pathway Development Lead	>10
								Lead Clinical Stroke Nurse Specialist	5-10
Northern Ireland	4	ASU	NIHSS, GCS, AVPU	No	Receives	No	1	Clinical Services Manager/ Stroke Improvement Manager	<5
	5	HASU	NIHSS, GCS, AVPU	No	Provides	No	1	Stroke Consultant/ Consultant Geriatrician	>10
	6	ASU	NIHSS, GCS, AVPU	No	Provides & Receives	No	1	Advanced Stroke Specialist Nurse	>10
England	7 Pilot	ASU	NIHSS, GCS, AVPU, Other (unspecified)	No	Both	Yes	1	Consultant Nurse/ Clinical Lead	>10
	8	Both	NIHSS, GCS, AVPU, SNOBS	Yes	No	No	2	Stroke Nurse	5-10
								Stroke Specialist Nurse	5-10
	9	Both	NIHSS, GCS, AVPU, SNOBS	No	No	No	2	Specialist Doctor Stroke Registrar	<5
								Stroke Speciality Doctor	5-10
	10	Both		Yes	Both	Yes	2	Consultant	5-10

Geographical Location (GL)	Site No	Unit Type	Scale/s used	Thrombectomy	Telemedicine	Hemi-craniectomy	No of interviews	Participant Job Title/s	Time in stroke services
			NIHSS, mNIHSS, GCS, AVPU, SNOBS					Specialist physiotherapist	<5
	11	Both	NIHSS, GCS, AVPU, FAST	No	No	Yes	2	Trust Stroke Registrar	5-10
								Sister/ Team Leader	>10
	12	Both	NIHSS, mNIHSS, GCS, AVPU	No	No	No	2	Deputy Sister	10
								Consultant Nurse	<5
	13	Both	NIHSS, GCS, AVPU (NEWS)	Yes	Both	No	1	Advanced Practitioner Role-Physio	<5
	14	HASU	NIHSS, GCS, AVPU	No	Receive	Yes	2	Stroke Nurse Practitioner/ Stroke Research Nurse/ Lead Stroke Educator	>10
								Consultant Stroke Medicine	>10

Results in the form of themes and sub-themes (often mirroring NPT components) are presented in sections under the four core components of NPT. To provide an indication of levels of data for each theme the numbers of interviews that contained relevant data and the total numbers of references coded to that theme or sub-theme are reported. Key quotes are presented within each analytic narrative of the theme to add illustration to the stories.

6.5.1 Coherence

Coherence in NPT is around the sense -making work that people do in relation to a new set of practices.

6.5.1.1 Differentiation *(8 interviews, 15 references)*

This theme looked at perceived differences between old and new systems of work and the consequences they have for how people operate in practice. There was a strong feeling amongst some interviewees that current practice was not ideal. Differing opinions about what can and should be done for patients who deteriorate can lead to variation in neurological assessment and monitoring practice. The reliance on the GCS was deemed inappropriate by some because of its insensitivity to change in stroke. There was a general awareness that practice changes can be difficult with individuals liking what they are used to and systems being resistant to change.

> *"It's difficult with the GCS because it's not very sensitive for strokes but is something we use because everybody uses it."* (Consultant Nurse/Clinical Lead, Site 7).

6.5.1.2 Communal Specification

This section focuses on collective agreement about the purpose and function of neurological assessment and monitoring and how it should /will work. It contains three distinct sub-themes that came out of the analysis: Current Practice, Deterioration, and Explanations. These sub-themes relate to research questions rather than NPT.

Current Practice (23 interviews, 193 references)

The decision processes regarding practice are mainly reported as medically led. Although a few described a multidisciplinary team (MDT) approach there was a distinct lack of nurse autonomy in decisions.

> *"it will be the doctors selecting how frequently the observations need to be done."* (Lead Clinical Stroke Nurse Specialist, Site 3).

Frequency of assessment and monitoring may be pre-defined by protocols for patients receiving specific treatments, such as thrombolysis or infusions for hypertension. Most

interviewees reported having a protocol for thrombolysis patients. There was variation in the assessment scale or tool used with Stroke Thrombolysis Observation Charts (STOC), the GCS, and the Standardised Nursing Observations for Stroke (SNOBS) all being mentioned. Some variation in assessment frequency was evident though several interviewees could not recall the specifications and it was difficult to clarify the range.

> *"so it is every 15 minutes for the first hour I believe, and then the frequency decreases over time, as the patient should be less vulnerable I guess from the effects of the thrombolysis."* (Specialist physiotherapist, Site 10).

Frequency of assessment and monitoring for patients not covered by specific treatment protocols were affected by multiple factors. Some were prescriptive such as the National Early Warning Score (NEWS) recommendations, although this is mainly driven by physiological parameters, and some more subjective based on the patient's condition. Four hourly neurological monitoring was commonly advocated especially for patients deemed stable. Although one participant challenged the usefulness of this frequency:

> *"what are you measuring in 4 hourly neuro obs? What you are looking for change [sic] so if you are only looking at them once every 4 hours if you are looking for neurological deterioration you need to be doing it at much more frequent intervals"* (Physio Team Leader/ Stroke Pathway Development Lead, Site 3)

Mirroring the survey results (Chapter 5), the NIHSS was mostly completed by doctors and senior nurses on arrival to the hospital or stroke unit. Although some participants described it as being used in doctor assessments, frequency varied from ad hoc, to when changes occur, and to every ward round. Repeating the NIHSS was primarily undertaken for patients who had received thrombolysis or thrombectomy treatments. The timing and frequency of this repeat assessment also varied, though 2 and 24 hours were the most common reflecting the Sentinel Stroke National Audit Programme (SSNAP) audit requirements.

Two interviewees, at one site, reported using the SNOBS for routine monitoring. However, the non-stroke-specific GCS and NEWS were the scales most commonly used for regular monitoring with 19 out of 23 interviewees reporting use of the GCS. However, two interviewees only used the GCS at the request of the neurosurgical teams if the patient was being considered for intervention.

> *"the Glasgow Coma Scale will be the primary kind of neurological assessment to determine conscious level so to speak. And, all of our patients will have an assessment and recording of the pupillary reaction to light, and size"* (Consultant, Site 10).

There was limited information on where and how neurological assessment and monitoring was recorded. Both paper and electronic documentation were used across sites. Inconsistencies in record-keeping were voiced by some:

> *"You will find some, some documentation of it somewhere but as I say it is a bit kind of disjointed it is yes,"* (Stroke Research Nurse, Site 1).

> *"we record all the neurological assessment. Sometimes it is in the computer, we keep it sometimes to the handwritten notes"* (Specialist Doctor Stroke Registrar, Site 9).

Variation was also reported in how the patient's neurological status was handed over between staff. Some reported written handovers and others purely verbal. One participant reported bedside handovers and advocated these to ensure subtle changes were not missed. Some sites reported handing over the neurological status of all patients but more common was to only include it if the patient had deteriorated. However, some expressed awareness that they were not sure but assumed deterioration is included in handovers.

> *"I don't think it is a routinely handed over piece of information so, if there has been deterioration, I would hope that people would be handing that over in the clinical progress part. But it is not always clearly documented"* (Consultant, Site 10).

Deterioration (22 interviews, 92 references)

Approximately two-thirds of interviewees mentioned drop in the GCS as an indicator of deterioration, though there was inconsistency in terms of what score change should be the marker. Other scales changes mentioned as markers of deterioration were NEWS, NIHSS, and STOC. A small proportion expressed concerns that subtle changes are not picked up by scales and that signs of deterioration were easier to spot when there is continuity of care. However, others reported that subtle changes can be challenging to escalate because some doctors were only responsive to changes in a scale score.

> *"I think there can be very subtle changes in stroke patients that would not be picked up by the GCS"* (Physio Team Leader/ Stroke Pathway Development Lead, Site 3).

> *"they go, "Oh the GCS hasn't changed it doesn't matter". And but actually, that is quite challenging sometimes to have that conversation. Actually, it doesn't matter that their GCS is no different but they are a completely different person and that has consequences so I would say that conversation happens a lot, and half of the consultants are receptive to it and half aren't."* (Specialist physiotherapist, Site 10).

An altered level of consciousness was the most-reported change associated with neurological deterioration. However, terminology used to describe altered level of consciousness varied widely and included 'drop-in conscious level', 'unresponsive', 'drowsy', 'sleepy', 'increased confusion', 'altered mental state', and 'changes in alertness'. Other key signs of deterioration mentioned were increased or worsening weakness in the arm, leg, or face, differences in speech, and seizure activity. As per the survey, physiological signs of deterioration were also reported as important markers. Change in blood pressure was the most commonly specified, by almost half of interviewees, as the physiological parameter that indicates deterioration, followed by pulse, respiratory rate, hypoxemia, or other potential complications. A small number of interviewees conveyed an understanding of the potential link between abnormal physiological parameters and neurological deterioration as well as the importance of managing both.

> *"Yeah, because it's not always your conscious level, it can be physical deterioration, as you will know that show signs of bleed and/or changes to the brain."* (Consultant Nurse/ Clinical Lead, Site 7).

> *"work out whether they have actually deteriorated neurologically or if it is just because they're haemodynamically unstable"* (Stroke Nurse Practitioner/ Stroke Research Nurse/ Lead Stroke Educator, Site 14).

It was felt by some that abnormal test results, including scan findings, can help to identify patients at risk of deterioration, although one interviewee expressed that you cannot always see deterioration coming. Alert patients or relatives of any patient may also notice a change in themselves or others and report deterioration directly to staff.

> *"I think there will always be those patients. Because sometimes neurological deterioration is going off the edge of a cliff and you can't always see it coming."* (Physio Team Leader/ Stroke Pathway Development Lead, Site 3).

Explanation (20 interviews, 34 references)

Explanations about why neurological assessment and monitoring is important were common. However, there was a split between those aimed at patients and carers as opposed to other clinicians. For patients and carers, it was felt important to explain why it was being carried out and that a proper explanation could be reassuring to those involved. Participants spoke about patients and carers being told that timely recognition of change and action where needed was vital to maintain safety. Generally, it was considered necessary to waken patients otherwise issues might be picked up too late which would prevent intervention.

> *"so the reason we are implementing those measures and waking them up so frequently is to make sure they are safe and if the change is identified that we act on it sooner rather than later."* (Lead Clinical Stroke Nurse Specialist, Site 3).

Explanations to other clinicians were also concerned with patient safety and stability and this was described over time. Some defined monitoring as providing a picture or the story of the patient's journey post-stroke and spoke of the importance of having a baseline from which to compare and identify fluctuation. Participants reported that the overall aim of monitoring was to detect deterioration, highlight problems that need to be escalated, and improve outcomes for patients. If deterioration was noted that there should be parameters around actions to be taken, such as changing management plans or prognosis. Variation in communication was also described in the interviews, though there was consensus that plain and simple communication to explain what was happening and why is key for both staff and patients.

> *"you know picking up, rapidly picking up either an improvement or deterioration in a patient, it gives you an overall picture perhaps over a longish period of time of any fluctuation in the patient,"* (Consultant Nurse, Site 12).

6.5.1.3 Individual Specification *(5 interviews, 7 references)*

This theme focuses on individuals' understanding of what the new practice requires of them. Variation in the perceived importance of neurological scales was the main theme identified and this was often linked to staff experience, with some feeling that they could detect change just through observation because they knew their patient and knew what to look for. Others highlighted levels of variation across clinical staff stating that those less experienced did not know what was normal or what to look for because of a lack of understanding of the disease process. However, it was suggested that training could potentially improve this.

"Yes you know your patient by looking at them" (Stroke Nurse, Site 8).

"I think there is varying levels of interpretation you know of what is norm [sic], and actually even what to be observing a patient for, I think sometimes people are doing it but they actually don't know," (Advanced Stroke Specialist Nurse, Site 6).

6.5.1.4 Internalisation *(23 interviews, 88 references)*

Internalisation covers the perceptions of the values, benefits, and importance of neurological assessment and monitoring, The value seemed to be widely accepted. Its use for the timely identification of change, specifically deterioration, was purported in many interviews, especially in the hyperacute setting. There was also a strong link to safety and quality requirements. Less pronounced but nonetheless present was that neurological assessment and monitoring can be used to guide the provision of care for individual patients.

"I suppose early detection of deterioration so, making sure that we are not missing anybody that deteriorates so that we can do something about it" (Stroke Nurse Practitioner/ Stroke Research Nurse/ Lead Stroke Educator, Site 14).

The differences in opinion around its true importance seemed to rest on the rationale for monitoring and knowing what can be done, or not, for patients. If it will impact management then there was strong agreement that it should be done, as the time invested in more frequent monitoring has the potential to improve outcomes and decrease length of stay. This must be balanced against the "why bother" argument if there is going to be no active treatment or change in management despite a change in neurological status. However, there was consensus amongst several participants that even when active treatment is not an option neurological assessment and monitoring can assist in the identification and management of patients with a poor prognosis. It can facilitate the preparation of patients and their families for end of life despite the often-uncertain trajectory after stroke. Planning around patient outcomes is also identified as important for the department and organisation.

"it is a means of potentially changing a management plan, and it is identifying prognosis" (Stroke Consultant/ Consultant Geriatrician, Site 5).

"for the organisation to understand how important the neurology examination is in determining the patient's outcomes. So, some patients stay longer, some patients stay

lesser, so it is impact on the department as well as the whole system so yes it is very important" (Specialist Doctor Stroke Registrar, Site 9).

6.5.2 Cognitive Participation

Within NPT this covers the relational work that people do to build and sustain the practice.

6.5.2.1 Initiation *(7 interviews, 10 references)*

This theme focuses on the identification of key individuals who drive neurological assessment and monitoring forward. Generally, neurological assessment and monitoring were seen as a team practice. However, nurses at the bedside were often reported as reluctant to instigate neurological monitoring without a senior decision, often consultant led. This is a particular issue in some units, where there is no 24-hour senior cover, meaning that patients may not be assessed or monitored until the next day. Senior nurses (ward managers and stroke specialists) or therapists may help drive the decision for increased frequency of monitoring for some patients. Experienced senior team members were also involved in teaching others knowledge and skills needed.

"getting the stroke consultant on call… to be getting like that senior advice because I mean I don't want to be saying as a nurse." (Deputy Charge Nurse, Site 2).

"nurses won't necessarily instigate neurological monitoring, it would need to be asked of them to step that in …. and that is a time of day of arrival thing as well. If the patient is admitted at 4 o'clock in the afternoon, they have missed our consultants for the day, they won't be reviewed formally by a stroke consultant until the next morning" (Physio Team Leader/ Stroke Pathway Development Lead, Site 3).

6.5.2.2 Enrolment *(16 interviews, 17 references)*

This looks at whether people agree that neurological assessment and monitoring should be part of their work. It includes both engagement with and reluctance towards the practice. Overall, it was agreed that neurological assessment and monitoring were an everyday part of the work of the unit. Key phrases around acceptance and "needs to be done" were repeated, potentially indicating that in some cases it could be more ritualistic than applied practice. The keenness of individual clinicians, as well as experience and comfort with the practice, were professed to assist engagement. Reluctance, especially about selected intensity, stemmed from it being deemed a heavy workload and intrusive to patients.

"it is part of your job if you don't do it I think ((laughs)) yes you basically fail at your job." (Trust Stroke Registrar, Site 11).

"it is quite intrusive, erm… to be on you know for example hourly neuro obs for 48/ 72 hours. That is quite hard going for the patients." (Consultant Stroke Medicine, Site 14).

6.5.2.3 Legitimation *(23 interviews, 116 references)*

Legitimation explores the governance procedures that surround neurological assessment and monitoring practice. Protocols were more prevalent for specific stroke types and when individuals were receiving certain treatments. Twenty participants stated their site had a protocol for thrombolysis, eight for thrombectomy with another in development, six for ICH patients with one being written, and one for patients with large MCA infarcts. Written protocols also existed for blood pressure management and other specific characteristics. There were no protocols reported by participants to guide assessment and monitoring practice for all patient groups. Four units talked about improved guidance for HASUs, but no detail was given as to what this included. One interviewee spoke of using the NEWS protocol, despite this mainly being based on physiological parameters. Only two interviewees stated they had a protocol about what should happen if deterioration was noted.

"When it comes to the thrombolysis protocol it does give you the how often and how regular we need to be doing observations for the patient. For the other patients that I spoke about, that potentially wouldn't get there within the 4 hours but will still be in the acute phase there is no actual protocol to follow." (Sister/ Team Leader, Site 11).

Assumptions and vagueness on what neurological assessment and monitoring is required were noted. Representatives of several units admitted that nothing was written down. One stated they were sure everyone got the same as it was "engrained" whereas others talked about using "gut instinct" or "gauging" as to how regularly monitoring should occur. One highlighted there would be nothing to demonstrate what should or does occur if organisations such as the Quality Care Commission (CQC) audited. Another unit reported the development of a draft standard operating procedure after completion of this study's SNOBSS survey (Chapter 5) as they realised previously it was just assumed monitoring was done.

"not really a specific set remit just kind of gut instinct of those patients" (Stroke Nurse Practitioner/ Stroke Research Nurse/ Lead Stroke Educator, Site 14).

There was also inconsistency in awareness of National guidelines in this area. Although 10 of the interviewees knew that there was no general guidance others assumed there was, or even presumed that their practice was international standard. There was consensus that there should be specific guidelines on how to complete neurological monitoring, who to do them on, and how frequently. Several participants agreed that there should be extended protocols to include all patients, even the unusual ones. Guidance that could be adapted to the patient's requirements and context of the unit would be preferable to participants in this study than the current "case by case" processes. One commented that guidelines in clinical practice were often commenced informally before being formally written and agreed upon due to the time to get things passed through governance procedures.

"I don't think we have any other guidelines, specifically on how to do neurological observations, who to do them on, and how frequently to do them on, to be honest." (Consultant, Site 10).

"Yeah, I think I think there is something about, … guidelines are great, you know, love them. But they have to work within the context of where you work" (Senior Clinical Lecturer and Honorary Consultant Physician, Site 1).

6.5.2.4 Activation *(18 interviews, 38 references)*

This theme looks at whether people work together or not and also highlights pathway issues related to neurological assessment and monitoring practice. Despite consensus that monitoring was undertaken, units had different strategies for instigating it and ensuring completion which resulted in different team and pathway challenges. Teamwork and working together were identified as important by many. Having a common agreement about what should be done, and good communication were acknowledged as central to ensuring effective neurological assessment and monitoring practice. Team dynamics varied across units with some describing partnership working between all team members and others challenging medical decision making. In most units decisions around neurological monitoring were made by medical staff. Some nurses were described as lacking confidence in instigating monitoring. In one unit, monitoring was not started unless specifically requested and this caused problems out of hours when there was no medical cover.

"think it is probably a joint partnership if I am honest with you." (Consultant Nurse, Site 12).

"we do challenge each other, and we do talk about that, and I have a really good relationship with our consultants to feel that, that challenge." (Physio Team Leader/ Stroke Pathway Development Lead, Site 3).

Intuitive work practices in well-established teams were discussed, and some believed that this meant they pick up on change in a more enhanced way than assessment alone permits. However, there was an awareness that not all staff had those skills, especially junior staff, and that formalised pathways should be developed. Pathways were understood as means to standardise care, establish patient safety, and ensure appropriate treatments are provided.

"how we work as a team and how in tune we are with what our patients are doing, that makes that monitoring more enhanced, more than you could ever write down I think…. our team are so intuitive at managing these patients that they know before they have written it down that the patient is deteriorating and needs a review…. but I think the pathway approach would strengthen that even more and give us some backing that that's what we are delivering…. if we want to standardise how we care for our patients, and make sure that patient safety is paramount, and every patient gets the right treatment, for them, I think there is a need for pathways…. the pathway is needed because we can't have variation based on things like time of day of arrival." (Physio Team Leader/ Stroke Pathway Development Lead, Site 3).

Whether pathways can be expanded beyond the stroke unit is unclear. Issues often occur outside the stroke unit and relationships need to be built with external teams, especially ED, neurosurgery, and radiology. Generally, it appeared that stroke-specific assessments do not occur until stroke teams were involved. In some areas, it was a conscious decision not to burden external teams with stroke-specific assessment as it maintained more conducive working relationships. However, not having a joined-up pathway across the hospital risks missing changes in neurological condition, especially during transfers. One interviewee described how they had attempted to stop using the GCS across the whole stroke pathway but ended up going back to it, including on the unit, as that was what other areas use and understand. One participant questioned whether monitoring recorded on paper records was perceived as less important than that recorded on electronic systems.

"the ones [medics on call] that are on overnight on their own often… unless they have done a rotation in our trust already they will be like what the hell is this piece of paper like… and maybe there is less importance to it because it is not electronic, whereas they are used to everything else being electronic."

(Stroke Nurse Practitioner/ Stroke Research Nurse/ Lead Stroke Educator, Site 14)

6.5.3 Collective action

This section of themes is around the operational work that people do in relation to neurological assessment and monitoring practice and highlights the barriers and facilitators discussed within clinical practice.

6.5.3.1 Interactional workability *(23 interviews, 261 references)*

This theme explores whether staff and patients can perform the tasks required by them. Multiple factors that influenced neurological assessment and monitoring being done effectively were identified. Staffing levels, particularly shortages, was the most commonly reported factor. The requirement for higher staffing to support enhanced monitoring in HASU is well established but this was not always deliverable. Without adequate staffing resource, there were consequences in terms of what can be achieved. With adequate staffing, assessments could occur at required frequencies, on time and this was seen to impact both quality of care and patient safety. Higher staff levels enhance the potential of continuity of care, which a few identified as optimal as it allows deterioration to potentially be seen straight away without repeated formal assessment.

"cutting the cloth to suit what we had, we thought, well we need to just be very mindful what we are asking people to do, given the nursing resource that we have."

(Senior Clinical Lecturer and Honorary Consultant Physician, Site 1).

As well as staff numbers, skill mix was also considered important, and the need for experienced staff regardless of grade was emphasised. Experienced staff are more likely to pick up on subtle changes, complete assessments correctly, escalate change, and generally reduce risk to the patients. One felt that speed of assessment improves with experience. Staff rotation to other wards has negative consequences in terms of less experienced staff looking after stroke patients. Senior support for decision-making can vary as many units do not have twenty-four-seven stroke-specific medical cover.

"So not just staffing levels but actually stroke trained staffing levels. Experienced staff you know. Not that you know [sic] junior staff can be trained, but if they pluck

somebody out of a different ward and throw them into stroke for the day just for the crack, ((laughs)) it is not as helpful as having a very experienced stroke nurse that just knows there is something wrong with their patient before they even do their observations." (Stroke Nurse Practitioner/ Stroke Research Nurse/ Lead Stroke Educator, Site 14).

Other resources that impact what can be achieved were identified. The clinical environment in terms of physical space, acuity range of patients, and other demands on staff time all impact delivery. The work is seen as time-consuming by many and can cause pressure on staff when increased frequencies are requested.

"time factors and demands on the girls' time, you know for actually being able to do the frequency of monitoring that is required because of the busy ward and there is a lot of patients, and they are understaffed, and there is sick leave, and you know all of those factors would definitely have an impact on people being able to carry out the frequency of monitoring that is required." (Advanced Stroke Specialist Nurse, Site 6).

Training was considered vital in terms of ensuring neurological assessment and monitoring was understood and completed correctly. Despite the desire for ongoing continuous renewal of training this often could not happen because of staff shortages, lack of time or funding to train, and more recently the pandemic preventing training. Online training and certification were encouraged particularly for the NEWS and NIHSS. Certification in the NIHSS was expected but was often not checked unless there was a research requirement.

"I think they should have it, would be ideal if they could have an ongoing rolling programme, maybe once every 3 months, where they have the, are given the time, to go over neurological assessments with the patients. Unfortunately, because of staffing levels and time constraints that is not possible," (Stroke Specialist Nurse, Site 8).

"the priority should be that you are up-to-date with your training for clinical reasons and then indirectly it means you are up to speed for the research but actually it is the research that seems to lead when people update their training." (Senior Clinical Lecturer and Honorary Consultant Physician, Site 1).

Informal training was claimed to occur at many sites though few interviewees could be sure and stipulate what is involved. Experiential learning on the ward seemed common, many

advocated its importance and value in terms of then being able to effectively perform neurological assessment and monitoring. Ensuring competency was discussed as an essential principle of practice but very few units had formal competency assessments in place. One unit had developed a competency assessment, but it had not been implemented, mainly due to the pandemic, and another had one in development.

> *"the staff that I work with I have shown them how to do it. Kind of more on an informal basis"* (Advanced Practitioner Role-Physio, Site 13).

> *"I think we are working on something for that but there isn't one yet."* {Competency} (Deputy Sister, Site 12).

It was understood that commonly used monitoring tools, especially the GCS, did not pick up neurological change well. There was also agreement that the subjectivity of assessments leads to wide variability in assessment process and outcome. Despite these misgivings, a structured approach to assessment, potentially involving protocols, was supported by several interviewees.

> "*I mean ultimately GCS is such a rudimentary tool, that doesn't really pick up much of the neurological change that happens in a lot of our stroke patients."* (Consultant, Site 12).

> *"GCS can be a bit subjective, when people, different people get different values when they assess, it can be quite subjective that is the only hindrance I can think of… also depends on the level of experience. I have also seen a lot of subjectiveness assessments NIHSS sometimes, I would see exaggerations."* (Stroke Speciality Doctor, Site 9).

There was ambiguity in terms of when monitoring should be discontinued. Some had set criteria, e.g., leaving the HASU, whereas others had no set pattern. Two interviewees described how unless they were prompted to stop, monitoring was continued unnecessarily. Another would carry on for longer if staffing levels would allow. Handing over monitoring results and patient status, especially if change had occurred, was generally advocated as it could allow better prioritisation of workload. This could be in verbal, written, or one-to-one formats. However, there was an assumption this was not done by all staff.

"I sometimes worry that we do observations for the sake of doing observations." (Consultant Nurse, Site 12).

"I'm not quite sure everybody always does that, hands it over, if I'm honest." (Consultant Nurse/ Clinical Lead, Site 7).

Monitoring was generally regarded as intrusive to the patient which could cause variation in whether it is completed. Although the majority expressed that monitoring should be completed even if that involved waking the patient, a few expressed it as potentially inappropriate. It is possible that some staff, in some units, avoid waking patients even if frequent overnight neurological monitoring has been requested. There was some awareness that interrupting sleep affects the assessment itself and that for some individuals or groups intensity of monitoring is inappropriate. One unit reported they were selective about who gets a higher intensity of monitoring.

"they are getting woke up every hour, they are really, really tired, you are going to feel like the neurology is worsening but it is because they are overtired, their brain needs time to rest and we are waking them up every hour, they are not getting that rest so I feel like that can impact it." (Stroke Nurse, Site 8).

6.5.3.2 Relational Integration *(21 interviews, 78 references)*

Relational integration looks at how staff trust each other's work and expertise. Trusting each other's work and expertise in neurological assessment and monitoring was highlighted as an issue for participants. The subjectivity of assessment was recognised. Intrinsic variation in the way assessments, including both the GCS and NIHSS are applied across both individuals and professional groups was also acknowledged by participants. This was considered particularly problematic with less skilled members of the workforce both in terms of variation in interpretation but also in knowledge around why they are completing it.

"different people get different values, when they assess it can be quite subjective" (Stroke Speciality Doctor, Site 9).

Experience and skills were important for the development of trust. There was some disparity in terms of trust related to professional groups. Some medics were not as confident with nursing staff assessments whereas others preferred experienced nurses over junior doctors. However,

experience and trust in the individual was not linked to professional grouping. Some interviewees reported confidence in healthcare assistants (HCAs) as they were with patients all the time. Bank and agency staff as well as rotating junior medics were not as trusted as established stroke team members.

> *"what I would feel more comfortable with is a good, trained nurse speaking to me, than necessarily a rotating medic you know that probably wouldn't have the same access to, you know the assessment skills, or the sort of language that we talk you know."* (Stroke Consultant/ Consultant Geriatrician, Site 5).

The development of trust can also be based on the experience of working together as a team. There seemed to be a great degree of comfort and confidence that comes with effective and prolonged teamwork. Within static teams, the continuity allows the development of good working relationships which brought confidence in others' skills and competence.

> *"just because we have all worked together so long, you know that's what we all do"* (Consultant Stroke Medicine, Site 14).

Effective training can also be seen as a precursor to trust development. One interviewee felt that doctors tended to learn assessments and then may digress from the standard assessment whereas nurses may be more prescriptive. Individuals, including experienced nurses, sometimes lack trust in their findings and will either repeat the assessment or get others to check their interpretation before escalating change. Despite the awareness of variation and subjectivity there generally was confidence that changes are being identified and reported.

> *"so if I am concerned that the way I am scoring a patient might be different to somebody else, I will ask somebody to come and do it with me just so I can see if there is an agreement, if that makes sense."* (Deputy Sister, Site 12).

6.5.3.3 Skill set workability *(22 interviews, 101 references)*

This theme is about who does the work involved and how it is allocated in relation to neurological assessment and monitoring practice There was a tendency for the professional groups responsible for assessment to determine the scale being used. Doctors and senior nurses (band 6s, clinical nurse specialists, and nurse consultants) mainly completed the NIHSS, and ward nurses and HCAs primarily undertook the GCS, although this varied across units. In

three interviews the GCS was only completed by band 6 nurses, and in one was only used by doctors to assess deterioration. One interviewee voiced that the reliance on seniority to complete the NIHSS was de-skilling the junior members of the nursing workforce. There was some discussion that nursing staff are as competent if not more so than some medical staff at NIHSS completion.

> *"Yes, they would either be a very senior nurse or well, probably not so much one of the junior medics because I don't think they are as skilled at doing it if I am honest. I think I think it just takes a heck of a lot of practice, and understanding of what you are really looking for, and I am not sure that the medics have always got it if I am honest."* {NIHSS} (Consultant Nurse, Site 12).

There were further points made about professional group responsibility in neurological assessment and monitoring. Thirteen interviews identified that nurses take the responsibility of regular monitoring and escalating deterioration identified. In a couple of cases, this was described as their [nurses'] "burden" and another spoke of it being nurses' responsibility to defend patients from unnecessary assessment burden. Doctors were generally reported as having the role of further assessment when changes were noted in clinical condition and leading decisions on how to manage deterioration. However, in some units out of hours, where stroke-specific medical cover is not available, specialist nurses took on these roles. Therapists although not formally completing regular assessments and monitoring were recognised for their skills in noticing changes in patients.

> *"Assessment of GCS, pupillary reflexes will be done by the nursing staff or the care support workers on the hyperacute unit…. the assessments by the doctors will probably happen as and when they might be asked to see patients because of changes in clinical condition"* (Consultant, Site 10).

> *"I think sometimes the therapists too can be quite good. This person is flat today, she is not really engaging, they flag that up"* (Stroke Speciality Doctor, Site 9).

Experience was seen as integral irrespective of what professional group staff belonged to. The importance of a stroke specialist team managing care was highlighted. When less experienced staff are involved, it was perceived to cause issues and increase risk. In some places, experience was linked with hierarchy of staff in terms of professional group and grading. Whereas in others, experience and skill was considered in relation to training and exposure,

with HCAs relied upon to complete the neurological assessment and monitoring. One participant voiced that any member of staff could undertake neurological assessment and monitoring with the right training.

> *"if you are reliant on locum agency staff, people who aren't inducted, relatively junior members of staff who are being tasked to do something that they are not familiar with, that presents a concern to me."* (Stroke Consultant/ Consultant Geriatrician, Site 5).

6.5.3.4 Contextual Integration *(19 interviews, 69 references)*

This theme discusses whether participants feel adequately supported by their organisation around neurological assessment and monitoring practice. Participants were asked about whether they felt supported in terms of being able to complete neurological assessment and monitoring. Generally, they described that support came from within the stroke team itself rather than the wider organisation. Clinical stroke service leaders, especially those passionate about services, were seen by many as integral support. Perceived support from organisations varied. Some reported feeling supported whereas others felt that senior management did not understand the speciality and its specific challenges. It was felt that senior organisational management often compared stroke units with care of the elderly wards despite stroke patients requiring more intensive care. Despite the unpredictable nature of admissions and acute interventions within stroke services staff are regularly moved to other areas. This can cause tension between the stroke service and the organisational management teams.

> *"I think management support for stroke isn't always there because stroke is more difficult than a lot of other aspects of medicine. It is more resource-intensive, it is much less predictable."* (Senior Clinical Lecturer and Honorary Consultant Physician, Site 1).

Organisational management styles affected how supported teams felt. With a reactive management style, staff reported feeling unsupported. Even where stroke was seen as a key priority, general bureaucracy and management scrutiny can cause serious delays in things being agreed and actioned. Resource issues were highlighted especially concerning the management of deterioration. This included access to funding, scans, and specialist staff especially out of hours.

> *"The other problem we have is that our consultants out of hours are only commissioned to be called about thrombolysis and thrombectomy. So although they do accept exceptions and we do contact them for patients that are deteriorating and are*

not candidates for thrombolysis or thrombectomy they are not paid, they are not commissioned" (Interview 18).

Organisational support was also needed to address practical issues such as the provision of electronic systems. This varied across units, but one reported they have been waiting three years for neurological parameters to be added to their system.

The structural layout of the unit itself can impact the delivery of neurological assessment and monitoring. One interviewee praised individual rooms as being conducive and preventing disturbing other patients whereas others placed those that required frequent monitoring together to facilitate better oversight.

"in individual rooms, so you can go in and shine lights in their eyes and make noise without waking up everyone else on the ward and that is a that is a big advantage" (Interview 16).

Although most interviewees mentioned training it was variable across sites and professional groups depending upon the organisational setup. There was little understanding across professional groups about what training others receive. Access to training can be inconsistent and there was a lack of competency assessments. Resource issues relating to releasing staff for training were also common. Staff were often reliant on in-house or experiential learning, and some had no formal training either because it was not available or not deemed part of their professional remit.

"The nursing staff themselves might have separate kind of training programmes which they do." (Consultant, Site 10).

6.5.3.5 Patient Groups *(23 interviews, 132 references)*

This theme is not derived from NPT but was created from inductive engagement with the data. It specifically explores what specific stroke types or characteristics impact on neurological assessment and monitoring practice. There was a range of opinions, however, the key groups identified as needing more frequent monitoring were those receiving thrombolysis or thrombectomy, those with ICH, and those who have had large infarcts, particularly those at risk of developing malignant middle cerebral artery (MCA). This perceived requirement for increased frequency appears to be guidance-driven in that if there were set protocols or management plans for what to do if deterioration is noted then these protocols supported and encouraged monitoring.

> *"we probably wouldn't do it as frequently as we do for patients who might be a candidate for neurosurgical intervention, or other kinds of intervention."* (Consultant, Site 10).

> *"try to be reasonably selective of that, in that there is not a huge amount to gain in intensive monitoring of someone who you are not able to do anything about if things got worse".* (Consultant Stroke Medicine, Site 14).

Additionally, within the ICH group, specific indicators of risk or signs of deterioration such as large bleeds or headaches and nausea were mentioned as factors that might impact decisions about increasing frequency of monitoring. Less common groups, that were felt to warrant increased monitoring, included those with fluctuating or stuttering symptoms, or who had a staggered onset, younger patients, and those with seizures. Stopping or reducing the frequency of monitoring was only described for palliative patients. Although some interviewees would stop monitoring in this group others continued a reduced frequency to understand the patient's trajectory allowing them to keep relatives informed.

> *"it can be intrusive.…what is the benefit of actually monitoring that patient, his comfort care would be the priority there"* (Advanced Stroke Specialist Nurse, Site 6).

6.5.4 Reflexive monitoring

Reflexive monitoring encompasses the appraisal work around neurological assessment and monitoring practice.

6.5.4.1 Systemization *(23 interviews, 88 references)*

This theme includes all the information that people collect around the impact of neurological assessment and monitoring practice. The main impacts reported by participants were concerning the actions that occur if deterioration is identified. As previously mentioned, a drop in the GCS was cited as the most common trigger for action. The change in score that would elicit a response varied, between 1 and 4 points. The key action described was escalation to medical staff. This occurred either directly or via senior stroke nursing staff dependent upon the unit and whether the deterioration occurred in or out of hours. In a small number of sites, action sometimes involved requesting additional assistance from critical care outreach or even making a peri-arrest call.

"the way it is escalated is different if it is in-hours when there is a medic on the ward, and if it is a mild deterioration the medics on the ward will see the patients first. If it is a severe deterioration they will still be, well obviously asked to come and review the patient but then a peri-arrest call would be activated as well. And the consultant may be asked to come to the ward as soon as they can if the patient is deteriorating quickly" (Lead Clinical Stroke Nurse Specialist, Site 3).

Other actions mentioned were re-scanning patients and increasing the frequency of monitoring. The decision to re-scan appeared to vary across sites with some advocating it in all cases, whereas in others it would only be warranted depending on the patient and specific situation. One interviewee mentioned that this might be the time that the decision to start end of life care might happen negating the need for further scanning. Increasing frequency of monitoring could be guided by either the medical review or linked to the NEWS score guidance. Over half of respondents talked of repeating the assessment to check that it was a true deterioration. For some, this was prompted by looking for any patterns or fluctuations whereas for others it was linked to confidence in others' assessment abilities and checking whether there was actually a change.

"So, I would, first of all, establish that there is change, is it fluctuant, is it static, is it on a trajectory towards erm… severe impairment" (Physio Team Leader/ Stroke Pathway Development Lead, Site 3).

"sometimes when they do a GCS it is so off. I will say they have got a GCS of 12 and someone will say it is 9 on the ward and I am going no, no, no and you have to go through it" (Stroke Research Nurse, Site 1).

6.5.4.2 Communal appraisal (3 interviews, 4 references)

Communal appraisal is about whether people collectively evaluate neurological assessment and monitoring as worthwhile. There was variation in whether it was evaluated as being worthwhile. One medical consultant discussed the disempowerment felt when members of the team repeatedly completed neurological monitoring and instigated escalations only to be told no response was required. Another participant felt that regular monitoring was worthwhile as the regular attendance at the bedside meant patients' other nursing needs were being attended to. Importance was linked to clinical incident reporting with success being few

or no incidents in which neurological assessment and monitoring had been missed or completed incorrectly.

> *"a success is that observations were carried out at the right time, and things were escalated because I have been involved with SAIs (Serious Adverse Incidents) whereby things weren't escalated."* (Stroke Consultant/ Consultant Geriatrician, Site 5).

6.5.4.3 Patients and Carers *(23 interviews, 58 references)*

This additional theme, derived from the data and not NPT, includes the opinions of participants on whether patients and carer evaluate neurological assessment and monitoring as worthwhile. Most interviewees felt that patients and carers found monitoring reassuring despite it also being reported as disruptive, disturbing, frustrating, or similar for the individual. Participants reported that severity of stroke and frequency of monitoring impacted tolerance by patients. The importance of communication was strongly advocated to increase patient understanding and tolerance, but it was also accepted that communication could often be improved. Regular disturbances will effect patients' sleep and potentially the assessment itself but this can vary between individuals as being disturbed and the effect of sleep deprivation impact on some more than others. Only one participant talked about not disturbing patients who were asleep. Overall, there was reported acceptance by patients, but some became agitated, affecting compliance, or refused altogether.

> *"We just talk about close monitoring and looking for any change and acting on that if it happens. So there is probably an element from a carer's perspective and patient that they don't really know what we are doing. Just that we are being reassuring."* (Physio Team Leader/ Stroke Pathway Development Lead, Site 3).

> *"for some people obviously that is just discomfort, a bit of sleep deprivation. For some people, that's a lot though you know if you are a bit delirious and confused and feeling a bit paranoid, then that has a really negative impact on you that kind of thing."* (Consultant Stroke Medicine, Site 14).

6.5.4.4 Individual appraisal *(3 interviews, 4 references)*

This theme focuses on whether individuals evaluate neurological assessment and monitoring as worthwhile. Participants appreciated neurological assessment and monitoring practice because it identifies and allow reported of deterioration. However, there was variation in how

they felt deterioration should be identified and communicated. Two felt that criteria set out in standard operating procedures (SOPs) or set score changes were the important markers. The other felt change was about more than a score.

> *"I would get phoned by our nursing staff, you know about patients… but again some of that might just be the conversation that we have and if they are telling me, that it is a left, [sic] I don't know I like to get a score on the door, you know so a score on the door is the most important thing."* (Stroke Consultant/ Consultant Geriatrician, Site 5).

> *"the fact that it is actually the change rather than the actual numbers that are important."* (Consultant, Site 10).

6.5.4.5 Reconfiguration *(22 interviews, 129 references)*

This final theme collates all the suggestions for modifications or improvement in neurological assessment and monitoring practice based on their current evaluation. There were numerous suggestions to improve neurological assessment and monitoring practice. Several interviewees reiterated that it should be for all patients across the whole pathway and not just those receiving hyperacute treatments. National standards/guidelines with better indications of what to do, when, and for which patients were considered key to reducing variation. However, there was also support for adaptation to the needs of the specific service. Response to deterioration is important and the use of a protocol was advocated to ensure an appropriate plan was instigated when change was noted.

> *"and also, if there was a protocol for post-stroke monitoring but you know for those patients that are not receiving hyperacute treatment. One just for the standard, if there is such a thing as a standard stroke."* (Advanced Stroke Specialist Nurse, Site 6).

A few interviewees advocated wider use of the NIHSS in practice. However, many felt that current scales were not appropriate, with two specifically stating that scales can miss subtle signs of deterioration and three wanted simpler, easier to use assessments. Better communication and documentation around neurological assessment and monitoring was suggested, but there were conflicting opinions about whether to use electronic systems. Some cited they were useful for reminders of when observations were due and automatic interpretation and escalation could reduce clinician workload. Others argued that they remove clinical interpretations skills and can cause additional workload by escalating unnecessarily.

"The scales that we use and is there scope to make, is there scope for something that is appropriate" (Physio Team Leader/ Stroke Pathway Development Lead, Site 3).

"Make it simpler" {neurological assessment and monitoring} (Deputy Sister, Site 12).

"one of the potential benefits of an electronic system is that it would trigger an alarm so for the less experienced member of staff, less experienced members of staff there is an automatic trigger." (Stroke Consultant/ Consultant Geriatrician, Site 5).

The additional workload of repeated assessments was seen as a key barrier especially if staffing levels were inadequate. Recommendations to overcome resistance to change included using quality improvement and change management strategies. Suggestions included using patient and family feedback, exemplars of best practice, and the communication and education of the rationale for change. Learning from incidents was also described as being able to drive change.

"patient stories are really big in terms of families' feedback, and patients' experience. And then unfortunately incident reporting. As soon as there is more than one incident related to something you can guarantee management will jump on that and allow you to put some time and effort into some sort of change management." (Specialist physiotherapist, Site 10).

Education and training were the most talked-about elements. They are recognised as crucial to ensuring change is successfully implemented, despite potential challenges with resources including releasing staff time. Comments fell into two broad themes: what should be included and how it should be designed and delivered. Suggestions for inclusion: the rationale behind assessment allowing clinicians to understand why it is important and empower those involved, how to do the assessments and ensure standardisation, especially with items that are perceived difficult to assess such as cognition and level of consciousness, and recognition of deterioration and what changes mean.

"having some clarity as to what the point is in general, and perhaps particularly if you are doing more intensive observations what the point is specifically for this person and why it is worth doing." (Consultant Stroke Medicine, Site 14).

> *"I think one thing that I find really difficult to interpret is people's assessment of conscious level which obviously makes a big difference doesn't it."* (Consultant Stroke Medicine, Site 14).

There were multiple suggestions around how training should be delivered from national competency and training standards, to learning through experience. One interviewee felt that all staff working in stroke should have stroke-specific mandatory training like other specialities such as critical and coronary care. Tailoring was suggested so that staff who might need more input e.g., newly qualified or overseas nurses can receive it. Interprofessional training was also suggested so that individuals learn from one another as well as other professions. There was a split between preferring face to face, where it is easier to motivate and ensure understanding, and online training, which is easier to access and could be available any time. Key agreements were that the training should be repeated regularly and that competencies should be achieved and maintained. Other suggestions to encourage and ensure participation were accreditation, auditing, and the use of an educational facilitator to make sure it happens.

> *"if we had access to kind of ready-made training programmes, on how to do neurological assessments what changes in neurological assessments mean, that the staff had time to do as well, to keep their competencies up I think that would be great too."* (Consultant, Site 10).

> *"I think they should have it would be ideal if they could have an ongoing rolling programme, maybe once every 3 months, where they have the, are given the time, to go over neurological assessments with the patients."* (Stroke Specialist Nurse, Site 8).

6.6 Discussion

The aims of these interviews were to determine knowledge, understanding, and acceptability of neurological assessment and monitoring after stroke and to explore the barriers and facilitators to implementing a change to practice. Extensive data were obtained, some provided validation of the survey findings (Chapter 5) such as the different scales used and the reasons for use. Other data provided more in-depth explanation of factors impacting on neurological assessment and monitoring practice including those identified in the survey discussion as needing more exploration. Multiple impacting factors were identified across individual, unit, and organisational levels so it is not possible to discuss them all. Fundamental resource issues that impact practice such as staffing levels are acknowledged but this

discussion focuses on factors that were evident across multiple coding themes, indicating potential importance and impact, or those that showed disparity.

Overall, neurological assessment and monitoring is reported as being an important part of stroke care however there were multiple differences in opinions, attitudes and beliefs that could be impacting on variation in practice. The interviews highlighted that clinicians felt that neurological assessment and monitoring should be done but that there are a complex range of factors that impact on its completion. Although it is generally agreed neurological assessment and monitoring should be completed, several participants indicated an assumption that it was which might indicate further variation in practice than reported. All participants provided suggestions for improvements in practice.

Experience of staff was reported to impact multiple aspects of neurological assessment and monitoring practice despite inconsistency in terms of what constitutes experience. Experience is generally considered important for both individual clinicians and unit teams. In some places, experience was linked with hierarchy of staff in terms of professional group and grading. However, overall perceived experience appears more important than profession in the development of trust in others' abilities. Many felt that experienced staff bring better understanding, in terms of knowing what to look for and assisted with engagement with neurological assessment and monitoring as well as reducing risk. However, others felt that experience could breed complacency increasing variation in assessment and potentially decision-making. The disparities in terms of experience and the associated trust could come from a mismatch that can arise between having the knowledge and skills to complete the intervention and understanding the theory behind its completion (Benner, 1982).

Perceived experience will depend on the role undertaken. From the data, it was clear that doctors are the primary decision-makers in both monitoring frequency and actions if deterioration is noted. However, variation in access to medical staff especially out of hours can cause variation in care provision. Some units have addressed this by allowing other members of the MDT to work collaboratively with the doctors by assuming complementary roles and sharing responsibility for making decisions to formulate and carry out plans for patient care (Fagin, 1992). This collaborative approach is only possible where teamwork is valued and there is mutual trust and respect (O'Daniel & Rosenstein, 2008). Trust can be earned from others using skilful communication and showing confidence in your actions and decisions (Markley & Winbery, 2008). However, this may be harder to establish and maintain with external teams

especially as prolonged working together was seen to increase confidence in others' skills and competency.

Communication was described as key to effective teamwork and collaboration and was seen to lead to an improvement in decision making and risk reduction. However, it is known that communication, collaboration, and teamwork do not always occur in clinical settings (O'Daniel & Rosenstein 2008). Poor communication can have profound impacts on practice and has been linked to adverse clinical events and outcomes (Interprofessional Education Collaborative Expert Panel 2011; Sutcliffe et al., 2004). Team structure will influence communication, but it can also be influenced by a range of social, relational, and organisational structures (Rabøl et al., 2012). Neurological assessment and monitoring like most healthcare practices relies on the communication of information. This can occur in multiple forms, such as verbal or written charts and documentation, in real-time or not (Conn et al., 2009). Communication needs to be effective both between clinicians and with patients or carers. Despite a clear understanding of its importance wide variation is reported within and across units.

Effective communication relies on healthcare literacy, cultural competency, and removal of language barriers (Ranta, 2019). Without a common understanding of the aims and responsibilities of the practice, communication can be limited (Rabøl et al., 2012). It was felt that the development of a common agreed plain and simple language would enhance good communication. Also, the practice and use of standard procedures or frameworks could support improved communication (Foronda et al., 2016). The development of a shared understanding and a culture of working together has the potential to improve care provision and patient safety (Green et al., 2017).

Good record keeping is part of effective communication and is a fundamental part of delivering safe patient care (Norris, 2009: Royal College of Nursing, 2017). Variation in written documentation processes was highlighted through the data, including mismatch of electronic and paper-based methods. Standardisation and agreement of methods have the potential to reduce this variation. Consistent documentation including relevant findings decisions made and actions agreed is recommended (General Medical Council, 2013). However, this data supports other literature that clinicians often fail to provide comprehensive information in an accurate and timely manner and that this should be addressed (Bjerkan et al., 2021; Vermeir et al., 2015).

Documentation might also include actions to be taken if deterioration is noted however the use of this strategy was not reported by many individuals. This could be because the main

action if deterioration is noted, is escalation to medical teams for further assessment and decision making. Although response to deterioration is seen as crucial there is real inconsistency in what constitutes deterioration. The main indicator reported was related to a change in the LOC although the communication and assessment of this showed wide variation. There was limited agreement on what change, and level of change should warrant escalation and potential intervention with a range of GCS score changes being quoted. Little was said about early stroke specific signs of deterioration. The widespread mandated use of the NEWS directs focus onto physiological parameters and less focus on specific disease clinical assessment (Nielsen et al., 2022). The NEWS has been shown to help identification of deterioration and improve communication with medical teams in other areas (Mohammmed Iddrisu et al., 2018). However, the organisational priority of using the NEWS could be in conflict with providing important condition specific monitoring, especially where there are staffing issues.

Where scales are used, stroke-specific or otherwise, there is common understanding of the subjectivity based on the individual doing the assessment. The data also indicates that there is limited knowledge about subtle changes in patient condition and the links between physiological and neurological deterioration (Helleberg et al., 2014). Despite all the potential issues around identification of deterioration, there was voiced confidence that changes are recognised and reported.

Recognition of deterioration and what the changes mean was seen as an important aspect of training that clinicians should receive training as it is deemed important for individuals to understand what they are doing and why. Current training was described as inconsistent despite its perceived value and importance. Although the project did not gather in-depth information about the structure of current training it was understood to include both formal and informal elements. Training was viewed as being able to overcome differences between professional groups and integral to the development of trust between individuals as it can address gaps in experience. Training provision was affected by resource issues such as funding and staff shortages. Still, participants recommended that training is continuous or regular and leads to the development of competency, though there was a reported lack of competency assessments across units.

There was a suggestion that national standards and competencies supported by interprofessional training would be beneficial. This supports other literature promoting standardised training that is interactive and multidisciplinary (Jones et al., 2018). National

minimum training standards have been adopted in other healthcare specialities (Skills for Care & Skills for Health, 2013).

Although most participants describe neurological assessment and monitoring as an important and integral part of stroke care whether it is valued as worthwhile seems to not be as clear-cut. There is an awareness that it is time-consuming and an additional workload especially if staffing levels are inadequate. Although generally reported as an established element of care, for some there was an underlying sense that it is only useful when something can be done. Specific patient groups, especially those with protocols were prioritised over others. If the value of completion of neurological assessment and monitoring practice was established for all patients and there were guidelines and protocols that reflected current levels of variation would hopefully reduce for some patient groups.

The interviews using NPT explored new, and reinforced existing barriers to be considered and overcome to make changes to neurological assessment and monitoring practice, as well as facilitators that may assist. NPT was chosen as it has been shown that combining it with a realist approach has been shown to add explanatory power to the understanding of implementation (Dalkin et al., 2021). Analysis was pre-defined so that it would be consistent, coherent, and useful (Nowell et al., 2017). A blended approach to coding, also known as abduction, was applied, which combines deductive and inductive elements and ensured movement back and forth between data and theory (Pierce, 1978). This approach suited the application of theory to obtain a more complete understanding and explanation of implementation issues whilst also being open to other elements and staying loyal to the data (Skjott Linneberg & Korsgaaard, 2019). Adopting this approach supported the identification of multiple interacting or contradictory factors and potential implementation challenges to future changes in practice.

Despite participants reporting that change is need resistance can still be a major barrier (DuBose & Mayo, 2020). Resistance to change stems from a human's basic need for a stable environment (Hogan, 2007). Health care organisations and clinicians within them are constantly facing change due to technological advancements, ageing populations, changing disease patterns and new discoveries for the treatment of diseases (Drotz & Poksinska, 2014; Hansson et al., 2008; Rafferty & Griffin, 2006) Findings from the interviews indicate that explaining the rationale behind making changes to neurological assessment and monitoring practice could help foster understanding and engagement. Training to ensure staff are aware of the importance and potential impacts of neurological assessment and monitoring practice is

therefore one strategy that should be utilised to overcome resistance to change. This could be strengthened if clinicians understood that changes to practice to improve identification of neurological deterioration were important to guide improved safety and quality implications for the care provided. Other literature supports this by showing changes to practice are more likely to succeed if those involved can influence the change, feel prepared for the change, and recognize the value of the change, including perceiving the benefit of the change for patients (Nilsen et al., 2020). Future implementation strategies for the SNOBSS will need to reflect the complexity and adaptability of the system, respect its resilient features, and put clinicians at the heart of the change (Braithwaite, 2018).

The data obtained in these interviews will be useful in developing recommendations for practice and further research but there is an awareness that it does not provide a whole picture. The concept of data saturation also described as information redundancy is defined as the point at which no new information, codes or themes are yielded from the data and evolved from theoretical saturation in grounded theory (Lincoln & Guba 1985; Braun &Clarke, 2021b). Data saturation is used as a criterion of quality in some qualitative research checklists such as Consolidated Criteria for Reporting Qualitative Research (COREQ) and the Critical Appraisal Skills Programme (CASP) 10-item checklist for qualitative research (CASP, 2018; Tong et al., 2007). This has meant that it has for many become an implicit assumption that data saturation is implicit of good practice even though within the literature it is often based on arbitrary and unexplained criteria (Braun &Clarke, 2021b). Data saturation suggests completeness of understanding and an implication of a determinable fixed point for data collection (Dey, 1999). This is in complete divergence to the author's theoretical perspective for the thesis of critical realism which supports the view that there is potentially always something new to be explored. Saturation is a logical fallacy as there are always new theoretical insights to be made if more data was collected and analysed (Low, 2019).

6.7 Strengths and Limitations

The interviews only represented two units that purely identified as a HASU. However, the selection criteria were pre-determined before the full survey results were obtained. From the final survey data, only three per cent of units reported being purely a HASU so having more joint HASU, and ASU units meant the selection was more representative of the range of UK stroke units. Although recruitment met planned numbers there is an awareness that only 14 units were represented across the interviews from a potential 158. Although this limits the generalisability of the results there was validation across survey data and between participants about some of the factors that impact neurological assessment and monitoring practice. This

illustrates that the data has utility in providing deeper insight into the knowledge, understanding and acceptability of neurological assessment and monitoring practice as well as the barriers and facilitators that impact on it.

Participant recruitment was potentially open to selection bias as it was reliant on individuals to self-select and overall, only senior members of staff were recruited. The opportunity to participate should have been made available to all eligible members of the stroke team but the author had no control over who how and to whom the contacts advertised the study. The limited participant range could have been due to the way the opportunity was advertised or potentially because senior staff were more motivated and engaged or junior staff lacked confidence in getting involved. Thus, the findings from the interviews may not be fully representative of the views of all staff members. Ideally, the interviews would have been completed with a more diverse range of staff, this may have provided different insights into knowledge, understanding, and acceptability of neurological assessment and monitoring after stroke and raised different barriers and facilitators to its use in practice. Nonetheless, there was good engagement from a range of staff with different professional backgrounds despite ongoing clinical pressures in the NHS. Although several interviews were delayed or had to be re-arranged due to workload, the staff remained committed to participation.

To ensure participation from a wide geographical area telephone methods were utilised. There are limitations to these including the lack of visual cues and restrictions on the development of a relationship and rapport between the researcher and participant. However, these did not appear to cause any issues which could be due to the professional nature of the interview content. There is a risk that interview responses may not represent participants' real opinions (Hootkoop-Steenstra, 2000). Again, this was considered very low risk due to the nature of the topic and its lack of sensitivity. There was only one interview that was not audio recorded due to lack of consent, but this was incorporated in the analysis using extensive field notes.

NPT was chosen as the theoretical approaches to these interviews; however, many other approaches are available that might have been expedient (Nilsen, 2015). NPT was useful as it helped identify factors that will aid the development of the Standardised Neurological OBservation Schedule for Stroke (SNOBSS) and plans for implementation in the future (May et al., 2018). However, these interviews have focused on exploring current practice and not directly evaluating and understanding implementation processes. This would account for why some NPT components had reduced data collected, such as individual specification as there was no new practice for the participants to consider. Implementation of the SNOBSS in the

future will require further exploration work to be completed to confirm the themes identified within the interviews.

Semi-structured interviews are interactive which can result in alteration in the questioning. To reduce the risk of overlooking important themes the interview schedule provided structured open question sequences to ensure all participants were provided the opportunity to talk about all key topics. The author was sensitive to not leading the interviews as this was an explorative process, and her grasp of the subject matter and use of field notes helped ensure that topics within the interview schedule were covered thoroughly. The reliability, and trustworthiness of the analysis were maintained by using a secondary coder to achieve consistency (Church et al., 2019). This was important to help prevent the author's opinions and biases from other parts of the study from influencing the results.

6.8 Chapter Summary

The interviews allowed deeper exploration of a complex element of care within and across several different stroke services. Through the application of NPT theory, barriers and facilitators that could impact the future implementation of the SNOBSS at individual, unit, and organisational levels were identified. The chapter further evidences the interconnectedness of factors that can impact elements of care such as neurological assessment and monitoring. These factors should be considered when developing future changes and implementation strategies to support them. Strategies for change need to accept the complexity and adaptability of healthcare systems and ensure clinicians are an integral part of any change to ensure strategies are as successful as possible.

The next chapter (chapter 7) describes the development of the SNOBSS by an expert stakeholder group. Key information from chapters 4, 5, and 6 were presented to the group and through an iterative process of discussion and review of information using consensus techniques the SNOBSS and associated documentation were created. The SNOBSS was then presented to interested clinicians to obtain opinions feedback on face validity and applicability to their pathways.

Chapter 7- Development of the Standardised Neurological OBservation Schedule (SNOBSS)

This chapter describes the design and development of the Standardised Neurological OBservation Schedule (SNOBSS) for acute stroke using consensus methods by a stakeholder expert group. It outlines the group development, activity, and outputs. Key findings from the reviews (chapter 4), survey (chapter 5), and interviews (chapter 6) were collated to provide background and information to drive and where possible inform the group's activity. The use of mixed methods within the project has allowed comprehensive exploration of current practice and factors that affect neurological assessment and monitoring practice. The SNOBSS once developed and agreed upon by the expert group was presented to interested clinicians to obtain feedback on face validity and applicability to their pathways.

7.1 Aim

To agree the content and design of the Standardised Neurological Observation Schedule for Stroke (SNOBSS).

7.2 Rationale for Design

A collaborative co-design approach by an expert group was selected for the development of the Standardised Neurological OBservation Schedule (SNOBSS) for acute stroke. The notion of co-production and co-design has emerged from the participatory approach, and refers to the joint working of people, not in the same organization, to produce goods or services (Durose et al., 2017; Ostrom, 1996). Consensus-building techniques were chosen as they are well-established ways of promoting discussion, inclusion, and participation in situations where there may be multiple perspectives. They also have the potential to rapidly generate a consistent approach (van der Scheer et al., 2021). There are multiple consensus-building methods available including the nominal group technique (NGT) (Harvey & Holmes, 2012), the consensus development conference (Lomas et al., 1988), the RAND/UCLA appropriateness method (Fitch et al., 2001), and the Delphi method (Dalkey & Helmer, 1963). Findings around the reliability, validity, and impact of these different strategies have been mixed (Black et al., 1999; Fink et al., 1984; Raine et al., 2005). Nonetheless, these approaches have been shown to help develop shared understanding, include varied expertise, and produce agreements on process improvements (van der Scheer et al., 2021).

Several factors influence which consensus approach is best suited including the geographical scope of the expertise required, the focus and subject matter being developed, the population affected by the recommendations, the quality of the available evidence; and the time and resource constraints involved (European Centre for Disease Prevention and Control, 2011). Hybrid approaches that draw upon the advantages of various methods, both formal and informal, can also be used (Black et al., 1999; Hutchings et al., 2006). The development of the SNOBSS utilised an informal NGT approach where participants came together, everyone had a voice, and discussion was used to reduce misunderstandings and expose reasons for differences of opinion (Harvey & Holmes, 2012). Consensus-building methods are useful, as in this case, where agreement is needed despite there being multiple and sometimes elusive options for process improvement (van der Scheer et al., 2021).

7.3 Methods

7.3.1 Nominal Group Technique (NGT)

NGT was developed as a procedure to facilitate effective group decision-making in social psychological research (Van de Ven & Delbecq, 1971). Although more traditionally used as a method for data collection or determining research priorities it is an appropriate method here to obtain the views of experts on a given topic (Harvey & Holmes. 2012). NGT is being used as a tool for developing a consensus on the content, delivery, and application of the SNOBSS and decision flowchart. It was the preferred consensus method for several reasons: it facilitates equal participation and allows all opinions to be respectfully considered (Carney et al., 1996), and it requires less time and resources than other methods (Delbecq et al. 1975).

The use of NGT aligns well with the explorative theoretical underpinning of critical realism. It will allow prioritisation of problems and issues through group discussion and allowing different ideas and opinions to be expressed and collated with a view to identifying areas of consensus and establishing priorities for change (Harvey & Holmes. 2012). Through comparison of priorities across different individuals within a group it can help identify divergence in views and support exploration of specific topics (Cantrill et al., 1996).

There are varying methods of NGT in terms of procedures and analysis (McMilan et al., 2014). The aim was to use the NGT five-point checklist described by Potter et al., 2004: Introduction and explanation, silent generation of ideas, sharing ideas round Robin, group discussion clarifying, and voting in ranking. Although all these elements were included there was variation in the way it was applied (described in more detail in Section 7.4). There ended up being an

extended introduction phase but once everyone shared a common aim, they were more willing to discuss and negotiate what they believed were the possible options for the SNOBSS. The silent generation of ideas phase was mainly planned to be around what items should be included and although there was a private method of allocation of votes, due to time constraints, it was completed outside of the group sessions. The sharing ideas and group discussion clarifying ended up almost simultaneous with a focus on both clarification and elimination. Voting in ranking had no stipulation on the number of items and agreement was reached through debate of what was reasonable and useful to be included in a regularly repeated assessment. Due to time constraints secondary ranking was not feasible but this was deemed reasonable as it was felt that the experts involved would not feel any social pressure to conform. Overall, the principles and ethos behind NGT were applied as all experts had the opportunity of equal representation and the environment was conducive to the initiation of change (Davis et al., 1998).

7.3.2 Group Development

Stakeholders are defined as "individuals, organizations or communities that have a direct interest in the process and outcomes of a project, research or policy endeavour" (Deverka et al., 2012). They needed to be able to interpret the information provided and apply their experience to the development of the SNOBSS. A group of individuals who had expertise in neurological assessment and monitoring practice after acute stroke were identified through discussion with the supervisory team (de Vet et al., 2011). Experts were needed that had knowledge of the subject area as well as competence in the practical application of the knowledge (Bojke et al., 2021). The experts were also required to have adaptive skills to elicit decisions as the evidence around neurological assessment and monitoring practice is less developed than in other areas of acute stroke care (O'Hagan et al., 2006).

Key criteria for inclusion were recognition by peers, specialist knowledge and clinical experience, and known research in the topic area (Bojke et al., 2010; Fischer et al., 2013; Leal et al., 2007; Sperber et al., 2013).

The aim was to convene an expert stakeholder group with up to 12 participants. Ideal group size for NGT is usually 6–12 people (Pastrana et al., 2010). Ten external experts (8 doctors, 2 nurses) and the four PhD supervisors (1 doctor, 3 nurses) were selected by the author and supervision team and invited to be part of the group. Approaching a large group was useful in terms of ensuring that there was sufficient group discussion even if some experts were unable to attend. Those selected and invited to take part had a wealth of acute stroke clinical and

research knowledge and experience and came from a range of professional backgrounds and geographical locations. All potential participants were provided with information explaining the purpose of the group, what their involvement would be, and how much time commitment would be required.

7.3.3 Planned Meetings

The purpose of these meetings was to evaluate the key findings of the results from the reviews, survey, and interviews and develop the SNOBSS and decision flowchart for future testing. The formation of the expert group and the development of the SNOBSS did not require specific ethical approval as per the Health Research Authority (HRA) 'Is my study research?' decision tool (HRA, 2022b). The expert group meetings were scheduled once data from the reviews, survey, and interviews were available. Due to the delays in data collection reducing the timeframe available for the SNOBSS development two separate two-hour group meetings were held on 16/09/2021 and 4/10/2021. Due to the ongoing concerns with the COVID pandemic and to make the sessions more accessible to busy clinicians the meetings were held virtually via Microsoft Teams. Consent was obtained to record the sessions for later review of the content. Sessions were planned to make the most of the time available and to facilitate the sessions being as efficient and as effective as possible (Appendix 7.1). Activities included reviewing pertinent results from the reviews (Chapter 4) and the UK-wide survey (Chapter 5) to stimulate discussion and ranking exercises to trigger debate and assist decision-making.

The meetings were purposively planned to be interactive with the participants having to work together to explore the evidence and develop the intervention (SNOBSS) (Pavelin et al., 2014). A participatory ethos was encouraged as this can increase acceptability, uptake, and impact of process improvement (van der Scheer et al., 2021). The sessions were facilitated by the author guided by key components of NGT and value-focused thinking (Keeney, 1996). The meetings were iterative and did not also follow the session plan dependent upon the flow of the group discussion.

All members were encouraged to contribute to the discussion. To prevent dominance individual opinions from all members were sought and encouraged. Insight into every stage of the decision process was sought and alternatives were explored. Divergent views were highlighted and explored as this provided a more systematic and potentially useful way to search for creative alternatives in the decision process (Black et al., 1999). Although an overall agreement was eventually sought this encouraged open and constructive debate (Pagliari et al., 2001).

Each session was closed with a summary of the agreements made and an agreement on next steps. After both sessions participants agreed to receive additional documents to work through to guide the author in the development of the SNOBSS, notes for completion and training, and its decision flowchart. The results from these activities were fed back to participants for reconsideration and to guide further discussion. Transparency of the decision-making process was maintained throughout (World Health Organisation, 2014). The recordings provide a clear record of proceedings in the sessions and all group members had access to all documents for comment and change before final approval. Once the prototype of the SNOBSS was developed and agreed upon by the expert group wider evaluation from clinicians was sought.

7.3.4 SNOBSS Evaluation

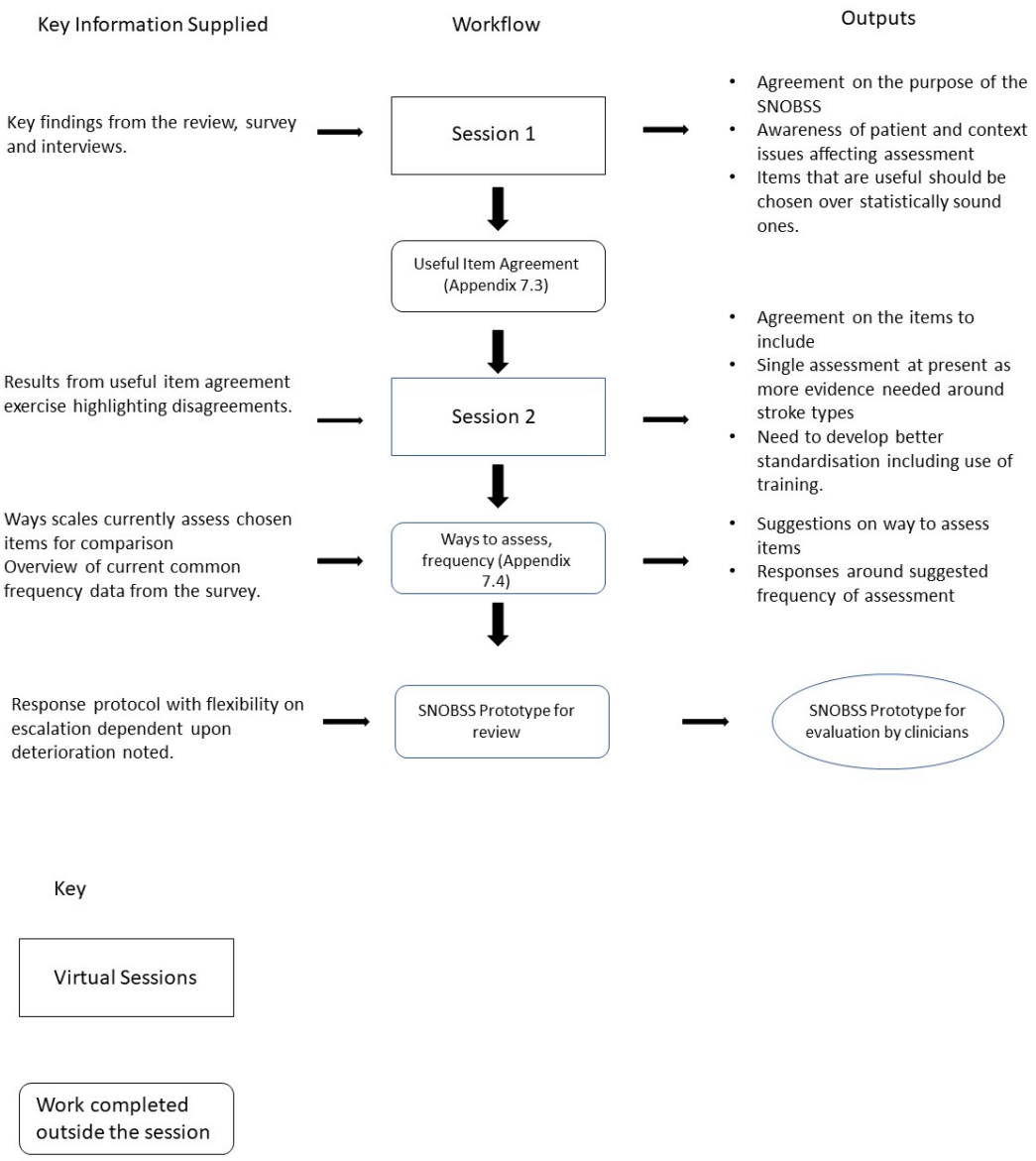
Clinical teams were approached to provide feedback and complete an initial face validity check of the SNOBSS. Requests for volunteers to review were circulated via the FutureNHS Collaboration Platform Stroke Forum and the National Stroke Nurses Forum (NSNF) membership. These sessions aimed to obtain evaluative feedback on comprehensibility, acceptability, and feasibility of the SNOBSS and identify issues for its implementation to aid future development and testing. The sessions were designed to be interactive with a small group of clinicians being introduced to the project and reviewing the SNOBSS. A series of questions were developed to guide the evaluation and identify specific local pathways issues that could impact implementation (Appendix 7.2). Sessions were planned to take less than an hour to complete virtually via Microsoft Teams. Consent was obtained to record the sessions for later analysis of content.

7.4 Results SNOBSS Development

Eight external experts and three supervisors (8 doctors, 3 nurses) engaged in the group meetings. A further nurse due to unforeseen circumstances was unable to attend the sessions but engaged with the prioritising and evaluating work completed after the sessions. Of the other two externals one did not reply to the invitation and the other was keen to be involved but changed job roles during the set up and was subsequently unable to attend.

It is not possible due to word count within this thesis to outline all the content and discussion that occurred during the expert group process. Therefore, a visual overview of the work of the group is displayed (Figure 7.1), followed by key points of discussion and agreement that guided and informed development of the agreed SNOBSS.

Figure 7.1 Overview of the expert group work and development of the SNOBSS.



It was intended in the first session to achieve a ranking of essential, desirable, and omit for all items identified from the scales in the reviews. It was then planned to use this data to decide on the most important items to identify deterioration and then discuss whether it is possible to develop a single scale that can be used for all stroke patients. However, the first session ended up focusing on group discussion to develop and agree on a shared understanding of what we wanted to measure (construct), in whom, and for what purpose. This was important as the content and layout of any assessment are shaped not just by the items it measures but by the construct under examination and factors such as who will be administering it (de Vet et al. 2011).

It was agreed that the SNOBSS needed to be designed to be used regularly with stroke patients to identify meaningful change (specifically Early Neurological Deterioration (END)) and that it could be used quickly and easily by any member of the multi-disciplinary team (MDT). It should guide the user to consider the patient's neurological status and function and where appropriate escalate change so that action can be taken as quickly as possible to improve patient outcomes. It was also agreed that the SNOBSS was not being designed as a scale and that no quantification or scoring of items would be warranted.

7.4.1 Selection of items

Between the meetings the expert group members rated all 40 items that were included in any scales in chapter 4 as either essential, desirable, or should be omitted (Appendix 7.3). Participants were also asked to identify other items that they felt could or should be included and to provide justifications behind their decisions. In session two the 18 items that had received any number of essential votes and four additional item suggestions were discussed and debated item by item. An agreement was then reached on the six items that the group felt should be included in the final SNOBSS (Figure 7.2 pgs. 234-235).

The group discussions were pragmatic looking to select the items that were useful to detect END. There was a unified understanding that there is a difference between what is clinically useful and statistically reliable. For example, Pupillary Response has been shown to be reliably assessed but is not felt to be clinically useful in the identification of END. The group continually returned to the question of whether it was felt that changes in the item would represent END. Other key considerations were around whether items could be practically and frequently repeated. There was an awareness that some items are impractical for repeated use at the bedside such as dysphagia. Although these items did not warrant inclusion in the SNOBSS assessment they should be seen as a crucial part of stroke care and management. The whole MDT should possess or develop a culture of awareness of other signs of deterioration in patient function and status.

The ordering of the chosen items was later agreed to represent both prioritisation of the items, with level of consciousness (LOC) deemed the most important, and a practical and sensible order in which to complete at the bedside. Despite the agreement on the six items, there was an appreciation that these decisions were based on knowledge and experience and that the evidence-base as to which items would best detect deterioration, in which patient groups, needs further exploration beyond the six items selected.

The items chosen were selected to be independently useful to trigger a response if change is noted. The SNOBSS is therefore a multi-dimensional assessment. However, it is likely that the items are likely to interrelate for example, in a brainstem stroke there might be an alteration in Gaze before a change in LOC. Identifying and stratifying the links between items would also require further testing as identification of the items that identify early changes should be prioritised over those that detect later changes.

Identification of seizure activity was highlighted as important by the group with some suggesting that it be included as an item in the SNOBSS. It was eventually not included as after extensive discussion it was decided that seizure activity would be the cause of the deterioration and not the marker of it. It was also felt that as there are multiple types of seizure activity that could present post stroke it would be too complex for a range of staff to quickly assess at the bedside.

7.4.2 Assessment of Items

Due to the time constraints on the sessions, the overall decisions on how best to describe and operationalise the items were made after session two (Appendix 7.4). It had been agreed in the session that the SNOBSS needed to focus on crude detection of clinically significant change and not be overpowered by assessing the level of severity. As per chapter 4, it was acknowledged that dichotomous systems do not allow for any categorisation but if too many options are supplied reliability is reduced as subjectivity increases.

The aim was to have instructions that allow meaningful and useful assessment in the clinical setting when used by a range of staff. Some key assessment criteria were agreed upon in the session such as it is important to assess both affected and non-affected sides when assessing Motor Power to ensure any loss of power detected was unilateral, stroke-related, and not due to generalised weakness. It was also acknowledged that some items will be untestable especially in unresponsive patients however as the SNOBSS is interested in change and does not assign scoring this should not cause issues in practice. However, it is recognised that alterations to the assessment may need to be made and tested for patients with limited cognition to ensure consistency in application.

The SNOBSS shows the categories agreed upon by the expert group but as per the choice of items, these are based on expert opinion. Again, there is currently no evidence to highlight how to best operationalise the items so that they are reliably applied and detect true change in all patients.

7.4.3 Target population

There is an awareness that neurological assessment and monitoring is important for all stroke patients. However, the data from the survey and interviews have shown that certain patient groups are currently prioritised in this element of care. Different patient characteristics, such as pathologic type or vascular territory affected could lead to different patterns of deterioration. At present repeated measures information is lacking and we don't know enough about which items get worse, when, and in whom.

The sessions attempted to challenge the experts and get them to think about specific patient groups in the choice of items. The SNOBSS was developed pragmatically attempting to be a single assessment tool useful across the whole stroke population but with the caveat that there might be specifics required for different patients and groups. Future scenario-based evaluation is needed to clarify whether a single SNOBSS will be suitable for all stroke types or whether specific patient groups will need amended versions. However, caution needs to be exercised on the use of multiple assessments based on different clinical scenarios as these would need to be well implemented and managed to ensure the right assessments are being completed for the right people. If variation in assessment, based on characteristics was chosen in the future, it would need to be based on stroke type or other factors that are well documented or communicated to the whole MDT.

7.4.4 Frequency

Throughout the sessions, the desired frequency was alluded to and discussed but this proved one of the most difficult considerations on which to achieve agreement. After session two the group was provided with an overview of the data from the survey on the most commonly used frequency across the range of patient groups and time periods. There was a strong feeling amongst the experts that frequency should be driven based on the potential benefit to the patient. Frequency should be increased in those that have the most to lose or when there is cause for concern.

The overall guidance in the SNOBSS is that all patients should receive a minimum of hourly monitoring on admission for the first four hours. Overall decision on frequency would be led by specialist assessment and/or discussion about patient risk and appropriate monitoring selected. Any changes and decisions about ongoing frequencies would also be driven by the

specialist stroke teams until further evidence is available including information on patient and staff acceptability.

7.4.5 Decision Flowchart

It was agreed that the aim was to develop an action plan-based monitoring schedule. Currently, the identification of deterioration primarily leads to escalation to the medical team for further assessment. Therefore, it was agreed that the SNOBSS decision flowchart should prioritise escalation of change but that it should allow some flexibility to patient condition and what could potentially be done if deterioration is noted. Like frequency, this would be decided based on a specialist local assessment of the patient's condition.

The SNOBSS currently incorporates three levels of response: change in any item, change in LOC, or any two other items, and not to escalate. This allows differentiation in response based on patient need and whether anything can be done for the patient in the event of deterioration. The SNOBSS should ensure a better understanding of the patient's condition amongst the whole team whilst also utilising resources such as staff time in a meaningful way. The decision flowchart highlights the value of repeating the assessment, with others, if necessary, to ensure confirmation of change. It also encourages the completion of a full set of physiological observations alongside the assessment to provide a more thorough handover whilst escalating and to identify or rule out a physiological basis for the deterioration. It is hoped that in the future other actions and interventions could be added to the decision flowchart once these are better understood and evidenced.

7.4.6 Layout format

The developed SNOBSS (Figure 7.2 pgs.234-235) is a multi-page document incorporating information, advice on how to use it, and the decision flowchart. The assessment itself was consciously designed to fit on one page that was adaptable to multiple frequency options. The group felt it was important that all pertinent information was together, and it allowed previous assessments to be reviewed alongside completion so that change should be recognised instantly. The SNOBSS is currently in a paper-based only format but could be integrated into electronic record systems in the future.

7.4.7 Education and Training

The expert group agreed that education and training were extremely important to successfully and effectively implementing the SNOBSS into practice. This was strongly mirrored in the

survey and interview data. There were multiple item specific completion and training notes highlighted by the expert groups and these have been included with the developed SNOBSS (Figure 7.2 pgs. 234-235). These included specific prompts and advice around the completion of all the items within the assessment.

The expert group also felt that all members of the MDT should have an awareness that multiple factors other than the items in the SNOBSS are important in terms of patient progress and overall outcome. The SNOBSS has a specific role to play in overall care provision. However, wider stroke care involves awareness and attentiveness to the monitoring of global function in all stroke patients. New or worsening symptoms could be important indicators of deterioration external to the SNOBSS items and are important to note, such as changes in swallow or respiratory pattern. In terms of neurological assessment and monitoring practice, there will always be subjectivity, but effective training should reduce it as much as possible.

7.5 *SNOBSS Evaluation Results*

Fourteen separate expressions of interest to take part in the evaluation were received. Due to the time constraints to complete and competing obligations of the clinicians three small group sessions and two individual sessions were held. Significant statements of contributors will be presented under four themes: Overall evaluation, items selected, specific changes or additions suggested, and issues for implementation.

7.5.1 Overall evaluation

This was positive with most contributors preferring the SNOBSS to the Glasgow Coma Scale (GCS) as it is a stroke-specific assessment. This was especially highlighted in terms of assessment of communication which in the GCS is inappropriate for a stroke population. It was generally felt that the removal of scoring from the assessment was good with some commenting that it reminded them of the Stroke Thrombolysis Observation (STOC) Chart so the change in a specific item by a box or more is not a new concept and one that could be quickly understood and adopted.

There were encouraging reactions to the flexibility within the SNOBSS. The ability to prioritise LOC and select different escalation processes was welcomed. It was also felt that the decision-making procedures were useful because they would provide a clear indication of what was wanted for an individual patient. This was felt to be particularly useful in units where there is currently disagreement between consultants as to frequencies of monitoring required.

Generally, it was felt that the SNOBSS provided a meaningful assessment through which deterioration could be identified and escalated.

7.5.2 Items selected

Although there was unanimous agreement on the prioritisation of LOC several contributors voiced that the AVPU assessment detail might not be sensitive enough and could require more depth. Some felt that the assessment of Orientation and Confusion should be included. However, others welcomed the use of the AVPU as it fits with the NEWS schedule and is accessible to all staff. One unit reported that if they identify a change or concern with the AVPU assessment then they complete the first three items of the NIHSS to gather more information. Multi-level assessment could be trialled in the future for effectiveness.

There were other differences of opinion as to what items to include. One contributor was unsure what assessment of Facial Paresis would add. Another, like members of the expert group, felt that respiration rate and pattern needed greater emphasis. Another contributor felt that Sensation should be added to make the assessment more stroke specific, but this was debated with the team as others felt that having too many items would lengthen the assessment too much and that there should be limits on the number of items included.

As in the expert group, there was debate about the inclusion of seizure activity. One contributor initially felt it was important to include. However, after discussion with their colleagues, they decided that it could not be as there are so many different presentations e.g. rigidity, absence, etc that it can be hard to distinguish. Their conclusion was that the escalation of change or unusual activity was the key priority and not the diagnosis of the underlying cause.

One consultant felt strongly that there should be different versions for different patient contexts. He felt that there should be at least two versions of the SNOBSS available one for ischaemic stroke and one for ICH. When pressed about what items are currently missing from the SNOBSS to accommodate this he spoke about the assessment of raised intracranial pressure (ICP). As this is something that constitutes a cause for deterioration rather than a sign and that cannot be quickly and easily assessed at the bedside it could not be easily integrated into the current assessment. However, it could be that particular changes or combinations of changes in items considered important for specific groups e.g., rising blood pressure or nausea in ICH could be indicated within the SNOBSS for particular attention.

Groups and individuals generally understood the removal of certain items once they comprehended that the SNOBSS was not designed to replace a full neurological assessment but to allow frequent assessment to identify change. It was explained that an important part of future testing would be to provide evidence to support the choice of items and methods of assessment in the identification of change, specifically END.

7.5.3 Specific changes or additions suggested

Specific changes requested were to make it clearer where to record left and right sides in motor power assessment and to be clear on whether the patients must be seated or lying to complete. Additions suggested included finding a way to ensure that an adequate baseline indicating previous function could be added and readily accessible. It is important to know if there are previous issues that would affect the assessment such as weakness from any cause, communication difficulties, longstanding confusion, issues with cognition, or visual deficits. A notes box was suggested although that would be reliant on effective history taking and documenting any issues. There might be a need for more guidance to complete this. Another suggestion was the addition of a prompt to ask the patient, where applicable if anything had changed. It was felt this could be especially useful for patients with symptoms such as headache and dizziness.

7.5.4 Issues for implementation

Resources around staffing were the key concerns raised regarding the implementation of a change in neurological assessment and monitoring practice. Issues that were identified in the survey and interview were mirrored here in that staff can be re-deployed from the unit and that experienced staff are not always available to complete assessments. This adds further justification to creating the SNOBSS to be accessible to a range of staff. Time taken to complete the assessment was raised as a potential issue that would need to be assessed but adds weight to an argument of not including too many items in the assessment. The development of different versions could also reduce assessment time in the future once more is known about which items are best for detecting deterioration in specific stroke subtypes.

The other concern raised was also about staffing but related to who should be making the decisions associated with the frequency and required escalation response for the patient. The consensus was that in hours this would be consultant led but several contributors were aware that out of hours it would be difficult to have appropriate cover. Some suggested that there would be a need to develop some guidelines to ensure standardisation around these decisions

based on stroke type and other patient characteristics. This could lead to variation between units before the evidence supports these decisions but should result in less variation within units.

Figure 7.2 Standardised Neurological OBservation Schedule for Stroke (SNOBSS)

Standardised Neurological Observation Schedule for Stroke (SNOBSS)

This observation schedule is designed to be regularly used with stroke patients to identify meaningful change (specifically deterioration). It can be used by any member of the multi-disciplinary team and guides the user to consider the patient's neurological status and function and where appropriate escalate change so that action can be taken as quickly as possible to improve patient outcomes.

It is recommended that all stroke patients initially receive at least hourly monitoring for a minimum of the first 4 hours. This covers the period where most hyperacute treatments would be instigated and allows establishment of a strong baseline of patient status. All patients must receive a full set of physiological observations at the same time as this assessment. (*more guidance on integration with NEWS)*

Patients who receive thrombolysis or thrombectomy should be monitored using agreed local protocols. There may be other patients such as those on blood pressure alteration who also warrant more frequent monitoring as per local protocols.

A senior member of staff (as per local procedure- *anticipated to be medical team, stroke nurse consultant or stroke specialist nurse level*) on admission will decide the escalation policy. There are three options for medical escalation:

- Any change in any item A-F
- Any change in Item A: Level of Consciousness or change in two items B-F
- Not to escalate from any changes in any items A-F (but continue standard NEWS observations unless instructed otherwise by clinical staff)
  (NOTE Physiological changes should still warrant escalation to the medical team if appropriate e.g., high temperature and altered respiratory rates unless decision has been made not for escalation in any capacity).

When escalating patients who have changed then report changes in both neurological and physiological parameters to provide a clearer clinical picture.

The same or another senior member of staff should decide whether the monitoring frequency will be reduced after 4 hours if the patient remains stable. It is anticipated that this decision will be based on patient condition and characteristics and local policies and procedures. For example, it is anticipated that patients with intracerebral haemorrhage or those likely to be considered for hemi-craniectomy would remain on hourly neurological monitoring for a longer period.

Standardised Neurological Observation Schedule for Stroke (SNOBSS) Prototype

Patient Label or Details

Escalation to medical team policy	
Change in any one item A-F	
Change in item A: Level of Consciousness or change in two items B-F	
Not to escalate from change in items A-F	
Signature Time & Date	

ITEMS A-F (below)	Date / Time												
A: Level of Consciousness													
Alert		X											
Voice													
Pain													
Unresponsive													
B: Communication													
Normal: no communication difficulty													
Mild communication difficulty													
Moderate difficulty, no proper sentences													
Severe difficulty, 1 or 2 words or less													
C: Gaze													
No gaze palsy													
Gaze palsy present													
Eyes deviated at rest													
D: Facial Paresis													
Normal													
Asymmetry on forced grimace													
Asymmetry or drooping at rest													
E: Motor Power-Arms													
No drift for 10 secs													
Drift but does not hit bed													
Drift, hits bed													
Some effort against gravity													
No effort against gravity													
No movement													
F: Motor Power- Legs													
No drift for 5 secs													
Drift but does not hit bed													
Drift, hits bed													
Some effort against gravity													
No effort against gravity													
No movement													
Initials													

Time/ Date of first item deterioration:		Escalated: Yes/No
Time/ Date of second item deterioration:		Escalated: Yes/No
When to reduce to 4 Hourly monitoring if no deterioration noted:		

Notes for completion/ Training considerations

Complete using a tick or cross in the box.

Showing the patient what is required is allowed.

Level of consciousness

- If patient is sleeping, they will need to be woken for the assessment
- Initially it may be difficult to differentiate between a sleepy patient and one who is drowsy due to a reduction in consciousness. Complete the whole assessment to allow time to determine the LOC.
- Need to apply a strong tactile stimulus (required to be able to differentiate between Pain and Unresponsive categories.

Speech

- General conversation and response to questions during assessment.
- The assessment is looking at communication generally and the overall aim is to identify any change.
- Distinction between types of speech difficulties is not required. (Examiners need an awareness of assessing for slurring but not formal dysphasic assessment).

Gaze

- Observe resting eye position on approaching the patient.
- To assess horizontal range of eye movements, steady the head and ask the patient to follow the examiner's finger whilst moving the finger from the left to the right, then vice versa.
- Establishing eye contact and then moving around the patient from side to side will occasionally clarify the presence of a partial gaze palsy.

Facial Paresis

- Need to be aware that most people are not purely symmetrical
- Patient's face is assessed while talking or smiling
- Only muscles in the lower half of the face.

Motor Power- Arms

- Nonparetic side first, each arm in turn
- If patient led place the arms extended to 45 degrees
- If patient sat place the arms extended to 90 degrees
- Drift is scored if the arm falls before 10 seconds but varies depending on whether the arm hits the bed or other support
- Some effort against gravity; limb cannot get to or maintain (if cued) 90 (or 45) degrees, drifts down to bed, but has some effort against gravity
- No effort against gravity the limb will just fall
- Only in the case of amputation or joint fusion at the shoulder would it be untestable (UN)

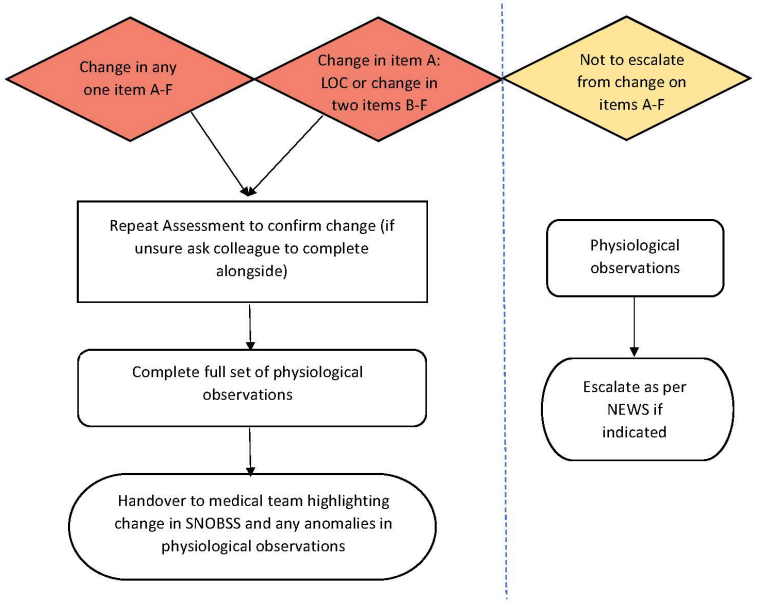
Motor Power- Legs

- Nonparetic side first, each leg in turn
- Test with patient led down
- Place leg at 30 degree and ask patient to hold for 5 seconds
- Drift is scored if the leg falls before 5 seconds but varies on whether the leg hits the bed or not
- Some effort against gravity; leg falls to bed by 5 seconds, but has some effort against gravity
- No effort against gravity the leg falls to bed immediately.
- Only in the case of amputation or joint fusion at the hip would it be untestable (UN)

Standardised Neurological Observation Schedule for Stroke (SNOBSS) Decision Flowchart

Initial monitoring schedule is a minimum of hourly for 4 hours (unless another schedule selected for the patient e.g., thrombolysis). Senior clinician needs to decide whether monitoring should continue after four hours and at what frequency dependent upon patient condition and characteristics as per local policies and procedures.

Senior clinician to decide on the escalation policy (as per shaded decision diamonds below).



Standardised Nursing Observation Schedule for Stroke (SNOBSS) Escalation Guide

Escalation Policy	Actions if changed noted	Escalation
Change in any one item A-F	• Repeat SNOBSS to check change, if unsure ask colleague to complete alongside • Complete full set of physiological observations	Handover to medical team highlighting change in SNOBSS and any anomalies in physiological observations
Change in Item A: Level of Consciousness or change in two items B-F		
Not to escalate from change in items A-F	Physiological observation	Escalate as per NEWS if indicated.

7.6 Discussion

Despite the complexity of the task the group quickly developed the SNOBSS and associated documentation. Although the consensus-building techniques did not require unanimous decisions there was no evidence of small group dynamics and overall disagreement was limited. This was helped by the shared agreement on the purpose of the SNOBSS which repeatedly guided decision-making processes by the group. The discussion included information from the findings of the reviews, survey, interviews, and expert opinions to develop a hierarchy of perceived importance and apply this to real world problems in stroke practice (Harvey & Holmes, 2012).

Collaborative research practices are well established within health care quality improvement research and practice (Locock & Boaz, 2019). Using consensus-building techniques allowed the incorporation of data collected and the knowledge and expertise of the group participants to develop the SNOBSS and decision flowchart. It also allowed consideration of barriers and facilitators to implementation. The major advantage of using this approach is that the SNOBSS should be of higher quality, more clinically relevant, credible, and hence easier to implement (Barber et al., 2011; Reis et al., 2016; Stewart & Liabo, 2012).

Key challenges to using NGT have been identified as recruitment, lack of engagement, and facilitation (Harvey & Holmes, 2012). Recruitment was not an issue in this group with only one invited participant not responding. However, engagement in terms of group attendance varied due to participants having other commitments. Engagement was encouraged through the development of a positive environment with mutual respect and all individuals' views having an equal voice within the process (Dingwell, 1992; Tollyfield, 2014). The author was inexperienced in facilitation, but the group was engaged and supportive of her attempts to guide and support the creation of the SNOBSS (Sanders & Stappers, 2008). The author ensured discussions were closed and summarised and ensured agreements were reaffirmed. Her lack of experience was balanced by the seniority and experience of the group participants.

In co-design and consensus approaches diversity is encouraged to reduce the risk of bias and provide a wider range of views (Boers et al., 2014). However, the deliberate selection of senior clinicians and researchers with extensive experience was justified in this group. Not only are senior staff the ones who would understand and be called on to manage deterioration but also the level of expertise of the group can be crucial to the success and validity of the data it generates (Cook & Birrell, 2007).

Due to the seniority and other workloads of the invited participants, the group design and planning tried to limit the amount of time the participants need to provide. The development of a shared agreement around what the SNOBSS was aiming to achieve was vital if unstructured work. Once everyone understood and shared the common aim it promoted discussion and negotiation. More time should have been allocated to this vital step in assessment development (de Vet et al. 2011). It is an accepted challenge of these processes that they can be time-consuming often taking longer than expected (Concannon et al., 2012; Kok et al., 2016).

The group was accommodating of the contrast between the time available and what needed to be achieved. They were open and receptive to the work required outside the group sessions to bring structure to the construction of the SNOBSS. Despite this additional workload the process was time-efficient and supported reviewing a substantial amount of information and development of the SNOBSS in a short space of time.

NGT techniques are useful in the development of solutions and establishing priorities for action (Harvey and Holmes, 2012). The development of the SNOBSS, although not a formal NGT approach utilised key features such as structured face-to-face meetings. These structured interactions encouraged the generation of ideas (Murphy et al., 1998). By utilising informal consensus-building techniques all members of the group could voice their opinion and contribute equally to the decision-making (World Health Organisation, 2014). This allowed the development of collaborative relationships that supported voting and ranking. Anonymous voting on the inclusion of items allowed identification of diverse viewpoints that were then used to structure discussion and compromise to reach a group consensus of an acceptable decision rather than unanimous agreement. All decisions reached were group decisions and not attributable to one individual.

Agreement can be difficult to reach, in cases such as this where the issues are complex, and the evidence is sparse (World Health Organisation, 2014). However, the group relatively quickly managed to agree on the purpose, content, and format of the SNOBSS. This was supported by them developing a common understanding of the purpose and priorities of what the SNOBSS needed to achieve. Several decisions were not straightforward as multiple factors are known to impact each other in such a complex element of care. For example, the usefulness of an item to detect change was prioritised above reliability but training needs to be developed to ensure consistency of assessment and improve reliability. Utilising key

information from the survey and interviews the group developed the SNOBSS and decision flowchart to try and be acceptable to all acute stroke services in the UK.

Stakeholder engagement has shown to play an important contribution within research guided by critical realism as it focuses understanding by individuals involved in the systems (Williams et al., 2017). The theoretical underpinning of this study supported the exploration for broader approaches to better suit the complexity of neurological assessment and monitoring within stroke practice. It was known from the beginning that a single assessment might not be possible. However, the group has developed a flexible approach that can be adjusted to stroke unit contexts to reduce current variation but that supports the acquisition of an evidence base to further develop this area of practice.

More formal and organised testing is required across which items best detect deterioration, the optimal way to describe and assess individual items and the frequencies of monitoring that should be employed in which patients. Without an evidence base to support these decisions the SNOBSS has been created to allow flexibility and adaptation to local context whilst also trying to address issues in current practice such as the use of non-stroke specific assessments and extensive variation heightened by lack of common language and poor documentation.

7.7 Future Plans

The expert group successfully developed the SNOBSS and decision flowchart which was well received by a wider group of clinicians. Extensive further research is needed to advance this important element of care and develop an evidence base for its use. However well-constructed and accepted, the SNOBSS and decision flowchart will not successfully change practice unless it is adopted by all staff. An implementation schedule should be developed informed by implementation science literature and the assessment and clinical practice barriers and facilitators identified within this research (chapters 4, 5, & 6). Future development and research should include more junior clinicians and those not specialising in stroke to ensure guidance is produced that is useful and that can be applied by a range of staff regardless of background

7.8 Strengths and Limitations

A crucial strength of the approach was that it was possible to undertake this work relatively rapidly. The original plan was to have more meetings of the expert group throughout the

whole project to oversee the data in more depth, but this was not possible due to the delays in data collection.

Heterogeneity of the group was achieved in terms of professional background, clinical & research experience, and geographical location. However, selection bias of the expert group may still have been present. Potential bias could have been reduced through the use of an initial longlist of potential participants for consideration or the use of a profile matrix in order to show that all essential and desirable characteristics required were met (Bolger & Rowe, 2015). However, the group membership were experts with a great knowledge of the subject domain who were also competent in the practical application of that knowledge (Boijke et al., 2021). This is important, as in consensus methods the level of expertise in the group impacts on its success and the validity of the decisions it makes (Cook & Birrell, 2007). The SNOBSS is clinically relevant not only from the choice of experts chosen but due to the inclusion of key information from the surveys and interviews.

All members were not able to attend all sessions which limited the interaction which potentially could have impacted the decision-making processes. To counteract this every group member had the opportunity to check everything agreed and comment both within and outside the sessions and there were no major differences of opinion. The invited participant who did not respond to the invitation was an international stroke physician. This is not necessarily a limitation as the data and project has been UK-centric and the SNOBSS has been developed with those systems and processes in mind. However, it is important to note that the SNOBSS would need further evaluation before being adopted outside the UK.

The clinical discussions went some way to ensuring that other professionals understood the SNOBSS and that it was workable in UK practice. However, this was only completed with a limited number of staff representing a small number of sites. Formal pilot testing would be needed to check comprehensibility, acceptability, and feasibility. The SNOBSS needs to be evaluated by a full range of staff that would be involved in completion, not just the potentially more engaged senior professionals.

A major limitation of the project is that it did not involve stroke survivors or carers in the development of the SNOBSS. Patients and the public are key stakeholders, and it would have been preferable if their perspectives could have been included (Marjanovic et al., 2019). It was initially planned that this group would include a minimum of two PPI members so their views

on importance, acceptability, and tolerance of neurological monitoring were included in the design process.

Before the COVID pandemic groups were attended in Preston, Warrington, and Aintree to try to recruit PPI members. The opportunity to be involved was also advertised on the people@UCLan website. Twenty-one individuals expressed potential interest in being involved in the study moving forward. Due to delays in data collection and other limitations, it was not possible within the timeframe to re-establish the involvement with individuals and groups to facilitate their involvement in the group. Given the time pressures of the work, their involvement might also have been untenable. The project data has shown that there can be a reluctance to "disturb" patients with neurological assessment and monitoring. From the introductory meetings and discussions with potential PPI members, it was clear that patients did not remember interventions such as neurological monitoring in the acute phase. There was also an assumption that if something was being done it was for a good reason. Future development and testing of the SNOBSS should include PPI involvement to assess acceptability, assist communication, and support implementation.

7.9 Chapter Summary

Using informal nominal group techniques an expert stakeholder group developed and agreed on a Standardised Neurological OBservation Schedule for Stroke (SNOBSS). This was informed by key findings from the reviews, survey, interviews, and the knowledge and skills of all group members. The SNOBSS represents a stroke specific assessment that could be used quickly and easily by a range of staff frequently to identify meaningful change (specifically Early Neurological Deterioration (END)). It guides the user to consider the patient's neurological status and function and where appropriate escalate change so that action can be taken as quickly as possible to improve patient outcomes.

There is currently wide variation in practice and a lack of evidence to support guidelines for neurological assessment and monitoring practice. Extensive further research is required both to test the SNOBSS and other important elements to develop an evidence base for neurological assessment and monitoring practice. The development of the SNOBSS represents a move towards more consistent and stroke specific monitoring to identify change in an acute stroke population. It is hoped that with further testing and development meaningful national guidance can be created to support neurological assessment and monitoring.

This chapter has described the development of the SNOBSS and decision flowchart. The next chapter (chapter 8) provides the overall discussion of the entire project, recommendations for both practice and future research, and a short overall conclusion of the project and thesis.

Chapter 8- Discussion

This mixed-methods thesis has explored different aspects of neurological assessment and monitoring practice in acute stroke. It has led to the development of the Standardised Neurological OBservation Schedule for Stroke (SNOBSS). This chapter begins by presenting summaries of the key findings from each chapter before a brief synthesis of the overall study findings. It then discusses the implications and presents the recommendations for current practice to improve standardisation and reduce variation and the recommendations for future research as the next steps needed to continue developing a consistent plan of how the neurological effects of stroke should be assessed, recorded, and monitored over time. The research recommendations focus on ways to address uncertainties so that evidence-based guidance on what should be done, when, and for whom to detect and respond to early neurological deterioration after stroke can be developed in the future. Finally, the chapter describes the strengths and limitations of the overall project before a short conclusion that outlines its original contribution to knowledge.

8.1. Summary of key findings by chapter

Chapter 4 identified the scales used in neurological assessment and monitoring and presented the series of reviews to identify, collate and evaluate the clinimetric evidence of these scales. The results allowed comparison between scales across clinimetric properties for which there was data. The data was generally limited, and it showed such variability that conclusions regarding which scale was superior were hard to reach. Questions about the effectiveness of scales to identify END in clinical practice, especially in more severe strokes were raised. Non-stroke-specific scales such as the Glasgow Coma Scale (GCS) are not sensitive to change in patients following a stroke and there are indications that they are not fit for practice in this population.

Reliability across items was found to vary with certain items such as Ataxia, Visual Fields, Gaze, and Extinction having lower reliability statistics than items such as Motor Power and Level of Consciousness (LOC)- Commands. There were small indications that training can improve the reliability of item ratings. However, reliability is only one property of an assessment and does not indicate the usefulness of the item for its purpose, in this case, identification of change. Different items are needed to detect deterioration across multiple stroke types and territories and need to be chosen for their usefulness and responsiveness regardless of their reliability. However, improved assessment descriptors and training could be used to increase reliability.

It was concluded from the data that assessment and identification within individual items is more receptive to the identification of change than the quantification of a total scale score which could prevent change from being identified. Overall, the methodological quality of the scales was poor especially in relation to testing within a whole stroke population meaning that the clinimetric properties have not been robustly tested. Furthermore, the widely accepted NIHSS training and certification allowed variation in scoring which could be further amplified in clinical practice.

The clinimetric properties of the scales did not provide any key revelations in terms of what should be used in assessment and monitoring practice. The findings indicated that was needed was a new tool containing a series of items to direct rather than quantify identification of change. However, clinimetric properties remain an important aspect of any measurement tool and should help guide future development and testing. In terms of identification of change responsiveness, measurement error, and reliability are key properties for future consideration. Time to complete assessments although not a clinimetric property is important due to its fundamental impact on workload. To be more useful in clinical practice, acceptable levels of properties for assessments need to be agreed.

The UK wide survey of stroke units in Chapter 5 established current practice and briefly explored clinicians' experiences of neurological assessment and monitoring in the acute phase of stroke whilst clarifying the current level of variation. Despite a general agreement that neurological assessment and monitoring is important for all patients following a stroke, practice was found to vary and this was greater than anticipated. Unwarranted clinical variation is occurring both within and across stroke units including differences in frequency and length of time of monitoring, identification and action on deterioration, and training provided.

A key finding of the survey is that clinicians want change in this area. Currently, there is a reliance on the non-stroke specific GCS and NEWS for monitoring and there is also a lack of guidance to support practice. There is a call for the development of protocols that are specific to stroke, achievable in busy clinical environments, and result in appropriate action if deterioration is noted. The results also directed the conclusion that the whole stroke MDT should be involved in neurological assessment and monitoring practice.

Chapter 6 reports on the series of semi-structured interviews completed to provide more depth of knowledge, understanding, and acceptability of neurological assessment and

monitoring and explored the barriers and facilitators to its implementation in clinical practice. The interviews validated many of the survey findings and provided further explanation of the quantitative results (Creswell, 2003). Using Normalisation Process Theory (NPT) for the development and analysis of the interviews meant the results provided insights into the barriers and facilitators that affect this area of practice. To increase the chance of successful future change, these factors need to be considered in both design and execution of implementation strategies.

The interviews identified factors that impact neurological assessment and practice, such as staffing levels, experience, trust, team structure, and subjectivity in assessment. Although the importance of monitoring for all stroke patients is widely purported, certain stroke types and treatments were prioritised, and this appears driven by the presence of protocols and the potential for intervention. Delivery and prioritisation of assessment and monitoring were often described as tailored to patient characteristics but there was little detail provided as to how these decisions were made indicates there might not be clear pathways even within a single unit. There was an awareness that out of hours decision-making processes may be even less robust which also suggests that variation could be wider than reported.

The development of the SNOBBS and associated decision flowchart is outlined in Chapter 7. Although further research is needed the SNOBSS represents a key step towards enabling stroke clinical practice to change and reduce variation in neurological assessment and monitoring. It could result in the systematic and effective identification of functionally meaningful changes in neurological status when used frequently by a range of clinicians. The SNOBSS centres on being able to focus and standardise the detection of change whilst allowing flexibility across different contexts such as patient characteristics. The consensus-building approach applied to the development of the SNOBSS also expediated the classification of key uncertainties that need to be answered to allow future development of evidence-based guidance in neurological assessment and monitoring practice.

8.2. Brief synthesis of overall findings

Critical realism has provided a valuable theoretical lens for this mixed-methods research and has supported greater exploration of neurological assessment and monitoring after stroke (Minger, 2004; Sayer, 2000). The methodological pluralism has supported the development of different types of knowledge. Around frequency of neurological assessment and monitoring, for example, this thesis has created new knowledge around what is done in practice, the

variation within it, and some of the interconnected factors that can impact on its delivery across multiple layers of reality both seen and unseen (Bhaskar, 1978). Neurological assessment and monitoring sits in the domain of the empirical because it can be observed and experienced but the factors that impact on its completion and accuracy can exist in and interconnect across different layers of reality. For instance, staffing levels are observable events (empirical) but mechanisms such as cultural attitudes towards the completion of neurological assessment and monitoring (real) could impact on motivation and diligence to undertake it (actual).

Critical realism has been useful as it supports the exploration of complexity, such as the complex adaptive systems where neurological assessment and monitoring practice resides, whilst seeking answers to real problems (Syed et al, 2010; Mingers, 2011). Critical realism opens up the space between empiricism and interpretivism (Mingers, 2004). This thesis through adopting a critical realist lens in both the design and delivery has provided multiple insights into contextual differences that impact on neurological assessment and monitoring within stroke services. This approach supported the development of an awareness and better understanding of factors and highlighted the interconnectedness at an individual, unit, and organisational level that impact on neurological assessment and monitoring practice in multiple different ways.

The impacts of the factors on neurological assessment and monitoring practice can vary depending upon whether they coincide with others and whether they are complementary or conflicting. Factors can therefore cause numerous and different impacts on delivery, sometimes in unpredictable ways. A simple example of a different effect of factors found within the research would be that in some units the lack of out of hours stroke medical cover increases variation whereas in others it has meant other members of the stroke team are involved in decision making and ensuring parity of service for patients regardless of time of day and medical staffing cover available. A seemingly independent factor has different effects because of other, seen and unseen, interconnected factors that impact on, with, or against it. In this example, individual experience, staffing levels, and trust in staff have been reported to impact on the out of hours decision processes if there is a lack of medical cover.

Critical realism affirms that much of reality exists and operates independently of our knowledge and awareness of it, however it also recognises the social nature of the world and that our knowledge of things is experienced and known through human minds. Within this thesis it has enabled development of new knowledge and a future research agenda whilst also

considering ways to intervene and change reality to abolish or at least mitigate factors that could negatively impact on neurological assessment and monitoring practice (Danermark, 2019). However, the author is aware that what has been created is best knowledge at this time and not a mirror of reality as it is not possible to observe and realise every factor that impacts on this or any element of care (ibid).

The research overall has enabled a clearer understanding of multiple factors that need to be considered and addressed to reduce variation in current practice and successfully implement change moving forward. Although the research had provided real insight into neurological assessment and monitoring the critical realist lens also suggests that there are factors that influence this area of care within their complex adaptive systems that might never be known. Future implementation work needs to be open, flexible, and adaptive where possible to allow for both known and unknown influences.

8.3. Recommendations

The programme of research presented in this thesis has allowed the identification of key recommendations. These are presented under two headings: practice and research recommendations.

8.3.1. Practice Recommendations

Based on the study findings these recommendations outline things that could be changed, adjusted, or improved in current practice to improve standardisation and reduce variation. Adoption of these recommendations should help ensure equitable and timely care provision within units. This in turn should help reduce unwarranted clinical variation and improve the prompt identification of Early Neurological Deterioration (END).

- **Develop agreed local procedures and proformas for neurological assessment and monitoring**

  The survey results indicated that many units (n=100) had local protocols or guidelines relating to neurological assessment and monitoring of patients within 72 hours (section 5.4.3.3). However, from the small number of clinical documents supplied (section 5.4.10) and the interview data (6.5.2.3) it became clear that these documents did not provide real detail in terms of what should be happening, when and for which patients in terms of neurological assessment and monitoring.

Multiple things could be done to enable and empower the workforce through operational models of service (Hargroves & Lowe, 2022). Agreed and consistent monitoring practices that include the whole stroke population could help reduce local variation within units. The use of protocols could ensure all staff are aware of what is expected in terms of assessment and intervention. (Tomson & van der Veer, 2013).

Protocols would need to be easily accessible and available to all those involved in neurological assessment and monitoring practice. A well-developed protocol would provide staff with more autonomy and ensure more consistency in practice. The introduction of local guidance could help facilitate the commencement and completion of monitoring as well as support the escalation of change, specifically deterioration. This should reduce variation in practice especially based on time and day of admission as reported in this study. These protocols should be incorporated into new starter and bank staff inductions to raise awareness of the importance of neurological assessment and monitoring practice.

It is felt that non-stroke specialists who support the acute stroke care pathway should also have access to these proformas to support stroke care provision across the pathway and to outliers from stroke services. This would hopefully support a reduction in variation especially out of hours. However, this will need a careful and planned introduction utilising strategies to engage the whole team as this project has shown that even where guidelines and protocols exist there is variation in both the assessments used and the frequency of monitoring. Where protocols currently exist, it appears to support engagement and completion of monitoring, but training and system processes need to be developed to ensure it happens in an effective and timely manner.

- **Ensure where possible stroke services have adequate staffing levels. Avoid moving staff from stroke services.**

  Multiple results from the survey and interviews highlighted that without adequate staffing levels stroke teams are unable to maintain standards of care including completion of neurological assessment and monitoring (Sections 5.4.5, 5.4.8.8, & 6.5.3.1). Inadequate staffing levels are known to have implications for care provision (Royal College of Nursing, 2017). It is also acknowledged that low staffing contributes to poor recognition and management of deteriorating patients. (Bray et al., 2014; McGaughey et al., 2017). However, it is recognised that this is not a simple solution as

stroke services currently have a lack of person-power to deliver services (Hart et al.2019; King's College London, 2021). The NHS Long Term Plan (NHS, 2019) identifies the need to build staff numbers and the skill mix necessary to support stroke pathways. However, this might require the blurring of professional boundaries (CordisBright, 2018). The use of standardised neurological assessments should help services identify and manage deteriorating patients. However, organisation management should be aware of the acuity of stroke services and the specialist skills required and as well as promoting recruitment they should limit staff rotation to other areas where possible.

- **Development of a team culture that values and supports neurological assessment and monitoring practice**

  The interviews highlighted the importance of trust between staff undertaking neurological assessment and monitoring (Section 6.5.3.2). Collaborative cultures with trusting relationships and reflective team learning are essential for the formation of effective teamwork (NHS England, 2014). The importance of team culture and leadership in successful stroke teams is acknowledged (Hargroves & Lowe, 2022). The generation of a culture where neurological assessment and monitoring is seen as an essential element of care should be a priority. From the survey data, there seemed to be an assumption that all members of the stroke MDT understand the importance of this element of care. However, it also highlighted some disparity in terms of perceived importance and whether the practice was seen as worthwhile indicating there might be greater differences in the prioritisation of this workload than reported. Development of a culture with a shared understanding that neurological assessment and monitoring practice is an essential element of care for all stroke patients would be valuable. Establishing a shared purpose serves as a critical driver for success for teams undertaking change (NHS England, 2012). Strategies to improve this could include feedback on score completion, audits, and case discussions to highlight the importance of neurological assessment and monitoring. A stronger and more cohesive team approach to monitoring delivery and associated decision-making could increase engagement, ownership, and trust (Rosen et al., 2018).

- **Develop stronger communication to support neurological assessment and monitoring practice**

Effective communication between healthcare professionals and with patients and carers is important (Ratna, 2019; Vermeir et al., 2015). The importance of communication was a recurring theme throughout the survey and interviews. However, there was variation in both verbal and written communication. The current lack of consistency means that staff are not clear about procedures. There was a lack of awareness about whether information on a patient's neurological status was included in handovers. There were also reported discrepancies in the completion of record-keeping in this area and where information is recorded. Additional consideration and planning will be required for those areas that currently have a mixture of paper-based and electronic systems for recording assessment, monitoring, and patient records. Furthermore, if units have different staff completing physiological observations and neurological monitoring, systems need to be in place to ensure any parameter or item changes are shared. Having agreed processes in terms of what needs to be recorded, when it should be included in handover, and how and where it should be documented could help reduce this variation and promote better consistency of communication.

Improving communication links strongly with the development of a team culture as it would help support shared understanding. Improved communication between clinicians would also potentially enhance assessment and monitoring completion through better prioritisation and planning of workload. Although it was generally accepted that communication with patients and carers was important there was variation in whether and how this was completed. Simple communication strategies are warranted to better inform patients and carers of the importance of neurological assessment and counsel them on the implications, especially in terms of disturbance. This could encourage enhanced cooperation with neurological assessment and monitoring practices.

- **Use of stroke specific assessments**

The reviews highlighted limitations of none-stroke-specific scales especially the GCS (Chapter 4). The survey and interview findings also show that clinicians feel that scales in common use, such as the GCS are not fit for purpose and that subjectivity in assessment can accentuate problems such as reliability with their use (Table 5.13 & Section 6.5.3.1). Despite the awareness of many of the limitations of the GCS it is the

most widely used scale in practice for ongoing neurological monitoring after stroke (Table 5.2).

This thesis advocates the abolition of the GCS in stroke practice and the use of stroke-specific assessments such as the newly developed SNOBSS. The GCS is a level of consciousness scale that was designed for use with traumatic brain injury patients. Although from the results, specifically the survey, a reduced level of consciousness is the most widely recognised sign of deterioration it represents a late sign. Stroke-specific assessments have the potential to pick up on earlier subtle signs of change, such as alterations in speech or limb power, and may help improve outcomes in patients with a treatable cause of END and guide management decisions for all patients.

- **Enhanced training provision to improve engagement with and quality of care provision in neurological assessment and monitoring**

  It is acknowledged that a skilled workforce is critical to maintaining and improving the quality of stroke care and that training can enable and empower individuals and teams (Hargroves & Lowe, 2022). For many years it has been recommended that staff working in stroke services receive stroke specialist training (Fisher et al., 2011; Royal College of Physicians, 2016a). Although the training sections within the survey were overall poorly completed (Section 5.4.9), it appears there is a lack of formalised accredited inter and transdisciplinary training avenues for neurological assessment and monitoring. Although the NIHSS training appears to be universally accepted there are no indications that clinicians are aware of the limitation of the training due to the variation allowed within the certification scoring (Chapter 4). Completion, therefore, does not indicate conformity in scoring and competency as assumed by many in clinical practice.

  Currently, the majority of training, in terms of knowledge and skills are provided in-house, over time, and with experience in the role which is consistent with training in other aspects of stroke services (Jones et al., 2018; Smith et al., 2008). However, this research advocates that all stroke team members should receive regular and repeated training in neurological assessment and monitoring practices. Involvement of all members of the MDT is advocated to support shared understanding and practice and to improve communication. However, the thesis also advocates a better emphasis on content and delivery methods.

Nationally recognised, quality-assured, and transferable education programmes in this element of stroke care should be established through the Stroke Specific Education Framework (SSEF) to develop more consistency of practice (Health Education England, 2022). Training that encourages and supports consistency of assessment approach and that helps reduce variation would be classed as effective. Training should also be deemed more important than perceived experience as the interview findings indicated that variation in assessment is currently seen and reported across all staff members regardless of their experience. Training should be made as accessible as possible with a mixture of face to and online training available.

<u>Specific suggestions about training content that have been deduced or taken from the research findings:</u>

- Rationale behind why neurological assessment and monitoring is important for all stroke patients. Developing an understanding in those trained that it should be a prioritised element of care because accurate and timely completion, and where possible intervention, has the potential to improve patient outcomes (Section 6.5.1.4 & 6.5.4.5).
- Regardless of the choice of assessment those using it should understand the justification and know why it is important to assess specific aspects of the patient's status. Training should provide clear explanations and demonstrations on how to assess all items. Extra focus should be provided on items that are difficult to assess especially in unresponsive or uncooperative patients (Section 6.5.4.5). Feedback should be provided to participants to ensure they learn from the experience and to support improved consistency of assessment and identification of change in an item. Reliance on scale total scores should be eradicated as these could be concealing important changes in a patient's condition (Chapter 4).
- Enhanced explanation of why it is important to wake patients to complete assessments. Patient and carer perspectives could be included to provide reassurance about the importance of these assessments. However, training should include an understanding and examples of how disturbed sleep can directly impact on the assessment itself to maintain consistency (Section 6.5.3.1).
- Knowledge about what represents a meaningful change in condition or function and how this should be escalated for further assessment. The

development of confidence in detecting change could be supported through the use of real-world case presentations (Chapter 7).

- Awareness that subtle changes in patient condition and function could be a sign of deterioration (Chapter 5).
- An understanding of the links between physiological and neurological monitoring as well as the importance of managing both (Section 5.4.6).
- Training that develops a more unified understanding amongst professionals should improve consistency of assessment and improve communication (Section 6.5.4.5).

- **Development of and adherence to competency assessments to maintain high standards of assessment and promote trust in neurological assessment and monitoring practice**

  Forty-eight percent of respondents to the survey reported having competency assessment in place (Section 5.4.9). However, it became clear that completion of the NIHSS training was deemed by some as a competency. Training should ensure individuals have the knowledge, skills, and attitudes to complete neurological assessment and monitoring. There was a desire within the data for more and better competencies to support neurological assessment and monitoring practice (Section 6.5.4.5). Performance-based competency assessments should be developed and accomplished by all staff members. Practice should be regularly audited to ensure continues proficient practice by all MDT members. Adherence to competency assessments should promote a consistent approach as well as confidence and trust in their own and other abilities to complete neurological assessment and monitoring. Trust is a key driver of behaviour within teams and there is some evidence to illustrate that trust between staff can be more important than structures in care delivery (Imison, 2016).

8.3.2. Recommendations for future research

The following research recommendations outline key elements of work that need to be done to take the next steps in developing a consistent plan of how the neurological effects of stroke should be assessed, recorded, and monitored over time. They are presented under two broad headings of effectiveness and implementation and acceptability. They outline the priority questions that need addressing to inform practice, produce evidence-based guidelines, and

support successful implementation and evaluation of the SNOBSS and other changes in neurological assessment and monitoring.

This thesis has identified multiple uncertainties concerning neurological assessment and monitoring practice and highlighted areas where evidence is fundamentally lacking. It has emphasised the complexity of neurological assessment and monitoring as an intervention delivered within complex adaptive systems. Although complexity in healthcare is generally well recognised health services research often continues to operate in a paradigm looking at linear causal effects.

To address the uncertainties that exist in monitoring and clinical response future research needs to be devised from a complex intervention perspective using appropriate frameworks to ensure the evidence answers the relevant questions and gaps in knowledge (Skivington et al., 2021b). These recommendations outline the questions that are most useful to decision makers and not those that can be answered with the greatest certainty. To address these and generate meaningful findings future studies need to be developed that offer a flexible and emergent approach to exploring them (Greenhalgh & Papoutsi, 2018).

8.3.2.1. Effectiveness

These questions focus on ensuring that neurological assessment and monitoring will have a meaningful effect on patients in normal clinical conditions (Burches & Burches, 2020).

- **Which items are the most useful to detect END after stroke and specifically which are the most useful based on specific stoke types, severity or other patient characteristics? What is the optimum frequency to complete neurological monitoring?**

  The whole study team, interview participants, and expert group members were aware that it may not be possible to identify a one size fits all SNOBSS and decision flowchart (Chapters 3 & 7). Different stroke subtypes may manifest differently in terms of clinical change before and during deterioration. Fundamentally we need a practical and acceptable selection of items, or series of items, to use at the bedside that has adequate responsiveness to change to identify END, ideally across a range of patients. The development of the SNOBSS addressed this, but the selection of items was

pragmatic and based on consensus approaches. There remains a real awareness that the best items to use are unknown.

The SNOBSS could have additional items added to ascertain which items best detect change, specifically deterioration. Extensive data collected across the whole heterogenous stroke population may lead to identification of different items being most responsive to change dependent on stroke type, severity, and other patient characteristics. Greater agreement on what represents an important change in specific items needs to be developed which would in time help with the development of a clinical and useable definition of END.

Real world testing of the effectiveness of SNOBSS and other items in detecting change is essential. The review chapter highlighted that most assessments were tested under experimental and not clinical conditions. It also indicated that there can be a pronounced difference in measurement properties between experimental and real-world conditions. A large scale stepped wedge evaluation could detect if deployment was associated with better outcomes. Effectiveness to detect change can be assessed alongside inter-rater and intra-rater agreement, across a whole stroke population when assessed by multiple different professionals.

Evidence collected from real world application of the SNOBSS linked to outcome data (e.g., SSNAP) could highlight differences across sub-populations. Analysis of this big data could lead to the development of different versions of the SNOBSS. Targeted assessment has the potential to reduce workload and resource use whilst resulting in more effective identification of meaningful change. However, this would need further research in terms of implementation to ensure staff knew and used the correct version on the right patient. If there were too many versions, it could be confusing and possibly not practically applicable.

As well as evidence for which items are the most useful to identify meaningful change, we need data to support the optimal timings of monitoring. This could vary depending upon stroke type, severity or other patient characteristics. The most common parameters identified within the survey data might help devise ranges and limits for the testing approach. It would probably not be feasible to robustly test every potential and actual neurological deficit that could be attributed to stroke in all patients at each observation time point depending upon the frequency decided (Ayis et al., 2013).

Again, large-scale real-world evaluations would be needed to devise evidence for the most effective monitoring schedules.

Development of the SNOBSS into an electronic version could support even more in-depth data collection and analysis. In the longer-term algorithms could be developed to support identification of meaningful change based on time since stroke, stroke type, baseline severity, and other factors.

- **When should neurological monitoring be discontinued?**
  In terms of discontinuation, the survey data was limited (Section 5.4.4), and no clear patterns were found. There were some indications that clinicians are unsure about when to stop with real ranges of time periods of neurological monitoring reported. This provides a real equipoise in that it is unclear for how long monitoring should be continued. The balance is between maintaining safety and detecting deterioration against unnecessary resource use and disturbance of the patient. Robust big data should allow identification of the point at which real change is unlikely to occur for different stroke populations.

- **What is the best way to assess specific items?**
  Differences in methods and descriptors of assessment are an important a source of variation and need to be eradicated (Powell et al., 2003). The way that an item is assessed could affect important clinimetric properties such as inter-rater reliability. As this thesis and the SNOBSS advocate change in an item rather than total score change other important clinimetric properties such as responsiveness, and measurement error will not be calculable. However, the theoretical grounding of these properties need to be considered and balanced with other factors including understanding and acceptability (Chapter 4). Different approaches and instruction formats should be tested against each other with stakeholder involvement to ensure they are practical and feasible to be adopted into practice. Stakeholders involved need to be broad to cover different professional groups, grades, and levels of experience as well as patients. Testing of different assessment methods for items will need to be completed to ascertain key differences in application and results. Once the key items and their descriptors are agreed a robust training schedule would ned to be developed to support delivery of the SNOBSS in practice.

Whether the SNOBSS does truly identify a change in condition and maintains patient safety is crucial data. However, effectiveness is only a small portion of what needs to be known to support the adoption of SNOBSS into routine practice across the UK. Implementation studies will be required to assess and evaluate other key concepts of the SNOBSS such as outlined below.

8.3.2.2. Implementation/ Acceptability

- **How well is the SNOBSS adhered to in practice?**

  Pilot testing should test the feasibility of SNOBSS and allow the collection of fidelity data on how well the SNOBSS is adhered to. Analysis of this could occur in several ways which both have merit. Descriptive with inferential methods applied to aid interpretation (e.g., are there certain times of day where observations are more often missed) or more complex examination of coverage, that is the extent to which eligible patients receive the intervention. It would allow exploration of whether decision-making and escalation of change are undertaken appropriately. Given the flexibility built into the SNOBSS it would be important to investigate its application across multiple settings to ensure that patients are receiving the appropriate monitoring. As the SNOBSS has been developed to address variation in practice ensuring equity of care within specific groups is an essential aim.

- **How best to implement the SNOBSS or other changes in neurological assessment and monitoring practice?**

  The thesis has provided some insights into the different contexts that impact on neurological assessment and monitoring practice. However complex intervention research will be needed to ask broad questions about how the context in which the intervention is implemented and how the CAS and the intervention adapt to each other (Skivington et al. 2021b). Changes will occur as a result of the implementation, but these can be both intentional and unintentional. The updated Medical Research Council guidance for developing and evaluating complex interventions draws heavily on realist principles and emphasises the importance of context, the development of programme theories, and gaining an understanding of the interaction between context and causal mechanisms in generating outcomes (ibid). Implementation research should be designed to uncover what works, for whom, under what circumstances, and to what extent.

This thesis obtained some information on the use of electronic observation systems for neurological monitoring. Although it ascertained that uptake is mixed across units there was not enough information to assess the impact of electronic systems for neurological assessment after stroke. Based on the results of the thesis a key recommendation would be that services should assess the impact these systems have on neurological assessment and monitoring practice. These systems could support the delivery of high-quality service and release more time for care (Hargroves & Lowe, 2022) or they might worsen the very problems they were introduced to solve (Greenhalgh & Papoutsi, 2018).

Implementation theory, such as NPT, could be used to create and apply knowledge to improve the implementation process (May et al., 2018). Successful implementation should go beyond intervention fidelity and embrace tailoring and adaption with key stakeholders to attend to rather than control for complexity (Braithwaite et al., 2018). Implementation plans should be flexible and be refined based on findings to enhance successful implementation and sustained embedding as an intervention is transferred across contexts.

- **How acceptable is the SNOBSS to clinicians and patients?**
  Chapter 7 outlined some preliminary assessments of acceptability with clinicians; however, this would need further testing in practice with both professionals and patients. Frequency of completion will also impact resource issues with the use of the SNOBSS. Feasibility testing and evaluation will allow identification of problems and gather suggestions in terms of what could be improved and how wider implementation could be supported. The SNOBSS could be updated as evidence becomes available to support better development. Flexibility is not a flaw if the SNOBSS is delivering its key function to detect change (Hawe et al., 2004).

- **Is the SNOBSS cost-effective?**
  The Medical Research Council (MRC) guidance for developing and evaluating complex interventions supports early consideration of economic analysis. Careful planning would be needed to support cost-effectiveness analysis (Ramsey et al., 2005). Economic analysis should use a broad perspective with an understanding that the time over which it is undertaken will impact the results (Skivington et al. 2021a&b).

The research recommendations have focused on identifying and prioritising answerable research questions and not the methods that should be used. Different methods have different strengths and weaknesses (McKee et al., 1999). Future research needs to develop methods to maximise their usefulness in contributing to decision making and health improvement (Skivington et al., 2021b). Adopting flexible methods that have a deliberate approach to achieve usefulness would also theoretically reduce research waste (Chalmers & Glasziou, 2009).

The current lack of neurological assessment and monitoring guidance needs addressing. Only with the development of a stronger evidence base can clear guidance be created on what should be used, when, how often, and for which patients and response if deterioration is noted. The development and application of clinical guidelines is complex (Plesk & Greenhalgh, 2001). National Clinical Guideline Methodology such as the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) principles should be utilised in the development (Guyatt et al., 2008). Once clear guidance is available development of or inclusion in existing auditing processes should be used to ensure adherence to the standards.

8.4. Strengths and Limitations

The thesis has met its aim and objectives (sections 1.7.1 & 1.7.2) to begin the development of a consistent plan of how the neurological effects of stroke should be assessed, recorded, and monitored over time through development of the Standardised Neurological Observation Schedule for Stroke (SNOBSS). This research has provided attention not only to the design of the SNOBSS but also to the conditions and resources that could impact on the implementation of change in this area. It used multi-methods research to explore various aspects of neurological assessment and monitoring practice. A key strength of this research was the engagement it received from clinicians which highlighted the perceived importance of the topic and the desire for change. Although the clinimetric review had limitations this was the first time that the evidence on multiple clinimetric properties across multiple scales was collated to allow comparison. The UK wide survey was also the first of its kind and provided much needed insights into current practice and the variation that presently exists.

Healthcare is complex and ever-changing, so the data collected technically only reflects the time of collection. However, the data was collected across multiple geographical areas and the replication of data via multiple methods provides a strong account of neurological assessment and practice. Without major drivers for change the landscape in this element of care is unlikely

to change rapidly or drastically. However, there is an awareness that all factors and contextual issues that affect this element of practice will not have been captured; practically because not all services were able to contribute to the data collection, and fundamentally because, as critical realism explains, not all factors can be seen and examined but also because complex adaptive systems are in a constant state of flux and are therefore never completely stable or knowable. However, the realist position through acknowledging these caveats aims to undertake the best possible research to produce credible findings, explanatory accounts and recommendations that are true to the context.

Both the survey and the interviews included a range of professionals' viewpoints. Although these cannot be claimed to be generalisable, there was concordance in several areas. That said, most of the staff involved were senior members of the stroke team and the findings may not represent the full views of all team members. Although the involvement of a greater range of professionals is a necessity for further work many of the factors identified as impacting on neurological assessment and monitoring are likely to persist regardless of different viewpoints. It is also likely that some of the practice recommendations would likely apply to other complex interventions within acute stroke care and not just neurological assessment and monitoring.

As this was a PhD programme of research some broad resource limitations could have impacted the overall project. This included the reviews being completed by the student as a lone reviewer and restrictions on the number of interviews that could be feasibly completed. However, systems and processes were in place to limit impact. These included completing the reviews in a systematic manner and the supervisory team having oversight and involvement where possible such as in agreeing the interviews data analysis. Although there were some potential issues because of the COVID-19 pandemic most of these were managed and the survey and interviews recruited well. The expert consensus group proceedings were squeezed because of the time limitations and potentially other methods, or drawing on wider opinions, could have strengthened the proceedings had time allowed. The experts were drawn from the UK and Ireland which was justified as these represented the services and systems under consideration. However, if experts from other healthcare systems had been included there may have been different opinions based on different pathways, resource levels, or experiences.

The author is a nurse with clinical and research experience in stroke care which was advantageous in terms of the author's knowledge of the subject area and context. However, she had to be mindful throughout the project that her experience in terms of variation in

practice might not have been reflected in other areas and that her task was an insightful enquiry. However, her previous research experience has taught her that there can be a large void between what areas think they do and the reality of their practice. Although the project has highlighted extensive variation there is potentially wider variation in practice which needs to be considered in future work.

8.5. Conclusion

This PhD has highlighted the current lack of evidence and widespread variation in stroke neurological assessment and monitoring practice. The research completed has extended the knowledge base around stroke neurological assessment and monitoring practice and made an original contribution in multiple ways. It has created an overview of clinimetric properties data across the range of stroke scales which allows comparison between scales and items. The UK wide survey and interviews have explored and described current practice, clinician experience and identified contextual factors that impact this element of care. The thesis has shown that current scales in widespread use are not fit for purpose and that there is need and desire for a stroke specific assessment like the SNOBSS to allow more timely and consistent identification of END. The exploration of the wider contexts and factors that impact this area of care has also allowed the identification of the uncertainties and gaps in evidence that need addressing to allow the development of clear guidelines in the future.

The findings along with expert knowledge and skills have been applied to develop a Standardised Neurological OBservation Schedule for Stroke (SNOBSS) and associated decision flowchart. Although more evidence is needed to support the development of evidence-based guidelines to guide what we should be doing when, how often, and for which patients the SNOBBS has the potential to detect meaningful changes in neurological status at the bedside systematically and effectively when used frequently by a range of clinicians. Further development of the SNOBSS and the evidence base behind stroke neurological assessment and monitoring could lead to better standardisation of processes which has the potential to reduce unwarranted clinical variation and ultimately improve outcomes for patients.

References

Adams, R. J., Meador, K. J., Sethi, K. D., Grotta, J. C., & Thomson, D. S. (1987). Graded neurologic scale for use in acute hemispheric stroke treatment protocols. *Stroke, 18*(3), 665-669.

Alasheev, A. M., Andreev, A. Y., Gonysheva, Y. V., Lagutenko, M. N., Lipin, G. I., Lokteva, E. E., Lutskovich, O. Yu., Mamonova, A. V., Prazdnichkova, E. V., & Belkin, A. A. (2017). Comparison of bedside and remote assessments on the NIHSS scale in patients in the acute period of stroke. *Neuroscience and Behavioral Physiology, 47*(7), 828-832.

Albanese, M. A., Clarke, W. R., Adams Jr, H. P., & Woolson, R. F. (1994). Ensuring reliability of outcome measures in multicenter clinical trials of treatments for acute ischemic stroke: The program developed for the Trial of ORG 10172 in Acute Stroke Treatment (TOAST). *Stroke, 25*(9), 1746-1751.

Alexandrov, A. V., & Grotta, J. C. (2002). Arterial reocclusion in stroke patients treated with intravenous tissue plasminogen activator. *Neurology, 59*(6), 862-867.

Alexandrov, A. V., Molina, C. A., Grotta, J. C., Garami, Z., Ford, S. R., Alvarez-Sabin, J., Montaner, J., Saqqur, M., Demchuk, A. M., Moye, L. A., Hill, M. D., & Wojner, A. W. (2004). Ultrasound-enhanced systemic thrombolysis for acute ischemic stroke. *N Engl J Med, 351*(21), 2170-2178.

Anderson, E. R., Smith, B., Ido, M., & Frankel, M. (2011). Remote assessment of stroke using the iPhone 4. *Journal of Stroke & Cerebrovascular Diseases, 22*(4), 340-344.

André, C. (2002). The NIH Stroke Scale is unreliable in untrained hands. *Journal of Stroke & Cerebrovascular Diseases, 11*(1), 43-46.

Appleton, J. P., Sprigg, N., & Bath, P. M. (2016). Blood pressure management in acute stroke. *Stroke Vasc Neurol, 1(*2), 72-82.

Arenillas, J. F., Rovira, A., Molina, C. A., Grivé, E., Montaner, J., & Alvarez-Sabín, J. (2002). Prediction of early neurological deterioration using diffusion- and perfusion-weighted imaging in hyperacute middle cerebral artery ischemic stroke. *Stroke, 33*(9), 2197-2203.

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. International *Journal of Social Research Methodology, 8*(1), 19-32.

Ashcraft, S., Wilson, S. E., Nyström, K. V., Dusenbury, W., Wira, C. R., & Burrus, T. M. (2021). Care of the patient with acute ischemic stroke (Prehospital and acute phase of care): Update to the 2009 Comprehensive Nursing Care Scientific Statement: A Scientific Statement From the American Heart Association. *Stroke, 52*(5), e164-e178.

Aslanyan, S., Weir, C. J., Diener, H.-C., Kaste, M., Lees, K. R., GAIN International Steering Committee and Investigators. (2004). Pneumonia and urinary tract infection after acute ischaemic stroke: a tertiary analysis of the GAIN International trial. *European Journal of Neurology, 11*(1), 49-53.

Aslanyan, S., Weir, C. J., Johnston, C. S., Krams, M., Grieve, A. P., & Lees, K. R. (2007). The association of post-stroke neurological improvement with risk of subsequent deterioration due to stroke events. *European Journal of Neurology, 14*(1), 1-6.

Astbury, B., & Leeuw, F. L. (2010). Unpacking black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation, 31*(3), 363-381.

Ayis, S. A., Coker, B., Rudd, A. G., Dennis, M. S., & Wolfe, C. D. (2013). Predicting independent survival after stroke: a European study for the development and validation of standardised stroke scales and prediction models of outcome. *Journal of Neurology, Neurosurgery & Psychiatry, 84*(3), 288-296.

Azzimondi, G., Bassein, L., Nonino, F., Fiorani, L., Vignatelli, L., Re, G., & D'Alessandro, R. (1995). Fever in acute stroke worsens prognosis. A prospective study. *Stroke, 26*(11), 2040-2043.

Baatiema, L., Otim, M. E., Mnatzaganian, G., de-Graft Aikins, A., Coombes, J., & Somerset, S. (2017). Health professionals' views on the barriers and enablers to evidence-based practice for acute stroke care: a systematic review. *Implementation science, 12*(1), 74-74.

Bagot, K. L., Cadilhac, D. A., Bladin, C. F., Watkins, C. L., Vu, M., Donnan, G. A., Dewey, H. M., Emsley, H., Davies, D. P., Day, E., Ford, G. A., Price, C. I., May, C. R., McLoughlin, A., Gibson, J., Lightbody, C. E., & VST and ASTUTE investigators (2017). Integrating acute stroke telemedicine consultations into specialists' usual practice: a qualitative analysis comparing the experience of Australia and the United Kingdom. *BMC health services research*, *17*(1), 751.

Baizabal-Carvallo, J. F., Alonso-Juarez, M., & Samson, Y. (2014). Clinical deterioration following middle cerebral artery hemodynamic changes after intravenous thrombolysis for acute ischemic stroke. *Journal of Stroke and Cerebrovascular Diseases, 23*(2), 254-258.

Barber, M., Wright, F., Stott, D. J., & Langhorne, P. (2004). Predictors of early neurological deterioration after ischaemic stroke: a case-control study. *Gerontology, 50*(2), 102-109.

Barber, R., Beresford, P., Boote, J., Cooper, C., & Faulkner, A. (2011). Evaluating the impact of service user involvement on research: a prospective case study. *International Journal of Consumer Studies, 35*(6), 609-615.

Baron, J. C., Macrae, I. M., Adams, H. P., & Dirnagl, U. (2013). ESC-BRAIN: Experimental and clinical stroke research - Do they connect?: Meeting report of the ESC-BRAIN joint symposium held in London and Shanghai in May 2013. *Cerebrovascular Diseases, 36*(4), 306-321.

Bartel, A. P., Beaulieu, N. D., Phibbs, C. S., & Stone, P. W. (2014). Human capital and productivity in a team environment: Evidence from the healthcare sector. *American Economic Journal: Applied Economics, 6*(2), 231-259.

Basauhra Singh, H. K., Chong, M. C., Thambinayagam, H. C., Zakaria, M. I., Cheng, S. T., Tang, L. Y., & Azahar, N. H. (2016). Assessing nurses knowledge of glasgow coma scale in emergency and outpatient department. *Nurs Res Pract, 2016*, 8056350.

Bath, P. M., Appleton, J. P., Krishnan, K., & Sprigg, N. (2018). Blood pressure in acute stroke: To treat or not to treat: That is still the question. *Stroke, 49*(7), 1784-1790.

Baxter, S. K., & Brumfitt, S. M. (2008). Professional differences in interprofessional working. *J Interprof Care, 22*(3), 239-251.

Begun, J. W., Zimmerman, B., & Dooley, K. J. (2003). Health care organizations as complex adaptive systems. In S. S. Mick, & M. E. Wyttenbach (Eds.), *Advances in health care organization theory (Ed.1) (*pp. 253-288). Jossey-Bass.

Benner, P. (1982). From novice to expert. *The American journal of nursing*, *82*(3), 402–407.

Bergin, M., Wells, J. S. G., & Owen, S. (2008). Critical realism: a philosophical framework for the study of gender and mental health. *Nursing Philosophy, 9*(3), 169-179.

Berthier, E., Decavel, P., Vuiller, F., Verlut, C., Moulin, T., & De Bustos, E. M. (2012). Review: Reliability of NIHSS by telemedicine. *European Research in Telemedicine, 1*(3-4), 111-114.

Berthier, E., Decavel, P., Vuillier, F., Verlut, C., Moulin, T., & de Bustos, E. M. (2013). Reliability of NIHSS by telemedicine in non-neurologists. *International Journal of Stroke, 8*(4), E11.

Bessenyei, M., Fekete, I., Csiba, L., & Bereczki, D. (2001). Characteristics of 4 stroke scales for the detection of changes in clinical signs in the acute phase of stroke. *Journal of Stroke & Cerebrovascular Diseases, 10*(2), 70-78.

Bhaskar, R. (1978). *A realist theory of science*. Harvester Press.

Bhaskar, R. (1979). *Philosophy and the human sciences.* Harvester Press.

Bhaskar, R. (1989). *Reclaiming reality: a critical introduction to contemporary philosophy.* Verso.

Binz, S., Khatri, R., Carroll, J., Prusakov, P., & Chang, F. L. (2013). The feasibility of IPAD technology for remote acute ischemic stroke assessment. *Neurology. Conference: 65th American Academy of Neurology Annual Meeting. San Diego, CA United States. Conference Publication, 80* (1 Meeting Abstracts).

Birschel, P. (2005). *Progressing stroke: definitions, risk factors and relation to physiological instability.* [Unpublished doctoral thesis]. Newcastle University.

Birschel, P., Ellul, J., & Barer, D. (2004). Progressing stroke: towards an internationally agreed definition. *Cerebrovascular Diseases, 17*(2-3), 242-252.

Bjerkan, J., Valderaune, V., & Olsen, R. M. (2021). Patient safety through nursing documentation: Barriers identified by healthcare professionals and students. *Frontiers in Computer Science, 3* (51): 624555

Black, N., Murphy, M., Lamping, D., McKee, M., Sanderson, C., Askham, J., & Marteau, T. (1999). Consensus development methods: a review of best practice in creating clinical guidelines. *J Health Serv Res Policy, 4*(4), 236-248.

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*(8476), 307-310.

Boers, M., Kirwan, J. R., Wells, G., Beaton, D., Gossec, L., d'Agostino, M. A., Conaghan, P. G., Bingham, C. O., 3rd, Brooks, P., Landewé, R., March, L., Simon, L. S., Singh, J. A., Strand, V., & Tugwell, P. (2014). Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *Journal of clinical epidemiology*, *67*(7), 745–753.

Bojke, L., Claxton, K., Bravo-Vergel, Y., Sculpher, M., Palmer, S., & Abrams, K. (2010). Eliciting distributions to populate decision analytic models. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, *13*(5), 557–564.

Bojke, L., Soares, M., Claxton, K., Colson, A., Fox, A., Jackson, C., Jankovic, D., Morton, A., Sharples, L., & Taylor, A. (2021). Developing a reference protocol for structured expert elicitation in health-care decision-making: a mixed-methods study. *Health technology assessment (Winchester, England)*, *25*(37), 1–124.

Bolger, F., & Rowe, G. (2015). The aggregation of expert judgment: do good things come to those who weight?. *Risk analysis : an official publication of the Society for Risk Analysis*, *35*(1), 5–11.

Boone, W. J. (2016). Rasch Analysis for Instrument Development: Why, When, and How?. *CBE life sciences education*, *15*(4), rm4.

Boulanger, J. M., Lindsay, M. P., Gubitz, G., Smith, E. E., Stotts, G., Foley, N., Bhogal, S., Boyle, K., Braun, L., Goddard, T., Heran, M., Kanya-Forster, N., Lang, E., Lavoie, P., McClelland, M., O'Kelly, C., Pageau, P., Pettersen, J., Purvis, H., Shamy, M., Tampiere, D., vanAdel, B., Verbeek, R., Blacquiere, D., Casaubon, L., Ferguson, D., Hegedus, Y., Jacquin, G. J., Kelly, M., Namal, N., Linkewich, B., Lum, C., Mann, B., Milot, G., Newcommon, N., Poirier, P., Simpkin, W., Snieder, E., Trivedi, A., Whelan, R., Eustace, M., Smitko, E., & Butcher, K. (2018). Canadian stroke best practice recommendations for acute stroke management: Prehospital, emergency department, and acute inpatient stroke care, 6th Edition, Update 2018. *International journal of stroke: official journal of the International Stroke Society*, *13*(9), 949–984.

Boutot F., Milojevic K., Parisse S., Lorilloux S., Sar A., Dalle L., Koukabi, M., & Lambert, Y. (2013). Is Mobile intensive care unit worthwhile for suspected acute stroke? *Cerebrovascular Diseases*, 35(Suppl. 3), 684.

Braithwaite, J. (2018). Changing how we think about healthcare improvement. *BMJ (Clinical research ed.)*, *361*, k2014.

Braithwaite, J., Churruca, K., Long, J. C., Ellis, L. A., & Herkes, J. (2018). When complexity science meets implementation science: a theoretical and empirical analysis of systems change. *BMC Medicine, 16*(1), 63.

Braun, V., & Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.

Braun, V., & Clarke, V. (2021a) One size fits all? What counts as quality practice in (reflexive) thematic analysis?, Qualitative Research in Psychology, 18(3), 328-352.

Braun, V., & Clarke, V. (2021b) To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health*, 13(2), 201-216.

Bray, B. D., Ayis, S., Campbell, J., Cloud, G. C., James, M., Hoffman, A., Tyrrell, P. J., Wolfe, C. D., & Rudd, A. G. (2014). Associations between stroke mortality and weekend working by stroke specialist physicians and registered nurses: prospective multicentre cohort study. *PLoS medicine*, *11*(8), e1001705.

Britton, M., & Rödén, A. (1985). Progression of stroke after arrival at hospital. *Stroke, 16*(4), 629-632.

Brott, T., Adams, H. P., Jr, Olinger, C. P., Marler, J. R., Barsan, W. G., Biller, J., Spilker, J., Holleran, R., Eberle, R., & Hertzberg, V. (1989). Measurements of acute cerebral infarction: a clinical examination scale. *Stroke*, *20*(7), 864–870.

Brouwers, H. B., & Goldstein, J. N. (2012). Therapeutic strategies in acute intracerebral hemorrhage. *Neurotherapeutics, 9*(1), 87-98.

Brown, E. B., Tietjen, G .E., Deveshwar, R. K., Ramadan, N. M., Levine, S. R., Dietrich, K. L., Nazareno, F., & Welch, K. M. A. (1990). Clinical stroke scales: an intra- and inter-scale evaluation. *Neurology* 40 (Suppl 1), 352.

Bruce, B., Fries, J., Lingala, B., Hussain, Y. N., & Krishnan, E. (2013). Development and assessment of floor and ceiling items for the PROMIS physical function item bank. *Arthritis Research & Therapy, 15*(5), R144. doi:10.1186/ar4327

Bugnicourt, J.-M., Roussel, B., Garcia, P.-Y., Canaple, S., Lamy, C., & Godefroy, O. (2011). Aspirin non-responder status and early neurological deterioration: A prospective study. *Clinical Neurology and Neurosurgery, 113*(3), 196-201.

Burches, E., & Burches, M. (2020) Efficacy, effectiveness and efficiency in the health care: The need for an agreement to clarify its meaning. *Int Arch Public Health Community Med 4*:035.

Burns, M., Somers, K., McElwaine, P. and Harbison, J. (2014). (Letter) National Institutes of Health Stroke Scale object naming test in a non-American English-speaking population. Int J Stroke, 9: 35.

Busto, R., Dietrich, W. D., Globus, M. Y., Valdés, I., Scheinberg, P., & Ginsberg, M. D. (1987). Small differences in intraischemic brain temperature critically determine the extent of ischemic neuronal injury. *J Cereb Blood Flow Metab, 7*(6), 729-738.

Busto, R., Globus, M. Y., Dietrich, W. D., Martinez, E., Valdés, I., & Ginsberg, M. D. (1989). Effect of mild hypothermia on ischemia-induced release of neurotransmitters and free fatty acids in rat brain. *Stroke, 20*(7), 904-910.

Cabal, M., Holes, D., Kral, J., Vaclavik, D., Klecka, L., Mikulik, R., & Bar, M. (2018). Inter-rater reliability between paramedics and neurologists in assessment of hemiplegia in acute stroke did not improve after re-education. *European Stroke Journal, 3 (*Suppl. 1*)*, 96.

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding In-depth semistructured Interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research, 42*(3), 294-320.

Campbell, M. K., Piaggio, G., Elbourne, D. R., Altman, D. G., & CONSORT Group (2012). Consort 2010 statement: extension to cluster randomised trials. *BMJ (Clinical research ed.)*, *345*, e5661.

Cantrill, J., Sibbald, B., & Buetow, S. (1996), The Delphi and nominal group techniques in health services research. *International Journal of Pharmacy Practice*, 4: 67-74.

Carney, O., McIntosh, J., & Worth, A. (1996). The use of the Nominal Group Technique in research with community nurses. *Journal of advanced nursing*, *23*(5), 1024–1029.

CASP (Critical Appraisal Skills Programme). (2018) *CASP Qualitative Checklist*. Retrieved March 12, 2023, from https://casp-uk.net/casp-tools-checklists/

Castillo, J. (1999). Deteriorating stroke: Diagnostic criteria, predictors, mechanisms and treatment. *Cerebrovascular Diseases, 9* (Suppl. 3), 1-8.

Castillo, J., Dávalos, A., Marrugat, J., & Noya, M. (1998). Timing for fever-related brain damage in acute ischemic stroke. *Stroke, 29*(12), 2455-2460.

Castillo, J., Dávalos, A., & Noya, M. (1997). Progression of ischaemic stroke and excitotoxic aminoacids. *Lancet, 349*(9045), 79-83.

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *Lancet, 374*(9683), 86-89.

Chapman Smith, S. N., Govindarajan, P., Padrick, M. M., Lippman, J. M., McMurry, T. L., Resler, B. L., Keenan, K., Gunnell, B. S., Mehndiratta, P., Chee, C. Y., Cahill, E. A., Dietiker, C., Cattell-Gordon, D. C., Smith, W. S., Perina, D. G., Solenski, N. J., Worrall, B. B., Southerland, A. M., & As the iTREAT Investigators (2016). A low-cost, tablet-based option for prehospital neurologic assessment: The iTREAT Study. *Neurology*, *87*(1), 19–26.

Chen, Q.T. (1995) Scoring standard of clinical neurological deficit in patients with stroke. Chin J Neurol, 29, 381–3.

Choi, K. H., Park, M. S., Kim, J. T., Chang, J., Nam, T. S., Choi, S. M., Lee, S. H., Kim, B. C., Kim, M. K., & Cho, K. H. (2012). Serum triglyceride level is an important predictor of early prognosis in patients with acute ischemic stroke. *Journal of the neurological sciences*, *319*(1-2), 111–116.

Choudhry, N. K., Fletcher, R. H., & Soumerai, S. B. (2005). Systematic review: the relationship between clinical experience and quality of health care. *Annals of internal medicine*, *142*(4), 260–273.

Christensen, H., Boysen, G., Johannesen, H. H., Christensen, E., & Bendtzen, K. (2002). Deteriorating ischaemic stroke. cytokines, soluble cytokine receptors, ferritin, systemic blood pressure, body temperature, blood glucose, diabetes, stroke severity, and CT infarction-volume as predictors of deteriorating ischaemic stroke. *J Neurol Sci, 201*(1-2), 1-7.

Church, S., Dunn, M., & Prokopy, L. (2019). "Benefits to qualitative data quality with multiple coders: Two case studies in multi-coder data Analysis." *Journal of Rural Social Sciences*, 34(1): Article 2.

Clarke, D. J., & Forster, A. (2015). Improving post-stroke recovery: the role of the multidisciplinary health care team. *Journal of multidisciplinary healthcare, 8*, 433-442.

Clarke, D. J., Godfrey, M., Hawkins, R., Sadler, E., Harding, G., Forster, A., McKevitt, C., Dickerson, J., & Farrin, A. (2013). Implementing a training intervention to support

caregivers after stroke: a process evaluation examining the initiation and embedding of programme change. *Implementation science*, *8*, 96.

Clarkson, J., Dean, J., Ward, J., Komashie, A., & Bashford, T. (2018). A systems approach to healthcare: from thinking to practice. *Future healthcare journal, 5*(3), 151-155.

Coghlan, D., & Brydon-Miller, M. (Eds.) (2014). (Vols. 1-2). SAGE Publications Ltd, Retrieved March 25, 2023, from https://doi.org/10.4135/9781446294406

Concannon, T. W., Meissner, P., Grunbaum, J. A., McElwee, N., Guise, J. M., Santa, J., Conway, P. H., Daudelin, D., Morrato, E. H., & Leslie, L. K. (2012). A new taxonomy for stakeholder engagement in patient-centered outcomes research. *Journal of general internal medicine*, *27*(8), 985–991.

Conn, L. G., Lingard, L., Reeves, S., Miller, K. L., Russell, A., & Zwarenstein, M. (2009). Communication channels in general internal medicine: a description of baseline patterns for improved interprofessional collaboration. *Qual Health Res, 19*(7), 943-953.

Cook, S., & Birrell, M. (2007). Defining an occupational therapy intervention for people with psychosis. *British Journal of Occupational Therapy, 70*(3), 96-106.

CordisBright. (2018). *What are the key factors for successful multidisciplinary team working?* Retrieved June 20, 2022, from https://www.cordisbright.co.uk/admin/resources/05-hsc-evidence-reviews-multidisciplinary-team-working.pdf

COSMIN. (COnsensus-based Standards for the selection of health status Measurement INstruments) (n.d.). *Search filters*. Retrieved August 15, 2020, from https://www.cosmin.nl/tools/pubmed-search-filters/

Côté, R., Battista, R. N., Wolfson, C., Boucher, J., Adam, J., & Hachinski, V. (1989). The Canadian Neurological Scale: Validation and reliability assessment. *Neurology, 39*(5), 638-643.

Côté, R., Battista, R. N., Wolfson, C. M., & Hachinski, V. (1988). Stroke assessment scales: guidelines for development, validation, and reliability assessment. *Can J Neurol Sci, 15*(3), 261-265.

Côté, R., Hachinski, V. C., Shurvell, B. L., Norris, J. W., & Wolfson, C. (1986). The Canadian Neurological Scale: A preliminary study in acute stroke. *Stroke, 17*(4), 731-737.

Craig, P., Di Ruggiero, E., Frohlich, K. L., Mykhalovskiy, E., White, M., Campbell, R., Cummins, S., on behalf of the Canadian Institutes of Health Research (CIHR)–National Institute for Health Research (NIHR) Context Guidance Authors Group (2018). *Taking account of context in population health intervention research: guidance for producers, users and funders of research.* Southampton: NIHR Evaluation, Trials and Studies Coordinating Centre.

Cramm, J. M., & Nieboer, A. P. (2011). Professionals' views on interprofessional stroke team functioning. *International journal of integrated care, 11*, e081-e081.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Sage.

Creswell, J.W. (2006). Chapter 4. *Choosing a mixed methods design*. Retrieved June 10, 2022, from https://www.sagepub.com/sites/default/files/upm-binaries/10982_Chapter_4.pdf

Creswell, J.W. (2007). Qualitative Inquiry & Research Design. London: Sage Publications.

Creswell, J. W. (2009). Editorial: Mapping the Field of Mixed Methods Research. *Journal of Mixed Methods Research, 3*(2), 95-108.

Criteria for the degree of clinical neurological Impairment in stroke patients. (1995). Chinese Journal of Neurology, 1996(6), 62-4.

Cruickshank, J. (2012). Positioning positivism, critical realism and social constructionism in the health sciences: a philosophical orientation. *Nursing Inquiry, 19*(1), 71-82.

Cryer, P. E. (2007). Hypoglycemia, functional brain failure, and brain death. *The Journal of clinical investigation, 117*(4), 868-870.

Cuadrado-Godia, E., Jimena, S., Ois, A., Rodríguez-Campello, A., Giralt-Steinhauer, E., Soriano-Tarraga, C., Jiménez-Conde, J., Martínez-Rodríguez, J. E., Capellades, J., & Roquer, J. (2013). Factors associated with early outcome in patients with large-vessel carotid strokes. *Journal of neurology, neurosurgery, and psychiatry*, *84*(3), 305–309.

Cui, Y., Meng, W. H., & Chen, H. S. (2022). Early neurological deterioration after intravenous thrombolysis of anterior vs posterior circulation stroke: a secondary analysis of INTRECIS. *Scientific reports, 12*(1), 3163-3163.

Dalkey, N., & Helmer, O. (1963). An experimental application of the DELPHI Method to the use of experts. *Management Science, 9*(3), 458-467.

Dalkin, S. M., Greenhalgh, J., Jones, D., Cunningham, B., & Lhussier, M. (2015). What's in a mechanism? Development of a key concept in realist evaluation. *Implementation science, 10*, 49-49.

Dalkin, S. M., Hardwick, R. J. L., Haighton, C. A., & Finch, T. L. (2021). Combining realist approaches and Normalization Process Theory to understand implementation: a systematic review. *Implementation Science Communications, 2*(1), 68.

Dall'Ora, C., Griffiths, P., Hope, J., Briggs, J., Jeremy, J., Gerry, S., & Redfern, O. C. (2021). How long do nursing staff take to measure and record patients' vital signs observations in hospital? A time-and-motion study. *International journal of nursing studies*, *118*, 103921.

Danermark, B. (2019) Applied interdisciplinary research: a critical realist perspective. Journal of critical realism, 18(4), 368-382.

Darehed, D., Norrving, B., Stegmayr, B., Zingmark, K., & Blom, M. C. (2017). Patients with acute stroke are less likely to be admitted directly to a stroke unit when hospital beds are scarce: A Swedish multicenter register study. *European Stroke Journal, 2*(2), 178-186.

Dávalos, A., Cendra, E., Teruel, J., Martinez, M., & Genís, D. (1990). Deteriorating ischemic stroke. *Risk factors and prognosis, 40*(12), 1865-1865.

Dávalos, A., Toni, D., Iweins, F., Lesaffre, E., Bastianello, S., & Castillo, J. (1999). Neurological deterioration in acute ischemic stroke: potential predictors and associated factors in the European cooperative acute stroke study (ECASS) I. *Stroke, 30*(12), 2631-2636.

Davie, C., Hunter, R. M., Mountford, J., & Morris, S. (2013). *London's hyperacute stroke units improve outcomes and lower costs*. Retrieved June 1, 2022, from https://hbr.org/2013/11/londons-hyperacute-stroke-units-improve-outcomes-and-lower-costs

Davis, D. C., Rhodes, R., & Baker, A. S. (1998). Curriculum revision: reaching faculty consensus through the Nominal Group Technique. *The Journal of nursing education*, *37*(7), 326–328.

Deaton, A., Cartwright, N. (2018) Understanding and misunderstanding randomized controlled trials. *Soc Sci Med*, 210, 2–21.

DeCoster, J. (1998). *Overview of Factor Analysis*. Retrieved June 23, 2022, from http://www.stat-help.com/notes.html

De Leon Benedetti, A. M., Bhatia, R., Ancheta, S. R., Romano, J. G., & Koch, S. (2021). How well do neurochecks perform after stroke?. *Stroke*, *52*(3), 1094–1097.

De Leon-Benedetti, A., Campo-Bustillo, I., Romano, J., & Koch, S. (2019). Neuro-checks and the national institute of health stroke scale in acute stroke. *Neurology. Conference: 71st Annual Meeting of the American Academy of Neurology, AAN, 92* (15 Suppl.1).

de Sousa, I., & Woodward, S. (2016). The Glasgow Coma Scale in adults: doing it right. *Emerg Nurse, 24*(8), 33-39.

de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in Medicine: A Practical Guide*. Cambridge University Press.

DeGraba, T. J., Hallenbeck, J. M., Pettigrew, K. D., Dutka, A. J., & Kelly, B. J. (1999). Progression in acute stroke: value of the initial NIH stroke scale score on patient stratification in future trials. *Stroke, 30*(6), 1208-1212.

Delbecq, A.L., Van de Ven, A.H. and Gustafson, D.H. (1975) Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes. Scott, Foresman Glenview.

Demaerschalk, B. M., Vegunta, S., Vargas, B. B., Wu, Q., Channer, D. D., & Hentz, J. G. (2012). Reliability of real-time video smartphone for assessing National Institutes of Health Stroke Scale scores in acute stroke patients. *Stroke, 43*(12), 3271-3277.

Demeestere, J., Garcia-Esperon, C., Lin, L., Bivard, A., Ang, T., Smoll, N. R., Garnett, A., Loudfoot, A., Miteff, F., Spratt, N., Parsons, M., & Levi, C. (2017). Validation of the National Institutes of Health Stroke Scale-8 to Detect Large Vessel Occlusion in Ischemic Stroke. *Journal of stroke and cerebrovascular diseases: the official journal of National Stroke Association*, *26*(7), 1419–1426.

Department of Health (2007) *National stroke strategy*. Department of Health, London.

Deverka, P. A., Lavallee, D. C., Desai, P. J., Esmail, L. C., Ramsey, S. D., Veenstra, D. L., & Tunis, S. R. (2012). Stakeholder participation in comparative effectiveness research: defining

a framework for effective engagement. *Journal of comparative effectiveness research, 1*(2), 181-194.

Dewey, H. M., Donnan, G. A., Freeman, E. J., Sharples, C. M., Macdonell, R. A., McNeil, J. J., & Thrift, A. G. (1999). Interrater reliability of the National Institutes of Health Stroke Scale: rating by neurologists and nurses in a community-based stroke incidence study. *Cerebrovascular Diseases, 9*(6), 323-327.

Dey, I. (1999) *Grounding grounded theory: Guidelines for qualitative inquiry*. Academic Press: San Diego.

Dingwall, R. (1992). Don't mind him—He's from Barcelona: Qualitative methods in health studies. In J. Daly, I. McDonald, & E. Willis (Eds.), *Researching health care* (pp. 161–175). Tavistock/Routledge

Doods, J., Botteri, F., Dugas, M., & Fritz, F. (2014). A European inventory of common electronic health record data elements for clinical trial feasibility. *Trials, 15*(1), 18.

Downey, C. L., Tahir, W., Randell, R., Brown, J. M., & Jayne, D. G. (2017). Strengths and limitations of early warning scores: A systematic review and narrative synthesis. *International Journal of Nursing Studies, 76*, 106-119.

Drotz, E., & Poksinska, B. (2014). Lean in healthcare from employees' perspectives. *J Health Organ Manag, 28*(2), 177-195.

DuBose, B. M., & Mayo, A. M. (2020). Resistance to change: A concept analysis. *Nurs Forum, 55*(4), 631-636.

Durose, C., Needham, C., Mangan, C., & Rees, J. (2017). Generating 'good enough' evidence for co-production. *Evidence & Policy, 13*(1), 135-151.

Eccles, M. P., & Mittman, B. S. (2006). Welcome to Implementation Science. *Implementation Science, 1*(1), 1.

Edwards, D. F., Chen, Y., Diringer, M. N., Edwards, D. F., Chen, Y. W., & Diringer, M. N. (1995). Unified Neurological Stroke Scale is valid in ischemic and hemorrhagic stroke. *Stroke, 26*(10), 1852-1858.

Etikan, I., Musa, S.A., Alkassim, R.S. (2016) Comparison of convenience sampling and purposive sampling. American Journal of Theoretical and Applied Statistics. 5(1)1-4.

Engelter, S. T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., Gutzwiller, F., & Lyrer, P. A. (2006). Epidemiology of aphasia attributable to first ischemic stroke: incidence, severity, fluency, etiology, and thrombolysis. *Stroke*, *37*(6), 1379–1384.

European Centre for Disease Prevention and Control (ECDC). (2011) *Evidence-based methodologies for public health – how to assess the best evidence when time is limited and there is a lack of sound evidence.* Retrieved June 3, 2022, from https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/1109_TER_evidence_based_methods_for_public_health.pdf

European Commission. Ethics for researchers. Facilitating research excellence in FP7. (2013) 37Retrieved Feburary 12, 2023, from

https://ec.europa.eu/research/participants/data/ref/fp7/89888/ethics-for-researchers_en.pdf

Evans, R. S. (2016). Electronic health records: Then, now, and in the future. *Yearbook of medical informatics, Suppl 1*(Suppl 1), S48-S61.

Fagin, C. M. (1992). Collaboration between nurses and physicians: no longer a choice. *Acad Med, 67*(5), 295-303.

Faigle, R., Butler, J., Carhuapoma, J. R., Johnson, B., Zink, E. K., Shakes, T., Rosenblum, M., Saheed, M., & Urrutia, V. C. (2020). Safety Trial of Low-Intensity Monitoring After Thrombolysis: Optimal Post Tpa-Iv Monitoring in Ischemic STroke (OPTIMIST). *The Neurohospitalist*, *10*(1), 11–15.

Feinstein, A. R. (1987). Clinimetric perspectives. *Journal of Chronic Diseases, 40*(6), 635-640.

Feinstein, A. R., Josephy, B. R., & Wells, C. K. (1986). Scientific and clinical problems in indexes of functional disability. *Annals of internal medicine*, *105*(3), 413–420.

Fens, M., van Heugten, C. M., Beusmans, G. H., Limburg, M., Haeren, R., Kaemingk, A., & Metsemakers, J. F. (2013). Not as transient: patients with transient ischaemic attack or minor stroke experience cognitive and communication problems; an exploratory study. *Eur J Gen Pract, 19*(1), 11-16.

Ferdinand, P., & Roffe, C. (2016). Hypoxia after stroke: a review of experimental and clinical evidence. *Experimental & translational stroke medicine, 8*, 9-9.

Fink, A., Kosecoff, J., Chassin, M., & Brook, R. H. (1984). Consensus methods: characteristics and guidelines for use. *Am J Public Health, 74*(9), 979-983.

Finocchi, C., Balestrino, M., Malfatto, L., Mancardi, G., Serrati, C., & Gandolfo, C. (2018). National Institutes of Health Stroke Scale in patients with primary intracerebral hemorrhage. *Neurological Sciences, 39*(10), 1751-1755.

Fischer, K., Lewandowski, D., & Janssen, M. P. (2013). Estimating unknown parameters in haemophilia using expert judgement elicitation. *Haemophilia : the official journal of the World Federation of Hemophilia*, *19*(5), e282–e288.

Fisher, R. J., Gaynor, C., Kerr, M., Langhorne, P., Anderson, C., Bautz-Holter, E., Indredavik, B., Mayo, N. E., Power, M., Rodgers, H., Rønning, O. M., Widén Holmqvist, L., Wolfe, C. D., & Walker, M. F. (2011). A consensus on stroke: early supported discharge. *Stroke*, *42*(5), 1392–1397.

Fitch, K., Bernstein, S. J., Aguilar, M. D., Burnand, B., Lacalle, J. R., Lazaro, P., van het Loo, M., McDonnell, J., Vader, J., & Kahan, J. P. (2001). *The RAND/UCLA Appropriateness Method User's Manual*. Retrieved June 12, 2022, from https://www.rand.org/content/dam/rand/pubs/monograph_reports/2011/MR1269.pdf

Flemming, K. D., Wijdicks, E. F., St Louis, E. K., & Li, H. (1999). Predicting deterioration in patients with lobar haemorrhages. *Journal of neurology, neurosurgery, and psychiatry, 66*(5), 600-605.

Flores, G., Laws, M. B., Mayo, S. J., Zuckerman, B., Abreu, M., Medina, L., & Hardt, E. J. (2003). Errors in medical interpretation and their potential clinical consequences in pediatric encounters. *Pediatrics*, *111*(1), 6–14.

Flores G. (2005). The impact of medical interpreter services on the quality of health care: a systematic review. *Medical care research and review : MCRR*, *62*(3), 255–299.

Foronda, C., MacWilliams, B., & McArthur, E. (2016). Interprofessional communication in healthcare: An integrative review. *Nurse Educ Pract, 19*, 36-40.

Fusch, P., & Ness, L. (2015) Are we there yet? Data saturation in qualitative research. *Qualitative Report*, 20, 1408-1416.

Gallagher, M., Hares, T., Spencer, J., Bradshaw, C., & Webb, I. (1993). The nominal group technique: a research tool for general practice?. *Family practice*, *10*(1), 76–81.

Garnett, A. R., Marsden, D. L., Parsons, M. W., Quain, D. A., Spratt, N. J., Loudfoot, A. R., Middleton, P. M., Levi, C. R., & Rural PAST Protocol Steering Group. (2010). The rural Prehospital Acute Stroke Triage (PAST) trial protocol: a controlled trial for rapid facilitated transport of rural acute stroke patients to a regional stroke centre. *International Journal of Stroke, 5*(6), 506-513.

Geisler, F., Kunz, A., Winter, B., Rozanski, M., Waldschmidt, C., Weber, J. E., Wendt, M., Zieschang, K., Ebinger, M., Audebert, H. J., & the Stroke Emergency Mobile (STEMO) Consortium. (2019). Telemedicine in Prehospital Acute Stroke Care. *Journal of the American Heart Association, 8*(6).

Gelderblom, M., Leypoldt, F., Steinbach, K., Behrens, D., Choe, C. U., Siler, D. A., Arumugam, T. V., Orthey, E., Gerloff, C., Tolosa, E., & Magnus, T. (2009). Temporal and spatial dynamics of cerebral immune cell accumulation in stroke. *Stroke*, *40*(5), 1849–1857.

Gelmers, H. J., Gorter, K., Weerdt, C. J. d., & Wiezer, H. J. (1988). Assessment of interobserver variability in a Dutch multicenter study on acute ischemic stroke. *Stroke, 19*(6), 709-711.

General Medical Council (GMC). (2013). *The duties of a doctor registered with the General Medical Council*. Retrieved June 15, 2022, from www.gmc-uk.org/guidance/good_medical_practice/duties_of_a_doctor.asp.

Geng, H. H., Wang, Q., Li, B., Cui, B. B., Jin, Y. P., Fu, R. L., Zhang, Q., Wang, J. J., & Wang, P. X. (2017). Early neurological deterioration during the acute phase as a predictor of long-term outcome after first-ever ischemic stroke. *Medicine*, *96*(51), e9068.

Georgiadis, D., Schwarz, S., Aschoff, A., & Schwab, S. (2002). Hemicraniectomy and moderate hypothermia in patients with severe ischemic stroke. *Stroke, 33*(6), 1584-1588.

Gerry, S., Bonnici, T., Birks, J., Kirtley, S., Virdee, P. S., Watkinson, P. J., & Collins, G. S. (2020). Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. *BMJ (Clinical research ed.)*, *369*, m1501.

Gill, M. & Turbin, V. (1999) Evaluating 'realistic evaluation': evidence from a study of CCTV. Crime *Prevention Studies 10* (1), 179-199.

Gillespie, B. M., Harbeck, E., Lavin, J., Gardiner, T., Withers, T. K., & Marshall, A. P. (2018). Using normalisation process theory to evaluate the implementation of a complex intervention to embed the surgical safety checklist. *BMC Health Serv Res, 18*(1), 170.

Ginsberg, M. D., Sternau, L. L., Globus, M. Y., Dietrich, W. D., & Busto, R. (1992). Therapeutic modulation of brain temperature: relevance to ischemic brain injury. *Cerebrovasc Brain Metab Rev, 4*(3), 189-225.

Gocan, S., & Fisher, A. (2008). Neurological assessment by nurses using the National Institutes of Health Stroke Scale: implementation of best practice guidelines. *Canadian Journal of Neuroscience Nursing, 30*(3), 31-42.

Goldstein, L. B. (1995). Common drugs may influence motor recovery after stroke. The Sygen In Acute Stroke Study Investigators. *Neurology, 45*(5), 865-871.

Goldstein, L. B., Bertels, C., & Davis, J. (1989). Interrater reliability of the NIH stroke scale. *Archives of Neurology, 46*(6), 660-662.

Goldstein, L. B., Samsa, G. P., Goldstein, L. B., & Samsa, G. P. (1997). Reliability of the National Institutes of Health Stroke Scale. Extension to non-neurologists in the context of a clinical trial. *Stroke, 28*(2), 307-310.

Gomez, C. R. (1993). Editorial: Time is brain! *J Stroke Cerebrovasc Dis, 3*(1), 1-2

Gonzalez, M. A., Hanna, N., Rodrigo, M. E., Satler, L. F., & Waksman, R. (2011). Reliability of prehospital real-time cellular video phone in assessing the simplified National Institutes of Health Stroke Scale in patients with acute stroke: a novel telemedicine technology. *Stroke*, *42*(6), 1522–1527.

Gotoh, F., Terayama, Y., Amano, T., Gotoh, F., Terayama, Y., & Amano, T. (2001). Development of a novel, weighted, quantifiable stroke scale: Japan stroke scale. *Stroke, 32*(8), 1800-1807.

Gottesman, R. F., & Hillis, A. E. (2010). Predictors and assessment of cognitive dysfunction resulting from ischaemic stroke. *The Lancet. Neurology, 9*(9), 895-905.

Govan, L., Langhorne, P., & Weir, C. J. (2007). Does the prevention of complications explain the survival benefit of organized inpatient (stroke unit) care?: further analysis of a systematic review. *Stroke, 38*(9), 2536-2540.

Govindarajan, P., Resler, B., Keenan, K., Beylin, M., Fu, J., Dietiker, C., Grimes, B., & Smith, W. (2015). Teleconsultation by emergency medical services for acute stroke recognition: A feasibility study. *Academic Emergency Medicine, 22*(5), S60.

Greatbatch, D. on behalf of the Council of Deans of Health. (2016). *A false economy. Cuts to continuing professional development funding for nursing, midwifery and the allied health professions in England*. Retrieved June 1, 2022, from https://www.councilofdeans.org.uk/wp-content/uploads/2016/09/19092016-A-False-Economy-CPD-cuts-in-England-2016-17-.pdf

Green, B., Oeppen, R. S., Smith, D. W., & Brennan, P. A. (2017). Challenging hierarchy in healthcare teams - ways to flatten gradients to improve teamwork and patient care. *Br J Oral Maxillofac Surg, 55*(5), 449-453.

Greenbaum, T. (2003). The gold standard? why the focus group deserves to be the most respected of all qualitative research tools. *Quirk's Marketing Research Review, 17*, 22–27.

Greener, I., & Mannion, R. (2009). A realistic evaluation of practice-based commissioning. *Policy & Politics, 37*(1), 57-73.

Greenhalgh, T., & Papoutsi, C. (2018). Studying complexity in health services research: desperately seeking an overdue paradigm shift. *BMC Medicine, 16*(1), 95.

Greer, D. M., Funk, S. E., Reaven, N. L., Ouzounelli, M., & Uman, G. C. (2008). Impact of fever on outcome in patients with stroke and neurologic injury: a comprehensive meta-analysis. *Stroke, 39*(11), 3029-3035.

Guest, G., Namey, E., Taylor, J., Eley, N., & McKenna, K. (2017) Comparing focus groups and individual interviews: findings from a randomized study, *International Journal of Social Research Methodology*, 20(6), 693-708.

Gur, A. Y., Lampl, Y., Gross, B., Royter, V., Shopin, L., & Bornstein, N. M. (2007). A new scale for assessing patients with vertebrobasilar stroke-the Israeli Vertebrobasilar Stroke Scale (IVBSS): inter-rater reliability and concurrent validity. *Clinical Neurology & Neurosurgery, 109*(4), 317-322.

Guterud, M., Bugge, H. F., Eriksen, E., Ingebrigtsen, S. H., Bache, K. G., Roislien, J., Thorsen, K., Kramer-Johansen, J., Sandset, E. C., & Hov, M. R. (2019). Training paramedics in full-scale NIHSS-a pilot study. *European Stroke Journal, 4 (Suppl. 1)*, 383.

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ, 336*(7650), 924-926.

Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol, 61*(Pt 1), 29-48.

Hahn, E. A., Cella, D., Chassany, O., Fairclough, D. L., Wong, G. Y., Hays, R. D., & Clinical Significance Consensus Meeting Group (2007). Precision of health-related quality-of-life data compared with other clinical measures. *Mayo Clinic proceedings*, *82*(10), 1244–1254.

Hajat, C., Hajat, S., & Sharma, P. (2000). Effects of poststroke pyrexia on stroke outcome : a meta-analysis of studies in patients. *Stroke, 31*(2), 410-414.

Handschu, R., Littmann, R., Reulbach, U., Gaul, C., Heckmann, J. G., Neundorfer, B., & Scibor, M. (2003). Telemedicine in emergency evaluation of acute stroke: interrater agreement in remote video examination with a novel multimedia system. *Stroke, 34*(12), 2842-2846.

Hansson, A.-S., Vingård, E., Arnetz, B. B., & Anderzén, I. (2008). Organizational change, health, and sick leave among health care employees: A longitudinal study measuring stress markers, individual, and work site factors. *Work & Stress, 22*(1), 69-80.

Hantson, L., De Weerdt, W., De Keyser, J., Diener, H. C., Franke, C., Palm, R., Van Orshoven, M., Schoonderwalt, H., De Klippel, N., & Herroelen, L. (1994). The European Stroke Scale. *Stroke*, *25*(11), 2215–2219.

Hargroves, D., & Lowe, D. (2022). NHS England. *Getting it right first time. Stroke. Grift programme national speciality report.* Retrieved June 16, 2022, from https://gmnisdn.org.uk/wp-content/uploads/2022/04/Stroke-GiRFT-report-Apr-22.pdf

Harrison, J. K., McArthur, K. S., & Quinn, T. J. (2013). Assessment scales in stroke: clinimetric and clinical considerations. *Clinical Interventions in Aging, 8*, 201-211.

Harrison, R., Manias, E., Mears, S., Heslop, D., Hinchcliff, R., & Hay, L. (2019). Addressing unwarranted clinical variation: A rapid review of current evidence. *Journal of Evaluation in Clinical Practice, 25*(1), 53-65.

Hart, S., Lowe, D., Hargroves, D., & Doubal, F. on behalf of the British Association of Stroke Physicians BASP/NHS Improvement GIRFT. (2019). *Meeting the future consultant workforce challenges: Stroke medicine consultant workforce requirements 2019 – 2022*. Retrieved June 1, 2022, from https://biasp.org/wp-content/uploads/2022/03/BASP-Stroke-Medicine-Workforce-Requirements-Report-FINAL.pdf.pagespeed.ce.K6If1Ae4ai.pdf

Harvey, N., & Holmes, C. A. (2012). Nominal group technique: An effective method for obtaining group consensus. *International Journal of Nursing Practice, 18*(2), 188-194.

Hawe, P., Shiell, A., & Riley, T. (2004). Complex interventions: how "out of control" can a randomised controlled trial be? *BMJ (Clinical research ed.), 328*(7455), 1561-1563.

Health Education England. (2021) Working differently together: *Progressing a one workforce approach. Multidisciplinary team toolkit.* Retrieved June 20, 2022, from https://www.hee.nhs.uk/sites/default/files/documents/HEE_MDT_Toolkit_V1.1.pdf

Health Education England. (2022) *Stroke-Specific Education Framework (SSEF).* Retrieved June 2, 2022, from https://stroke-education.org.uk/

Health Research Authority (HRA). (2022a) *GDPR guidance for researchers and study coordinators.* Retrieved June 15, 2022, from https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/data-protection-and-information-governance/gdpr-guidance/

Health Research Authority (HRA). (2022b) *Is my study research? Decision tool.* Retrieved March 25, 2023, from https://www.hra-decisiontools.org.uk/research/

Heikinheimo, T., Broman, J., Haapaniemi, E., Kaste, M., Tatlisumak, T., & Putaala, J. (2013). Preceding and poststroke infections in young adults with first-ever ischemic stroke: effect on short-term and long-term outcomes. *Stroke, 44*(12), 3331-3337.

Helleberg, B. H., Ellekjaer, H., & Indredavik, B. (2016). Outcomes after early neurological deterioration and transitory deterioration in acute ischemic stroke patients. *Cerebrovascular Diseases, 42*(5-6), 378-386.

Helleberg, B. H., Ellekjær, H., Rohweder, G., & Indredavik, B. (2014). Mechanisms, predictors and clinical impact of early neurological deterioration: the protocol of the Trondheim early neurological deterioration study. *BMC Neurology, 14*(1), 201.

Herman, B. C. (2015). The influence of global warming science views and sociocultural factors on willingness to mitigate global warming. *Science Education, 99*(1), 1-38.

Hills, N. K., Josephson, S. A., Lyden, P. D., & Johnston, S. C. (2009). Is the NIHSS certification process too lenient? *Cerebrovascular Diseases, 27*(5), 426-432.

Hodgson, T., Burton-Jones, A., Donovan, R., & Sullivan, C. (2021). The role of electronic medical records in reducing unwarranted clinical variation in acute health care: Systematic review. *JMIR medical informatics, 9*(11), e30432-e30432.

Hogan, R. (2007). *Personality and the fate of organizations*. Lawrence Erlbaum Associates.

Holland, J., & Christian, L. M. (2009). The influence of topic interest and interactive probing on responses to open-ended questions in web surveys. *Social Science Computer Review, 27*, 196-212.

Honebein, P. C. (1996). *Seven goals for the design of constructivist learning environments*. Retrieved June 24, 2022 from http://studentcenteredlearning.pbworks.com/f/DesignConstructivistHonebein.pdf

Houston, S. (2001). Beyond social constructionism: critical realism and social work. *The British Journal of Social Work, 31*(6), 845-861.

Houtkoop-Steenstra, H. (2000). *Interaction and the Standardized Survey Interview: The Living Questionnaire*. Cambridge University Press.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, *15*(9), 1277–1288.

Hunter, R. M., Davie, C., Rudd, A., Thompson, A., Walker, H., Thomson, N., Mountford, J., Schwamm, L., Deanfield, J., Thompson, K., Dewan, B., Mistry, M., Quoraishi, S., & Morris, S. (2013). Impact on clinical and cost outcomes of a centralized approach to acute stroke care in London: a comparative effectiveness before and after model. *PloS one*, *8*(8), e70420.

Hunter, R. M., Fulop, N. J., Boaden, R. J., McKevitt, C., Perry, C., Ramsay, A. I. G., Rudd, A. G., Turner, S. J., Tyrell, P. J., Wolfe, C. D. A., & Morris, S. (2018). The potential role of cost-utility analysis in the decision to implement major system change in acute stroke services in metropolitan areas in England. *Health Research Policy and Systems, 16*(1), 23.

Hutchings A, R. R., Sanderson C, Black N. (2006). A comparison of formal consensus methods used for developing clinical guidelines. *Journal of Health Services Research & Policy, 11*(4), 218-224.

Iacono, L. A., Wells, C., & Mann-Finnerty, K. (2014). Standardizing neurological assessments. *Journal of Neuroscience Nursing, 46*(2), 125-132.

ICH (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use). (2016) *E6 Guideline for Good Clinical Practice (GCP*). Retrieved June 15, 2022, from https://database.ich.org/sites/default/files/E6_R2_Addendum.pdf

Imison, C. (2016) *Trust in transformation change. Nuffield Trust*. Retrieved June, 20 2022, from https://www.nuffieldtrust.org.uk/news-item/trust-in-transformational-change#better-team-working

Imison, C., Sonola, L., Honeyman, M., & Ross, S. (2014) The King's Fund. *The reconfiguration of clinical services. What is the evidence?* Retrieved June 14, 2022, from https://www.kingsfund.org.uk/sites/default/files/field/field_publication_summary/Reconfiguration-of-clinical-services-kings-fund-nov-2014.pdf

Inatomi, Y., Yonehara, T., Omiya, S., Hashimoto, Y., Hirano, T., & Uchino, M. (2008). Aphasia during the acute phase in ischemic stroke. *Cerebrovascular diseases*, 25(4), 316–323.

Indredavik, B., Rohweder, G., Naalsund, E., & Lydersen, S. (2008). Medical complications in a comprehensive stroke unit and an early supported discharge service. *Stroke, 39*(2), 414-420.

Information Services Division (ISD) Scotland. (2017) *Scottish stroke care audit, Scottish stroke improvement programme report-2016*. Retrieved October 24, 2021, from http://bit.ly/2lckiKz

Institute of Medicine. (1997) *The computer-based patient record: an essential technology for health care.* Revised Edition. National Academy Press.

Interprofessional Education Collaborative. (2011). *Core competencies for interprofessional collaborative practice: Report of an expert panel.* Retrieved June 24, 2022, from https://www.aacom.org/docs/default-source/insideome/ccrpt05-10-11.pdf

Iramaneerat, C., Smith, E. V., Jr., Millis, S. R., & Lyden, P. D. (2009). Selecting a measurement model for the analysis of the National Institutes of Health Stroke Scale. *International Journal of Neuroscience, 119*(7), 1042-1059.

Isahaya, K. (2017). Effects of a training video on the assessment of NIH stroke scale score via telestroke. *Journal of the Neurological Sciences, 381 (Supplement 1)*, 169.

Itoh, H., Shioi, M., Oshida, N., & Nagai, A. (1996). [Progressive stroke during the acute state of cerebral infarction in the territory of the internal carotid artery]. *No To Shinkei, 49*(2), 138-144.

Izumi-Richards, N., & Simon, C. (2013). Neurological assessment. *InnovAiT, 6*(7), 405-415.

Jauch, E. C., Saver, J. L., Adams, H. P., Jr, Bruno, A., Connors, J. J., Demaerschalk, B. M., Khatri, P., McMullan, P. W., Jr, Qureshi, A. I., Rosenfield, K., Scott, P. A., Summers, D. R., Wang, D. Z., Wintermark, M., Yonas, H., American Heart Association Stroke Council, Council on Cardiovascular Nursing, Council on Peripheral Vascular Disease, & Council on Clinical Cardiology (2013). Guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, *44*(3), 870–947.

Jobber, D., Saunders, J., & Mitchell, V.-W. (2004). Prepaid monetary incentive effects on mail survey response. *Journal of Business Research, 57*(4), 347-350.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher, 33*(7), 14-26.

Johnston, M. J., Arora, S., King, D., Bouras, G., Almoudaris, A. M., Davis, R., & Darzi, A. (2015). A systematic review to identify the factors that affect failure to rescue and escalation of care in surgery. *Surgery, 157*(4), 752-763.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374*(2065), 20150202.

Jones, S. P., Miller, C., Gibson, J. M. E., Cook, J., Price, C., & Watkins, C. L. (2018). The impact of education and training interventions for nurses and other health care staff involved in the delivery of stroke care: An integrative review. *Nurse Educ Today, 61*, 249-257.

Jones, T. L., Baxter, M. A. J., & Khanduja, V. (2013). A quick guide to survey research. *Annals of the Royal College of Surgeons of England, 95*(1), 5-7.

Jørgensen, H., Nakayama, H., Raaschou, H. O., & Olsen, T. S. (1994). Stroke in patients with diabetes. The Copenhagen Stroke Study. *Stroke, 25*(10), 1977-1984.

Josephson, S. A., Hills, N. K., & Johnston, S. C. (2006). NIH Stroke Scale reliability in ratings from a large sample of clinicians. *Cerebrovascular Diseases, 22*(5-6), 389-395.

Julnes, G., Mark, M. M., & Henry, G. T. (1998). Review: Promoting realism in evaluation: Realistic evaluation and the broader context. *Evaluation, 4*(4), 483-504.

Kadojić, D., Bijelić, B. R., Radanović, R., Porobić, M., Rimac, J., & Dikanović, M. (2012). Aphasia in patients with ischemic stroke. Acta clinica Croatica, 51(2), 221–225.

Kasner, S. E. (2006). Clinical interpretation and use of stroke scales. *Lancet Neurology, 5*(7), 603-612.

Keeney, R. L., & Keeney, R. L. (2009). *Value-focused thinking: A path to creative decisionmaking*: Harvard University Press.

Kennedy, P. J., Leathley, C. M., & Hughes, C. F. (2010). Clinical practice variation. *Med J Aust, 193*(S8), S97-99.

Khatri, P., Conaway, M. R., & Johnston, K. C. (2012). Ninety-day outcome rates of a prospective cohort of consecutive patients with mild ischemic stroke. *Stroke, 43*(2), 560-562.

Kidd, P. S., & Parshall, M. B. (2000). Getting the focus and the group: enhancing analytical rigor in focus group research. *Qualitative health research*, *10*(3), 293–308.

Kim, J. T., Park, M. S., Chang, J., Lee, J. S., Choi, K. H., & Cho, K. H. (2013). Proximal arterial occlusion in acute ischemic stroke with low NIHSS scores should not be considered as mild stroke. *PLoS ONE, 8*(8), e70996.

Kim, S. J., Moon, G. J., & Bang, O. Y. (2013). Biomarkers for stroke. *Journal of Stroke, 15*(1), 27-37.

King's College London. Sentinel Stroke National Audit Programme on behalf of the Intercollegiate Stroke Working Party (2022) *Sentinel Stroke National Audit Programme (SSNAP) Acute Organisational Audit 2021.* Retrieved June 16, 2022, from Ref.-346-SSNAP-AOA-Report-FINAL.pdf (hqip.org.uk)

Kline, P. (1994). *An Easy Guide to Factor Analysis* (1st ed.). Routledge.

Kok, G., Gottlieb, N. H., Peters, G. J., Mullen, P. D., Parcel, G. S., Ruiter, R. A., Fernández, M. E., Markham, C., & Bartholomew, L. K. (2016). A taxonomy of behaviour change methods: an Intervention Mapping approach. *Health psychology review*, *10*(3), 297–312.

Kondracki, N. L., Wellman, N. S., & Amundson, D. R. (2002). Content analysis: review of methods and their applications in nutrition education. Journal of nutrition education and behavior, 34(4), 224–230.

Krathwohl, D. R., & Smith, N. L. (2005). *How to prepare a dissertation proposal: Suggestions for students in education & the social and behavioral sciences*. Syracuse University Press.

Kuriakose, D., & Xiao, Z. (2020). Pathophysiology and Treatment of Stroke: Present Status and Future Perspectives. *International journal of molecular sciences, 21*(20), 7609.

Kwan, J., & Hand, P. (2006). Early neurological deterioration in acute stroke: clinical characteristics and impact on outcome. *Qjm, 99*(9), 625-633.

LaMonte, M. P., Xiao, Y., Hu, P. F., Gagliano, D. M., Bahouth, M. N., Gunawardane, R. D., MacKenzie, C. F., Gaasch, W. R., & Cullen, J. (2004). Shortening time to stroke treatment using ambulance telemedicine: TeleBAT. *Journal of stroke and cerebrovascular diseases : the official journal of National Stroke Association*, *13*(4), 148–154.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Langhorne, P., Audebert, H. J., Cadilhac, D. A., Kim, J., & Lindsay, P. (2020). Stroke systems of care in high-income countries: what is optimal? *The Lancet, 396*(10260), 1433-1442.

Langhorne, P., Lewsey, J. D., Jhund, P. S., Gillies, M., Chalmers, J. W., Redpath, A., Briggs, A., Walters, M., Capewell, S., McMurray, J. J., & MacIntyre, K. (2010). Estimating the impact of stroke unit care in a whole population: an epidemiological study using routine data. *Journal of neurology, neurosurgery, and psychiatry*, *81*(12), 1301–1305.

Langhorne, P., & Pollock, A. (2002). What are the components of effective stroke unit care? *Age Ageing, 31*(5), 365-371.

Lau, L.-H., Lew, J., Borschmann, K., Thijs, V., & Ekinci, E. I. (2019). Prevalence of diabetes and its effects on stroke outcomes: A meta-analysis and literature review. *Journal of Diabetes Investigation, 10*(3), 780-792.

Le Lagadec, M. D., & Dwyer, T. (2017). Scoping review: The use of early warning systems for the identification of in-hospital patients at risk of deterioration. *Aust Crit Care, 30*(4), 211-218.

Leal, J., Wordsworth, S., Legood, R., & Blair, E. (2007) Eliciting expert opinion for economic models: an applied example. *Value Health*, 10(3), 95-203.

Lee, C. F., Venketasubramanian, N., Wong, K. S., Chen, C. L., & Investigators, C. S. (2016). Comparison between the original and shortened versions of the National Institutes of

Health Stroke Scale in ischemic stroke patients of intermediate severity. *Stroke, 47*(1), 236-239.

Lee, T. K. P., Kitchell, A. K. B., Siu, A. Y. C., & Chen, N. K. (2017). Validation of the Full Outline of Unresponsiveness score coma scale in patients clinically suspected to have acute stroke in the emergency department. *Hong Kong Journal of Emergency Medicine, 24*(5), 230-236.

Leira, R., Dávalos, A., Silva, Y., Gil-Peralta, A., Tejada, J., Garcia, M., & Castillo, J. (2004). Early neurologic deterioration in intracerebral hemorrhage: predictors and associated factors. *Neurology, 63*(3), 461-467.

Levers, M. J. D. (2013). *Philosophical Paradigms, Grounded Theory, and Perspectives on Emergence.* SAGE Open.

Li, W., Xiao, W. M., Luo, G. P., Liu, Y. L., Qu, J. F., Fang, X. W., Wang, F., & Chen, Y. K. (2020). Asymmetrical cortical vein sign predicts early neurological deterioration in acute ischemic stroke patients with severe intracranial arterial stenosis or occlusion. *BMC neurology*, *20*(1), 331.

Liang, D., Bhatta, S., Gerzanich, V., & Simard, J. M. (2007). Cytotoxic edema: mechanisms of pathological cell swelling. *Neurosurgical focus, 22*(5), E2-E2.

Liman, T. G., Winter, B., Waldschmidt, C., Hufnagl, P., Zerbe, N., Audebert, H. J., & Endres, M. (2012). Telestroke ambulances in prehospital stroke management-are they ready yet for clinical implementation? *Cerebrovascular Diseases, 33*, 589-590.

Lin, K., Zink, W. E., Tsiouris, A. J., John, M., Tekchandani, L., & Sanelli, P. C. (2012). Risk assessment of hemorrhagic transformation of acute middle cerebral artery stroke using multimodal CT. *Journal of Neuroimaging, 22*(2), 160-166.

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. Beverly Hills, CA: Sage.

Lindenstrøm, E., Boysen, G., Waage Christiansen, L., à Rogvi Hansen, B., & Würtzen Nielsen, P. (1991). Reliability of Scandinavian Neurological Stroke Scale. *Cerebrovascular Diseases, 1*(2), 103-107.

Linfante, I., Llinas, R. H., Schlaug, G., Chaves, C., Warach, S., & Caplan, L. R. (2001). Diffusion-Weighted Imaging and National Institutes of Health Stroke Scale in the Acute Phase of Posterior-Circulation Stroke. *Archives of Neurology, 58*(4), 621-628.

Locock, L., & Boaz, A. (2019). Drawing straight lines along blurred boundaries: Qualitative research, patient and public involvement in medical research, coproduction and co-design. *Evidence & Policy*, *15*(3), 409-421.

Lomas, J., Anderson, G., Enkin, M., Vayda, E., Roberts, R., & MacKinnon, B. (1988). The role of evidence in the consensus process. Results from a Canadian consensus exercise. *JAMA, 259*(20), 3001-3005.

Lord, A. S., Gilmore, E., Choi, H. A., & Mayer, S. A. (2015). Time course and predictors of neurological deterioration after intracerebral hemorrhage. *Stroke, 46*(3), 647-652.

Low, J. (2013). Unstructured and semi-structured interviews in health research. *Researching health: Qualitative, quantitative and mixed methods, 2*, 87-105.

Low, J. (2019). A pragmatic definition of the concept of theoretical saturation. *Sociological Focus,* 52(2), 131-139.

Lowry, M. (1999). The Glasgow Coma Scale in clinical practice: a critique. *Nurs Times, 95*(22), 40-42.

Lyden, P. (2017). Using the National Institutes of Health Stroke Scale: A Cautionary Tale. *Stroke, 48*(2), 513-519.

Lyden, P., Brott, T., Tilley, B., Welch, K. M., Mascha, E. J., Levine, S., Haley, E. C., Grotta, J., & Marler, J. (1994a). Improved reliability of the NIH Stroke Scale using video training. NINDS TPA Stroke Study Group. *Stroke*, *25*(11), 2220–2226.

Lyden, P. D., & Lau, G. T. (1991). A critical appraisal of stroke evaluation and rating scales. *Stroke, 22*(11), 1345-1352.

Lyden, P., Lu, M., Jackson, C., Marler, J., Kothari, R., Brott, T., & Zivin, J. (1999). Underlying structure of the National Institutes of Health Stroke Scale: results of a factor analysis. NINDS tPA Stroke Trial Investigators. *Stroke, 30*(11), 2347-2354.

Lyden, P. D., Lu, M., Levine, S. R., Brott, T. G., Broderick, J., & Group, NINDS rtPA Stroke Study Group. (2001). A modified National Institutes of Health Stroke Scale for use in stroke clinical trials: preliminary reliability and validity. *Stroke, 32*(6), 1310-1317.

Lyden, P., Raman, R., Liu, L., Emr, M., Warren, M., & Marler, J. (2009). National Institutes of Health Stroke Scale certification is reliable across multiple venues. *Stroke*, *40*(7), 2507–2511.

Lyden, P., Raman, R., Liu, L., Grotta, J., Broderick, J., Olson, S., Shaw, S., Spilker, J., Meyer, B., Emr, M., Warren, M., & Marler, J. (2005). NIHSS training and certification using a new digital video disk is reliable. *Stroke*, *36*(11), 2446–2449.

Lyden, P. D., Zweifler, R., Mahdavi, Z., & Lonzo, L. (1994b). A rapid, reliable, and valid method for measuring infarct and brain compartment volumes from computed tomographic scans. *Stroke, 25*(12), 2421-2428.

Manning, K. (1997). Authenticity in constructivist inquiry: Methodological considerations without prescription. Qualitative inquiry, 3(1), 93-115.

Mahon-Haft, T. A., & Dillman, D. A. (2010). Does visual appeal matter? Effects of web survey aesthetics on survey quality. Survey *Research Methods, 4*(1), 43-59.

Maramattom, B. V., Bahn, M. M., & Wijdicks, E. F. M. (2004). Which patient fares worse after early deterioration due to swelling from hemispheric stroke? *Neurology, 63*(11), 2142-2145.

Marjanovic, S., Harshfield, A., Carpenter, A., Bertscher, A., Punch, D., & Ball, S. (2019) *Involving NHS staff in research. Healthcare Improvement Studies.* Retrieved June 3, 2022, from https://www.thisinstitute. cam.ac.uk/research-articles/involving-nhs-staff-in-research/

Markley, J., & Winbery, S. (2008). Communicating with physicians: How agencies can be heard. *Home Health Care Management & Practice, 20*(2), 161-168.

Martin, A. J., & Price, C. I. (2018). A systematic review and meta-analysis of molecular biomarkers associated with early neurological deterioration following acute stroke. *Cerebrovascular Diseases, 46*(5-6), 230-241.

Mathew, N., Rivera, V., Meyer, J., Charney, J., & Hartmann, A. (1972). Double-blind evaluation of glycerol therapy in acute cerebral infarction. *The Lancet, 300*(7791), 1327-1329.

Massey, D., Aitken, L. M., & Chaboyer, W. (2015). The impact of a nurse led rapid response system on adverse, major adverse events and activation of the medical emergency team. *Intensive Crit Care Nurs, 31*(2), 83-90.

May, C. R., Cummings, A., Girling, M., Bracher, M., Mair, F. S., May, C. M., Murray, E., Myall, M., Rapley, T., & Finch, T. (2018). Using Normalization Process Theory in feasibility studies and process evaluations of complex healthcare interventions: a systematic review. *Implementation Science, 13*(1), 80.

May, C., & Finch, T. (2009). Implementing, embedding, and integrating practices: An outline of Normalization Process Theory. *Sociology, 43*(3), 535-554.

May, C., Rapley, T., & Finch, T. (2020) Normalization Process Theory. In: P. Nilsen, & S. Birken S (Eds.), *International Handbook of Implementation Science*. (pp. 144-167). Edward Elgarp.

May, C., Rapley, T., Mair, F. S., Treweek, S., Murray, E., et al: (2021) Normalization Process Theory on-line user's manual, toolkit, and NoMAD instrument. Retrieved October 27, 2021 from http://www.Normalizationprocess.org.

Mayer, S. A., Sacco, R. L., Shi, T., & Mohr, J. P. (1994). Neurologic deterioration in noncomatose patients with supratentorial intracerebral hemorrhage. *Neurology, 44*(8), 1379-1384.

McEvoy, P., & Richards, D. (2006). A critical realist rationale for using a combination of quantitative and qualitative methods. *Journal of Research in Nursing, 11*(1), 66-78.

McGaughey, J., O'Halloran, P., Porter, S., Trinder, J., & Blackwood, B. (2017). Early warning systems and rapid response to the deteriorating patient in hospital: A realist evaluation. *Journal of Advanced Nursing, 73*(12), 3119-3132.

McHugh M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, *22*(3), 276–282.

McKee, M., Britton, A., Black, N., McPherson, K., Sanderson, C., & Bain, C. (1999). Methods in health services research. Interpreting the evidence: choosing between randomised and non-randomised studies. *Bmj, 319*(7205), 312-315.

Mcloughlin, A., Olive, P., & Lightbody, L. (2022) *How reliable is the National Institutes for Health Stroke Scale (NIHSS) when rated from telemedicine recordings? British Journal of Neuroscience Nursing*, 18(Sup5). 3-10.

Meurer, W. J., Majersik, J. J., Frederiksen, S. M., Kade, A. M., Sandretto, A. M., & Scott, P. A. (2011). Provider perceptions of barriers to the emergency use of tPA for acute ischemic stroke: a qualitative study. *BMC Emerg Med, 11*, 5.

Meyer, B. C., Hemmen, T. M., Jackson, C. M., & Lyden, P. D. (2002). Modified National Institutes of Health Stroke Scale for use in stroke clinical trials: prospective reliability and validity. *Stroke, 33*(5), 1261-1266.

Meyer, B. C., Lyden, P. D., Al-Khoury, L., Cheng, Y., Raman, R., Fellman, R., Beer, J., Rao, R., & Zivin, J. A. (2005). Prospective reliability of the STRokE DOC wireless/site independent telemedicine system. *Neurology*, *64*(6), 1058–1060.

Meyer, B. C., Raman, R., Chacon, M. R., Jensen, M., & Werner, J. D. (2008). Reliability of site-independent telemedicine when assessed by telemedicine-naive stroke practitioners. *Journal of Stroke & Cerebrovascular Diseases, 17*(4), 181-186.

Middleton, S., McElduff, P., Ward, J., Grimshaw, J. M., Dale, S., D'Este, C., Drury, P., Griffiths, R., Cheung, N. W., Quinn, C., Evans, M., Cadilhac, D., Levi, C., & QASC Trialists Group (2011). Implementation of evidence-based treatment protocols to manage fever, hyperglycaemia, and swallowing dysfunction in acute stroke (QASC): a cluster randomised controlled trial. *Lancet (London, England)*, *378*(9804), 1699–1706.

Millis, S. R., Straube, D., Iramaneerat, C., Smith, E. V., Jr, & Lyden, P. (2007). Measurement properties of the National Institutes of Health Stroke Scale for people with right- and left-hemisphere lesions: further analysis of the clomethiazole for acute stroke study-ischemic (class-I) trial. *Archives of physical medicine and rehabilitation*, *88*(3), 302–308.

Mingers, J. (2004). Real-izing information systems: critical realism as an underpinning philosophy for information systems. *Information and Organization, 14*(2), 87-103.

Mingers, J. (2011) The contribution of systemic thought to critical realism, *Journal of Critical Realism,* 10(3), 303-330.

Mingers, J., Mutch, A., & Willcocks, L. P. (2013) *Critical realism in information systems research. MIS Quarterly: Management Information Systems, 37*(3), 795-802.

Miyamoto, N., Tanaka, R., Ueno, Y., Watanabe, M., Kurita, N., Hira, K., Shimada, Y., Kuroki, T., Yamashiro, K., Urabe, T., & Hattori, N. (2017). Analysis of the usefulness of the WORSEN score for predicting the deterioration of acute ischemic stroke. *Journal of stroke and cerebrovascular diseases: the official journal of National Stroke Association*, *26*(12), 2834–2839.

Mohammmed Iddrisu, S., Hutchinson, A. F., Sungkar, Y., & Considine, J. (2018). Nurses' role in recognising and responding to clinical deterioration in surgical patients. *Journal of Clinical Nursing, 27*(9-10), 1920-1930.

Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research, 27*(5), 1171-1179.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010a). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation*, *19*(4), 539–549.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010b). The COSMIN study reached international consensus on

taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of clinical epidemiology*, *63*(7), 737–745.

Mokkink, L. B., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., & Terwee, C. B. (2017). *COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs).* Retrieved June 24, 2022, from https://cosmin.nl/wp-content/uploads/COSMIN_manual_syst-review-PROMs_V1.0.pdf

Moore, A. (2022). *Exclusive: sharp drop in quality of stroke care*. Retrieved June 1, 2022, from https://www.hsj.co.uk/quality-and-performance/exclusive-sharp-drop-in-quality-of-stroke-care/7032477

Morgan, D. (1998). *Planning focus groups*. Thousand Oaks, CA: Sage.

Muir, K. W., Grosset, D. G., & Lees, K. R. (1994). Interconversion of stroke scales: Implications for therapeutic trials. *Stroke, 25*(7), 1366-1370.

Murphy, M. K., Black, N. A., Lamping, D. L., McKee, C. M., Sanderson, C. F., Askham, J., & Marteau, T. (1998). Consensus development methods, and their use in clinical guideline development. *Health Technol Assess, 2*(3), i-iv, 1-88.

Murta, V., & Ferrari, C. C. (2013). Influence of Peripheral inflammation on the progression of multiple sclerosis: evidence from the clinic and experimental animal models. *Mol Cell Neurosci, 53*, 6-13.

Myint, P. K., Staufenberg, E. F., & Sabanathan, K. (2006). Post-stroke seizure and post-stroke epilepsy. *Postgrad Med J, 82*(971), 568-572.

Nanri, Y., Yakushiji, Y., Hara, M., Eriguchi, M., Okada, R., Yukitake, M., & Hara, H. (2013). Stroke scale items associated with neurologic deterioration within 24 hours after recombinant tissue plasminogen activator therapy. *Journal of Stroke & Cerebrovascular Diseases, 22*(7), 1117-1124.

Nemoto, E. M., & Frankel, H. M. (1970). Cerebral oxygenation and metabolism during progressive hyperthermia. *Am J Physiol, 219*(6), 1784-1788.

NICE (National Institute for Health and Care Excellence). (2019) *Stroke and transient ischaemic attack in over 16s: diagnosis and initial management. NICE guideline [NG128].* Retrieved June 4, 2022, from https://www.nice.org.uk/guidance/ng128/chapter/Recommendations#blood-pressure-control-for-people-with-acute-intracerebral-haemorrhage

NICE (National Institute for Health and Care Excellence. (2021) *Exceptional surveillance of stroke and transient ischaemic attack in over 16s (NICE guideline NG128)* Retrieved June 4, 2022, from https://www.nice.org.uk/guidance/ng128/resources/2021-exceptional-surveillance-of-stroke-and-transient-ischaemic-attack-in-over-16s-nice-guideline-ng128-pdf-11757288440005

National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. (1995) Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med, 333*(24), 1581-1587.

National Records of Scotland. (2016). *Deaths by sex, age and cause, Scotland, 2015*. Retrieved December 20, 2018, from http://bit.ly/2h671Pu.

NHS. (2019). *The NHS long term plan*. Retrieved June 16, 2022, from https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf

NHS. (2020). *We are the NHS: People plan for 2020/21- action for us all.* Retrieved June 16, 2022, from https://www.england.nhs.uk/wp-content/uploads/2020/07/We-Are-The-NHS-Action-For-All-Of-Us-FINAL-March-21.pdf

NHS Digital. (2018). *Mortality from stroke*. Retrieved January 31, 2019, from https://digital.nhs.uk/data-and-information/publications/clinical-indicators/compendium-of-population-health-indicators/compendium-mortality/current/mortality-from-stroke

NHS England. (2012). *Change Model*. Retrieved June 21, 2022, from https://www.england.nhs.uk/sustainableimprovement/change-model/

NHS England. (2014). *MDT development- working towards an effective multidisciplinary/multiagency team*. Retrieved June 20, 2022, from https://www.england.nhs.uk/wp-content/uploads/2015/01/mdt-dev-guid-flat-fin.pdf

NHS England and NHS Improvement. (2018). *Patient Safety Alert: Resources to support the safe adoption of the revised National Early Warning Score (NEWS2).* Retrieved June 1, 2022, from https://www.england.nhs.uk/wp-content/uploads/2019/12/Patient_Safety_Alert_-_adoption_of_NEWS2.pdf

NHS England and NHS Improvement. (2020) National Stroke Service Model. Retrieved June 14, 2022, from https://www.england.nhs.uk/wp-content/uploads/2021/05/national-stroke-service-model-integrated-stroke-delivery-networks-may-2021.pdf

NHS RightCare. (2017) *RightCare Pathway: Stroke*. Retrieved June 3, 2022, from https://www.england.nhs.uk/rightcare/wp-content/uploads/sites/40/2017/09/stroke-pathway.pdf

Nielsen, P. B., Langkjær, C. S., Schultz, M., Kodal, A. M., Pedersen, N. E., Petersen, J. A., Lange, T., Arvig, M. D., Meyhoff, C. S., Bestle, M. H., Hølge-Hazelton, B., Bunkenborg, G., Lippert, A., Andersen, O., Rasmussen, L. S., & Iversen, K. K. (2022). Clinical assessment as a part of an early warning score-a Danish cluster-randomised, multicentre study of an individual early warning score. *The Lancet. Digital health*, *4*(7), e497–e506.

Nilanont, Y., Komoltri, C., Saposnik, G., Côté, R., Di Legge, S., Jin, Y., Prayoonwiwat, N., Poungvarin, N., & Hachinski, V. (2010). The Canadian Neurological Scale and the NIHSS: development and validation of a simple conversion model. *Cerebrovascular diseases (Basel, Switzerland)*, *30*(2), 120–126.

Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science, 10*(1), 53.

Nilsen, P., Seing, I., Ericsson, C., Birken, S. A., & Schildmeijer, K. (2020). Characteristics of successful changes in health care organizations: an interview study with physicians, registered nurses and assistant nurses. *BMC Health Services Research, 20*(1), 147.

Norris, B. (2009). Human factors and safe patient care. *Journal of Nursing Management, 17*(2), 203-211.

Northern Ireland Statistics and Research Agency (NISRA). (2016). *Deaths, by sex, age and cause, 2015.* Retrieved December 20, 2019 from http://bit.ly/29ffuON

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1609406917733847.

Nunnally, J. C. (1994). *Psychometric theory*. McGraw-Hill.

Nye, B. R., Hyde, C. E., Tsivgoulis, G., Albright, K. C., Alexandrov, A. V., & Alexandrov, A. W. (2012). Slim stroke scales for assessing patients with acute stroke: ease of use or loss of valuable assessment data? *American Journal of Critical Care, 21*(6), 442-447.

O'Daniel, M., & Rosenstein, A. H. (2008). Professional Communication and Team Collaboration. In R. G. Hughes (Ed.), *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*. Agency for Healthcare Research and Quality (US) (chapter 33) Retrieved June 1, 2022, from https://www.ncbi.nlm.nih.gov/books/NBK2637/

Office for National Statistics. (2016). *Deaths registered in England and Wales: 2015*. Retrieved December 20, 2019 from http://bit.ly/2h7Om7P

O'Hagan, A., Buck, C., Daneshkhah, A., Eiser, J., Garthwaite, P., Jenkinson, D., Oakley, J., & Rakow, T. (2006) Uncertain judgements: Eliciting experts' probabilities. Chichester: John Wiley & Sons.

O'Reilly, M., & Parker, N. (2013). 'Unsatisfactory saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research*, 13(2), 190-197.

O'Rourke, F., Chan, D. K., Chan, D. L., & Man Ding, X. (2013). Preferences, barriers and facilitators for establishing comprehensive stroke units: a multidisciplinary survey. *Aust Health Rev, 37*(3), 318-323.

Olivato, S., Nizzoli, S., Cavazzuti, M., Casoni, F., Nichelli, P. F., & Zini, A. (2016). e-NIHSS: an expanded National Institutes of Health Stroke Scale weighted for anterior and posterior circulation strokes. *Journal of Stroke and Cerebrovascular Diseases, 25*(12), 2953-2957.

Oltmann, C., & Boughey, C. (2012). Using critical realism as a framework in pharmacy education and social pharmacy research. *Res Social Adm Pharm, 8*(4), 333-337.

Orgogozo, J. M., Asplund, K., & Boysen, G. (1992). A unified form for neurological scoring of hemispheric stroke with motor impairment. *Stroke, 23*(11), 1678-1679.

Orgogozo, J. M., Capildeo, R., Anagnostou, C. N., Juge, O., Péré, J. J., Dartigues, J. F., Steiner, T. J., Yotis, A., & Rose, F. C. (1983). Mise au point d'un score neurologique pour l'évaluation clinique des infarctus sylviens [Development of a neurological score for the clinical evaluation of sylvian infarctions]. *Presse medicale (Paris, France : 1983)*, *12*(48), 3039–3044.

Ornstein, S. M., Oates, R. B., & Fox, G. N. (1992). The computer-based medical record: current status. *The Journal of family practice*, *35*(5), 556–565.

Ostrom, E. (1996). Crossing the great divide: Coproduction, synergy, and development. *World Development, 24*(6), 1073-1087.

Ovesen, C., Christensen, A. F., Havsteen, I., Krarup Hansen, C., Rosenbaum, S., Kurt, E., & Christensen, H. (2015). Prediction and prognostication of neurological deterioration in patients with acute ICH: a hospital-based cohort study. *BMJ Open, 5*(7), e008563.

Oxbury, J. M., Greenhall, R. C., & Grainger, K. M. (1975). Predicting the outcome of stroke: acute stage after cerebral infarction. *Br Med J, 3*(5976), 125-127.

Pagliari, C., Grimshaw, J., & Eccles, M. (2001). The potential influence of small group processes on guideline development. *Journal of Evaluation in Clinical Practice, 7*(2), 165-173.

Panteli, D., Legido-Quigley, H., Reichebner, C., Ollenschlager, G., Schafer, C., & Busse, R. Clinical Practice Guidelines as a quality strategy. In: Busse R, Klazinga N, Panteli D, & Quentin W. (Eds.) Improving healthcare quality in Europe: Characteristics, effectiveness and implementation of different strategies. European Observatory on Health Systems and Policies.(chapter 9) Retrieved June 1, 2022, from https://www.ncbi.nlm.nih.gov/books/NBK549283/

Paradis, E., O'Brien, B., Nimmon, L., Bandiera, G., & Martimianakis, M. A. (2016). Design: Selection of Data Collection Methods. *Journal of graduate medical education*, *8*(2), 263–264.

Pastrana, T., Radbruch, L., Nauck, F., Höver, G., Fegg, M., Pestinger, M., Ross, J., Krumm, N., & Ostgathe, C. (2010). Outcome indicators in palliative care--how to assess quality and success. Focus group and nominal group technique in Germany. *Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer*, *18*(7), 859–868.

Patel, A., Berdunov, V., King, D., Quayyum, Z., Wittenberg, R., & Knapp, M. (2017) *Executive summary part 2: societal cost of stroke in the next 20 years and potential returns from increased spending on research*. Stroke Association. Retrieved June 2, 2022, from https://www.stroke.org.uk/sites/default/files/costs_of_stroke_in_the_uk_report_-executive_summary_part_2.pdf

Patel, M., Coshall, C., Rudd, A. G., & Wolfe, C. D. (2003). Natural history of cognitive impairment after stroke and factors associated with its recovery. *Clinical Rehabilitation, 17*(2), 158-166.

Pavelin, K., Pundir, S., & Cham, J. A. (2014). Ten simple rules for running interactive workshops. *PLoS Comput Biol, 10*(2), e1003485.

Pawson, R. & Tilley, N. (1997) *Realistic evaluation*. Sage.

Pelletier, L. R. (2010). *Information-Enabled Decision-Making in Health Care: EHR-Enabled Standardization, Physician Profiling and Medical Home*. Retrieved June 2, 2022, from https://web.wpi.edu/Pubs/ETD/Available/etd-042510-120618/unrestricted/Dissertation_Lori_Pelletier_FINAL.pdf .

Peters, C., O'Callaghan, C., O'Callaghan, S., O'Hanlon, S., O'Malley, G., O'Connor, M., & Lyons, D. (2012). App deployment of NIHSS (National Institutes of Health Stroke Scale) is faster than paper format in clinical practice. *Irish Journal of Medical Science, 181*, S246.

Peirce, C. S. (1978). "Pragmatism and abduction." In C. Hartshorne & P. Weiss (Eds.), *Collected papers of Charles Sanders Peirce* (pp180-212). Harvard University Press.

Physio—pedia. (2023) Stroke: Assessment. Miami Emergency Neurological deficit (MEND). Retrieved February 21, 2023, from https://www.physio-pedia.com/Stroke:_Assessment.

Plsek, P. E., & Greenhalgh, T. (2001). Complexity science: The challenge of complexity in health care. *BMJ (Clinical research ed.), 323*(7313), 625-628.

Pope, C., Ziebald, S., Mays, N. (2000). Qualitative Research in Health Care, Analysing qualitative data. British Medical Journal, 320,114-116.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: applications to practice* (Vol. 892): Pearson/Prentice Hall Upper Saddle River, NJ.

Potter, M., Gordon, S., & Hamer, P. (2004) The nominal group technique: A useful consensus methodology in physiotherapy research. *New Zealand Journal of Physiotherapy,* 32(3), 126-130.

Powell, A. E., Davies, H. T., & Thomson, R. G. (2003). Using routine comparative data to assess the quality of health care: understanding and avoiding common pitfalls. *Qual Saf Health Care, 12*(2), 122-128.

Powers, J. H., & Scheld, W. M. (1996). Fever in neurologic diseases. *Infect Dis Clin North Am, 10*(1), 45-66.

Powers, W. J., Rabinstein, A. A., Ackerson, T., Adeoye, O. M., Bambakidis, N. C., Becker, K., Biller, J., Brown, M., Demaerschalk, B. M., Hoh, B., Jauch, E. C., Kidwell, C. S., Leslie-Mazwi, T. M., Ovbiagele, B., Scott, P. A., Sheth, K. N., Southerland, A. M., Summers, D. V., & Tirschwell, D. L. (2019). Guidelines for the Early Management of Patients With Acute Ischemic Stroke: 2019 Update to the 2018 Guidelines for the Early Management of Acute Ischemic Stroke: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke*, *50*(12), e344–e418.

Prescott, R. J., Garraway, W. M., & Akhtar, A. J. (1982). Predicting functional outcome following acute stroke using a standard clinical examination. *Stroke, 13*(5), 641-647.

Price, C. J., Seghier, M. L., & Leff, A. P. (2010). Predicting language outcome and recovery after stroke: the PLORAS system. Nature reviews. Neurology, 6(4), 202–210.

Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res, 27*(5), 1147-1157.

Przelomski, M. M., Roth, R. M., Gleckman, R. A., & Marcus, E. M. (1986). Fever in the wake of a stroke. *Neurology, 36*(3), 427-429.

Purrucker, J. C., Härtig, F., Richter, H., Engelbrecht, A., Hartmann, J., Auer, J., Hametner, C., Popp, E., Ringleb, P. A., Nagel, S., & Poli, S. (2017). Design and validation of a clinical

scale for prehospital stroke recognition, severity grading and prediction of large vessel occlusion: the shortened NIH Stroke Scale for emergency medical services. *BMJ open*, *7*(9), e016893.

Purvis, T., Moss, K., Denisenko, S., Bladin, C., & Cadilhac, D. A. (2014). Implementation of evidence-based stroke care: enablers, barriers, and the role of facilitators. *J Multidiscip Healthc, 7*, 389-400.

Pype, P., Mertens, F., Helewaut, F., & Krystallidou, D. (2018). Healthcare teams as complex adaptive systems: understanding team behaviour through team members' perception of interpersonal interaction. *BMC Health Services Research, 18*(1), 570.

Qureshi, A. I., Ali, Z., Suri, M. F., Shuaib, A., Baker, G., Todd, K., Guterman, L. R., & Hopkins, L. N. (2003). Extracellular glutamate and other amino acids in experimental intracerebral hemorrhage: an in vivo microdialysis study. *Critical care medicine*, *31*(5), 1482–1489.

Rabøl, L. I., McPhail, M. A., Østergaard, D., Andersen, H. B., & Mogensen, T. (2012). Promoters and barriers in hospital team communication. A focus group study. *Journal of Communication in Healthcare, 5*(2), 129-139.

Rafferty, A. E., & Griffin, M. A. (2006). Perceptions of organizational change: a stress and coping perspective. *J Appl Psychol, 91*(5), 1154-1162.

Raine, R., Sanderson, C., & Black, N. (2005). Developing clinical guidelines: a challenge to current methods. *Bmj, 331*(7517), 631.

Ramsey, S., Willke, R., Briggs, A., Brown, R., Buxton, M., Chawla, A., Cook, J., Glick, H., Liljas, B., Petitti, D., & Reed, S. (2005). Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, *8*(5), 521–533.

Ramzai, J. (2020) Clearly explained: Pearson V/S Spearman Correlation Coefficient. Retrieved June 23, 2022, from https://towardsdatascience.com/clearly-explained-pearson-v-s-spearman-correlation-coefficient-ada2f473b8

Ratna, H., 2019. *The Importance of Effective Communication in Healthcare Practice*. Harvard Public Health Review. Retrieved June 25, 2022, from https://hphr.org/23-article-ratna/

Redfern, S. (1996). Individualised patient care: Its meaning and practice in a general setting. *NT Research, 1*(1), 22-33.

Reeves, M., Khoury, J., Alwell, K., Moomaw, C., Flaherty, M., Woo, D., Khatri, P., Adeoye, O., Ferioli, S., Kissela, B., & Kleindorfer, D. (2013). Distribution of National Institutes of Health stroke scale in the Cincinnati/Northern Kentucky stroke study. *Stroke*, *44*(11), 3211–3213.

Reis, R. S., Salvo, D., Ogilvie, D., Lambert, E. V., Goenka, S., & Brownson, R. C. (2016). Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. *Lancet, 388*(10051), 1337-1348.

Reith, F. C., Van den Brande, R., Synnot, A., Gruen, R., & Maas, A. I. (2016). The reliability of the Glasgow Coma Scale: a systematic review. *Intensive Care Med, 42*(1), 3-15.

Richardson, J., Murray, D., House, C. K., & Lowenkopf, T. (2006). Successful implementation of the National Institutes of Health Stroke Scale on a stroke/neurovascular unit. *J Neurosci Nurs, 38*(4 Suppl), 309-315.

Rödén-Jüllig, Å., & Britton, M. (2000). Effectiveness of heparin treatment for progressing ischaemic stroke: before and after study. *Journal of Internal Medicine, 248*, 287-291.

Rödén-Jüllig, Å., Britton, M., Malmkvist, K., & Leijd, B. (2003). Aspirin in the prevention of progressing stroke: a randomized controlled study. *Journal of Internal Medicine, 254*(6), 584-590.

Roffe, C., Nevatte, T., Sim, J., Bishop, J., Ives, N., Ferdinand, P., & Gray, R. (2017). Effect of routine low-dose oxygen supplementation on death and disability in adults with acute stroke: The stroke oxygen study randomized clinical trial. *JAMA, 318*(12), 1125-1135.

Rohde, A., Worrall, L., Godecke, E., O'Halloran, R., Farrell, A., & Massey, M. (2018). Diagnosis of aphasia in stroke populations: A systematic review of language tests. *PloS one*, *13*(3).

Roquer, J., Rodríguez-Campello, A., Gomis, M., Jiménez-Conde, J., Cuadrado-Godia, E., Vivanco, R., Giralt, E., Sepúlveda, M., Pont-Sunyer, C., Cucurella, G., & Ois, A. (2008). Acute stroke unit care and early neurological deterioration in ischemic stroke. *Journal of neurology*, *255*(7), 1012–1017.

Rosen, M. A., DiazGranados, D., Dietz, A. S., Benishek, L. E., Thompson, D., Pronovost, P. J., & Weaver, S. J. (2018). Teamwork in healthcare: Key discoveries enabling safer, high-quality care. *The American psychologist, 73*(4), 433-450.

Rouse, W. B. (2008). Health Care as a Complex Adaptive System: Implications for Design and Management. *The Bridge, 38*(1), 17-25.

Royal College of Nursing (RCN). (2017) Safe and effective staffing: nursing against the odds. UK Policy Report. Retrieved June 15, 2022, from https://www.rcn.org.uk/professional-development/publications/pub-006415

Royal College of Physicians. (2012) *National Early Warning Score (NEWS): standardising the assessment of acute-illness severity in the NHS*. Report of a working party. London: RCP.

Royal College of Physicians. (2016a) *National clinical guidelines for stroke*. Prepared by the intercollegiate stroke working party (5th ed.) Retrieved June 25, 2022, from https://www.rcplondon.ac.uk/guidelines-policy/stroke-guidelines

Royal College of Physicians. (2016b) *Sentinel Stroke National Audit Programme (SSNAP). Acute organisational audit November 2016.* Retrieved May 31, 2022, from https://www.strokeaudit.org/Documents/National/AcuteOrg/2016/2016-AOANationalReport.aspx

Royal College of Physicians. (2017a) *National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity in the NHS.* Updated report of a working party. London: RCP.

Royal College of Physicians. (2017b) *Sentinel Stroke National Audit Programme (SSNAP). National clinical audit annual results portfolio March 2016-April 2017*. Retrieved May 9, 2022, from http://bit.ly/1NHYlqH

Ryglewicz, D., Hier, D. B., Wiszniewska, M., Cichy, S., Lechowicz, W., & Czlonkowska, A. (2000). Ischemic strokes are more severe in Poland than in the United States. *Neurology*, *54*(2), 513–515.

Salkind, N. J. (Ed.) (2010). *Encyclopedia of research design*. SAGE Publications, Inc.

Salvador-Oliván, J. A., Marco-Cuenca, G., & Arquero-Avilés, R. (2019). Errors in search strategies used in systematic reviews and their effects on information retrieval. *Journal of the Medical Library Association : JMLA, 107*(2), 210-221.

Sanders, E. B. N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *CoDesign, 4*(1), 5-18.

Sapsford, R. (2007). *Survey research*. SAGE Publications, Ltd.

Sare, G. M., Geeganage, C., & Bath, P. M. W. (2009). High Blood Pressure in Acute Ischaemic Stroke – Broadening Therapeutic Horizons. *Cerebrovascular Diseases, 27*(Suppl. 1), 156-161.

Sayer, A. (2000). *Realism and social science*. SAGE Publications Ltd.

Scheppers, E., van Dongen, E., Dekker, J., Geertzen, J., & Dekker, J. (2006). Potential barriers to the use of health services among ethnic minorities: a review. Family practice, 23(3), 325–348.

Schmülling, S., Grond, M., Rudolf, J., & Kiencke, P. (1998). Training as a prerequisite for reliable use of NIH Stroke Scale. *Stroke, 29*(6), 1258-1259.

Schwab, S., & Hacke, W. (2003). Surgical decompression of patients with large middle cerebral artery infarcts is effective. *Stroke, 34*(9), 2304-2305.

Seki, M., Hase, K., Takahashi, H., & Liu, M. (2014). Comparison of three instruments to assess changes of motor impairment in acute hemispheric stroke: the Stroke Impairment Assessment Set (SIAS), the National Institute of Health Stroke Scale (NIHSS) and the Canadian Neurological Scale (CNS). *Disability & Rehabilitation, 36*(18), 1549-1554.

Seners, P., Turc, G., Oppenheim, C., & Baron, J. C. (2015). Incidence, causes and predictors of neurological deterioration occurring within 24 h following acute ischaemic stroke: a systematic review with pathophysiological implications. *Journal of Neurology, Neurosurgery & Psychiatry, 86*(1), 87-94.

Shafqat, S., Kvedar, J. C., Guanci, M. M., Chang, Y., & Schwamm, L. H. (1999). Role for telemedicine in acute stroke. Feasibility and reliability of remote administration of the NIH stroke scale. *Stroke, 30*(10), 2141-2145.

Shahid, S., & Thomas, S. (2018). Situation, Background, Assessment, Recommendation (SBAR) Communication Tool for Handoff in Health Care – A Narrative Review. *Safety in Health, 4*(1), 7.

Shimoyama, T., Kimura, K., Uemura, J., Saji, N., & Shibazaki, K. (2014). Elevated glucose level adversely affects infarct volume growth and neurological deterioration in non-diabetic stroke patients, but not diabetic stroke patients. *European Journal of Neurology, 21*(3), 402-410.

Shinar, D., Gross, C. R., Mohr, J. P., Caplan, L. R., Price, T. R., Wolf, P. A., Hier, D. B., Kase, C. S., Fishman, I. G., & Wolf, C. L. (1985). Interobserver variability in the assessment of neurologic history and examination in the Stroke Data Bank. *Archives of neurology*, *42*(6), 557–565.

Siegler, J. E., Albright, K. C., George, A. J., Boehme, A. K., Gillette, M. A., Kumar, A. D., Aswani, M., & Martin-Schild, S. (2017). Time to Neurological Deterioration in Ischemic Stroke. Medical student research journal, 4, 18–24.

Siegler, J. E., Boehme, A. K., Kumar, A. D., Gillette, M. A., Albright, K. C., & Martin-Schild, S. (2013). What change in the National Institutes of Health Stroke Scale should define neurologic deterioration in acute ischemic stroke? *Journal of stroke and cerebrovascular diseases : the official journal of National Stroke Association, 22*(5), 675-682.

Siegler, J. E., & Martin-Schild, S. (2011). Early Neurological Deterioration (END) after Stroke: The END Depends on the Definition. *International Journal of Stroke, 6*(3), 211-212.

Silva, Y., Puigdemont, M., Castellanos, M., Serena, J., Suñer, R. M., García, M. M., & Dávalos, A. (2005). Semi-intensive monitoring in acute stroke and long-term outcome. *Cerebrovascular Diseases, 19*(1), 23-30.

Sim, J., Saunders, b., Waterfield, J., & Kingstone, T. (2018). Can sample size in qualitative research be determined apriori? *International Journal of Social Research Methodology, 21*(5), 619-634.

Sim, J., & Wright, C. C. (2005). The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy, 85*(3), 257-268.

Singer, O. C., Dvorak, F., du Mesnil de Rochemont, R., Lanfermann, H., Sitzer, M., & Neumann-Haefelin, T. (2005). A simple 3-item stroke scale: comparison with the National Institutes of Health Stroke Scale and prediction of middle cerebral artery occlusion. *Stroke, 36*(4), 773-776.

Skills for Care & Skills for Health. (2013). *Code of conduct for healthcare support workers and adult social care workers in England.* Retrieved June 15, 2022, from https://www.skillsforcare.org.uk/Documents/Standards-legislation/Code-of-Conduct/Code-of-Conduct.pdf

Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., Boyd, K. A., Craig, N., French, D. P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M., & Moore, L. (2021a). Framework for the development and evaluation of complex interventions: gap analysis, workshop and consultation-informed update. *Health technology assessment (Winchester, England)*, *25*(57), 1–132

Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., Boyd, K. A., Craig, N., French, D. P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M., & Moore,

L. (2021b). A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ (Clinical research ed.)*, *374*, n2061.

Skjott Linneberg, M., & Korsgaard, S. (2019). Coding qualitative data: a synthesis guiding the novice. *Qualitative Research Journal, 19*(3), 259-270.

Smith, L. N., Craig, L. E., Weir, C. J., & McAlpine, C. H. (2008). Stroke education for healthcare professionals: making it fit for purpose. *Nurse Educ Today, 28*(3), 337-347.

Sorimachi, T., & Fujii, Y. (2010). Early neurological change in patients with spontaneous supratentorial intracerebral hemorrhage. *Journal of Clinical Neuroscience, 17*(11), 1367-1371.

Specogna, A. V. (2013). *Evidence for the appropriate testing and implementation of new therapies for intracerebral hemorrhage in Canada.* [doctoral thesis] Retrieved October 20, 2019, from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2013-99120-337&site=ehost-live

Specogna, A. V., Patten, S. B., Turin, T. C., & Hill, M. D. (2013). The reliability and sensitivity of the National Institutes of Health Stroke Scale for spontaneous intracerebral hemorrhage in an uncontrolled setting. *PLoS ONE, 8*(12), e84702.

Sperber, D., Mortimer, D., Lorgelly, P., & Berlowitz, D. (2013). An expert on every street corner? Methods for eliciting distributions in geographically dispersed opinion pools. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, *16*(2), 434–437.

Spilker, J., Kongable, G., Barch, C., Braimah, J., Brattina, P., Daley, S., Donnarumma, R., Rapp, K., & Sailor, S. (1997). Using the NIH Stroke Scale to assess stroke patients. The NINDS rt-PA Stroke Study Group. *The Journal of neuroscience nursing: journal of the American Association of Neuroscience Nurses*, *29*(6), 384–392.

SSNAP (Sentinel Stroke National Audit Programme), School of Population Health and Environmental Sciences, King's College London on behalf of the Intercollegiate Stroke Working Party. (2019). *Clinical audit April 2013-March 2018. National Results*. Retrieved June 5, 2022, from https://www.strokeaudit.org/Documents/National/Clinical/Apr2017Mar2018/Apr2017Mar2018-AnnualReport.aspx

SSNAP (Sentinel Stroke National Audit Programme). (2021), *SSNAP Collaboration*. Retrieved June 5, 2022, from https://www.strokeaudit.org/Research/SSNAP-Collaboration.aspx

Stecksén, A., Lundman, B., Eriksson, M., Glader, E. L., & Asplund, K. (2014). Implementing thrombolytic guidelines in stroke care: perceived facilitators and barriers. *Qual Health Res, 24*(3), 412-419.

Steinke, W., & Ley, S. C. (2002). Lacunar stroke is the major cause of progressive motor deficits. *Stroke, 33*(6), 1510-1516.

Stewart, R., & Liabo, K. (2012). Involvement in research without compromising research quality. *Journal of Health Services Research & Policy, 17*(4), 248-251.

Streiner, D.L. and Norman, G.R. (2008) *Health measurement scales: A practical guide to their development and use*. 4th Edition, Oxford University Press, Oxford

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health Measurement Scales: A practical guide to their development and use: A practical guide to their development and use*. 5th Edition, Oxford University Press.

Stroke Unit Trialists' Collaboration, Organised inpatient (stroke unit) care for stroke. (2013). *Cochrane Database of Systematic Reviews, 2013*(9), Cd000197.

Stubits, D., Silliman, S., Hodges, W., & Johnson, Y. (2015). Post stroke assessment tool for nurses validated. *Stroke. Conference: American Heart Association/American Stroke Association, 46* (Suppl. 1).

Suh, S. W., Shin, B. S., Ma, H., Van Hoecke, M., Brennan, A. M., Yenari, M. A., & Swanson, R. A. (2008). Glucose and NADPH oxidase drive neuronal superoxide formation in stroke. *Ann Neurol, 64*(6), 654-663.

Sümer, M. M., & Özön, A. Ö. (2017). Progression in acute ischemic stroke: Is widespread atherosclerotic background a risk factor? *Turkish journal of physical medicine and rehabilitation, 64*(1), 46-51.

Summers, D., Leonard, A., Wentworth, D., Saver, J. L., Simpson, J., Spilker, J. A., Hock, N., Miller, E., Mitchell, P. H., & American Heart Association Council on Cardiovascular Nursing and the Stroke Council (2009). Comprehensive overview of nursing and interdisciplinary care of the acute ischemic stroke patient: a scientific statement from the American Heart Association. *Stroke*, *40*(8), 2911–2944.

Sun, W., Peacock, A., Becker, J., Phillips-Bute, B., Laskowitz, D. T., & James, M. L. (2012). Correlation of leukocytosis with early neurological deterioration following supratentorial intracerebral hemorrhage. *J Clin Neurosci, 19*(8), 1096-1100.

Sutcliffe, K. M., Lewton, E., & Rosenthal, M. M. (2004). Communication failures: an insidious contributor to medical mishaps. *Acad Med, 79*(2), 186-194.

Syed, J., Mingers, J., & Murray, P. A. (2010). Beyond rigour and relevance: A critical realist approach to business education. *Management Learning, 41*(1), 71-85.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics, 5th ed*. Allyn & Bacon/Pearson Education.

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education, 48*(6), 1273-1296.

Taj Alexander, M.-H., & Don, A. D. (2010). Does Visual Appeal Matter? Effects of Web Survey Aesthetics on Survey Quality. *Survey Research Methods, 4*(1).

Tang, E. Y. H., Price, C., Stephan, B. C. M., Robinson, L., & Exley, C. (2017). Gaps in care for patients with memory deficits after stroke: views of healthcare providers. *BMC Health Serv Res, 17*(1), 634.

Tang, H., Zhang, S., Yan, S., Liebeskind, D. S., Sun, J., Ding, X., . . . Lou, M. (2016). Unfavorable neurological outcome in diabetic patients with acute ischemic stroke is associated with

incomplete recanalization after intravenous thrombolysis. *J Neurointerv Surg, 8*(4), 342-346.

Tappen, R. M., Wolf, D. G., Rahemi, Z., Engstrom, G., Rojido, C., Shutes, J. M., & Ouslander, J. G. (2017). Barriers and facilitators to implementing a change initiative in long-term care Using the INTERACT® Quality improvement program. *The health care manager, 36*(3), 219-230.

Tarassenko, L., Hann, A., & Young, D. (2006). Integrated monitoring and analysis for early warning of patient deterioration. *Br J Anaesth, 97*(1), 64-68.

Tariq, S., & Woodman, J. (2013). Using mixed methods in health research. *JRSM short reports*, *4*(6), 2042533313479197.

Tashakkori, A., & Creswell, J. W. (2007). Editorial: The New Era of Mixed Methods. *Journal of Mixed Methods Research, 1*(1), 3-7.

Tashakkori, A., & Teddlie, C. (2010). *SAGE Handbook of Mixed Methods in Social & Behavioral Research.* (2 ed.). SAGE Publications, Inc.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education, 2*, 53-55.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness. A practical scale. *Lancet, 2*(7872), 81-84.

Teherani, A., Martimianakis, T., Stenfors-Hayes, T., Wadhwa, A., & Varpio, L. (2015). Choosing a Qualitative Research Approach. *Journal of graduate medical education*, *7*(4), 669–670.

Tei, H., Uchiyama, S., Ohara, K., Kobayashi, M., Uchiyama, Y., & Fukuzawa, M. (2000). Deteriorating ischemic stroke in 4 clinical categories classified by the Oxfordshire Community Stroke Project. *Stroke, 31*(9), 2049-2054.

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of clinical epidemiology*, *60*(1), 34–42.

Terwee, C. B., Jansma, E. P., Riphagen, II, & de Vet, H. C. (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res, 18*(8), 1115-1123.

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W. J. G., Bouter, L. M., & de Vet, H. C. W. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research, 21*(4), 651-657.

Terwee, C. B., Prinsen, C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation*, *27*(5), 1159–1170.

Theofanidis, D., Fitsioris, X., & Iakovos, T. (2015) Stroke scales and trajectory of recovery: A major concern for patients and nurses alike. *Health Science Journal, 9*(4).

Thomas, J., Dahm, M. R., Li, J., Smith, P., Irvine, J., Westbrook, J. I., & Georgiou, A. (2020). Variation in electronic test results management and its implications for patient safety: A multisite investigation. *Journal of the American Medical Informatics Association : JAMIA, 27*(8), 1214-1224.

Thompson, D. S., Fazio, X., Kustra, E., Patrick, L., & Stanley, D. (2016). Scoping review of complexity theory in health services research. *BMC Health Services Research, 16*(1), 87.

Tisdall, M., Crocker, M., Watkiss, J., & Smith, M. (2006). Disturbances of sodium in critically ill adult neurologic patients: a clinical review. *Journal of neurosurgical anesthesiology, 18*(1), 57-63.

Tollyfield, R. (2014). Facilitating an accelerated experience-based co-design project. *Br J Nurs, 23*(3), 136-141.

Tomson, C. R., & van der Veer, S. N. (2013). Learning from practice variation to improve the quality of care. *Clinical medicine (London, England), 13*(1), 19-23.

Tong, A., Sainsbury, P., & Craig, J. (2007) Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care,* 19(6), 349-357.

Toni, D., Fiorelli, M., Zanette, E. M., Sacchetti, M. L., Salerno, A., Argentino, C., Solaro, M., & Fieschi, C. (1998). Early spontaneous improvement and deterioration of ischemic stroke patients. A serial study with transcranial Doppler ultrasonography. *Stroke*, *29*(6), 1144–1148.

Treves, T. A., Karepov, V. G., Aronovich, B. D., Gorbulev, A. Y., Bornstein, N. M., & Korczyn, A. D. (1994). Interrater agreement in evaluation of stroke patients with the unified neurological stroke scale. *Stroke, 25*(6), 1263-1264.

Tysinger, E. L. (2015). *How Vital Are Vital Signs ? A Systematic Review of Vital Sign Compliance and Accuracy in Nursing.* Journal of Science & Medicine. Retrieved June 25, 2022, from https://docplayer.net/14504887-Original-science-how-vital-are-vital-signs-a-systematic-review-of-vital-sign-compliance-and-accuracy-in-nursing.html

University of Newcastle Library (2022) Research Methods: What are research methods? Retrieved June 6, 2022, from https://libguides.newcastle.edu.au/researchmethods#:~:text=Research%20methods%20are%20the%20strategies,different%20tools%20for%20data%20collection.

Van de Ven, A., & Delbecq, A. (1971) Nominal vs. interacting group process for committee decision-making effectiveness. *Academy of Management Journal,* 14**,** 203–212.

van der Scheer, J. W., Woodward, M., Ansari, A., Draycott, T., Winter, C., Martin, G., Kuberska, K., Richards, N., Kern, R., Dixon-Woods, M., Thiscovery Authorship Group, & Obstetric Emergency Consensus Authorship Group. (2021). How to specify healthcare process improvements collaboratively using rapid, remote consensus-building: a framework and a case study of its application. *BMC Medical Research Methodology, 21*(1), 103.

Van Der Weijden, T., Hooi, J. D., Grol, R., & Limburg, M. (2004). A multidisciplinary guideline for the acute phase of stroke: barriers perceived by Dutch neurologists. *J Eval Clin Pract, 10*(2), 241-246.

Venkatesh, V., Brown, S. A., & Bala, H. (2013). "Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems," *MIS Quarterly, 37*(1), 21-25.

Vermeir, P., Vandijck, D., Degroote, S., Peleman, R., Verhaeghe, R., Mortier, E., Hallaert, G., Van Daele, S., Buylaert, W., & Vogelaers, D. (2015). Communication in healthcare: a narrative review of the literature and practical recommendations. *International journal of clinical practice*, *69*(11), 1257–1267.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Family medicine, 37 5*, 360-363.

Vila, N., Castillo, J., Dávalos, A., & Chamorro, A. (2000). Proinflammatory cytokines and early neurological worsening in ischemic stroke. *Stroke, 31*(10), 2325-2329.

Vitt, J. R., Trillanes, M., & Hemphill, J. C., 3rd. (2019). Management of blood pressure during and after recanalization therapy for acute ischemic stroke. *Front Neurol, 10*, 138.

Wade, D. T. (2004). Assessment, measurement and data collection tools. *Clinical Rehabilitation, 18*(3), 233-237.

Wahlgren, N., Ahmed, N., Dávalos, A., Ford, G. A., Grond, M., Hacke, W., Hennerici, M. G., Kaste, M., Kuelkens, S., Larrue, V., Lees, K. R., Roine, R. O., Soinne, L., Toni, D., Vanhooren, G., & SITS-MOST investigators (2007). Thrombolysis with alteplase for acute ischaemic stroke in the Safe Implementation of Thrombolysis in Stroke-Monitoring Study (SITS-MOST): an observational study. Lancet (London, England), 369(9558), 275–282.

Walsh, D., & Evans, K. (2014). Critical realism: an important theoretical perspective for midwifery research. *Midwifery*, *30*(1), e1–e6.

Wang, J., Fields, J., Zhao, C., Langer, J., Thimmulappa, R. K., Kensler, T. W., Yamamoto, M., Biswal, S., & Doré, S. (2007). Role of Nrf2 in protection against intracerebral hemorrhage injury in mice. *Free radical biology & medicine*, *43*(3), 408–414.

Wang, S., Lee, S. B., Pardue, C., Ramsingh, D., Waller, J., Gross, H., Nichols, F. T., 3rd, Hess, D. C., & Adams, R. J. (2003). Remote evaluation of acute ischemic stroke: reliability of National Institutes of Health Stroke Scale via telestroke. *Stroke*, *34*(10), e188–e191.

Ward, J., & Wood, C. (2000). Education and training of healthcare staff: the barriers to its success. *European journal of cancer care, 9*(2), 80-85.

Wartenberg, K. E., Stoll, A., Funk, A., Meyer, A., Schmidt, J. M., & Berrouschot, J. (2011). Infection after acute ischemic stroke: Risk factors, biomarkers, and outcome. *Stroke Research and Treatment, 2011*, 830614.

Weimar, C., Mieck, T., Buchthal, J., Ehrenfeld, C. E., Schmid, E., Diener, H. C., & Collaboration, G. S. S. (2005). Neurologic worsening during the acute phase of ischemic stroke. *Archives of Neurology, 62*(3), 393-397.

Wells-Pittman, J., & Gullicksrud, A. (2020). Standardizing the frequency of neurologic assessment after acute stroke. *Am J Nurs, 120*(3), 48-54.

Wennberg, J. E. (2002). Unwarranted variations in healthcare delivery: implications for academic medical centres. *BMJ (Clinical research ed.), 325*(7370), 961-964.

Wennberg, J. E. (2011). Time to tackle unwarranted variations in practice. *BMJ (Clinical research ed.)*, *342*, d1513.

Westhorp, G., Prins, E., Kusters, C. S. L., Hultink, M., Guijt, I. M., & Brouwers, J. H. A. M. (2011). *Realist Evaluation: an overview*. Retrieved June 20, 2022, from https://edepot.wur.nl/173918

Wijdicks, E. F. M., Bamlet, W. R., Maramattom, B. V., Manno, E. M., & McClelland, R. L. (2005). Validation of a new coma scale: The FOUR score. *Annals of Neurology, 58*(4), 585-593.

Williams, J. M., Jude, M. R., & Levi, C. R. (2013). Recombinant tissue plasminogen activator (rt-PA) utilisation by rural clinicians in acute ischaemic stroke: a survey of barriers and enablers. *Aust J Rural Health, 21*(5), 262-267.

Williams, L., Rycroft-Malone, J., & Burton, C. R. (2017). Bringing critical realism to nursing practice: Roy Bhaskar's contribution. *Nursing Philosophy, 18*(2), e12130.

Willis, G. B. (2005). *Cognitive interviewing*. Sage Publications Ltd, Retrieved March 31, 2023, from https://methods.sagepub.com/book/cognitive-interviewing.

Wilson, V., & McCormack, B. (2006). Critical realism as emancipatory action: the case for realistic evaluation in practice development. *Nurs Philos, 7*(1), 45-57.

Woodruff, T. M., Thundyil, J., Tang, S. C., Sobey, C. G., Taylor, S. M., & Arumugam, T. V. (2011). Pathophysiology, treatment, and animal and cellular models of human ischemic stroke. *Mol Neurodegener, 6*(1), 11.

World Health Organization (WHO). (1988). MONICA Project (monitoring trends and determinants in cardiovascular disease): a major international collaboration. *J Clin Epidemiol, 41*(2), 105-114.

World Health Organization (WHO). (2014). *WHO handbook for guideline development-2nd edition.* Retrieved June 3, 2022, from http://apps.who.int/iris/bitstream/handle/10665/145714/9789241548960_%20eng.pdf?sequence=1

Wu, T. C., Nguyen, C., Ankrom, C., Yang, J., Persse, D., Vahidy, F., Grotta, J. C., & Savitz, S. I. (2014). Prehospital utility of rapid stroke evaluation using in-ambulance telemedicine: a pilot feasibility study. *Stroke*, *45*(8), 2342–2347.

Wu, T. C., Parker, S. A., Jagolino, A., Yamal, J. M., Bowry, R., Thomas, A., Yu, A., & Grotta, J. C. (2017). Telemedicine can replace the neurologist on a mobile stroke unit. *Stroke*, *48*(2), 493–496.

Xu, M. Y. (2019). Poststroke seizure: optimising its management. *Stroke and Vascular Neurology, 4*(1), 48.

Yamamoto, H., Bogousslavsky, J., & van Melle, G. (1998). Different predictors of neurological worsening in different causes of stroke. *Archives of Neurology, 55*(4), 481-486.

Yanko, J. R., & Lang, S. (2013). Abstract WP360: Utilization of the National Institutes of Health Stroke Scale: Challenges and Interventions. *Stroke, 44* (suppl_1), AWP360-AWP360.

Yu, W. M., Abdul-Rahim, A. H., Cameron, A. C., Kõrv, J., Sevcik, P., Toni, D., Lees, K. R., & SITS Scientific Committee* (2020). The incidence and associated factors of early neurological deterioration after thrombolysis: Results from SITS Registry. *Stroke*, *51*(9), 2705–2714.

Zandieh, A., Kahaki, Z. Z., Sadeghian, H., Pourashraf, M., Parviz, S., Ghaffarpour, M., & Ghabaee, M. (2012). The underlying factor structure of National Institutes of Health Stroke scale: an exploratory factor analysis. *International Journal of Neuroscience, 122*(3), 140-144.

Zhao, X. J., Li, Q. X., Liu, T. J., Wang, D. L., An, Y. C., Zhang, J., Peng, Y. B., Chen, R. Y., Chang, L. S., Wang, Y., Zhang, L., Fan, H. Y., Wang, X. J., & Zheng, F. X. (2018). Predictive values of CSS and NIHSS in the prognosis of patients with acute cerebral infarction: A comparative analysis. *Medicine*, *97*(39), e12419.

Zimmerman, B. J., Lindberg, C., & Plsek, P. E. (1998). *Edgeware: Insights from Complexity Science for Health Care Leaders.* Dallas, TX: VHA Publishing.

Zink, E. K., Helzer, J. M., Geocadin, R. G., Mirski, M. A., & Ziai, W. C. (2012). Utility of serial four score measurements by nurses in a neurosciences critical care unit. *Neurocritical Care, 17*, S243.

Appendices

Appendix 1.1 Definitions of Early Neurological Deterioration (END)

Examples of definitions, including time period, of early neurological deterioration by paper and scale.

Paper	Scale	Definition	Time Period
Arenillas 2002	NIHSS	Increase > 4 points	48 hours
Birschel 2004	SSS	2 or more-point worsening in either Level of Consciousness, Arm, Leg or Eye Movement scores, and/or a 3 or more-point worsening in Speech	72 hours
Bugnicourt 2011	NIHSS	≥ 4 point increase	72 hours
Cui 2022	NIHSS	≥ 4 point increase or death	24 hrs
Davalos 1999	SSS	A decrease of 2 or more points in Level of Consciousness or Motor Power or a decrease of 3 or more points in Speech scores	24 hours
Flemming 1999	GCS	Decrease in sum score by 2 points	48 hours
	Other indicators	New neurological deficit or clinical signs of brain herniation	
Geng 2017	NIHSS	An increment change of at least one point in Motor Power or total score deterioration 2 points	Within the first week
Helleberg 2014	NIHSS	2 point increase	Baseline to 72 hours
	SSS	2 point decrease	
Kwan 2006	NIHSS	≥ 2points	Between admission and Day 5
Leira 2004	CSS	Decrease of ≥ 1 point	48 hours
Maramattom 2004	GCS	Decrease of ≥ 2 points	Not clearly specified
Mayer 1994	GCS	Decrease of ≥ 2 points	24 hours and beyond
	Other	Increase of 1 or more point in the Stroke Data Bank (SDB) weakness score	
		New deficit, unrelated to medical or surgical complications	
Miyamoto 2017	NIHSS	≥ 4 point increase	1 week
Ovesen 2015	GCS	≥ 2 point decrease	24 hours
	SIP	≥ 4 points change	
Roden-Jullig 2003	SSS	≥ 2 points	5 days
Sorimachi 2010	NIHSS	Level of Consciousness and Motor Skills	24 hours
Sun 2012	GCS	≥ 3 points decrease or death	72 hours
Wei 2020	NIHSS	≥ 2 point increase	72 hours
Weimar 2005	NIHSS	≥ 1 point increase	72 hours
Wahlgren 2007	NIHSS	≥ 4 points change	Baseline to 24 hours

Key
CNS= Canadian Neurological Scale
GCS= Glasgow Coma Scale
NIHSS= National Institutes for Health Stroke Scale
SIP= Stroke in Progression Scale a shortened version of the SSS
SSS= Scandinavian Stroke Scale

Appendix 2.1 Study details of example END literature

Table showing study aims, numbers included, and stroke type of participants of studies included in Appendix 1.1 (examples of definitions of early neurological deterioration).

Paper	Aim	No.	Sampling	Stroke Type
Arenillas 2002	Assess the value of ultra-early MRI in the prediction of END after stroke	38	Consecutive (strict inclusion criteria)	MCA or ICA
Birschel 2004	Agreement on definitions of early deterioration episode (EDE) and stroke progressions (SP) and validation in an observational study	581	Consecutive	All
Bugincourt 2011	To test the hypothesis that biological aspirin non-responder status (ANRS) helps predict END	85	Consecutive	Acute ischaemic
Cui 2022	Secondary analysis of INTRECIS (Intravenous Thrombolysis Registry for Chinese Ischemic Stroke Within 4.5h Onset)	**1194** (ACS 942, PCS252)	ND-secondary analysis	Anterior circulation stroke (ACS) from posterior circulation stroke (PCS)
Davalos 1999	Secondary analysis of European Cooperative Acute Stroke Study (ECASS) I data to identify predictors of early and late progression after stroke to look at rates of END between two stroke types.	**615**	ND-secondary analysis	Ischaemic stroke (eligible for intravenous thrombolysis)
Flemming 1999	Study the clinical course and determine predictors of deterioration	**61**	Retrospective data analysis	Lobar haemorrhage
Geng 2017	Explore the association between END and long-term outcomes in patients	1064	Consecutive	Ischaemic stroke (first-ever)
Helleberg 2016	study outcome after END and transitory deterioration (TD).	368	Screened from larger protocol	Ischaemic stroke
Kwan 2006	Explore the frequency, clinical characteristics, and consequences of END during the acute recovery period	188	Consecutive	All stroke patients
Leira 2004	Identify potential predictors of and factors associated with END.	266	Selected	Intracerebral haemorrhage (ICH)
Maramattom 2004	Identify level of and features of those patients who deteriorated after coMCAI.	24	Selected consecutive	Complete middle cerebral artery infarction (coMCAI)

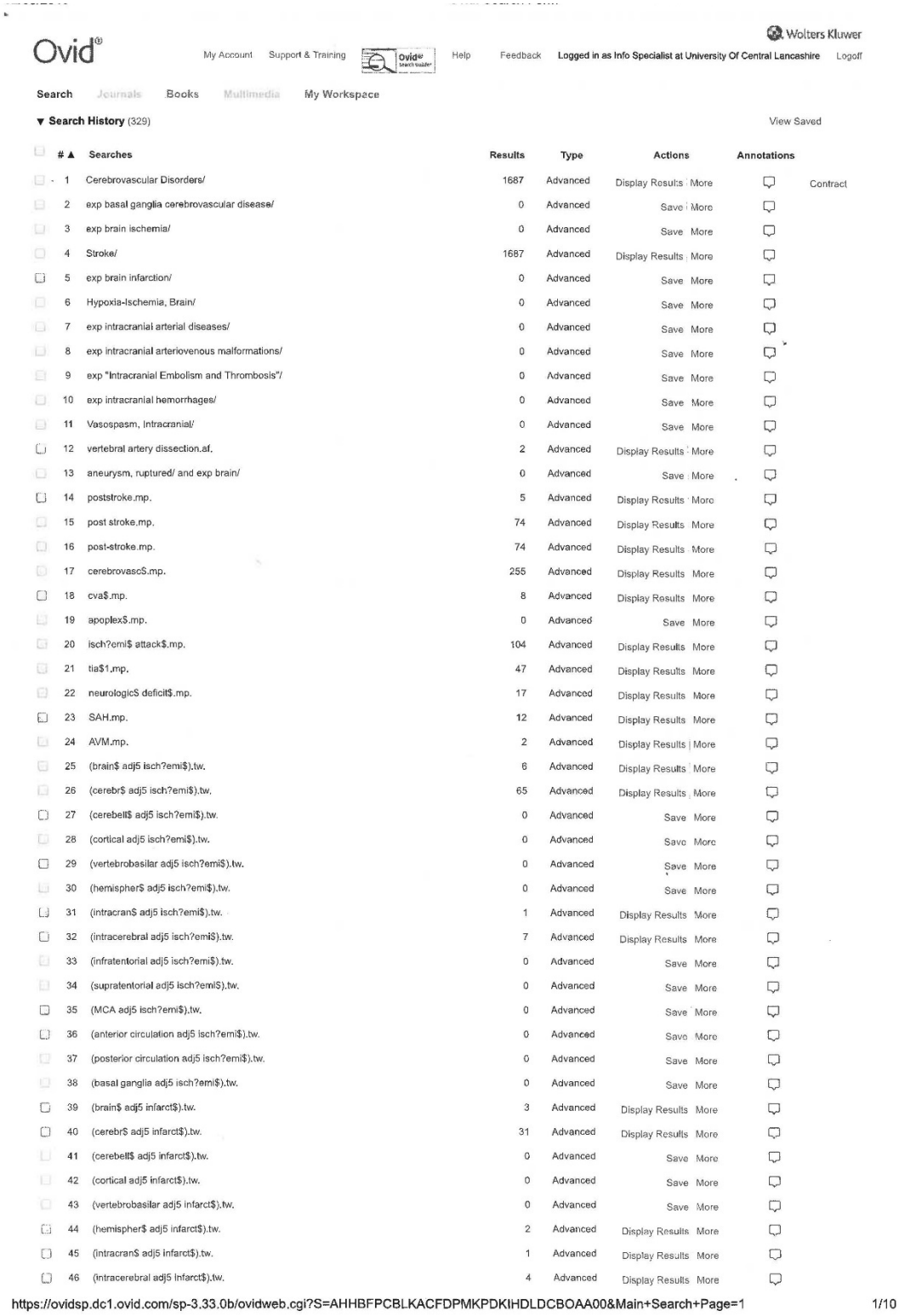
Mayer 1994	Determine the frequency, time course, and predictors of ND	**46**	Selected	ICH (non-comatose)
Miyamoto 2017	Analysis of the usefulness of the WORSEN Score for predicting deterioration during the week after stroke onset	**478** then 456	Retrospective data analysis	Ischaemic Stroke
Ovesen 2015	Establishing predictors of early and late neurological deterioration and the impact of neurological stability during the first week on long-term prognosis	**300**	Retrospective data analysis	ICH
Roden-Jullig 2003	Evaluate the efficacy of aspirin for prevention of stroke progression	441 (220 aspirin, 221 placebo)	Selected	Ischaemic Stroke (not complete paralysis)
Sorimachi 2010	Report the frequency and causes of neurological change within 24 hours	184	Selected	ICH
Sun 2012	Determine the factors associated with END	**83**	Retrospective data analysis-consecutive	ICH
Wahlgren 2007	Assess the safety and efficacy of intravenous alteplase as thrombolytic therapy in the first 3 hours on onset	6483	Cohort of existing register	Ischaemic Stroke
Weimar 2005	Identify patients at risk for neurologic worsening	1964	Consecutive	Ischaemic Stroke

Numbers in **bold** represent retrospective data collection

WORSEN Score derived from the following factors: wrong (poor) blood sugar control (W), old myocardial infarction (O), radiological findings (R), infarct size (S), elevated low-density lipoprotein cholesterol (E), and neurological findings (N).

Appendix 4.1 Scoping search example

This search strategy was to obtain relevant literature to identify scales available or used for neurological assessment and monitoring after acute stroke within the HMIC database.



Ovid® My Account Support & Training Ovid Search Toolbar Help Feedback Logged in as Info Specialist at University Of Central Lancashire Logoff

Wolters Kluwer

Search Journals Books Multimedia My Workspace

▼ Search History (329) View Saved

# ▲	Searches	Results	Type	Actions	Annotations
1	Cerebrovascular Disorders/	1687	Advanced	Display Results More	Contract
2	exp basal ganglia cerebrovascular disease/	0	Advanced	Save More	
3	exp brain ischemia/	0	Advanced	Save More	
4	Stroke/	1687	Advanced	Display Results More	
5	exp brain infarction/	0	Advanced	Save More	
6	Hypoxia-Ischemia, Brain/	0	Advanced	Save More	
7	exp intracranial arterial diseases/	0	Advanced	Save More	
8	exp intracranial arteriovenous malformations/	0	Advanced	Save More	
9	exp "Intracranial Embolism and Thrombosis"/	0	Advanced	Save More	
10	exp intracranial hemorrhages/	0	Advanced	Save More	
11	Vasospasm, Intracranial/	0	Advanced	Save More	
12	vertebral artery dissection.af.	2	Advanced	Display Results More	
13	aneurysm, ruptured/ and exp brain/	0	Advanced	Save More	
14	poststroke.mp.	5	Advanced	Display Results More	
15	post stroke.mp.	74	Advanced	Display Results More	
16	post-stroke.mp.	74	Advanced	Display Results More	
17	cerebrovascS.mp.	255	Advanced	Display Results More	
18	cva$.mp.	8	Advanced	Display Results More	
19	apoplexS.mp.	0	Advanced	Save More	
20	isch?emi$ attack$.mp.	104	Advanced	Display Results More	
21	tia$1.mp.	47	Advanced	Display Results More	
22	neurologicS deficit$.mp.	17	Advanced	Display Results More	
23	SAH.mp.	12	Advanced	Display Results More	
24	AVM.mp.	2	Advanced	Display Results More	
25	(brain$ adj5 isch?emi$).tw.	6	Advanced	Display Results More	
26	(cerebr$ adj5 isch?emi$).tw.	65	Advanced	Display Results More	
27	(cerebell$ adj5 isch?emi$).tw.	0	Advanced	Save More	
28	(cortical adj5 isch?emi$).tw.	0	Advanced	Save More	
29	(vertebrobasilar adj5 isch?emi$).tw.	0	Advanced	Save More	
30	(hemispher$ adj5 isch?emi$).tw.	0	Advanced	Save More	
31	(intracran$ adj5 isch?emi$).tw.	1	Advanced	Display Results More	
32	(intracerebral adj5 isch?emi$).tw.	7	Advanced	Display Results More	
33	(infratentorial adj5 isch?emi$).tw.	0	Advanced	Save More	
34	(supratentorial adj5 isch?emiS).tw.	0	Advanced	Save More	
35	(MCA adj5 isch?emi$).tw.	0	Advanced	Save More	
36	(anterior circulation adj5 isch?emi$).tw.	0	Advanced	Save More	
37	(posterior circulation adj5 isch?emi$).tw.	0	Advanced	Save More	
38	(basal ganglia adj5 isch?emi$).tw.	0	Advanced	Save More	
39	(brain$ adj5 infarct$).tw.	3	Advanced	Display Results More	
40	(cerebr$ adj5 infarct$).tw.	31	Advanced	Display Results More	
41	(cerebell$ adj5 infarct$).tw.	0	Advanced	Save More	
42	(cortical adj5 infarct$).tw.	0	Advanced	Save More	
43	(vertebrobasilar adj5 infarct$).tw.	0	Advanced	Save More	
44	(hemispher$ adj5 infarct$).tw.	2	Advanced	Display Results More	
45	(intracran$ adj5 infarct$).tw.	1	Advanced	Display Results More	
46	(intracerebral adj5 infarct$).tw.	4	Advanced	Display Results More	

https://ovidsp.dc1.ovid.com/sp-3.33.0b/ovidweb.cgi?S=AHHBFPCBLKACFDPMKPDKIHDLDCBOAA00&Main+Search+Page=1 1/10

47	(infratentorial adj5 infarct$).tw.	0	Advanced	Save	More	
48	(supratentorial adj5 infarct$).tw.	2	Advanced	Display Results	More	
49	(MCA adj5 infarct$).tw.	0	Advanced	Save	More	
50	(anterior circulation adj5 infarct$).tw.	1	Advanced	Display Results	More	
51	(posterior circulation adj5 infarct$).tw.	0	Advanced	Save	More	
52	(basal ganglia adj5 infarct$).tw.	0	Advanced	Save	More	
53	(brain$ adj5 thrombo$).tw.	3	Advanced	Display Results	More	
54	(cerebr$ adj5 thrombo$).tw.	7	Advanced	Display Results	More	
55	(cerebell$ adj5 thrombo$).tw.	0	Advanced	Save	More	
56	(cortical adj5 thrombo$).tw.	0	Advanced	Save	More	
57	(vertebrobasilar adj5 thrombo$).tw.	0	Advanced	Save	More	
58	(hemispher$ adj5 thrombo$).tw.	0	Advanced	Save	More	
59	(intracran$ adj5 thrombo$).tw.	1	Advanced	Display Results	More	
60	(intracerebral adj5 thrombo$).tw.	0	Advanced	Save	More	
61	(infratentorial adj5 thrombo$).tw.	0	Advanced	Save	More	
62	(supratentorial adj5 thrombo$).tw.	0	Advanced	Save	More	
63	(MCA adj5 thrombo$).tw.	0	Advanced	Save	More	
64	(anterior circulation adj5 thrombo$).tw.	0	Advanced	Save	More	
65	(posterior circulation adj5 thrombo$).tw.	0	Advanced	Save	More	
66	(basal ganglia adj5 thrombo$).tw.	0	Advanced	Save	More	
67	(brain$ adj5 emboli$).tw.	0	Advanced	Save	More	
68	(cerebr$ adj5 emboli$).tw.	6	Advanced	Display Results	More	
69	(cerebell$ adj5 emboli$).tw.	0	Advanced	Save	More	
70	(cortical adj5 emboli$).tw.	0	Advanced	Save	More	
71	(vertebrobasilar adj5 emboli$).tw.	0	Advanced	Save	More	
72	(hemispher$ adj5 emboli$).tw.	0	Advanced	Save	More	
73	(intracran$ adj5 emboli$).tw.	1	Advanced	Display Results	More	
74	(intracerebral adj5 emboli$).tw.	0	Advanced	Save	More	
75	(infratentorial adj5 emboli$).tw.	0	Advanced	Save	More	
76	(supratentorial adj5 emboli$).tw.	0	Advanced	Save	More	
77	(MCA adj5 emboli$).tw.	0	Advanced	Save	More	
78	(anterior circulation adj5 emboli$).tw.	0	Advanced	Save	More	
79	(posterior circulation adj5 emboli$).tw.	0	Advanced	Save	More	
80	(basal ganglia adj5 emboli$).tw.	0	Advanced	Save	More	
81	(brain$ adj5 occlus$).tw.	0	Advanced	Save	More	
82	(cerebr$ adj5 occlus$).tw.	0	Advanced	Save	More	
83	(cerebell$ adj5 occlus$).tw.	0	Advanced	Save	More	
84	(cortical adj5 occlus$).tw.	0	Advanced	Save	More	
85	(vertebrobasilar adj5 occlus$).tw.	0	Advanced	Save	More	
86	(hemispher$ adj5 occlus$).tw.	0	Advanced	Save	More	
87	(intracran$ adj5 occlus$).tw.	0	Advanced	Save	More	
88	(intracerebral adj5 occlus$).tw.	0	Advanced	Save	More	
89	(infratentorial adj5 occlus$).tw.	0	Advanced	Save	More	
90	(supratentorial adj5 occlus$).tw.	0	Advanced	Save	More	
91	(MCA adj5 occlus$).tw.	0	Advanced	Save	More	
92	(anterior circulation adj5 occlus$).tw.	0	Advanced	Save	More	
93	(posterior circulation adj5 occlus$).tw.	0	Advanced	Save	More	
94	(basal ganglia adj5 occlus$).tw.	0	Advanced	Save	More	
95	(brain$ adj5 hypox$).tw.	1	Advanced	Display Results	More	
96	(cerebr$ adj5 hypox$).tw.	1	Advanced	Display Results	More	
97	(cerebell$ adj5 hypox$).tw.	0	Advanced	Save	More	
98	(cortical adj5 hypox$).tw.	0	Advanced	Save	More	

99	(vertebrobasilar adj5 hypox$).tw.	0	Advanced	Save	More	
100	(hemispher$ adj5 hypox$).tw.	0	Advanced	Save	More	
101	(intracran$ adj5 hypox$).tw.	0	Advanced	Save	More	
102	(intracerebral adj5 hypox$).tw.	0	Advanced	Save	More	
103	(infratentorial adj5 hypox$).tw.	0	Advanced	Save	More	
104	(supratentorial adj5 hypox$).tw.	0	Advanced	Save	More	
105	(MCA adj5 hypox$).tw.	0	Advanced	Save	More	
106	(anterior circulation adj5 hypox$).tw.	0	Advanced	Save	More	
107	(posterior circulation adj5 hypox$).tw.	0	Advanced	Save	More	
108	(basal ganglia adj5 hypox$).tw.	0	Advanced	Save	More	
109	(brain$ adj5 vasospasm).tw.	0	Advanced	Save	More	
110	(cerebr$ adj5 vasospasm).tw.	1	Advanced	Display Results	More	
111	(cerebell$ adj5 vasospasm).tw.	0	Advanced	Save	More	
112	(cortical adj5 vasospasm).tw.	0	Advanced	Save	More	
113	(vertebrobasilar adj5 vasospasm).tw.	0	Advanced	Save	More	
114	(hemispher$ adj5 vasospasm).tw.	0	Advanced	Save	More	
115	(intracran$ adj5 vasospasm).tw.	0	Advanced	Save	More	
116	(intracerebral adj5 vasospasm).tw.	0	Advanced	Save	More	
117	(infratentorial adj5 vasospasm).tw.	0	Advanced	Save	More	
118	(supratentorial adj5 vasospasm).tw.	0	Advanced	Save	More	
119	(MCA adj5 vasospasm).tw.	0	Advanced	Save	More	
120	(anterior circulation adj5 vasospasm).tw.	0	Advanced	Save	More	
121	(posterior circulation adj5 vasospasm).tw.	0	Advanced	Save	More	
122	(basal ganglia adj5 vasospasm).tw.	0	Advanced	Save	More	
123	(brain$ adj5 obstruction).tw.	0	Advanced	Save	More	
124	(cerebr$ adj5 obstruction).tw.	0	Advanced	Save	More	
125	(cerebell$ adj5 obstruction).tw.	0	Advanced	Save	More	
126	(cortical adj5 obstruction).tw.	0	Advanced	Save	More	
127	(vertebrobasilar adj5 obstruction).tw.	0	Advanced	Save	More	
128	(hemispher$ adj5 obstruction).tw.	0	Advanced	Save	More	
129	(intracran$ adj5 obstruction).tw.	0	Advanced	Save	More	
130	(intracerebral adj5 obstruction).tw.	0	Advanced	Save	More	
131	(infratentorial adj5 obstruction).tw.	0	Advanced	Save	More	
132	(supratentorial adj5 obstruction).tw.	0	Advanced	Save	More	
133	(MCA adj5 obstruction).tw.	0	Advanced	Save	More	
134	(anterior circulation adj5 obstruction).tw.	0	Advanced	Save	More	
135	(posterior circulation adj5 obstruction).tw.	0	Advanced	Save	More	
136	(basal ganglia adj5 obstruction).tw.	0	Advanced	Save	More	
137	(brain$ adj5 vasculopathy).tw.	0	Advanced	Save	More	
138	(cerebr$ adj5 vasculopathy).tw.	0	Advanced	Save	More	
139	(cerebell$ adj5 vasculopathy).tw.	0	Advanced	Save	More	
140	(cortical adj5 vasculopathy).tw.	0	Advanced	Save	More	
141	(vertebrobasilar adj5 vasculopathy).tw.	0	Advanced	Save	More	
142	(hemispher$ adj5 vasculopathy).tw.	0	Advanced	Save	More	
143	(intracran$ adj5 vasculopathy).tw.	0	Advanced	Save	More	
144	(intracerebral adj5 vasculopathy).tw.	0	Advanced	Save	More	
145	(infratentorial adj5 vasculopathy).tw.	0	Advanced	Save	More	
146	(supratentorial adj5 vasculopathy).tw.	0	Advanced	Save	More	
147	(MCA adj5 vasculopathy).tw.	0	Advanced	Save	More	
148	(anterior circulation adj5 vasculopathy).tw.	0	Advanced	Save	More	
149	(posterior circulation adj5 vasculopathy).tw.	0	Advanced	Save	More	
150	(basal ganglia adj5 vasculopathy).tw.	0	Advanced	Save	More	

151	(lacunar adj5 infarct$).tw.	0	Advanced	Save | More	
152	(brain$ adj5 haemorrhage$).tw.	5	Advanced	Display Results | More	
153	(cerebr$ adj5 haemorrhage$).tw.	16	Advanced	Display Results | More	
154	(cerebelli$ adj5 haemorrhage$).tw.	0	Advanced	Save | More	
155	(intracerebral adj5 haemorrhage$).tw.	18	Advanced	Display Results | More	
156	(intracran$ adj5 haemorrhage$).tw.	16	Advanced	Display Results | More	
157	(parenchymal adj5 haemorrhage$).tw.	2	Advanced	Display Results | More	
158	(intraventricular adj5 haemorrhage$).tw.	9	Advanced	Display Results | More	
159	(infratentorial adj5 haemorrhage$).tw.	0	Advanced	Save | More	
160	(supratentorial adj5 haemorrhage$).tw.	0	Advanced	Save | More	
161	(basal ganglii$ adj5 haemorrhage$).tw.	0	Advanced	Save | More	
162	(subarachnoid adj5 haemorrhage$).tw.	19	Advanced	Display Results | More	
163	(putaminal adj5 haemorrhage$).tw.	0	Advanced	Save | More	
164	(putamen adj5 haemorrhage$).tw.	0	Advanced	Save | More	
165	(posterior fossa adj5 haemorrhage$).tw.	0	Advanced	Save | More	
166	(brain$ adj5 hemorrhage$).tw.	0	Advanced	Save | More	
167	(cerebr$ adj5 hemorrhage$).tw.	0	Advanced	Save | More	
168	(cerebell$ adj5 hemorrhage$).tw.	0	Advanced	Save | More	
169	(intracerebral adj5 hemorrhage$).tw.	0	Advanced	Save | More	
170	(intracran$ adj5 hemorrhage$).tw.	1	Advanced	Display Results | More	
171	(parenchymal adj5 hemorrhage$).tw.	0	Advanced	Save | More	
172	(intraventricular adj5 hemorrhage$).tw.	0	Advanced	Save | More	
173	(infratentorial adj5 hemorrhage$).tw.	0	Advanced	Save | More	
174	(supratentorial adj5 hemorrhage$).tw.	0	Advanced	Save | More	
175	(basal ganglii$ adj5 hemorrhage$).tw.	0	Advanced	Save | More	
176	(subarachnoid adj5 hemorrhage$).tw.	2	Advanced	Display Results | More	
177	(putaminal adj5 hemorrhage$).tw.	0	Advanced	Save | More	
178	(putamen adj5 hemorrhage$).tw.	0	Advanced	Save | More	
179	(posterior fossa adj5 hemorrhage$).tw.	0	Advanced	Save | More	
180	(brain$ adj5 haematoma$).tw.	0	Advanced	Save | More	
181	(cerebr$ adj5 haematoma$).tw.	0	Advanced	Save | More	
182	(cerebell$ adj5 haematoma$).tw.	0	Advanced	Save | More	
183	(intracerebral adj5 haematoma$).tw.	0	Advanced	Save | More	
184	(intracran$ adj5 haematoma$).tw.	7	Advanced	Display Results | More	
185	(parenchymal adj5 haematoma$).tw.	0	Advanced	Save | More	
186	(intraventricular adj5 haematoma$).tw.	0	Advanced	Save | More	
187	(infratentorial adj5 haematoma$).tw.	0	Advanced	Save | More	
188	(supratentorial adj5 haematoma$).tw.	0	Advanced	Save | More	
189	(basal ganglii$ adj5 haematoma$).tw.	0	Advanced	Save | More	
190	(subarachnoid adj5 haematoma$).tw.	0	Advanced	Save | More	
191	(putaminal adj5 haematoma$).tw.	0	Advanced	Save | More	
192	(putamen adj5 haematoma$).tw.	0	Advanced	Save | More	
193	(posterior fossa adj5 haematoma$).tw.	0	Advanced	Save | More	
194	(brain$ adj5 hematoma$).tw.	0	Advanced	Save | More	
195	(cerebr$ adj5 hematoma$).tw.	0	Advanced	Save | More	
196	(cerebell$ adj5 hematoma$).tw.	0	Advanced	Save | More	
197	(intracerebral adj5 hematoma$).tw.	0	Advanced	Save | More	
198	(intracran$ adj5 hematoma$).tw.	0	Advanced	Save | More	
199	(parenchymal adj5 hematoma$).tw.	0	Advanced	Save | More	
200	(intraventricular adj5 hematoma$).tw.	0	Advanced	Save | More	
201	(infratentorial adj5 hematoma$).tw.	0	Advanced	Save | More	
202	(supratentorial adj5 hematoma$).tw.	0	Advanced	Save | More	

203	(basal gangli$ adj5 hematoma$).tw.	0	Advanced	Save	More	
204	(subarachnoid adj5 hematomaS).tw.	0	Advanced	Save	More	
205	(putaminal adj5 hematoma$).tw.	0	Advanced	Save	More	
206	(putaminal adj5 hematoma$).tw.	0	Advanced	Save	More	
207	(posterior fossa adj5 hematoma$).tw.	0	Advanced	Save	More	
208	(brain$ adj5 bleed$).tw.	1	Advanced	Display Results	More	
209	(cerebr$ adj5 bleed$).tw.	5	Advanced	Display Results	More	
210	(cerebell$ adj5 bleed$).tw.	0	Advanced	Save	More	
211	(intracerebral adj5 bleed$).tw.	1	Advanced	Display Results	More	
212	(intracranS adj5 bleed$).tw.	4	Advanced	Display Results	More	
213	(parenchymal adj5 bleed$).tw.	0	Advanced	Save	More	
214	(intraventricular adj5 bleedS).tw.	0	Advanced	Save	More	
215	(infratentorial adj5 bleed$).tw.	0	Advanced	Save	More	
216	(supratentorial adj5 bleed$).tw.	0	Advanced	Save	More	
217	(basal gangliS adj5 bleed$).tw.	0	Advanced	Save	More	
218	(subarachnoid adj5 bleed$).tw.	1	Advanced	Display Results	More	
219	(putaminal adj5 bleed$).tw.	0	Advanced	Save	More	
220	(putamen adj5 bleed$).tw.	0	Advanced	Save	More	
221	(posterior fossa adj5 bleed$).tw.	0	Advanced	Save	More	
222	(brain adj5 aneurysm$).tw.	1	Advanced	Display Results	More	
223	(cerebral adj5 aneurysm$).tw.	1	Advanced	Display Results	More	
224	(intracranial adj5 aneurysm$).tw.	4	Advanced	Display Results	More	
225	(communicating adj5 aneurysm$).tw.	0	Advanced	Save	More	
226	(giant adj5 aneurysm$).tw.	0	Advanced	Save	More	
227	(basilar adj5 aneurysm$).tw.	0	Advanced	Save	More	
228	(vertebral artery adj5 aneurysm$).tw.	0	Advanced	Save	More	
229	(berry adj5 aneurysm$).tw.	0	Advanced	Save	More	
230	(saccular adj5 aneurysm$).tw.	1	Advanced	Display Results	More	
231	(ruptured adj5 aneurysm$).tw.	11	Advanced	Display Results	More	
232	vertebral artery dissection.tw.	2	Advanced	Display Results	More	
233	cerebral art$ disease$.tw.	0	Advanced	Save	More	
234	(brain adj5 (vascular adj5 disease$)).tw.	0	Advanced	Save	More	
235	(brain adj5 (vascular adj5 disorder)).tw.	0	Advanced	Save	More	
236	(brain adj5 (vascular adj5 accident)).tw.	0	Advanced	Save	More	
237	(brain adj5 (vascular adj5 injur$)).tw.	0	Advanced	Save	More	
238	(brain adj5 (vascular adj5 trauma$)).tw.	0	Advanced	Save	More	
239	(brain adj5 (vascular adj5 insult)).tw.	0	Advanced	Save	More	
240	(brain adj5 (vascular adj5 event)).tw.	0	Advanced	Save	More	
241	(intracranial adj5 (vascular adj5 disease$)).tw.	0	Advanced	Save	More	
242	(intracranial adj5 (vascular adj5 disorder)).tw.	0	Advanced	Save	More	
243	(intracranial adj5 (vascular adj5 accident)).tw.	0	Advanced	Save	More	
244	(intracranial adj5 (vascular adj5 injur$)).tw.	0	Advanced	Save	More	
245	(intracranial adj5 (vascular adj5 trauma$)).tw.	0	Advanced	Save	More	
246	(intracranial adj5 (vascular adj5 insult)).tw.	0	Advanced	Save	More	
247	(intracranial adj5 (vascular adj5 event)).tw.	0	Advanced	Save	More	
248	(basal ganglia adj5 (vascular adj5 disease$)).tw.	0	Advanced	Save	More	
249	(basal ganglia adj5 (vascular adj5 disorder)).tw.	0	Advanced	Save	More	
250	(basal ganglia adj5 (vascular adj5 accident)).tw.	0	Advanced	Save	More	
251	(basal ganglia adj5 (vascular adj5 injur$)).tw.	0	Advanced	Save	More	
252	(basal ganglia adj5 (vascular adj5 traumaS)).tw.	0	Advanced	Save	More	
253	(basal ganglia adj5 (vascular adj5 insult)).tw.	0	Advanced	Save	More	
254	(basal ganglia adj5 (vascular adj5 event)).tw.	0	Advanced	Save	More	

255	(lenticulostriate adj5 (vascular adj5 disease$)).tw.	0	Advanced	Save	More	
256	(lenticulostriate adj5 (vascular adj5 disorder)).tw.	0	Advanced	Save	More	
257	(lenticulostriate adj5 (vascular adj5 accident)).tw.	0	Advanced	Save	More	
258	(lenticulostriate adj5 (vascular adj5 injur$)).tw.	0	Advanced	Save	More	
259	(lenticulostriate adj5 (vascular adj5 traumaS)).tw.	0	Advanced	Save	More	
260	(lenticulostriate adj5 (vascular adj5 insult)).tw.	0	Advanced	Save	More	
261	(lenticulostriate adj5 (vascular adj5 event)).tw.	0	Advanced	Save	More	
262	(isch?emic adj5 event).tw.	5	Advanced	Display Results	More	
263	(isch?emic adj5 events).tw.	29	Advanced	Display Results	More	
264	(isch?emic adj5 insult).tw.	0	Advanced	Save	More	
265	(isch?emic adj5 attack$).tw.	106	Advanced	Display Results	More	
266	(apoplectic adj5 event).tw.	0	Advanced	Save	More	
267	(apoplectic adj5 events).tw.	0	Advanced	Save	More	
268	(apoplectic adj5 insult).tw.	0	Advanced	Save	More	
269	(apoplectic adj5 attack$).tw.	0	Advanced	Save	More	
270	(cerebral vein adj5 thrombo$).tw.	0	Advanced	Save	More	
271	(cerebral venous adj5 thrombo$).tw.	1	Advanced	Display Results	More	
272	(sinus adj5 thrombo$).tw.	0	Advanced	Save	More	
273	(sagittal adj5 thrombo$).tw.	0	Advanced	Save	More	
274	CVDST.tw.	0	Advanced	Save	More	
275	CVT.tw.	2	Advanced	Display Results	More	
276	(intracranial adj5 stenosis).tw.	0	Advanced	Save	More	
277	(intracranial adj5 isch?emia).tw.	0	Advanced	Save	More	
278	(intracranial adj5 insufficiency).tw.	0	Advanced	Save	More	
279	(intracranial adj5 arteriosclero$).tw.	0	Advanced	Save	More	
280	(intracranial adj5 atherosclero$).tw.	0	Advanced	Save	More	
281	(intracranial adj5 occlus$).tw.	0	Advanced	Save	More	
282	(cerebral art$ adj5 stenosis).tw.	0	Advanced	Save	More	
283	(cerebral art$ adj5 isch?emia).tw.	0	Advanced	Save	More	
284	(cerebral art$ adj5 insufficiency).tw.	0	Advanced	Save	More	
285	(cerebral art$ adj5 arteriosclero$).tw.	0	Advanced	Save	More	
286	(cerebral art$ adj5 atherosclero$).tw.	0	Advanced	Save	More	
287	(cerebral art$ adj5 occlus$).tw.	0	Advanced	Save	More	
288	(basilar art$ adj5 stenosis).tw.	0	Advanced	Save	More	
289	(basilar art$ adj5 isch?emia).tw.	0	Advanced	Save	More	
290	(basilar art$ adj5 insufficiency).tw.	0	Advanced	Save	More	
291	(basilar art$ adj5 arteriosclero$).tw.	0	Advanced	Save	More	
292	(basilar art$ adj5 atherosclero$).tw.	0	Advanced	Save	More	
293	(basilar art$ adj5 occlus$).tw.	0	Advanced	Save	More	
294	(vertebral art$ adj5 stenosis).tw.	0	Advanced	Save	More	
295	(vertebral art$ adj5 isch?emia).tw.	0	Advanced	Save	More	
296	(vertebral art$ adj5 insufficiency).tw.	0	Advanced	Save	More	
297	(vertebral art$ adj5 arteriosclero$).tw.	0	Advanced	Save	More	
298	(vertebral art$ adj5 atherosclero$).tw.	0	Advanced	Save	More	
299	(vertebral art$ adj5 occlus$).tw.	0	Advanced	Save	More	
300	(vertebrobasilar adj5 stenosis).tw.	0	Advanced	Save	More	
301	(vertebrobasilar adj5 isch?emia).tw.	0	Advanced	Save	More	
302	(vertebrobasilar adj5 insufficiency).tw.	2	Advanced	Display Results	More	
303	(vertebrobasilar adj5 arteriosclero$).tw.	0	Advanced	Save	More	
304	(vertebrobasilar adj5 atherosclero$).tw.	0	Advanced	Save	More	
305	(vertebrobasilar adj5 occlus$).tw.	0	Advanced	Save	More	
306	(vertebral basilar adj5 stenosis).tw.	0	Advanced	Save	More	

☐	307	(vertebral basilar adj5 isch?emia).tw.	0	Advanced	Save	More	
☐	308	(vertebral basilar adj5 insufficiency).tw.	0	Advanced	Save	More	
☐	309	(vertebral basilar adj5 arteriosclero$).tw.	0	Advanced	Save	More	
☐	310	(vertebral basilar adj5 atherosclero$).tw.	0	Advanced	Save	More	
☐	311	(vertebral basilar adj5 occlus$).tw.	0	Advanced	Save	More	
☐	312	(venous adj5 malformation$).tw.	0	Advanced	Save	More	
☐	313	(arteriovenous adj5 malformation$).tw.	0	Advanced	Save	More	
☐	314	(brain vasc$ adj5 malformation$).tw.	0	Advanced	Save	More	
☐	315	(brain adj5 angioma$).tw.	0	Advanced	Save	More	
☐	316	(brain adj5 hemangioma$).tw.	0	Advanced	Save	More	
☐	317	(brain adj5 haemangioma$).tw.	0	Advanced	Save	More	
☐	318	(cerebral adj5 angioma$).tw.	0	Advanced	Save	More	
☐	319	(cerebral adj5 hemangioma$).tw.	0	Advanced	Save	More	
☐	320	(cerebral adj5 haemangioma$).tw.	0	Advanced	Save	More	
☐	321	carotid$.tw.	93	Advanced	Display Results	More	
☐	322	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97 or 98 or 99 or 100 or 101 or 102 or 103 or 104 or 105 or 106 or 107 or 108 or 109 or 110 or 111 or 112 or 113 or 114 or 115 or 116 or 117 or 118 or 119 or 120 or 121 or 122 or 123 or 124 or 125 or 126 or 127 or 128 or 129 or 130 or 131 or 132 or 133 or 134 or 135 or 136 or 137 or 138 or 139 or 140 or 141 or 142 or 143 or 144 or 145 or 146 or 147 or 148 or 149 or 150 or 151 or 152 or 153 or 154 or 155 or 156 or 157 or 158 or 159 or 160 or 161 or 162 or 163 or 164 or 165 or 166 or 167 or 168 or 169 or 170 or 171 or 172 or 173 or 174 or 175 or 176 or 177 or 178 or 179 or 180 or 181 or 182 or 183 or 184 or 185 or 186 or 187 or 188 or 189 or 190 or 191 or 192 or 193 or 194 or 195 or 196 or 197 or 198 or 199 or 200 or 201 or 202 or 203 or 204 or 205 or 206 or 207 or 208 or 209 or 210 or 211 or 212 or 213 or 214 or 215 or 216 or 217 or 218 or 219 or 220 or 221 or 222 or 223 or 224 or 225 or 226 or 227 or 228 or 229 or 230 or 231 or 232 or 233 or 234 or 235 or 236 or 237 or 238 or 239 or 240 or 241 or 242 or 243 or 244 or 245 or 246 or 247 or 248 or 249 or 250 or 251 or 252 or 253 or 254 or 255 or 256 or 257 or 258 or 259 or 260 or 261 or 262 or 263 or 264 or 265 or 266 or 267 or 268 or 269 or 270 or 271 or 272 or 273 or 274 or 275 or 276 or 277 or 278 or 279 or 280 or 281 or 282 or 283 or 284 or 285 or 286 or 287 or 288 or 289 or 290 or 291 or 292 or 293 or 294 or 295 or 296 or 297 or 298 or 299 or 300 or 301 or 302 or 303 or 304 or 305 or 306 or 307 or 308 or 309 or 310 or 311 or 312 or 313 or 314 or 315 or 316 or 317 or 318 or 319 or 320 or 321	2102	Advanced	Display Results	More	
☐	323	neurologic examination.mp. or Neurologic Examination/	8	Advanced	Display Results	More	
☐	324	neurologic deterioration.mp.	1	Advanced	Display Results	More	
☐	325	deterio*.mp.	909	Advanced	Display Results	More	
☐	326	323 or 324 or 325	917	Advanced	Display Results	More	
☐	327	322 and 326	12	Advanced	Display Results	More	
☐	328	limit 327 to humans [Limit not valid; records were retained]	12	Advanced	Display Results	More	
☐	329	limit 328 to last 10 years	3	Advanced	Display Results	More	

Save Remove **Combine with:** AND OR

Save All Edit Create RSS View Saved

▼ **Search History** (41) View Saved

	# ▲	Searches	Results	Type	Actions	Annotations
	1	(instrumentation or methods).fs.	3902614	Advanced	Display Results More	
	2	(Validation Studies or Comparative Study).pt.	1909167	Advanced	Display Results More	
	3	exp Psychometrics/	71281	Advanced	Display Results More	
	4	psychometr*.ti,ab.	42453	Advanced	Display Results More	
	5	(clinimetr* or clinometr*).tw.	990	Advanced	Display Results More	
	6	exp "Outcome Assessment (Health Care)"/	1008558	Advanced	Display Results More	
	7	outcome assessment.ti,ab.	3553	Advanced	Display Results More	
	8	outcome measure*.tw.	210471	Advanced	Display Results More	
	9	exp Observer Variation/	40818	Advanced	Display Results More	
	10	observer variation.ti,ab.	1039	Advanced	Display Results More	
	11	exp Health Status Indicators/	287151	Advanced	Display Results More	
	12	exp "Reproducibility of Results"/	379489	Advanced	Display Results More	
	13	reproducib*.ti,ab.	151211	Advanced	Display Results More	
	14	exp Discriminant Analysis/	10054	Advanced	Display Results More	
	15	(reliab* or unreliab* or valid* or coefficient or homogeneity or homogeneous or "internal consistency").ti,ab.	1294914	Advanced	Display Results More	
	16	(cronbach* and (alpha or alphas)).ti,ab.	20402	Advanced	Display Results More	
	17	(item and (correlation* or selection* or reduction*)).ti,ab.	20238	Advanced	Display Results More	
	18	(agreement or precision or imprecision or "precise values" or test-retest).ti,ab.	391961	Advanced	Display Results More	
	19	(test and retest).ti,ab.	25146	Advanced	Display Results More	
	20	(reliab* and (test or retest)).ti,ab.	83420	Advanced	Display Results More	
	21	(stability or interrater or inter-rater or intrarater or intra-rater or intertester or inter-tester or intratester or intra-tester or interobserver or inter-observer or intraobserver or intraobserver or intra-individual or interparticipant or inter-participant or intraparticipant or intra-participant intraassay or intra-assay or interindividual or inter-individual or intraindividual or interexaminer or inter-examiner or intraexaminer or intra-examiner or interassay or interassay intertechnician or inter-technician or intratechnician or intra-technician or kappa or kappa's or kappas or repeatab*).ti,ab.	565301	Advanced	Display Results More	
	22	((replicab* or repeated) and (measure or measures or findings or result or results or test or tests)).ti,ab.	184957	Advanced	Display Results More	
	23	(generaliza* or generalisa* or concordance).ti,ab.	81675	Advanced	Display Results More	
	24	(intraclass and correlation*).ti,ab.	22591	Advanced	Display Results More	
	25	(discriminative or "known group" or factor analysis or factor analyses or dimension* or subscale*).ti,ab.	615469	Advanced	Display Results More	
	26	(multitrait and scaling and (analysis or analyses)).ti,ab.	138	Advanced	Display Results More	
	27	(item discriminant or interscale correlation* or error or errors or "individual variability").ti,ab.	288651	Advanced	Display Results More	
	28	(variability and (analysis or values)).ti,ab.	92165	Advanced	Display Results More	
	29	(uncertainty and (measurement or measuring)).ti,ab.	7285	Advanced	Display Results More	
	30	("standard error of measurement" or sensitiv* or responsive*).ti,ab.	1493284	Advanced	Display Results More	
	31	((minimal or minimally or clinical or clinically) and (important or significant or detectable) and (change or difference)).ti,ab.	216403	Advanced	Display Results More	
	32	(small* and (real or detectable) and (change or difference)).ti,ab.	6944	Advanced	Display Results More	
	33	(meaningful change or "ceiling effect" or "floor effect" or "item response model" or IRT or Rasch or "Differential item functioning" or DIF or "computer adaptive testing" or "item bank" or "cross-cultural equivalence").ti,ab.	12070	Advanced	Display Results More	
	34	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33	9073350	Advanced	Display Results More	
	35	exp Stroke/	122899	Advanced	Display Results More	
	36	34 and 35	57672	Advanced	Display Results More	
	37	NIHSS.mp.	4586	Advanced	Display Results More	
	38	NIH stroke scale.mp.	706	Advanced	Display Results More	

☐	39	National Institutes of Health Stroke Scale.mp.	4266	Advanced	Display Results More	☐
☐	40	37 or 38 or 39	7075	Advanced	Display Results More	☐
☐	41	36 and 40	4204	Advanced	Display Results More	☐

Save | Remove | Combine with: AND | OR

Save All | Edit | Create RSS | View Saved

Basic Search | Find Citation | Search Tools | Search Fields | **Advanced Search** | Multi-Field Search

1 Resource selected | Hide | Change

Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily 1946 to June 21, 2019

Enter keyword or phrase (* or $ for truncation) — ◉ Keyword ○ Author ○ Title ○ Journal — Search

▾ **Limits** *(close)* — ☐ Include Multimedia ☑ Map Term to Subject Heading

☐ Abstracts ☐ Structured Abstracts ☐ English Language
☐ Full Text ☐ Review Articles ☐ Humans
☐ Core Clinical Journals (AIM) ☐ Latest Update ☐ Pharmacologic Actions

Publication Year - - -

Additional Limits | Edit Limits

To search Open Access content on Ovid, go to Basic Search.

Options ◂ Print Email Export + My Projects Keep Selected ☐ All Range Clear 100 Per Page 1 Go Next ›

▾ **Search Information**

You searched:
36 and 40

Search terms used:
agreement
alpha
alphas
analyses
analysis
ceiling
effect
change
clinical
clinically
clinimetr*
clinometr*
coefficient
comparative study
computer adaptive testing
concordance
correlation*
cronbach*
cross-cultural
equivalence
detectable
dif
difference
differential item functioning
dimension*
discriminant
discriminative
error
errors
factor
findings
floor
generalisa*
generaliza*
health status indicators
homogeneity
homogeneous
important
imprecision
individual

1. **A longitudinal study investigating how stroke severity, disability, and physical function the first week post-stroke are associated with walking speed six months post-stroke.**
   + My Projects + Annotate
2. **Predictors of functional outcome among stroke patients in Lima, Peru.**
   + My Projects + Annotate
3. **Ischemia-modified albumin in acute stroke.**
   + My Projects + Annotate Article as PDF (130KB)
4. **Abciximab in acute ischemic stroke. A randomized, double-blind, placebo-controlled, dose-escalation study.**
   + My Projects + Annotate
5. **Nonalcoholic fatty liver disease in patients with acute ischemic stroke is associated with more severe stroke and worse outcome.**
   + My Projects + Annotate
6. **Matrix Metalloproteinase-9 and Recovery of Acute Ischemic Stroke.**
   + My Projects + Annotate
7. **National institutes of health stroke scale item profiles as predictor of patient outcome: external validation on independent trial data.**
   + My Projects + Annotate Article as PDF (655KB)

Appendix 4.2 Grey literature search strategy

Grey Literature Search Strategy
The aim of the grey literature search was to identify scales for neurological assessment and monitoring practice not published in peer reviewed journals and research in progress around the topic of interest. The term 'stroke' was used to search websites where possible and sections of websites (e.g. reports, publications, resources etc..) were searched for relevant documents. Searches were undertaken between 23rd January 2019 to 22nd March 2019 and involved searches across the following:

- **OpenGrey Database**
- **NIHR Funding and Awards**
- **Systematic Reviews:** Prospero; Cochrane Database of systematic reviews.
- **Research Registries:** International Standard Randomised Controlled Trial Number (ISRCTN), International Clinical Trials Registry Platform (ICTRP); ClinicalTrials.gov.
- **Experts and authors contact:** for information about unpublished or ongoing studies or systematic reviews in development
- **Professional Organisations:** British Association of Neuroscience Nursing; The British & Irish Association of Stroke Physicians; Royal College of Physicians; National Stroke Nursing Forum.
- **Practice Guidelines:** National Institute for Health and Care Excellence, (NICE); SIGN Scotland. Social Care Institute for Excellence.
- **Google search:** stroke* AND (neurological* OR deterioration* OR END* OR early neurological deterioration*) AND (guide* OR scale OR tool OR recommend* OR protocol OR practic* OR process* OR guidance OR policy OR policies OR rule OR instruction OR "scheme of work" OR standard OR manual OR assess*)

Appendix 4.3 Clinimetric search example

This search strategy was to identify relevant literature for the clinimetric properties of NIHSS (National Institutes for Health Research Stroke Scale) within the Medline database.

Ovid® My Account Support & Training Help Feedback Logged in as Info Specialist at University Of Central Lancashire Logoff

Wolters Kluwer

Search Journals Books Multimedia My Workspace

▼ Search History (41) View Saved

# ▲	Searches	Results	Type	Actions	Annotations
1	(instrumentation or methods).fs.	3902614	Advanced	Display Results More	
2	(Validation Studies or Comparative Study).pt.	1909167	Advanced	Display Results More	
3	exp Psychometrics/	71281	Advanced	Display Results More	
4	psychometr*.ti,ab.	42453	Advanced	Display Results More	
5	(clinimetr* or clinometr*).tw.	990	Advanced	Display Results More	
6	exp "Outcome Assessment (Health Care)"/	1008558	Advanced	Display Results More	
7	outcome assessment.ti,ab.	3553	Advanced	Display Results More	
8	outcome measure*.tw.	210471	Advanced	Display Results More	
9	exp Observer Variation/	40818	Advanced	Display Results More	
10	observer variation.ti,ab.	1039	Advanced	Display Results More	
11	exp Health Status Indicators/	287151	Advanced	Display Results More	
12	exp "Reproducibility of Results"/	379489	Advanced	Display Results More	
13	reproducib*.ti,ab.	151211	Advanced	Display Results More	
14	exp Discriminant Analysis/	10054	Advanced	Display Results More	
15	(reliab* or unreliab* or valid* or coefficient or homogeneity or homogeneous or "internal consistency").ti,ab.	1294914	Advanced	Display Results More	
16	(cronbach* and (alpha or alphas)).ti,ab.	20402	Advanced	Display Results More	
17	(item and (correlation* or selection* or reduction*)).ti,ab.	20238	Advanced	Display Results More	
18	(agreement or precision or imprecision or "precise values" or test-retest).ti,ab.	391961	Advanced	Display Results More	
19	(test and retest).ti,ab.	25146	Advanced	Display Results More	
20	(reliab* and (test or retest)).ti,ab.	83420	Advanced	Display Results More	
21	(stability or interrater or inter-rater or intrarater or intra-rater or intertester or inter-tester or intratester or intra-tester or interobserver or inter-observer or intraobserver or intraobserver or intra-individual or interparticipant or inter-participant or intraparticipant or intra-participant intraassay or intra-assay or interindividual or inter-individual or intraindividual or interexaminer or inter-examiner or intraexaminer or intra-examiner or interassay or interassay intertechnician or inter-technician or intratechnician or intra-technician or kappa or kappa's or kappas or repeatab*).ti,ab.	565301	Advanced	Display Results More	
22	((replicab* or repeated) and (measure or measures or findings or result or results or test or tests)).ti,ab.	184957	Advanced	Display Results More	
23	(generaliza* or generalisa* or concordance).ti,ab.	81675	Advanced	Display Results More	
24	(intraclass and correlation*).ti,ab.	22591	Advanced	Display Results More	
25	(discriminative or "known group" or factor analysis or factor analyses or dimension* or subscale*).ti,ab.	615469	Advanced	Display Results More	
26	(multitrait and scaling and (analysis or analyses)).ti,ab.	138	Advanced	Display Results More	
27	(item discriminant or interscale correlation* or error or errors or "individual variability").ti,ab.	288651	Advanced	Display Results More	
28	(variability and (analysis or values)).ti,ab.	92165	Advanced	Display Results More	
29	(uncertainty and (measurement or measuring)).ti,ab.	7285	Advanced	Display Results More	
30	("standard error of measurement" or sensitiv* or responsive*).ti,ab.	1493284	Advanced	Display Results More	
31	((minimal or minimally or clinical or clinically) and (important or significant or detectable) and (change or difference)).ti,ab.	216403	Advanced	Display Results More	
32	(small* and (real or detectable) and (change or difference)).ti,ab.	6944	Advanced	Display Results More	
33	(meaningful change or "ceiling effect" or "floor effect" or "item response model" or IRT or Rasch or "Differential item functioning" or DIF or "computer adaptive testing" or "item bank" or "cross-cultural equivalence").ti,ab.	12070	Advanced	Display Results More	
34	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33	9073350	Advanced	Display Results More	
35	exp Stroke/	122899	Advanced	Display Results More	
36	34 and 35	57672	Advanced	Display Results More	
37	NIHSS.mp.	4586	Advanced	Display Results More	
38	NIH stroke scale.mp.	706	Advanced	Display Results More	

ovidsp.dc2.ovid.com/sp-3.33.0b/ovidweb.cgi 1/10

☐	39	National Institutes of Health Stroke Scale.mp.	4266	Advanced	Display Results More	☐
☐	40	37 or 38 or 39	7075	Advanced	Display Results More	☐
☐	41	36 and 40	4204	Advanced	Display Results More	☐

Save Remove Combine with: AND OR

Save All Edit Create RSS View Saved

Appendix 4.4 Search numbers for clinimetric property searches

Data summarising the screening process for the clinimetric property searches. Reports number of articles found, excluded at different stages and total numbers included.

	Records identified through database searching	**Additional records identified**	**Duplicates removed**	**Records screened (title & abstract)**	**Records excluded**	**Full- text articles assessed for eligibility**	**Records excluded**	**Potential to include**	**Records Included**
Mathew	3	3	0	6	3	3	0	3	3
GCS	1219	1	149	1070	1011	60	58	2	2
Toronto	2	2	0	4	2	2	1	1	1
SSS	420	5	96	327	296	33	28	5	3
CNS	1317	1	143	1175	1144	31	24	7	5
Hemispheric	43	2	10	34	29	5	4	1	1
NIHSS	8786	6	149	8643	8318	325	262	63	49
MCANS	21	0	5	16	12	4	3	1	1
Orgogozo^	8	0	0	8	8	0	-	-	
Unified	1219	1	149	1070	1011	60	58	2	2
ESS	181	1	30	152	143	9	8	1	1
Chinese/	129	1	23	107	93	14	13	1	0
MESSS^	5	0	0	5	5	-	-	-	-
SNOBS	0	0	-	-	-	-	-	-	-
MEND	5	0	0	5	3	2	1	1	0
Japan	13	0	2	11	8	3	2	1	1
mNIHSS	200	1	36	165	152	13	8	6	5
NIHSS-11	63	0	5	58	55	3	2		
NIHSS-8	131	2	18	115	109	6	3	3	3*

	Records identified through database searching	Additional records identified	Duplicates removed	Records screened (title & abstract)	Records excluded	Full- text articles assessed for eligibility	Records excluded	Potential to include	Records Included
NIHSS-5	205	1	30	176	171	5	4	1	1
FOUR Score	353	1	37	317	304	13	9	4	3
IVBSS	2	0	1	1	0	1	0	1	1
sNIHSS	24	3	8	19	12	7	3	4	2
e-NIHSS	25	0	5	20	17	3	2	1	1
PSAT	17	0	5	12	11	1	1	1	1
sNIHSS-EMS	4	0	1	3	0	3	2	1	0

^ Data presented together as despite searches being run separately scales are the same

* Records split as NIHSS-8 and Hunter NIHSS- 8 found to be separate scales

Proforma used to extract data from original papers(first paper introducing scale and its development).

Standardised Neurological OBservation Schedule for Stroke (SNOBSS) Seminal (SCALE/TOOL) Paper/s Data Extraction Template.

General Information

Date form completed *(dd/mm/yyyy)*	
Name/ID of person extracting data	
Reference citation	
Study author contact details	
Publication type *(e.g. full report, abstract, letter)*	
Notes:	

Development

	Descriptions as stated in report/paper	**Location in text or source** *(pg & ¶/fig/table/other)*
Method of development		
Context of development (purpose, diagnostic group)		
Time taken to develop		
Scale Type (nominal/ordinal/ interval/ratio)		
Scoring type (total score/weighted/ reverse scoring)		

Method of administration/ Response format		
Number of domains/ questions covered		
Domains/ Questions covered		
Content Validity Reported (coverage/relevance/ representiveness)		
Structural Validity		
Internal Consistency		
Items excluded? If so justification for exclusion		
Language/ understandability? (idiomatic/ very specific)		
Training requirement stipulated on development (Yes/No and if yes what?		
Time to score/ complete the scale ***(note if not reported)***		
Clinically important difference (definition within the scale and how decided upon)		
Is tool/scale validated?	☐ Yes ☐ No ☐ Unclear	

Notes:

Appendix 4.6 Data Extraction Proforma- Later Papers

Proforma used to extract data from papers that report one or more clinimetric property of the scale.

Standardised Neurological OBservation Schedule for Stroke (SNOBSS) Data Extraction Template.

This form has been developed by adopting and customizing the "Data collection form for intervention review – RCTs and non-RCTs" of The Cochrane Collaboration. Some new sections have been added into this tool and the irrelevant sections have been removed from the original form. Information included on this form should be comprehensive and may be used in the text of the review.

General Information

Date form completed *(dd/mm/yyyy)*	
Name/ID of person extracting data	
Country in which the study conducted	

Characteristics of Included Studies

	Descriptions as stated in report/paper	**Location in text or source** *(pg & ¶/fig/table/other)*
Aim of study *(e.g. efficacy, equivalence, pragmatic)*		
Study Design		
Sampling Technique (if applicable)		
Start date/End date *(if applicable)*		

Duration of participation *(from recruitment to last follow-up)*			
Ethical approval needed/ obtained for study	☐ Yes ☐ No ☐ Unclear		
Notes:			

Population and setting

	Description		Location in text or source *(pg & ¶/fig/table/other)*
Population description *(from which study participants are drawn)*			
Setting			
Inclusion criteria			
Exclusion criteria			
Method of recruitment of participants			
Informed consent obtained	☐ Yes ☐ No ☐ Unclear		
Withdrawals and exclusions			
Number of Patients included			

Relevant sociodemographics *(age, race/ethnicity, severity- does it compare with a usual stroke population or are there obvious limitations)*		
Population (SCALE/TOOL) administered by:	Professional group/s	
	Numbers	
	Equality in administration (all assess the same patients etc)	
	Experience (if stated if not state not documented)	
Training Received		
Notes:		

Results

	Description as stated in report/paper	Location in text or source *(pg & ¶/fig/table/other)*
*Content Validity (Relevance, Comprehensiveness, Comprehensibility)		
*Structural Validity		
*Internal consistency		
Reliability		
Measurement error		
Criterion validity		
Construct validity *(convergent validity- comparison with other measures, comparison between sub-groups)*		
Responsiveness *(sensitivity, specificity and effectiveness of detection of change)*		
Time taken to complete		

Power *(e.g. power & sample size calculation, level of power achieved)*		
Missing Data *(how it was dealt with)*		
Notes:		

*likely only in seminal papers of tool/scale development

Limitation and Mitigation Strategy

	Description as stated in report/paper	Location in text or source *(pg & ¶/fig/table/other)*
Strengths		
Limitations		
Strategies to overcome the limitations		
Notes:		

Conclusions and other information

Key conclusions of study authors		
Study funding sources *(including role of funders)*		
Possible conflicts of interest *(for study authors)*		
References to other relevant studies		
Issues affecting directness *(Note any aspects of population, etc. that affect this study's direct applicability to the review question)*		
Notes:		

Quality Assessment as per COSMIN

	Score as calculated (n/a for studies not including specific clinimetric property	Notes
*Content Validity (Relevance, Comprehensiveness, Comprehensibility)		
*Structural Validity		
*Internal consistency		
Cross-cultural validity/Measurement invariance		
Reliability		
Measurement error		
Criterion validity		
Construct validity (convergent validity-comparison with other measures, comparison between sub-groups)		
Responsiveness		
Notes:		

Appendix 4.7 Date occurrence of scales within literature

Table showing most recent occurrence (year) of scales within literature indexed on PubMed (searches completed 20.03.23)

Scale	**Year of most recent occurrence within literature.**
Mathew	2001
GCS	2023
Toronto	1998
SSS	2022
CNS	2021
Hemispheric	2023 (language asset)
NIHSS	2023
MCANS/ Orgogozo	2001
Unified	2005
ESS	2007
CSS/MESSS	2006/2021
SNOBS	2005
MEND	2018
Japan	2013
mNIHSS / NIHSS-11	2022
NIHSS-8	2017
NIHSS-5	2018
FOUR Score	2022
IVBSS	2021
Hunter NIHSS-8	2010
sNIHSS	2021
e-NIHSS	2016
PSAT	2015
sNIHSS-EMS	2021

Appendix 4.8 Mathew Stroke Scale

MATHEW STROKE SCALE

Patient Name: _____
Rater Name: _____
Date: _____

Activity	Score
MENTATION	
Level of Consciousness 8 = Fully conscious 6 = Lethargic but mentally intact 4 = Obtunded 2 = Stuperous 0 = Comatose	_____
Orientation (time, place, person) 6 = Oriented x 3 4 = Oriented x 2 2 = Oriented x 1 0 = Disoriented	_____
SPEECH	
0-23, according to Reitan test	_____
CRANIAL NERVES	
Homonymous hemianopsia 3 = Intact 2 = Mild 1 = Moderate 0 = Severe	_____
Conjugate deviation of eyes 3 = Intact 2 = Mild 1 = Moderate 0 = Severe	_____
Facial Weakness 3 = Intact 2 = Mild 1 = Moderate 0 = Severe	_____
MOTOR POWER	
Right arm	_____
Right leg	_____
Left arm	_____
Left leg 5 = Normal strength 4 = Contracts against resistance 3 = Elevates against gravity 2 = Gravity eliminated 1 = Flicker 0 = No movements	_____

Provided by the Internet Stroke Center — www.strokecenter.org

PERFORMANCE, OR DISABILITY STATUS SCALE _____

28 = Normal
21 = Mild impairment
14 = Moderate impairment
7 = Severe impairment
0 = Death

REFLEXES _____

3 = Normal
2 = Asymmetrical or pathological reflexes
1 = Clonus
0 = No reflexes elicited

SENSATION _____

3 = Normal
2 = Mild
1 = Severe sensory abnormality
0 = No response to pain

TOTAL _____

Reference

Mathew NT, Rivera VM, Meyer JS, Charney JZ, Hartmann A. "Double-blind evaluation of glycerol therapy in acute cerebral infarction."
Lancet. 1972;2:1327-9.

Appendix 4.9 Glasgow Coma Scale (GCS)

GLASGOW COMA SCALE

Patient Name: _____
Rater Name: _____
Date: _____

Activity		Score
EYE OPENING		
None	1 = Even to supra-orbital pressure	
To pain	2 = Pain from sternum/limb/supra-orbital pressure	
To speech	3 = Non-specific response, not necessarily to command	
Spontaneous	4 = Eyes open, not necessarily aware	_____
MOTOR RESPONSE		
None	1 = To any pain; limbs remain flaccid	
Extension	2 = Shoulder adducted and shoulder and forearm internally rotated	
Flexor response	3 = Withdrawal response or assumption of hemiplegic posture	
Withdrawal	4 = Arm withdraws to pain, shoulder abducts	
Localizes pain	5 = Arm attempts to remove supra-orbital/chest pressure	
Obeys commands	6 = Follows simple commands	_____
VERBAL RESPONSE		
None	1 = No verbalization of any type	
Incomprehensible	2 = Moans/groans, no speech	
Inappropriate	3 = Intelligible, no sustained sentences	
Confused	4 = Converses but confused, disoriented	
Oriented	5 = Converses and oriented	_____

TOTAL (3–15): _____

References

Teasdale G, Jennett B. "Assessment of coma and impaired consciousness. A practical scale." *The Lancet* 13;2(7872):81-4, 1974.

Appendix 4.10 Scandinavian Stroke Scale (SSS)

SCANDINAVIAN STROKE SCALE

Patient Name: _____
Rater Name: _____
Date: _____

Function	Score	Prognostic Score	Long Term Score
Consciousness:			
-fully conscious	**6**	____	
-somnolent, can be awaked to full consciousness	**4**		
-reacts to verbal command, but is not fully conscious	**2**		
Eye movement:			
-no gaze palsy	**4**	____	
-gaze palsy present	**2**		
-conjugate eye deviation	**0**		
Arm, motor power *:			
-raises arm with normal strength	**6**	____	____
-raises arm with reduced strength	**5**		
-raises arm with flexion in elbow	**4**		
-can move, but not against gravity	**2**		
-paralysis	**0**		
Hand, motor power *:			
-normal strength	**6**		____
-reduced strength in full range	**4**		
-some movement, fingertips do not reach palm	**2**		
-paralysis	**0**		
Leg, motor power *:			
-normal strength	**6**	____	____
-raises straight leg with reduced strength	**5**		
-raises leg with flexion of knee	**4**		
-can move, but not against gravity	**2**		
-paralysis	**0**		

Provided by the Internet Stroke Center — www.strokecenter.org

Orientation:			
-correct for time, place and person	**6**		____
-two of these	**4**		
-one of these	**2**		
-completely disorientated	**0**		
Speech:			
-no aphasia	**10**		____
-limited vocabulary or incoherent speech	**6**		
-more than yes/no, but not longer sentences	**3**		
-only yes/no or less	**0**		
Facial palsy:			
-none/dubious	**2**		____
-present	**0**		
Gait:			
-walks 5 m without aids	**12**		____
-walks with aids	**9**		
-walks with help of another person	**6**		
-sits without support	**3**		
-bedridden/wheelchair	**0**		
Maximal Score	____	**22**	**48**

* Motor power is assessed only on the affected side.

Reference

Multicenter trial of hemodilution in ischemic stroke--background and study protocol. Scandinavian Stroke Study Group. **Stroke 1985 Sep-Oct;16(5):885-90**.

Appendix 4.11 Canadian Neurological Scale (CNS)

CANADIAN NEUROLOGICAL SCALE

Patient Name: _____
Rater Name: _____
Date: _____

Mentation			Score
	Level Consciousness	Alert	3.0
		Drowsy	1.5
	Orientation	Oriented	1.0
		Disoriented/NA	0.0
	Speech	Normal	1.0
		Expressive Deficit	0.5
		Receptive Deficit	0.0

TOTAL: _____

Section A1	Motor Functions	Weakness	Score
NO COMPREHENSION DEFICIT	**Face**	None	0.5
		Present	0.0
	Arm: Proximal	None	1.5
		Mild	1.0
		Significant	0.5
		Total	0
	Arm: Distal	None	1.5
		Mild	1.0
		Significant	0.5
		Total	0
	Leg: Proximal	None	1.5
		Mild	1.0
		Significant	0.5
		Total	0
	Leg: Distal	None	1.5
		Mild	1.0
		Significant	0.5
		Total	0

TOTAL: _____

Provided by the Internet Stroke Center — www.strokecenter.org

Section A2	Motor Functions	Weakness	Score
COMPREHENSION DEFICIT	**Face**	Symmetrical	0.5
		Asymmetrical	0.0
	Arms	Equal	1.5
		Unequal	0.0
	Legs	Equal	1.5
		Unequal	0.0

TOTAL: _____

References

Cote, R, Hachinski, V. C., Shurvell, B. L., Norris, J. W., and Wolfson, C. "The Canadian Neurological: Scale A preliminary study in acute stroke."
***Stroke* 1986; 17:731-737**

Cote R, Battista RN, Wolfson C, Boucher J, Adam J, and Hachinski VC. "The Canadian Neurological Scale: Validation and reliability assessment."
***Neurology* 1989; 39:638-643**

Cheryl D. Bushnell, MD; Dean CC, Johnston, FRCPC; Larry B. Goldstein, MD. "Retrospective Assessment of Initial Stroke Severity. Comparison of the NIH Stroke Scale and the Canadian Neurological Scale."
***Stroke* 2001;32:656**

Appendix 4.12 Hemispheric Stroke Scale

HEMISPHERIC STROKE SCALE

Patient Name: _____
Rater Name: _____
Date: _____

Scored to give 0 (= good) to 100 (= bad)

Score

LEVEL OF CONSCIOUSNESS
15 – Glasgow Coma Scale Score _____

LANGUAGE

Comprehension _____
Give three commands:
'Stick out your tongue' *or* "Close your eyes'
'Point to the door'
'Place left/right hand on left/right ear and then on left/right knee (using unaffected side)

Score on number correctly followed:
0 = 5
1 = 4
2 = 2
3 = 0

Naming _____
Ask patient to name the following items:
Watch *or* Belt
Watch strap *or* Belt buckle
Index finger *or* Ring finger

Score on number correctly named:
0 = 5
1 = 4
2 = 2
3 = 0

Repetition _____
Ask the patient to repeat the following:
A single word, such as 'dog' or 'cat'
'The president lives in Washington'
'No ifs, ands, or buts'

Score on number repeated:
0 = 5
1 = 4
2 = 2
3 = 0

Page 1 TOTAL _____

Provided by the Internet Stroke Center — www.strokecenter.org

Fluency _____
Score according to patient's spontaneous speech fluency, *or*
Ask patient to name as many words as he can within one minute beginning with the letter 'A' (excluding proper names)

Score as:
5 = Essentially no verbal output
3 = Moderately loss; inability to recognize stationary finger, sees moving finger
1 = Mild loss; defect to double simultaneous stimulation
0 = Normal

OTHER CORTICAL FUNCTIONS AND CRANIAL NERVES

Visual fields _____
Test clinically and score hemi-field loss as:
3 = Severe loss; inability to recognize moving hand, no response to threat
2 = Moderate loss; inability to recognize stationary finger, sees moving finger
1 = Mild loss: defect to double simultaneous stimulation
0 = Normal

Gaze _____
Score eye movements:
2 = Gaze play, or persistent deviation
1 = Gaze preference, or difficulty with far lateral gaze
0 = Normal

Facial expression _____
Score movement:
3 = Severe weakness; drooling
2 = Moderate loss; asymmetry at rest
1 = Mild weakness; asymmetry on similing
0 = Normal

Dysarthria _____
Score talking:
2 = Severe dysarthria
1 = Moderate dysarthria
0 = Normal

Dysphagia _____
Score swallow of glass water:
2 = Severe dysphagia
1 = Moderate dysphagia
0 = Normal

Neglect syndrome _____
Ask about weak limbs, *and* ask to bisect a line 7 inches (20 cm) long on piece of paper in visual midline

Score:
2 = Anosagnosia, or denial of body part
1 = Consistently bisects line towards 'good' side of body
0 = Bisects line in middle

Visual construction _____
Ask patient to copy three figure given, and score:
3 = Unable to copy any figure
2 = Can copy a square
1 = Can copy a 'Greek Cross' ('Cross of St. George')
0 = Can copy 3D drawing of cube

Page 2 TOTAL _____

MOTOR FUNCTION

Arm, proximal _____

Arm, distal _____

Leg, proximal _____

Leg, distal _____

All scored 0-7 as:
7 = No movement (MRC 0)
6 = Trace movement only (MRC 1)
5 = Motion without gravity only (MRC 2)
4 = Moves against gravity but not against resistance (MRC 3)
3 = Moderate weakness (MRC 4 -)
2 = Mild weakness (MRC 4)
1 = Positive drift of arm/leg (MRC 4 +)
0 = Normal (MRC 5)

Deep tendon reflexes _____
2 = Hypoactive *or* hyperactive
0 = Normal

Pathologic reflexes _____
2 = Babinski (plantar) *and* another abnormal
1 = Babinski (plantar) *or* another abnormal
0 = Normal

Muscle tone _____
2 = Increased *or* decreased
0 = Normal

Gait _____
Test ability to stand and walk, and score:
6 = Unable to stand unsupported *or* cannot evaluate
5 = Can stand with support but cannot walk
4 = Severely abnormal; walking distance limited even with support (from aid or person)
3 = Moderately abnormal; no assistance required (apart from a stick/cane), but distance limited
2 = Mildly abnormal (weak, uncoordinated); can walk independently but slowly
1 = Minimally abnormal, no reduction in speed or distance
0 = Normal

SENSORY

Primary modalities (of affected side only), arm _____
Test touch, pain and score as:
4 = Anaesthesia
3 = Severe hypaesthesia
2 = Moderate hypaesthesia or deficit only; *or* extinction to double simultaneous stimulation
1 = Mild hypaesthesia or dysaesthesia
0 = Normal

Page 3 TOTAL _____

Appendix 4.13 National Institutes for Health Stroke Scale (NIHSS)

N I H STROKE SCALE

Patient Identification. ___ ___-___ ___ ___-___ ___ ___

Pt. Date of Birth ___ ___/___ ___/___ ___

Hospital _____(___ ___-___ ___)

Date of Exam ___ ___/___ ___/___ ___

Interval: [ ] Baseline [ ] 2 hours post treatment [ ] 24 hours post onset of symptoms ±20 minutes [ ] 7-10 days
[ ] 3 months [ ] Other _____(___ ___)

Time: ___ ___:___ ___ [ ]am [ ]pm

Person Administering Scale _____

Administer stroke scale items in the order listed. Record performance in each category after each subscale exam. Do not go back and change scores. Follow directions provided for each exam technique. Scores should reflect what the patient does, not what the clinician thinks the patient can do. The clinician should record answers while administering the exam and work quickly. Except where indicated, the patient should not be coached (i.e., repeated requests to patient to make a special effort).

Instructions	Scale Definition	Score
1a. Level of Consciousness: The investigator must choose a response if a full evaluation is prevented by such obstacles as an endotracheal tube, language barrier, orotracheal trauma/bandages. A 3 is scored only if the patient makes no movement (other than reflexive posturing) in response to noxious stimulation.	0 = **Alert;** keenly responsive. 1 = **Not alert;** but arousable by minor stimulation to obey, answer, or respond. 2 = **Not alert;** requires repeated stimulation to attend, or is obtunded and requires strong or painful stimulation to make movements (not stereotyped). 3 = Responds only with reflex motor or autonomic effects or totally unresponsive, flaccid, and areflexic.	_____
1b. LOC Questions: The patient is asked the month and his/her age. The answer must be correct - there is no partial credit for being close. Aphasic and stuporous patients who do not comprehend the questions will score 2. Patients unable to speak because of endotracheal intubation, orotracheal trauma, severe dysarthria from any cause, language barrier, or any other problem not secondary to aphasia are given a 1. It is important that only the initial answer be graded and that the examiner not "help" the patient with verbal or non-verbal cues.	0 = **Answers** both questions correctly. 1 = **Answers** one question correctly. 2 = **Answers** neither question correctly.	_____
1c. LOC Commands: The patient is asked to open and close the eyes and then to grip and release the non-paretic hand. Substitute another one step command if the hands cannot be used. Credit is given if an unequivocal attempt is made but not completed due to weakness. If the patient does not respond to command, the task should be demonstrated to him or her (pantomime), and the result scored (i.e., follows none, one or two commands). Patients with trauma, amputation, or other physical impediments should be given suitable one-step commands. Only the first attempt is scored.	0 = **Performs** both tasks correctly. 1 = **Performs** one task correctly. 2 = **Performs** neither task correctly.	_____
2. Best Gaze: Only horizontal eye movements will be tested. Voluntary or reflexive (oculocephalic) eye movements will be scored, but caloric testing is not done. If the patient has a conjugate deviation of the eyes that can be overcome by voluntary or reflexive activity, the score will be 1. If a patient has an isolated peripheral nerve paresis (CN III, IV or VI), score a 1. Gaze is testable in all aphasic patients. Patients with ocular trauma, bandages, pre-existing blindness, or other disorder of visual acuity or fields should be tested with reflexive movements, and a choice made by the investigator. Establishing eye contact and then moving about the patient from side to side will occasionally clarify the presence of a partial gaze palsy.	0 = **Normal.** 1 = **Partial gaze palsy;** gaze is abnormal in one or both eyes, but forced deviation or total gaze paresis is not present. 2 = **Forced deviation,** or total gaze paresis not overcome by the oculocephalic maneuver.	_____

Rev 10/1/2003

NIH STROKE SCALE

Patient Identification. ___ ___-___ ___ ___-___ ___ ___

Pt. Date of Birth ___ ___/___ ___/___ ___

Hospital _____(___ ___-___ ___)

Date of Exam ___ ___/___ ___/___ ___

Interval: [ ] Baseline [ ] 2 hours post treatment [ ] 24 hours post onset of symptoms ±20 minutes [ ] 7-10 days
[ ] 3 months [ ] Other _____(___ ___)

3. Visual: Visual fields (upper and lower quadrants) are tested by confrontation, using finger counting or visual threat, as appropriate. Patients may be encouraged, but if they look at the side of the moving fingers appropriately, this can be scored as normal. If there is unilateral blindness or enucleation, visual fields in the remaining eye are scored. Score 1 only if a clear-cut asymmetry, including quadrantanopia, is found. If patient is blind from any cause, score 3. Double simultaneous stimulation is performed at this point. If there is extinction, patient receives a 1, and the results are used to respond to item 11.	0 = **No visual loss.** 1 = **Partial hemianopia.** 2 = **Complete hemianopia.** 3 = **Bilateral hemianopia** (blind including cortical blindness).	_____
4. Facial Palsy: Ask – or use pantomime to encourage – the patient to showteeth or raise eyebrows and close eyes. Score symmetry of grimace in response to noxious stimuli in the poorly responsive or non-comprehending patient. If facial trauma/bandages, orotracheal tube, tape or other physical barriers obscure the face, these should be removed to the extent possible.	0 = **Normal** symmetrical movements. 1 = **Minor paralysis** (flattened nasolabial fold, asymmetry on smiling). 2 = **Partial paralysis** (total or near-total paralysis of lower face). 3 = **Complete paralysis** of one or both sides (absence of facial movement in the upper and lower face).	_____
5. Motor Arm: The limb is placed in the appropriate position: extend the arms (palms down) 90 degrees (if sitting) or 45 degrees (if supine). Drift is scored if the arm falls before 10 seconds. The aphasic patient is encouraged using urgency in the voice and pantomime, but not noxious stimulation. Each limb is tested in turn, beginning with the non-paretic arm. Only in the case of amputation or joint fusion at the shoulder, the examiner should record the score as untestable (UN), and clearly write the explanation for this choice.	0 = **No drift;** limb holds 90 (or 45) degrees for full 10 seconds. 1 = **Drift;** limb holds 90 (or 45) degrees, but drifts down before full 10 seconds; does not hit bed or other support. 2 = **Some effort against gravity;** limb cannot get to or maintain (if cued) 90 (or 45) degrees, drifts down to bed, but has some effort against gravity. 3 = **No effort against gravity;** limb falls. 4 = **No movement.** UN = **Amputation** or joint fusion, explain: _____ **5a. Left Arm** **5b. Right Arm**	_____ _____
6. Motor Leg: The limb is placed in the appropriate position: hold the leg at 30 degrees (always tested supine). Drift is scored if the leg falls before 5 seconds. The aphasic patient is encouraged using urgency in the voice and pantomime, but not noxious stimulation. Each limb is tested in turn, beginning with the non-paretic leg. Only in the case of amputation or joint fusion at the hip, the examiner should record the score as untestable (UN), and clearly write the explanation for this choice.	0 = **No drift;** leg holds 30-degree position for full 5 seconds. 1 = **Drift;** leg falls by the end of the 5-second period but does not hit bed. 2 = **Some effort against gravity;** leg falls to bed by 5 seconds, but has some effort against gravity. 3 = **No effort against gravity;** leg falls to bed immediately. 4 = **No movement.** UN = **Amputation** or joint fusion, explain: _____ **6a. Left Leg** **6b. Right Leg**	_____

Rev 10/1/2003

NIH STROKE SCALE

Patient Identification. ___ ___-___ ___ ___-___ ___ ___

Pt. Date of Birth ___ ___/___ ___/___ ___

Hospital _____(___ ___-___ ___)

Date of Exam ___ ___/___ ___/___ ___

Interval: [ ] Baseline [ ] 2 hours post treatment [ ] 24 hours post onset of symptoms ±20 minutes [ ] 7-10 days
[ ] 3 months [ ] Other _____(___ ___)

7. Limb Ataxia: This item is aimed at finding evidence of a unilateral cerebellar lesion. Test with eyes open. In case of visual defect, ensure testing is done in intact visual field. The finger-nose-finger and heel-shin tests are performed on both sides, and ataxia is scored only if present out of proportion to weakness. Ataxia is absent in the patient who cannot understand or is paralyzed. Only in the case of amputation or joint fusion, the examiner should record the score as untestable (UN), and clearly write the explanation for this choice. In case of blindness, test by having the patient touch nose from extended arm position.	0 = **Absent.** 1 = **Present in one limb.** 2 = **Present in two limbs.** UN = **Amputation** or joint fusion, explain: _____	____
8. Sensory: Sensation or grimace to pinprick when tested, or withdrawal from noxious stimulus in the obtunded or aphasic patient. Only sensory loss attributed to stroke is scored as abnormal and the examiner should test as many body areas (arms [not hands], legs, trunk, face) as needed to accurately check for hemisensory loss. A score of 2, "severe or total sensory loss," should only be given when a severe or total loss of sensation can be clearly demonstrated. Stuporous and aphasic patients will, therefore, probably score 1 or 0. The patient with brainstem stroke who has bilateral loss of sensation is scored 2. If the patient does not respond and is quadriplegic, score 2. Patients in a coma (item 1a=3) are automatically given a 2 on this item.	0 = **Normal;** no sensory loss. 1 = **Mild-to-moderate sensory loss;** patient feels pinprick is less sharp or is dull on the affected side; or there is a loss of superficial pain with pinprick, but patient is aware of being touched. 2 = **Severe to total sensory loss;** patient is not aware of being touched in the face, arm, and leg.	____
9. Best Language: A great deal of information about comprehension will be obtained during the preceding sections of the examination. For this scale item, the patient is asked to describe what is happening in the attached picture, to name the items on the attached naming sheet and to read from the attached list of sentences. Comprehension is judged from responses here, as well as to all of the commands in the preceding general neurological exam. If visual loss interferes with the tests, ask the patient to identify objects placed in the hand, repeat, and produce speech. The intubated patient should be asked to write. The patient in a coma (item 1a=3) will automatically score 3 on this item. The examiner must choose a score for the patient with stupor or limited cooperation, but a score of 3 should be used only if the patient is mute and follows no one-step commands.	0 = **No aphasia;** normal. 1 = **Mild-to-moderate aphasia;** some obvious loss of fluency or facility of comprehension, without significant limitation on ideas expressed or form of expression. Reduction of speech and/or comprehension, however, makes conversation about provided materials difficult or impossible. For example, in conversation about provided materials, examiner can identify picture or naming card content from patient's response. 2 = **Severe aphasia;** all communication is through fragmentary expression; great need for inference, questioning, and guessing by the listener. Range of information that can be exchanged is limited; listener carries burden of communication. Examiner cannot identify materials provided from patient response. 3 = **Mute, global aphasia;** no usable speech or auditory comprehension.	____
10. Dysarthria: If patient is thought to be normal, an adequate sample of speech must be obtained by asking patient to read or repeat words from the attached list. If the patient has severe aphasia, the clarity of articulation of spontaneous speech can be rated. Only if the patient is intubated or has other physical barriers to producing speech, the examiner should record the score as untestable (UN), and clearly write an explanation for this choice. Do not tell the patient why he or she is being tested.	0 = **Normal.** 1 = **Mild-to-moderate dysarthria;** patient slurs at least some words and, at worst, can be understood with some difficulty. 2 = **Severe dysarthria;** patient's speech is so slurred as to be unintelligible in the absence of or out of proportion to any dysphasia, or is mute/anarthric. UN = **Intubated** or other physical barrier, explain: _____	____

Rev 10/1/2003

NIH STROKE SCALE

Patient Identification. ___ ___-___ ___ ___-___ ___ ___

Pt. Date of Birth ___ ___/___ ___/___ ___

Hospital _____(___ ___-___ ___)

Date of Exam ___ ___/___ ___/___ ___

Interval: [ ] Baseline [ ] 2 hours post treatment [ ] 24 hours post onset of symptoms ±20 minutes [ ] 7-10 days
[ ] 3 months [ ] Other _____(___ ___)

11. Extinction and Inattention (formerly Neglect): Sufficient information to identify neglect may be obtained during the prior testing. If the patient has a severe visual loss preventing visual double simultaneous stimulation, and the cutaneous stimuli are normal, the score is normal. If the patient has aphasia but does appear to attend to both sides, the score is normal. The presence of visual spatial neglect or anosagnosia may also be taken as evidence of abnormality. Since the abnormality is scored only if present, the item is never untestable.	0 = **No abnormality.** 1 = **Visual, tactile, auditory, spatial, or personal inattention** or extinction to bilateral simultaneous stimulation in one of the sensory modalities. 2 = **Profound hemi-inattention or extinction to more than one modality;** does not recognize own hand or orients to only one side of space.	____

____

____

Rev 10/1/2003

COOKIE
JAR

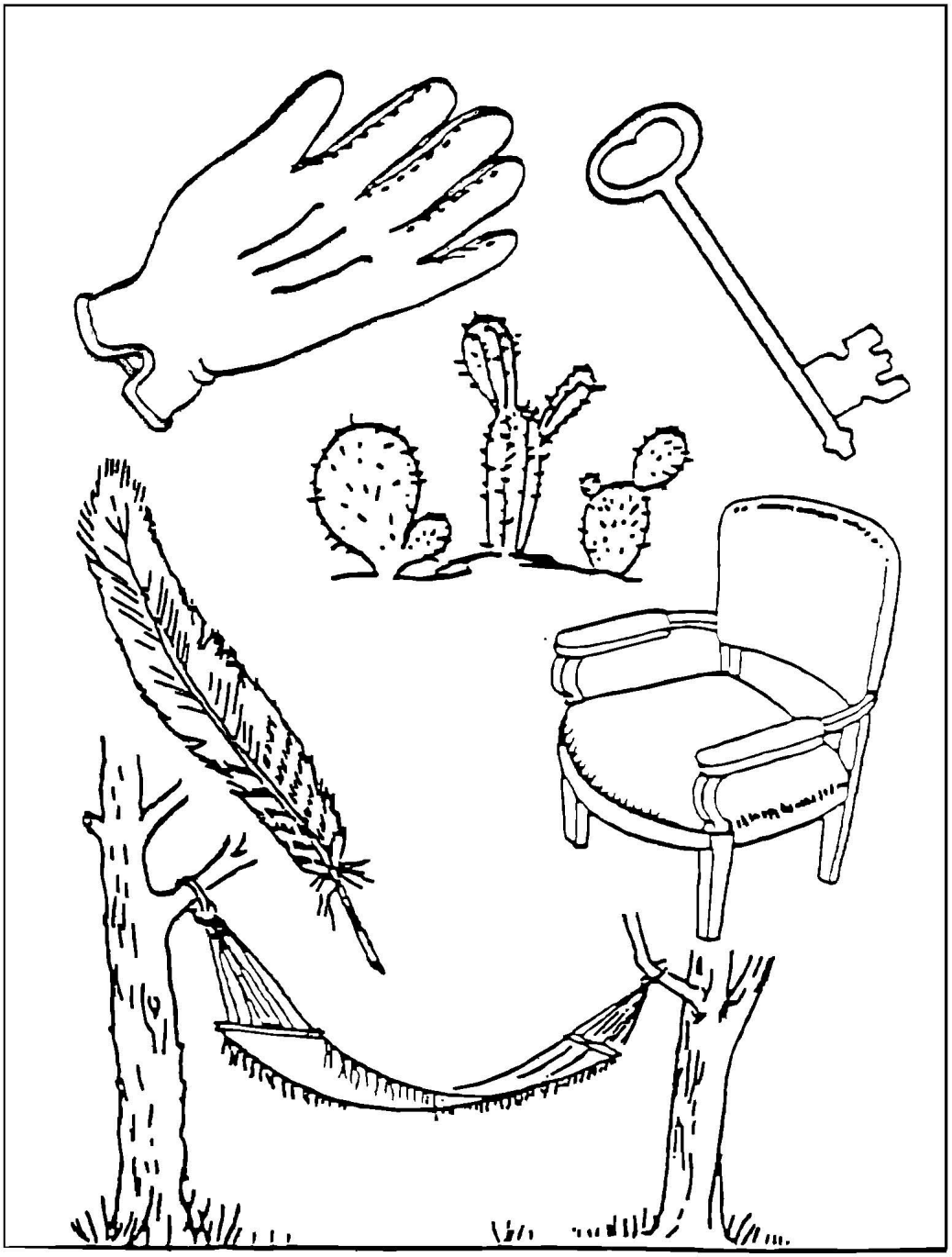
You know how.

Down to earth.

I got home from work.

Near the table in the dining room.

They heard him speak on the radio last night.



MAMA

TIP – TOP

FIFTY – FIFTY

THANKS

HUCKLEBERRY

BASEBALL PLAYER

Appendix 4.14 MCANS/Orgogozo Scale

ORGOGOZO STROKE SCALE

Patient Name: _____
Rater Name: _____
Date: _____

Activity	Score
CONSCIOUSNESS 0 = coma 5 = stupor 10 = drowsiness 15 = normal	_____
VERBAL COMMUNICATION 0 = impossible 5 = difficult 10 = normal	_____
EYES AND HEAD SHIFT 0 = forced 5 = gaze failure 10 = none	_____
FACIAL MOVEMENTS 0 = paralysis 5 = normal	_____
ARM RAISING 0 = impossible 5 = incomplete 10 = possible	_____
HAND MOVEMENTS 0 = useless 5 = useful 10 = skilled 15 = normal	_____
UPPER LIMB TONE 0 = increased or decreased 5 = normal	_____
LEG RAISING 0 = impossible 5 = gravity 10 = resistance 15 = normal	_____
FOOT DORSIFLEXION 0 = foot drop 5 = gravity 10 = resistance or normal	_____
LOWER LIMB TONE 0 = increased or decreased 5 = normal	_____
TOTAL (0–100):	_____

Provided by the Internet Stroke Center — www.strokecenter.org

(Orgogozo et al., 1983)

Appendix 4.15 European Stroke Scale

The European Stroke Scale

Overview :

The European Stroke Scale can be used to assess a patient who has recently had a stroke involving the distribution of a middle cerebral artery. This can be used to measure therapeutic efficacy and to match patients for comparison.

Parameters:

(1) level of consciousness

(2) comprehension: The patient is asked to follow these commands: (a) stick out tongue, (b) put a finger from the (unaffected) side on the nose, (c) close the eyelids. The examiner must not demonstrate the action.

(3) speech: The examiner makes general conversation with the patient.

(4) visual field: The examiner stands at arm's length and compares the patient's field of vision by advancing a moving finger from the periphery inwards. The patient is asked to fixate on the examiner's pupil. The test is done first with one eye open and the other closed, then the opposite.

(5) gaze: The examiner steadies the patient's head and asks the patient to follow the examiner's finger. The examiner observes the resting eye position and subsequently the full range of movements by moving the finger from the left to the right, then vice versa.

(6) facial movement: The patient's face is examined while talking and smiling, with any asymmetries noted. Only the muscles in the lower half of the face are assessed.

(7) arm in outstretched position: The patient is asked to close the eyes. The patient's arms are actively lifted into a 45° position relative to the horizontal plane, with both hands in mid position facing each other. The patient is asked to maintain this position for 5 seconds after the examiner withdraws support. Only the affected side is evaluated.

(8) arm raising: The patient's arm is rested next to the leg with the hand in mid-position. The patient is asked to raise the arm outstretched to 90° (vertical).

(9) extension of wrist: The patient is tested with the forearm supported. The hand is unsupported but relaxed in pronation. The patient is asked to extend the hand.

(10) fingers: The patient is asked to form a pinch grip with the thumb and forefinger and to resist a weak pull. The examiner assesses the strength of the pinch grip by pulling on the pinched fingers using one finger.

(11) leg maintained in position: The examiner actively lifts the patient's affected leg into position, with the thigh perpendicular to the bed and the lower leg parallel to the bed. The patient is asked to close the eyes and to maintain the leg in position for 5 seconds without support.

(12) leg flexing: The patient is supine with the leg outstretched. The patient is asked to flex the hip and knee.

(13) dorsiflexion of foot: The patient's leg is outstretched, with the patient asked to dorsiflex the foot.

(14) gait

Parameter	Finding	Points
level of consciousness	alert, keenly responsive	10
	drowsy but can be aroused by minor stimulation to obey, answer or respond	8

	requires repeated stimulation to attend, or is lethargic or obtunded, requiring strong or painful stimulation to make movements	6
	cannot be roused by any stimulation, does react purposefully to painful stimuli	4
	cannot be roused by any stimulation, does react with decerebration to painful stimuli	2
	cannot be roused by any stimulation, does not react to painful stimuli	0
comprehension	patient performs 3 commands	8
	patient performs 1 or 2 commands	4
	patient does not perform any command	0
speech	normal speech	8
	slight word-finding difficulty, conversation is possible	6
	severe word-finding difficulties, conversation is difficult	4
	only yes or no	2
	mute	0
visual field	normal	8
	deficit	0
gaze	normal	8
	median eye position, deviation to one side impossible	4
	lateral eye position, return to midline possible	2
	lateral eye position, return to midline impossible	0
facial movement	normal	8
	paresis	4
	paralysis	0
arm (ability to maintain outstretched position)	arm maintains position for 5 seconds	4
	arm maintains position for 5 seconds but affected hand pronates	3
	arm drifts before 5 seconds pass and maintains lower position	2
	arm can't maintain position but attempts to oppose gravity	1
	arm falls	0
arm (raising)	normal	4
	straight arm, movement not full	3
	flexed arm	2
	trace movements	1
	no movement	0
extension of the wrist	normal (full isolated movement, no decrease in strength)	8
	full isolated movement, reduced strength	6
	movement not isolated and/or full	4
	trace movements	2
	no movement	0
fingers	equal strength	8
	reduced strength on affected side	4
	pinch grip impossible on affected side	0
leg (maintain position)	leg maintains position for 5 seconds	4
	leg drifts to intermediate position by the end of 5 seconds	2
	leg drifts to bed within 5 seconds but not immediately	1
	leg falls to bed immediately	0
leg (flexing)	normal	4
	movement against resistance, reduced strength	3
	movement against gravity	2
	trace movements	1

	no movement	0
dorsiflexion of foot	normal (leg outstretched, full movement, no decrease in strength)	8
	leg outstretched, full movement, reduced strength	6
	leg outstretched, movement not full or knee flexed or foot in supination	4
	trace movements	2
	no movement	0
gait	normal	10
	gait has abnormal aspect and/or distance limited and/or speed limited	8
	patient can walk with aid	6
	patient can walk with physical assistance of one or more persons	4
	patient cannot walk but can stand supported	2
	patient cannot walk nor stand	0

European stroke score = SUM(points for all 14 parameters)

Interpretation:

- minimum score: 0
- maximum score: 100
- A completely normal person would have a score of 100.
- The maximally affected person has a score of 0.

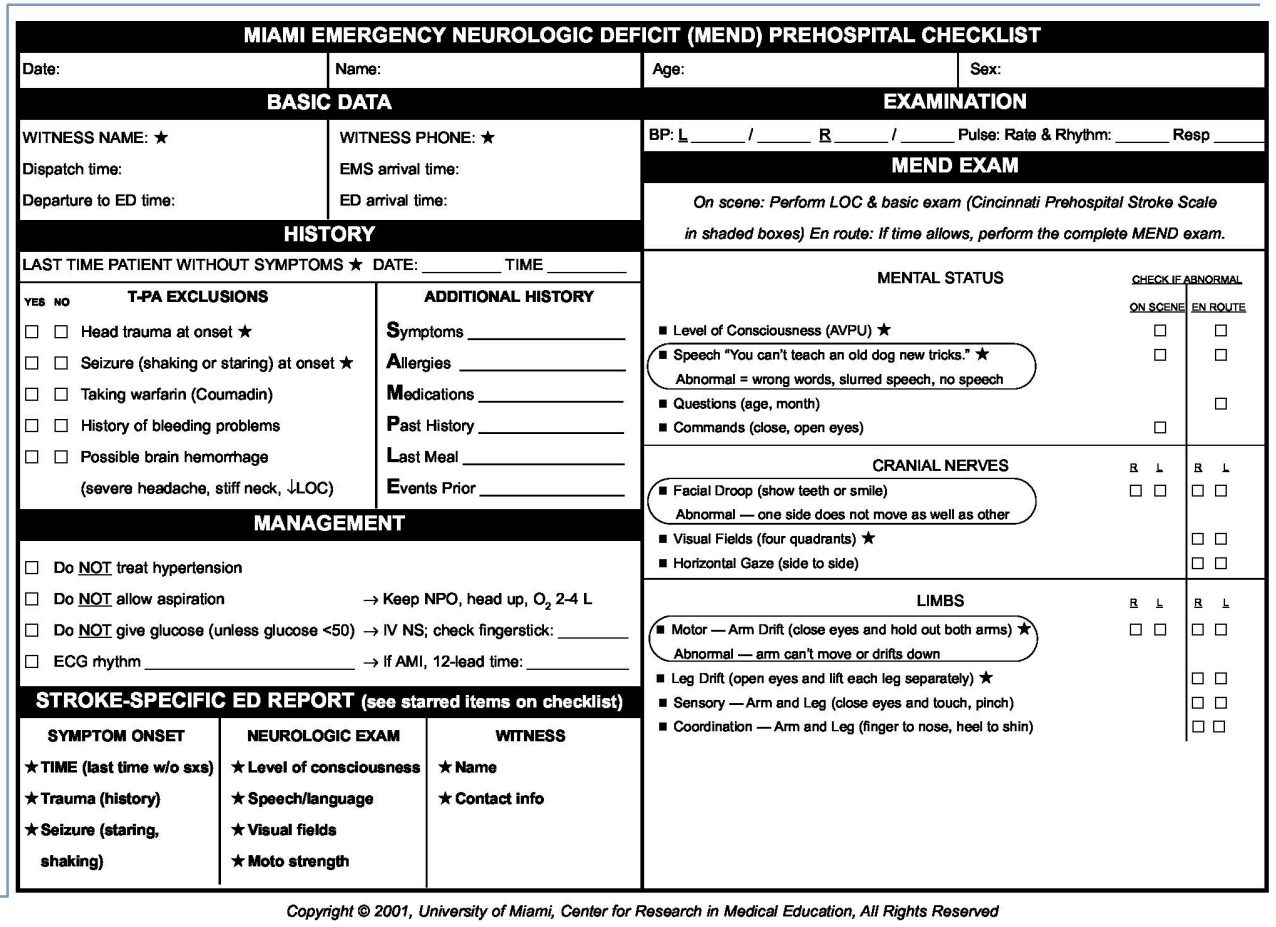
References:

Hantson L, De Weerdt W, et al. The European Stroke Scale. Stroke. 1994; 25: 2215-2219.

Appendix 4.16 Standardised Nursing OBservation Schedule (SNOBS)

SNOBS signs		*DAY 1*				*DAY 2*				*DAY 3*			
		1	2	3	4	1	2	3	4	1	2	3	4
Conscious level	Fully conscious, alert												
	Sleepy, can be awakened to full consciousness												
	Reacts to voice / stimulus, **cannot** be fully conscious												
	Coma: no response to stimulus												
Speech & communic -ation	Normal: no communication difficulty												
	Mild communication difficulty												
	Moderate difficulty, no proper sentences												
	Severe difficulty, 1 or 2 words or less												
Eye movements	Normal conjugate movement, eyes move L & R equally												
	Difficulty looking to affected side (lateral paresis)												
	Eyes deviated at rest (away from affected side)												
Arm	Raises arm with normal strength												
	[Raises arm with reduced strength, elbow straight]												
	Raises arm against gravity but with bent elbow												
	Can move arm but **not** against gravity												
	Paralysed, no movement												
Leg	Raises leg with normal strength												
	[Raises straight leg with reduced strength]												
	Raises leg against gravity but with bent knee												
	Can move leg but **not** against gravity												
	Paralysed, no movement												

Appendix 4.17 MEND



MIAMI EMERGENCY NEUROLOGIC DEFICIT (MEND) PREHOSPITAL CHECKLIST

Date: | Name: | Age: | Sex:

BASIC DATA

WITNESS NAME: ★ | WITNESS PHONE: ★

Dispatch time: | EMS arrival time:

Departure to ED time: | ED arrival time:

HISTORY

LAST TIME PATIENT WITHOUT SYMPTOMS ★ DATE: _____ TIME _____

YES	NO	T-PA EXCLUSIONS
☐	☐	Head trauma at onset ★
☐	☐	Seizure (shaking or staring) at onset ★
☐	☐	Taking warfarin (Coumadin)
☐	☐	History of bleeding problems
☐	☐	Possible brain hemorrhage (severe headache, stiff neck, ↓LOC)

ADDITIONAL HISTORY

Symptoms _____

Allergies _____

Medications _____

Past History _____

Last Meal _____

Events Prior _____

MANAGEMENT

☐ Do NOT treat hypertension

☐ Do NOT allow aspiration → Keep NPO, head up, O_2 2-4 L

☐ Do NOT give glucose (unless glucose <50) → IV NS; check fingerstick: _____

☐ ECG rhythm _____ → If AMI, 12-lead time: _____

STROKE-SPECIFIC ED REPORT (see starred items on checklist)

SYMPTOM ONSET	NEUROLOGIC EXAM	WITNESS
★ TIME (last time w/o sxs)	★ Level of consciousness	★ Name
★ Trauma (history)	★ Speech/language	★ Contact info
★ Seizure (staring, shaking)	★ Visual fields	
	★ Moto strength	

EXAMINATION

BP: L _____ / _____ R _____ / _____ Pulse: Rate & Rhythm: _____ Resp _____

MEND EXAM

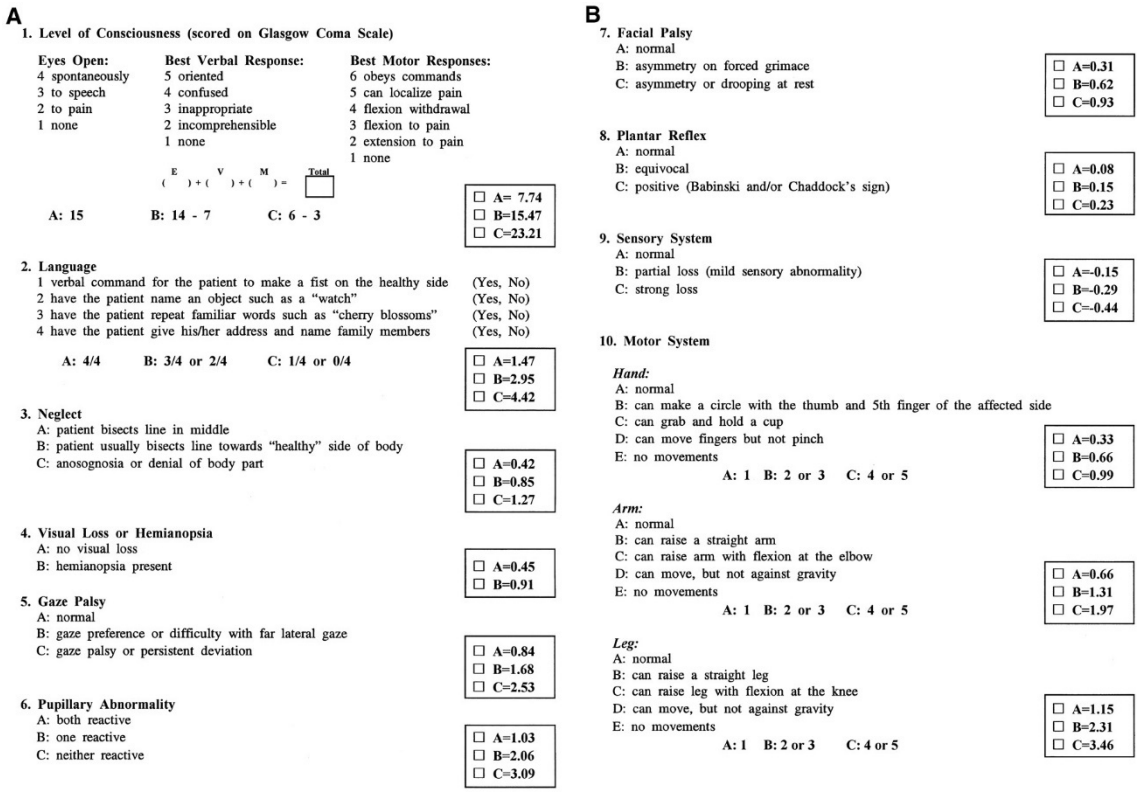
On scene: Perform LOC & basic exam (Cincinnati Prehospital Stroke Scale in shaded boxes) En route: If time allows, perform the complete MEND exam.

MENTAL STATUS	CHECK IF ABNORMAL ON SCENE	CHECK IF ABNORMAL EN ROUTE
■ Level of Consciousness (AVPU) ★	☐	☐
■ Speech "You can't teach an old dog new tricks." ★ Abnormal = wrong words, slurred speech, no speech	☐	☐
■ Questions (age, month)		☐
■ Commands (close, open eyes)	☐	

CRANIAL NERVES	R	L	R	L
■ Facial Droop (show teeth or smile) Abnormal — one side does not move as well as other	☐	☐	☐	☐
■ Visual Fields (four quadrants) ★			☐	☐
■ Horizontal Gaze (side to side)			☐	☐

LIMBS	R	L	R	L
■ Motor — Arm Drift (close eyes and hold out both arms) ★ Abnormal — arm can't move or drifts down	☐	☐	☐	☐
■ Leg Drift (open eyes and lift each leg separately) ★			☐	☐
■ Sensory — Arm and Leg (close eyes and touch, pinch)			☐	☐
■ Coordination — Arm and Leg (finger to nose, heel to shin)			☐	☐

Appendix 4.18 Japan Stroke Scale



A

1. Level of Consciousness (scored on Glasgow Coma Scale)

Eyes Open:	Best Verbal Response:	Best Motor Responses:
4 spontaneously	5 oriented	6 obeys commands
3 to speech	4 confused	5 can localize pain
2 to pain	3 inappropriate	4 flexion withdrawal
1 none	2 incomprehensible	3 flexion to pain
	1 none	2 extension to pain
		1 none

E ( ) + V ( ) + M ( ) = Total

A: 15 B: 14 - 7 C: 6 - 3

☐ A= 7.74
☐ B=15.47
☐ C=23.21

2. Language

1 verbal command for the patient to make a fist on the healthy side (Yes, No)
2 have the patient name an object such as a "watch" (Yes, No)
3 have the patient repeat familiar words such as "cherry blossoms" (Yes, No)
4 have the patient give his/her address and name family members (Yes, No)

A: 4/4 B: 3/4 or 2/4 C: 1/4 or 0/4

☐ A=1.47
☐ B=2.95
☐ C=4.42

3. Neglect

A: patient bisects line in middle
B: patient usually bisects line towards "healthy" side of body
C: anosognosia or denial of body part

☐ A=0.42
☐ B=0.85
☐ C=1.27

4. Visual Loss or Hemianopsia

A: no visual loss
B: hemianopsia present

☐ A=0.45
☐ B=0.91

5. Gaze Palsy

A: normal
B: gaze preference or difficulty with far lateral gaze
C: gaze palsy or persistent deviation

☐ A=0.84
☐ B=1.68
☐ C=2.53

6. Pupillary Abnormality

A: both reactive
B: one reactive
C: neither reactive

☐ A=1.03
☐ B=2.06
☐ C=3.09

B

7. Facial Palsy

A: normal
B: asymmetry on forced grimace
C: asymmetry or drooping at rest

☐ A=0.31
☐ B=0.62
☐ C=0.93

8. Plantar Reflex

A: normal
B: equivocal
C: positive (Babinski and/or Chaddock's sign)

☐ A=0.08
☐ B=0.15
☐ C=0.23

9. Sensory System

A: normal
B: partial loss (mild sensory abnormality)
C: strong loss

☐ A=-0.15
☐ B=-0.29
☐ C=-0.44

10. Motor System

Hand:

A: normal
B: can make a circle with the thumb and 5th finger of the affected side
C: can grab and hold a cup
D: can move fingers but not pinch
E: no movements

A: 1 B: 2 or 3 C: 4 or 5

☐ A=0.33
☐ B=0.66
☐ C=0.99

Arm:

A: normal
B: can raise a straight arm
C: can raise arm with flexion at the elbow
D: can move, but not against gravity
E: no movements

A: 1 B: 2 or 3 C: 4 or 5

☐ A=0.66
☐ B=1.31
☐ C=1.97

Leg:

A: normal
B: can raise a straight leg
C: can raise leg with flexion at the knee
D: can move, but not against gravity
E: no movements

A: 1 B: 2 or 3 C: 4 or 5

☐ A=1.15
☐ B=2.31
☐ C=3.46

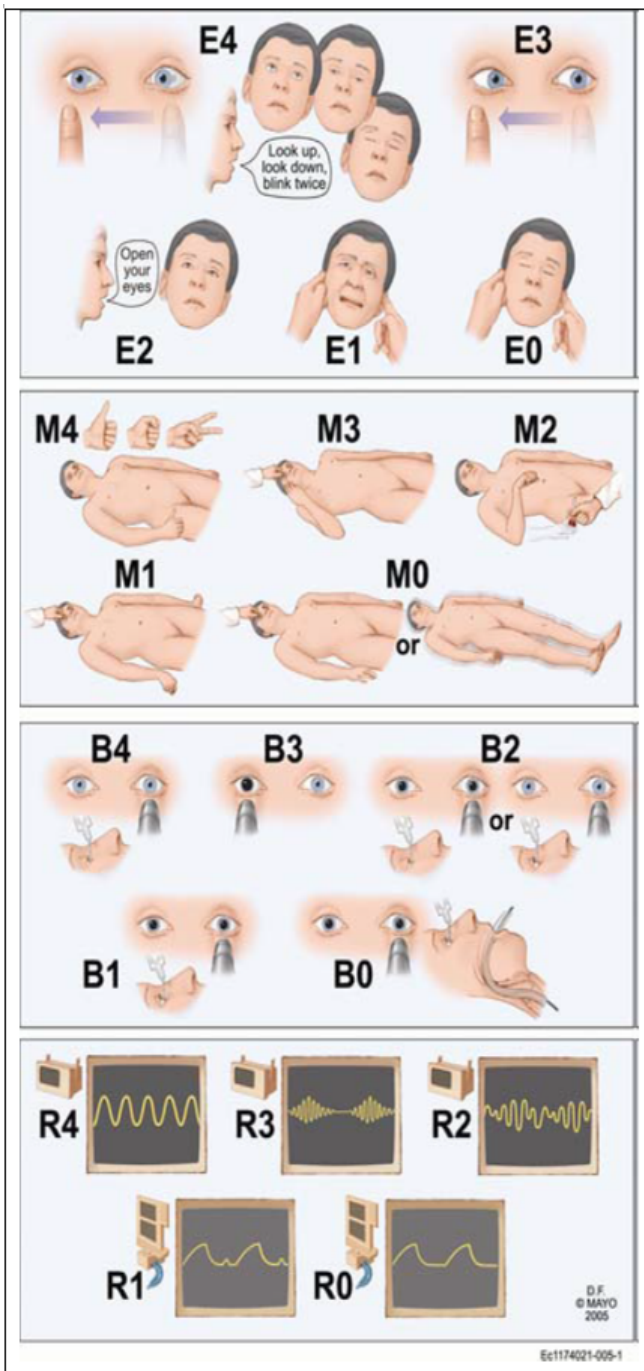
Appendix 4.19 mNIHSS/ NIHSS-11 (Modified National Institutes for Health Stroke Scale)

Item Number	Item Name	Scoring Guide	Patient Score
1B	LOC Questions	0=answers both correctly 1=answers one correctly 2=answers neither correctly	_____
1C	LOC Commands	0=performs both tasks correctly 1=performs one task correctly 2=performs neither task	_____
2.	Gaze	0=normal 1=partial gaze palsey 2=total gaze palsey	_____
3.	Visual Fields	0=no visual loss 1=partial hemianopsia 2=complete hemianopsia 3=bilateral hemianopsia	_____
5a.	Left Arm Motor	0=no drift 1=drift before 10 seconds 2=falls before 10 seconds 3=no effort against gravity 4=no movement	_____
5b.	Right Arm Motor	0=no drift 1=drift before 10 seconds 2=falls before 10 seconds 3=no effort against gravity 4=no movement	_____
6a.	Left Leg Motor	0=no drift 1=drift before 5 seconds 2=falls before 5 seconds 3=no effort against gravity 4=no movement	_____
6b.	Right Leg Motor	0=no drift 1=drift before 5 seconds 2=falls before 5 seconds 3=no effort against gravity 4=no movement	_____
8.	Sensory	0=normal 1=abnormal	_____
9.	Language	0=normal 1=mild aphasia 2=severe aphasia 3=mute or global aphasia	_____
11.	Neglect	0=normal 1=mild 2=severe	_____
		Score (out of 31):	_____

* Scoring from Original Scale

Obtained from Meyer et al., 2002

Appendix 4.20FOUR Score (Full Outline for UnResponsiveness)



EYE RESPONSE

4=Eyelids open or opened, tracking or blinking to command

3= Eyelids open but not to tracking

2=Eyelids closed but opens to loud voice

1=Eyelids closed but opens to pain

0=Eyelids remain closed with pain stimuli

MOTOR RESPONSE

4=Thumbs up, fist, or peace sign

3=Localizing to pain

2=Flexion response to pain

1=Extension response

0=No response to pain or generalized Myoclonus status

BRAINSTEM REFLEXES

4=Pupil and corneal reflexes present

3=One pupil wide and fixed

2=Pupil or corneal reflexes absent

1=Pupil and corneal reflexes absent

0=Absent pupil, corneal, or cough reflex

RESPIRATION

4=Regular breathing pattern

3=Cheyne-Stokes breathing pattern

2=Irregular breathing

1=Triggers ventilator or breathes above ventilator rate

0=Apnea or breathes at ventilator rate

Appendix 4.21 Structural Validity Data

Structural validity data by scale showing which papers completed item test statistics and their methodological quality. It shows the populations of study, the statistical method used, the number and descriptors of factors obtained, as well as the broad purpose for completion.

Scale	Paper	Population	Method	No of Factors	Factor Descriptions	Purpose
Mathew	Bessenyai 2001 (I)	77 patients in the first week after stroke	PCA	3	**Left hemispheric signs** Orientation, Speech, paresis of the right upper and lower extremities, overall disability, and Sensation. **Right hemispheric damage signs** paresis of the left extremities. **Reflexes**	Detection of change
SSS	Bessenyai 2001 (I)	77 patients in the first week after stroke	PCA	3	**motor function on the affected side** arm, hand, and leg motor power and gait **Orientation and Speech** **Facial Palsy**	Detection of change
Unified (MCANS and SSS)	Edwards 1995 (A)	84 patients (30 ICH, 15 SAH, 15 ischaemic stroke & 24 TBI)	CFA- SCM	2 SSS 3 MCANS	**Consciousness & Motor** **Consciousness, Motor- upper & Motor-lower**	Reliability and construct validity
NIHSS	Bessenyai 2001 (I)	77 patients in the first week after stroke	PCA	5	**Signs of extended damage** LOC and Gaze **Dominant hemispheric cortical signs** response to questions and the language items **Motor performance on the affected side** Motor skills of arms and legs **Ataxia** ***Dysarthria***	Detection of change

Scale	Paper	Population	Method	No of Factors	Factor Descriptions	Purpose
NIHSS	Lyden 1999 (VG)	284 (Part 1) 331 (Part 2) (ischaemic stroke)	EFA CFA (excluded LOC, Facial Palsy and Ataxia)	2 4	**Left and right brain function** (construct validity) **left brain cortical and motor function and right brain cortical and motor function**, respectively (2 extra items distal motor function in the affected arm)	Validity as an outcome measure in patients treated with thrombolysis
NIHSS	Lyden 2004 (VG)	1191 large acute ischaemic stroke within 12 hrs of onset	Repeat of above then new EFA and CFA (retained Facial Palsy)	2 4	**left and right hemispheres** (goodness of fit=0.97) **left hemisphere cortical function** **right hemisphere cortical function** **left hemisphere motor function** **right hemisphere motor function**	Validation of design in large strokes
NIHSS	Millis 2007 (A)	380 left and 347 right hemisphere ischaemic stroke within 12 hours of onset	RA DIF	2	**left hemisphere** **right hemisphere**	Improving the scale's sensitivity in detecting neurologic impairment
NIHSS	Zandieh 2012 (A)	152 consecutive patients with first time ischaemic stroke	EFA	4	**Left brain lesions** Consciousness, Gaze, Sensory Right-Sided Motor impairment, Language and Dysarthria **Right brain motor function** Left arm and Leg **Left brain lesions** Facial Palsy, Language and Right Arm and Leg Motor impairment **Posterior circulation** Visual Field and Limb Ataxia	Internal structure

Scale	Paper	Population	Method	No of Factors	Factor Descriptions	Purpose
MCANS/ Orgogozo	Bessenyai 2001 (I)	77 patients in the first week after stroke	PCA	2	**proximal and distal strength and tone of the upper and lower extremities.** **Signs of extended damage** includes the Level of Consciousness and Gaze	Detection of change
mNIHSS	Lyden 2001 (VG)	291 (Part 1) 333 (Part 2) NINDS rtPA stroke trial data (ischaemic stroke)	As Lyden 1999	4	**left brain cortical and motor function and right brain cortical and motor function** removed/condensed redundant and unreliable items	Determining whether the mNIHSS behaves like the NIHSS over serial examinations

Key

PCA- Principle Component Analysis
IRT- Item Response Theory
EFA- Exploratory Factor Analysis
CFA- Confirmatory Factor Analysis
RA- Rasch Analysis

DIF- Differential Item Functioning
SCM- Structural Equation Modelling
ICH- Intracerebral Haemorrhage
SAH- Sub-arachnoid Haemorrhage
TBI- Traumatic Brain Injury

Methodological quality assessment: Very Good (VG) Adequate (A) Doubtful (D) Inadequate (I)

Appendix 4.22 Statistical methods to calculate inter-rater reliability

The statistical methods used to calculate inter-rater reliability by scale and paper.

Scale	Paper	% Agreement	Kappa	Weighted kappa	Mean kappa	Modified kappa	Observed/Expected Agreement	ICC	Other reliability coefficient
Matthew	Gelmers 1998	-	✓*	-	-	-	✓	-	-
GCS	Lee 2017	-	-	✓	-	-	-	-	-
	Wijdicks 2005	-	-	✓	-	-	-	-	-
SSS	Edwards 1995	-	-	-	-	-	-	-	✓
	Lindenstrom 1991	-	✓*	✓*	-	-	-	-	✓
CNS	Cote 1986	-	✓*	✓*	-	-	-	-	✓*
	Cote 1989	-	✓*	✓*	-	-	-	✓	-
	Nilanot 2010	-	-	-	-	-	-	✓	-
	Specogna 2013 Thesis	-	-	-	-	-	-	✓	-
HSS	Adams 1987	-	-	-	-	-	-	✓	-
NIHSS	Alasheev 2017	-	✓	✓	-	-	-	-	-
	Albanese 1994	-	✓	-	-	-	-	✓	-
	Anderson 2011	-	✓	✓*	-	-	-	✓	-
	Berthier 2012 & 2013	-	-	✓*	-	-	-	-	✓
	Boutot 2013	✓	-	-	-	-	-	-	-
	Brott 1989	-	✓*	-	-	-	-	-	-
	Cabal 2018	-	✓*	-	-	-	-	-	-
	Chapman 2016	-	-	-	-	-	-	✓	-
	Demaerschalk 2012	-	-	✓	-	-	-	-	✓
	Dewey 1999	-	✓	✓	-	-	-	✓	-
	Geisler 2019	-	-	-	-	-	-	✓	-
	Goldstein 1989	-	✓*	-	-	-	-	-	-
	Goldstein 1997	-	-	-	-	-	-	✓	-
	Govindarajan 2015	-	-	-	-	-	-	✓	-
	Gur 2007	-	-	-	-	-	-	-	✓
	Josephson 2006	-	✓	-	-	✓	✓m	-	-
	LaMonte 2004	✓	✓*	-	-	-	-	-	-
	Liman 2012	-	-	✓	-	-	-	-	-
	Lyden 1994a	-	✓	-	-	-	✓	-	-
	Lyden 2005	-	✓	-	-	-	-	✓	-
	Lyden 2009	-	✓	-	-	-	-	✓	-
	Meyer 2002	-	-	✓	-	-	-	-	-
	Meyer 2005	-	-	✓	-	-	-	✓	-
	Meyer 2008	-	-	✓	-	-	-	✓	-
	Nanri 2013	-	✓~	-	-	-	-	-	-
	Nilanot 2010	-	-	-	-	-	-	✓	-

Scale	Paper	% Agreement	Kappa	Weighted kappa	Mean kappa	Modified kappa	Observed/Expected Agreement	ICC	Other reliability coefficient
	Peters 2012	✓	-	-	-	-	-	-	-
	Schmulling 1998	-	✓*	-	✓	-	-	-	-
	Shafqat 1999	✓	-	✓*	-	-	-	✓	-
	Singer 2005	-	-	-	-	-	-	✓	-
	Specogna 2013	-	✓	-	-	-	-	✓	-
	Specogna 2013 (thesis)	-	✓	-	-	-	-	✓	-
	Wang 2003	-	-	-	-	-	-	-	✓
	Wu 2014	-	✓	-	-	-	-	✓	-
	Wu 2017	-	-	✓^	-	-	-	✓	-
MCANS	Edwards 1995	-	-	-	-	-	-	-	✓
Unified	Treves 1994	-	-	-	-	-	-	-	✓
ESS	Hantson 1994	-	-	✓	-	-	-	-	-
Japan	Gotoh 2001			✓	✓				
mNIHSS	Lyden 2001*	-	✓$	-	-	-	-	-	-
	Meyer 2002	-	-	✓	-	-	-	-	-
	Meyer 2005	-	-	✓	-	-	-	✓	-
	Meyer 2008	-	-	✓	-	-	-	✓	-
NIHSS-8	Demeestere 2017	-	-	✓	-	-	-	-	-
FOUR Score	Lee 2017	-	-	✓	-	-	-	-	-
	Wijdicks 2005	-	-	✓	-	-	-	-	-
IVBSS	Gur 2007	-	-	✓	-	-	-	-	✓
sNIHSS	Gonzalez 2011	-	-	✓*	-	-	-	✓	-
e-NIHSS	Olivato 2016	-	✓	-	-	-	-	-	-

Key

✓ Items

*- not all items scored by this method

✓ Total score

✓ Both items and total score

NR= Not reported method of calculating ICC or other correlation co-efficient

[m] calculated differently basis of modified kappa used

~ - errors in reported kappa values for Language and Extinction/Neglect

^- grouped NIHSS scores (0– 5, 6–12, and >13)

$- Qualitative distribution of kappa (compares number of items by agreement levels between the NIHSS and MNIHSS)

Appendix 4.23 Details of reliability studies

Details of inter-rater and intra-rater reliability studies by scale and paper. (Binz et al., 2013 only *paper that presented purely intra-rater reliability data).* Specifics of the population are provided alongside the number and professional groups of examiners. Information on the number of times the patients are assessed including the time interval between assessments are provided as well as the setting in which the studies are completed. Details of any training the examiners received is given before the COSMIN methodological rating of the study. The key is located on Page. 310.

Paper	Population	Stroke Type	Timing since onset	No of examiners	Prof Group	No of times asessed	Time Interval	Setting	Training Received	COSMIN Rating
Mathew										
Gelmers 1998	S	12 Isc	3 to 11 months	4	senior neurologists	OER	All same morning	Nursing Home – part of seperate multicentre study	Sub-study so may have occurred under the other study	D
GCS										
Lee 2017	SS	67 Isc 11 ICH 27 NS	ND	TNU	1 nurse 1 doctor	2	Within 1 hr	Hospital Emergency Department (ED)-research	ND	D
Wijdicks 2005	IS	29 Mix 91 NS	ND	9	3 neuro nurses, 3 residents/ fellows, 3 neurointensivists	2	Within 1 hr	Hospital-neurointensive care unit and other ICUs	One-page handout with written instructions available during each examination practice graded a few patients prior.	A
SSS										
Lindenstrom 1991	S	28 Isc 4 ICH 11 Unk	Median 3 days after stroke	2	2 senior neurologists 7 residents	OER	Within 3 hrs	Hospital	Doctors received written instructions in which the criteria for assessing each item was precisely defined	D
CNS										
Cote 1986	S	28 Isc 4 ICH 2 TIA	Most within 2-3 days	4	1 neurologist 1 resident and 2 nurses	3 or 4 OER	Within 2-4 hrs	Hospital preliminary validation study	All examiners were given identical definitions and guidelines and instructed in the practical aspects of patient evaluation.	D
Cote 1989	S *	144 104 Isc 17 ICH 36 TIA	Less than 48 hrs	2	nurses	OER	Average 1.63hrs (+-2.7)	Neurocritical care or emergency room	All nurses involved in the study were trained	D

Nilanot 2010	S *	38 Mix	ND- Acute	4	2 stroke fellows 2 residents	Twice by all examiners	First videoed others rated, all repeated 3.5 wks later	Hospital not usual clinical practice	Trained to administer scale prior to the study	I I
Specogna 2013 Thesis	S	7 ICH	Within 1st week of admission	TNU	1 nurse 1 doctor	1-3 pairs daily ratings	Within 6 hours	Uncontrolled routine clinical practice.	No formal training on either scale.	I
HSS										
Adams 1987	S *	16 Isc	12 - 96 hours (mean 22)	TNU	neurologist or neurosurgeon	2	Within 5 hours. 11 re-rated 5-15 days later	6 centers in conjunction with a pilot study	None- recommended examiners instructed to indicate unassessable tasks to minimize future discrepancies	I
NIHSS										
Alasheev 2017	S *	81 Isc 2 ICH 6 TIA 1 NS	Within 48 hrs	6	senior neurologists	2	Within 30 minutes	Neurological unit comparing bedside and remote NIHSS assessments.	Assisting nurse, no previous experience and trained in administration of NIHSS.	D
Albanese 1994	S V	6 Isc	ND	75	physicians	2	Approx 1-3 weeks after first rating	Research- Org 1072 in Acute Stroke Treatment (TOAST) trial.	Detailed instructional manual. Videotape examination of 3 patients to complete checked with physician. Could repeat.	D I
Anderson 2011	S	20 Mix	ND- nonacute / stable	TNU	physicians	2	n/a simultaneous	Telemedicine	ND	D
Berthier et al., 2012 &2013	S	28 Mix	ND- reported acute	TNU	neurologist and non-neurologist (local), junior and senior neurologist (remote)	4	n/a simultaneous	Telemedicine	ND	I
Binz 2013	S	15 Isc	ND	1	neurologist	2	ND	Telemedicine	ND	I

Boutot 2013	SS	381 Isc 184 ICH 213 NS	ND- but during admission	TNU	unknown mobile intensive care unit (MICU) and stroke unit staff	2	ND	Comparison between MICU and standard ambulance	ND	I
Brott 1989	S	24 Isc	Within 1 week	4	Neurologist examined neuro house officer, neuro and ED nurse clinicans rated	Twice by all examiners	Within 24 hrs	Research- scale development Hospital setting not stated.	States requires little training. Detailed glossary provided. Examiners developed the scale and so were experienced.	D I
Cabal 2018	S	1st 435 2nd 71 Mix	ND	TNU	1 neurologist 1 paramedic	2	ND -pre hospital	Multiple hospitals	Paramedics only educated to diagnose mild or severe hemiparesis 1st phase Internet e-learning 2nd phase webinars and examination.	I
Chapman 2016	S A	(1) 6 (2) 15 Mix	ND	(1) 2 (2) 3	vascular neurologists	2 or 3	n/a simultaneous	Two prehospital settings	ND	I
Demaersc halk 2012	SS	47 Isc 7 ICH 9 TIA 37 NS	ND	2	vascular neurologists assisted by medical bedside aide	OER	n/a simultaneous	Telemedicine	Approximately 15 minutes on the use of the iPhone technology Wi-Fi connectivity and FaceTime functionality.	D
Dewey 1999	S	30 Mix	20 within 10 days, 10 1-9 mnths	4	2 neurologists and 1 of 2 research nurses	3 OER	Within range 2.7- 217.5 hours	Within a community based stroke incidence study	Research nurses received 1 months intensive training in stroke neurology and NIHSS administration	I
Geisler 2019	IS	90 Mix inc. NS	ND	2	2 neurologists assisted by radiology assistant	OER	n/a simultaneous assessments	Telemedicine- aboard the mobile stroke unit	ND	I
Goldstein 1989	S	20 Mix	ND- recent stroke	4	clinical stroke fellows	2 by 2 of 4 examiners	immediately after each other	Research- ongoing prospective stroke registry	ND	D

Goldstein 1997	S	4 Mix	ND	59	30 physicians and 29 non physician study cordinators	Twice by all examiners	3 months	Research- within a randomised trial of a new therapy (30 centres)	Trained using standardised videotaped patient exams. Detailed written instructions. Opportunity to discuss.	I D
Govindarajan 2015	S A	15 Mix	ND	2	vascular neurologists	OER	n/a simultaneous assessments	Telemedicine- simulated	ND	I
Gur 2007	S	(1)-18 (2)-15 Isc	Within 72 hrs	2 each sites	examiners	OER	Within 2 hrs	Hospital	NIHSS Certified	D
Josephson 2006	S V	11 Mix	ND	6268	3385 nurses 2131 doctors 63 masters 689 unknown	OER	n/a all examiners completed in own time	Certification assessment	ND assume the standard NIHSS pre-test education but no idea if taken or how many times	D
LaMonte 2004	S V	5 Mix	ND- ? thrombolysis eligible	2	stroke neurologists	Twice by each eamin er	At least 1 year between different formats.	Telemedicine- simulated	Standard NIHSS training	D D
Liman 2012	S A	12 Mix	ND	3	stroke physicians	OER	ND – 2 consecutive 3rd rated video	Telemedicine- real ambulances	ND	A
Lyden 1994a	S V	11 Mix	ND	162	trial investigators	OER	Tape 1 within 2-3 days then Tape 2 6 mnths	Training to improve consistentcy of scoring within acute therapy trial.	45min training tape outlining proper exam technique and correct scoring based on patient response.	D
Lyden 2005	S V	18 Mix	ND	112	29 nurses 38 ED/other physicians 45 neurologists	TNU	n/a all examiners completed 3-6 patients in own time	Training- replacing 10yr old training videotapes.	ND- Were provided training DVD.	D

Lyden 2009	S	18 Mix	ND	8214	2711 nurses 1889 ED /other physicians 3614 neurologists	TNU	n/a all examiners completed 3-6 patients in own time	Certification data-validation amongst general examiners	Standard training DVD.	D
Meyer 2002	S	42 Isc 3 ICH	ND- 18 acute inpatient 27 outpatient	4	stroke neurologists	2 by 2 of 4	n/a simultaneous assessments	Research- Hospital setting	ND	D
Meyer 2005	S	25 Mix	ND- chronic	4	Not clear- stroke practitioners	2 by 2 of 4	n/a simultaneous assessments	Telemedicine	ND	D
Meyer 2008	S	5 Mix	ND- mix of chronic and acute	3	2 stroke speciality fellows 1 senior stroke faculty member	2 by 2 of 3	n/a simultaneous assessments	Telemedicine (naïve practitioners)	ND	D
Nanri 2013	S V	8 Mix	ND- certification videos	TNU	neurologists	ND	ND completed in own time	Training video used within acute study.	ND	D
Nilanot 2010	S *	38 Mix	Within 48 hrs	4	2 stroke fellows and 2 internal medicine residents	twice by each rater	First videoed others rated, all repeated 3.5 wks later	Hospital not usual clinical practice	Trained to adminster both scales prior to study but no detail how	I I
Peters 2012	S	24 Mix	ND- inpatients	5	specialist registrars	2- by 2 out of 5	At least 1 hr apart	Research comparison of paper versus app NIHSS	ND	I
Schmulling 1989	S	18 Isc 4 ICH	More than 12 hrs	4	neurologists	OER	Within 90 minutes	Research to compare trained versus untrained examiners	Two examiners inexperienced in NIHSS were given the original NIHSS examination form	D
Shafqat 1999	S	20 Isc	ND- stable inpatients	2	neurologists	OER	ND but separately	Telemedicine- assisted by bedside nurse	ND	D

Singer 2005	SS	20 Mix	6hrs or less	2	neurologists	Not clear	ND	Substudy within the development and testing of the 3I-SS.	ND	I
Specogna 2013	S	38 ICH	Within 1 week	TNU	physicians and nurses (one of each for rating)	2	Less than 4 hrs	Routine clinical practice with typical examiners .	No formal training	D
Specogna 2013 Thesis	S	12 ICH	Within 1 week	TNU	nurses and doctors	1- 3 pairs of daily ratings	Within 6 hours	Uncontrolled routine clinical practice.	No formal training	D
Wang 2003	S	20 Isc	ND- acute	4	Neurologists	2 by 2 of 4	Within 1 hr	Telemedicine (more real-life network conditions).	ND	I
Wu 2014	S A	10 Mix	ND	20	10 vascular neurologists 1 stroke attending 9 stroke fellows	TNU	n/a- second assessment from video of first	Telemedicine feasibility in the field and ambulance	Training only on telemedicine	D
Wu 2017	SS	170 Mix	ND	18	8 vascular neurologists 10 fellows in vascular neurology training proramme	2	n/a simultaneous assessments-	Telemedicine- reviewed as quality improvement project	All examiners trained on the use of the telemedicine hardware and software and retrieval of CT images.	I

Unified										
Treves 1994	S	50 Mix	ND- inpatients	3	neurologists	OER	ND	Hospital setting prior to a large-scale clinical trial.	Not documented but concludes reliability can be improved by common training of examiners	I
ESS										
Hantson 1994	S *	74 Isc	Average 12.5 days (range 0-68 days	TNU	neurologists	2	Within 3 hours	Seminal Paper. Research.	ND	D I
Japan										
Gotoh 2001	S	56 Isc 6 ICH	ND- reported as new	TNU	physicians	2	ND	11 hospitals with revised scale	ND	D I
mNIHSS										
Meyer 2002	S	42 Isc 3 ICH	ND- 18 inpatient & 27 out-patients	4	stroke neurologists	2 by 2 of 4	n/a Simultaneous assessment	Research- Hospital setting	ND	D
Meyer 2005	S	25 Mix	ND- chronic	4	Not clear- stroke practitioners	2 by 2 out of 4	n/a simultaneous assessments	Telemedicine	ND	D
Meyer 2008	S	25 Mix	ND- mix of chronic and acute	3	2 stroke fellows 1 senior faculty member (bedside)	2	n/a simultaneous assessments	Telemedicine (naïve practitioners)	ND	D
NIHSS-8										
Demeestere 2017	SS	64 Isc	ND but acute on arrival to ED	TNU	Emergency medical services and stroke team	2 by 1 of each prof	Within 5 mins	ED setting real time acute patients.	EMS trained using online NIHSS training resources during a 1-week training session.	A
FOUR										
Lee 2017	SS	67 Isc 11 ICH 27 NS	ND	TNU	1 nurse 1 doctor	2	Within 1 hr of ED arrival	Hospital ED dept to test inter-rater reliability	Briefing sessions on the FOUR score. Written definition and	A

									illustrative diagrams provided in the study forms.	
Wijdicks 2005	IS	25 Mix 91 NS	ND	9	3 neuro nurses, 3 residents/ fellows, 3 neurointensivists	2	Within 1 hr	Hospital- Neuro Intensive Care Unit and other ICUs	20-minute instruction including patient videos. Practice gradings. One-page written instructions.	A
Zink 2012	IS	8 Mix 42NS	ND	TNU	nurses	Serial assessments	ND	Neurocritical Care Unit (NCCU) - detecting change.	ND	-
IVBSS										
Gur 2007	S	(1)-18 (2)-15 Isc	Within 72 hrs	2 each sites	examiners	OER	Within 2 hrs	Hospital	Repeated clinical examinations of all 11 IVBSS items and video training with illustrative cases	D
sNIHSS										
Gonzalez 2011	S A	1 Isc	ND	40	physicians	Twice by each rater	Within 5 minutes	Telemedicine	3 examiners not NIHSS certified had 5-minute training on how to administer sNIHSS	D
e-NIHSS										
Olivato 2016	S	47 Mix	Within 24 hrs	2	examiners	2	Unclear if simultaneous or seperate	Hospital Stroke Unit	ND	I

Key
ND- Not Documented

Population
S- Stroke population
* specifically selected only to be alert or drowsy
SS- Suspected stroke population
IS- Include stroke population
SA- Stroke scenarios performed by actors
SV- Videos of stroke patients

Stroke Type
(1)- first hospital setting
(2)- second hospital setting

Isc Ischaemic Stroke
ICH Intracebral Haemorrhage
Mix Mixed or Unknown Stroke Population
TIA Transient Ischaemic Attack (TIA)
NS None Stroke

Examiners / Assessments
TNU- Total Number Unknown
2 by 1 range of examiners involved
OER- once by each rater
within 5 minutes- stipulated independent

COSMIN Rating (of methodological quality)
Blackscale- Inter-rater reliability study

Greyscale- Intra-rater reliability study

VG- Very good
A- Adequate
D- Doubtful
I – Inadequate

Colour Key

	Examiners all doctors
	Examiners all nurses
	Examiners mix of doctors and other professions
	Training in scale
	Training not specifically in scale
	No training or not documented

Appendix 4.24 Correlation coefficients other than Intraclass Correlation Coefficients (ICCs) by item

Correlation coefficients other than ICCs by item and scale showing coefficient of agreement used and the paper in which reported.

Item	Scale	Result	Coefficient of Agreement	Paper
LOC	SSS	0.87	General coefficient of agreement	Birschel 2005
		0.90	Reliability Coefficient (R^2)- LISREL	Edwards 1995
		0.94	Validity Coefficient - LISREL	
	CNS	0.93	General coefficient of agreement	Birschel 2005
	NIHSS	0.83	General coefficient of agreement	Birschel 2005
	MCANS	0.85	Reliability Coefficient (R^2)- LISREL	Edwards 1995
		0.92	Validity Coefficient - LISREL	
	Unified	0.91	Kendall coefficients	Treves 1994
	ESS	0.85	General coefficient of agreement	Birschel 2005
Orientation	SSS	0.94	Reliability Coefficient (R^2)- LISREL	Edwards 1995
		0.97	Validity Coefficient - LISREL	
	CNS	0.979	Cronbach's Alpha	Cote 1986
	Unified	0.97	Kendall coefficients	Treves 1994
Speech (Dysphasia)	SSS	0.86	General coefficient of agreement	Birschel 2005
		1.00	Reliability Coefficient (R^2)- LISREL	Edwards 1995
		1.00	Validity Coefficient - LISREL	
	CNS	0.8	General coefficient of agreement	Birschel 2005
		1.00	Cronbach's Alpha	Cote 1986
	NIHSS	0.83	General coefficient of agreement	Birschel 2005
	MCANS	0.95	Reliability Coefficient (R^2)- LISREL	Edwards 1995
		0.99	Validity Coefficient - LISREL	
	Unified	0.94	Kendall coefficients	Treves 1994
	ESS	0.86	General coefficient of agreement	Birschel 2005
Speech (Dysarthria)	NIHSS	0.82	General coefficient of agreement	Birschel 2005
Eye Movements	SSS	0.91	General coefficient of agreement	Birschel 2005
		0.31	Reliability Coefficient (R^2)- LISREL	Edwards 1995
		0.56	Validity Coefficient - LISREL	
	NIHSS	0.9	General coefficient of agreement	Birschel 2005

Item	Scale	Result	Coefficient of Agreement	Paper
	MCANS	0.58	Reliability Coefficient (R^2)-LISREL	Edwards 1995
		0.61	Validity Coefficient - LISREL	
	Unified	0.87	Kendall coefficients	Treves 1994
	ESS	0.82	General coefficient of agreement	Birschel 2005
Facial Palsy	CNS	0.934	Cronbach's Alpha	Cote 1986
	MCANS	0.01	Reliability Coefficient (R^2)-LISREL	Edwards 1995
		0.12	Validity Coefficient - LISREL	
	Unified	0.93	Kendall coefficients	Treves 1994
Arm Motor	SSS	0.65	General coefficient of agreement	Birschel 2005
		0.93	Reliability Coefficient (R^2)-LISREL	Edwards 1995
		0.96	Validity Coefficient - LISREL	
	CNS	0.73	General coefficient of agreement	Birschel 2005
	MCANS	0.95	Reliability Coefficient (R^2)-LISREL	Edwards 1995
		0.96	Validity Coefficient - LISREL	
	Unified	0.97	Kendall coefficients	Treves 1994
	ESS	0.73	General coefficient of agreement	Birschel 2005
Proximal Arm Weakness	CNS	0.980	Cronbach's Alpha	Cote 1986
Distal Arm Weakness	CNS	0.969	Cronbach's Alpha	Cote 1986
Hand Power	SSS	0.92	Reliability Coefficient (R^2)-LISREL	Edwards 1995
		0.96	Validity Coefficient - LISREL	
	MCANS	0.91	Reliability Coefficient (R^2)-LISREL	
		0.96	Validity Coefficient - LISREL	
	Unified	0.96	Kendall coefficients	Treves 1994
Leg Motor	SSS	0.7	General coefficient of agreement	Birschel 2005
		0.91	Reliability Coefficient (R^2)-LISREL	Edwards 1995
		0.83	Validity Coefficient - LISREL	
	CNS	0.68	General coefficient of agreement	Birschel 2005
		0.896	Cronbach's Alpha	Cote 1986
	MCANS	0.92	Reliability Coefficient (R^2)-LISREL	Edwards 1995
		0.96	Validity Coefficient - LISREL	
	Unified	0.97	Kendall coefficients	Treves 1994
	ESS	0.63	General coefficient of agreement	Birschel 2005
Foot Dorsiflexion	MCANS	0.87	Reliability Coefficient (R^2)-LISREL	Edwards 1995

Item	Scale	Result	Coefficient of Agreement	Paper
		0.93	Validity Coefficient - LISREL	
	Unified	0.95	Kendall coefficients	Treves 1994
Upper Limb Tone	MCANS	0.78	Reliability Coefficient (R^2)- LISREL	Edwards 1995
		0.88	Validity Coefficient - LISREL	
	Unified	0.96	Kendall coefficients	Treves 1994
Lower Limb Tone	MCANS	0.78	Reliability Coefficient (R^2)- LISREL	Edwards 1995
		0.88	Validity Coefficient - LISREL	
	Unified	0.96	Kendall coefficients	Treves 1994
Gait	Unified	0.95	Kendall coefficients	Treves 1994

Edwards et al., 1995 for the SSS excluded Facial Palsy and Gait as their population was highly skewed with 91% having facial paralysis or marked paralysis and 81% were bedridden.

Appendix 4.25 Intra-rater reliability study data

Details for the papers by scale that reported intra-rater reliability including the method of calculation, brief population details, nos of examiners and time intervals as reported.

Scale	Paper	Method of calculating Intra-rater Reliability	Patient Details	No of Examiners	Time Interval
CNS	Nilanot 2010	ICC	38- only 1 of 4 examiners assessed at baseline all other ratings completed on recording of first assessment	4	3-5 weeks
NIHSS	Albanese 1994	ICC	6 videotaped patients	2	1-3 weeks
	Binz 2013	% agreement	15- assessed remotely and then at patient's bedside	1	Not documented
	Brott 1989	Mean kappa	24	4	Within a 24-hour interval
	Goldstein 1997	ICC	4 video cases	30	3 months
	LaMonte 2004	% agreement and kappa	5 same videos by Telebat (NIHSS) and then TV/VCR (mNIHSS)	2	Not documented
	Nilanot 2010	ICC	38- only 1 of 4 examiners assessed at baseline all other ratings completed on recording of first assessment	4	3-5 weeks
ESS	Hantson 1994	Kappa statistics	38 patients	Not stated	Between 1-2 hrs
Japan	Gotoh 2001	Cronbach's alpha	62 patients	2	Not documented

e-NIHSS (Olivato 2016) excluded at point of data extraction.

Appendix 4.26 Total score differences reported for the NIHSS and mNIHSS by paper

By paper shows the reported difference in points allowed or accepted in the total score between examiners without it being formally classified as measurement error.

Paper	Reported Difference (diff) in points
NIHSS	
Alasheev 2017	No more than 3 in 85.6% of patients
Anderson 2011	9 identical scores, 8 diff of one, 2 diff of two and 1 diff of three
Berthier 2012 & 2013	**Neurologists:** 9 identical scores, 4 diff two or less points **Non neurologists:** 5 identical, 6 diff of two, 5 diff of three
Chapman 2016	All within a two-point difference
Demaerschalk 2012	In 67% of assessments diff in one or less points. In 76% of assessments, the total of both scores differed by 2 or less points. (mathematical assumption in 24% of assessments more than diff of two) Remote assessments were considered equivalent to bedside assessment if the 95% limits of agreement were within three points
Geisler 2019	In 21 (23.3%) patience the MSU and remote neurologist disagreed by greater than one point in 10 patients (11.1%) diff was more than two points
Josephson 2006	7 of the 11 patients (64%) had a four or more-point difference in NIHSS score from the 5th to 95th percentile.
Peters 2012	In 2 patients (8%) there was a diff of 1 point and 2 points
Schmulling 1998	**Trained** 5 diff of three, 4 diff Two, 7 diff of one **Untrained**, 4 diff of four or more (max. ten points) Between trained and untrained examiners, the difference of total scores reached 4 or more points in 12 patients
Shafqat 1999	Examiners did not differ on any patient by > 3 points
Wang 2003	There was no difference of >3 points between bedside and remote evaluators but only agreed on 3 out of 20 patients
Wu 2014	Matching of real-time assessments occurred for 88% (30/34) of NIHSS scores by ±2 points
NIHSS and mNIHSS	
Meyer 2002	NIHSS no more than four points diff mNIHSS no more than two points diff
Meyer 2005	NIHSS no more than four points diff mNIHSS no more than three points diff
Meyer 2008	NIHSS no more than five points diff mNIHSS no more than four points diff

Appendix 4.27 Time taken to complete scales

By scale and paper this appendix shows the mean assessment time and the range of time taken to complete the assessment alongside the mode of delivery of the assessment.

Scale	Paper	Mean Assessment Time (mins)	Range (mins)	Mode of delivery
CNS	Cote 1986	-	5-10	
NIHSS	Alasheev 2017	8 Bed 6 Rem	IQR (7-9, 5-8)	Telemedicine assessment
	Anderson et al., 2011	8.45 Rem	5-15	Telemedicine assessment- calculated only on remote
	Berthier et al., 2012 & 2013	15.09 (4 examiners)	-	Combination local and remote telemedicine
	Brott 1989	6.6 ±1.3	-	Face to face one completing three observing
	Demaerschalk 2012	8.77 ±3.45 (excluded 1 min set up)	4-19	Telemedicine assessment
	Isahaya 2017	7.72 (463.2±54.2s) **Pre-training** 6.25 (374.7±64.1s) **Post-training**	-	Telemedicine assessment
	Peters 2012	5mins 45secs App 7mins 6secs Paper	-	Bedside comparisons of paper and App versions of the NIHSS.
	Shafqat et al., 1999	6.55 Bed 9.7 Rem	4-12 6- 18	Telemedicine assessment
	Wang et al., 2003	6.43 Bed 9.11 Rem	-	Telemedicine assessment (last 10 patients)
ESS	Hantson 1994	8.2	4 to 14 mins	Face to face
IVBSS	Gur 2007	5.5 ±1.5		Face to face
sNIHSS	Gonzalez 2011	2.9±0.8 Bed 3.4±0.8 VP	-	Telemedicine assessment

Key
Bed= bedside assessor
Rem= remote assessor
VP= cellular video phone

University of Central Lancashire
UCLan

Neurological Assessment Practices after Stroke

SURVEY TO EXPLORE CURRENT PRACTICE AND UNDERSTANDING OF NEUROLOGICAL ASSESSMENT IN UK STROKE UNITS

Reference Number:

Section 1 Unit Demographics

1. **Please provide your job title.**

☐	Registered Nurse	☐	Nurse Unit Manager
☐	Clinical Nurse Specialist	☐	Stroke Physician
☐	Nurse Practitioner	☐	Geriatrician in a stroke specific role
☐	Clinical Nurse Consultant	☐	Neurologist in a stroke specific role
☐	Clinical Nurse Educator	☐	Therapist (please specify below)
☐	Other (please specify below)		

_____

2. **What best describes your hospital setting?**

☐	Large tertiary hospital (takes referrals and patients from other hospitals)
☐	Non-tertiary, General, District or Community Hospital – Including an Emergency Department
☐	Non-tertiary, General, District or Community Hospital – no Emergency Department
☐	Rehabilitation or Sub-acute Hospital
☐	Other, please specify: _____

3. **How many beds does your unit have?**

In total:	
Dedicated to stroke:	

a. **Numbers of bed by type in the stroke unit:**

Type 1 (solely for the first 72 hours of care)	
Type 2 (solely for beyond first 72 hours of care)	
Type 3 (both first 72 hours of care and post 72 hours of care)	
Other (please specify)	

4. **Does your hospital provide the following specialist stroke services on site?**

	Yes	***No***
Hyper Acute Stroke Unit		
Acute Stroke Unit		

Thrombolysis		
Thrombectomy		
Carotid Endarterectomy		
Provides Telemedicine service to other hospitals:		
• If yes; 9-5 service		
Out of hours		
Uses Telemedicine service (provided from elsewhere):		
• If yes; 9-5 service		
Out of hours		
Rehabilitation unit (please specify type i.e. stroke specific, neuro or general)		
Other (please specify) (e.g. outreach to Emergency department)		

5. What was the latest score of your unit on the Sentinel Stroke National Audit Programme (SSNAP**) or Scottish Stroke Care Audit?**

_____

Section 2 Neurological Assessment/Monitoring Practices

6. **Who completes neurological assessment/ monitoring? (tick all that apply)**

	Doctors
	Physician Associates
	Nurses
	Specialist Nurses/ Nurse Practitioners
	Therapists (please specify)
	Healthcare Assistants
	Students
	Other (please specify)

7. **Which group completes neurological assessments/monitoring most regularly? (tick only the one that takes most responsibility for completion)**

	Doctors
	Physician Associates
	Nurses
	Specialist Nurses/ Nurse Practitioners
	Therapists (please specify)
	Healthcare Assistants
	Students
	Other (please specify)

8. **Are neurological observations generally completed at the same time as physiological ones?**

Yes	
No	

9. **Are neurological observations completed by the same person as the physiological observations?**

Yes	
No	

If no please briefly describe how and why:

10. **For all the following physiological observations can you indicate whether they are taken intermittently or continuously and where you record them?**

	How taken?		Where recorded?	
	Intermittent	Continuously	Paper	Electronic*
Blood pressure				
Heart Rate				
Oxygen Saturations				
Temperature				

***If electronic please specify the I.T.system make and model you use below:**

11. **Where are neurological assessments/observations documented?**

Please tick all that apply:

Dedicated neurological assessment form*	
Electronic notes system (please specify make)	
Paper based patient clinical/progress notes	
Care plan	
Clinical pathway	
Observation chart	
Not routinely documented	
Other, please specify:	

***If you use a dedicated neurological assessment form, if you are willing, please could you return a copy with the survey or email to** asrmcloughlin1@uclan.ac.uk

12. **Are results of neurological monitoring routinely included in your patient handovers?**

	Yes
	No
	Sometimes

If yes or sometimes please specify when, how and why?

13. **Does your stroke unit/ward or any of the areas where stroke or suspected stroke patients are managed (within 72 hrs of onset) have a protocol or guidelines relating to neurological assessment/monitoring?**

☐	No
☐	Yes – protocol/guideline specific to stroke patients
☐	Yes – general protocol/guideline for all patients regardless of condition
☐	Don't know

If YES, please answer the following and if willing enclose a copy of your protocol when returning the survey

To what extent do you think your stroke unit/ward adhere to the neurological assessment/monitoring protocol? (make a mark in the relevant section)

0% ***50%***

100%

14. **Which tool/scale(s) do you use for neurological assessment and monitoring? (please complete for all scales/tools listed).**

	Used?	**If Yes-** please specify when or how often (e.g. once on admission, every observation round e.t.c.)	**If No**- are you aware of the tool/scale
National Institute for Health Stroke Scale (NIHSS)	Y / N		Y / N
Modified NIHSS (mNIHSS)	Y / N		Y / N
Glasgow Coma Scale (GCS)	Y / N		Y / N
AVPU (alert, voice, pain, unresponsive)	Y / N		Y / N
Canadian Neurological Scale	Y / N		Y / N
Scandinavian Neurologic Stroke Scale	Y / N		Y / N
Standardised nursing	Y / N		Y / N

observations for stroke (SNOBS)			
FOUR Score (Full Outline of UnResponsiveness)	Y / N		Y / N

If you use a different tool/scale to those listed above please provide details below (Name of Tool, when and/or how often you use it)

_____

_____

***If you use a tool/scale not included in the table please could you supply a copy on return of the survey by post or email to** asrmcloughlin1@uclan.ac.uk

15. **Frequency of neurological observation/ monitoring for different patient groups. For each of the patient groups on the following pages please circle the most common frequency of monitoring for each of the time periods (as per example). Please note that for some patient groups the time periods may be extended.**

- **If your most typical frequency is not available to circle please write it out.**
- **If the patient group is not seen in your unit please circle n/a in the patient group column.**
- **If you do not have patients on your unit in a specific time period circle n/a in the time period row.**
- **If you discontinue neurological observation or monitoring for the patient group after a set time period please write it in the patients group column.**

So in the example table below for an ischaemic stroke patient the most common frequency of monitoring is four hourly in the first 24 hours then six hourly or four times a day for the next 24 hours (24-48 hrs) then eight hourly for the 48-72hr period and then twelve hourly after 72 hrs. Generally monitoring stops after 5 days as per handwriting in patient group column.

	0-24 hours **n/a**	**24-48 hours** **n/a**	**48-72 hours** **n/a**	**Beyond 72 hours** **n/a**
PATIENT GROUP	**Frequency**	**Frequency**	**Frequency**	**Frequency**
EXAMPLE Ischaemic Stroke **n/a** *Stop monitoring after 5 days*	15 mins 30 mins Hourly Two hourly Four Hourly (circled) Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) (circled) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) (circled) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD) (circled)

	Frequency	Frequency	Frequency
PATIENT GROUP	**0-8 hours** **n/a**	**8-16 hours** **n/a**	**16-24 hours** **n/a**
Post Thrombolysis **n/a**	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)
	24-48 hours **n/a**	**48-72 hours** **n/a**	**Beyond 72 hours** **n/a**
Post Thrombolysis **n/a**	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)

	Frequency	Frequency	Frequency
PATIENT GROUP	**0-8 hours** **n/a**	**8-16 hours** **n/a**	**16-24 hours** **n/a**
Post Thrombectomy **n/a**	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)
	24-48 hours **n/a**	**48-72 hours** **n/a**	**Beyond 72 hours** **n/a**
Post Thrombectomy **n/a**	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)

	0-24 hours **n/a**	**24-48 hours** **n/a**	**48-72 hours** **n/a**	**Beyond 72 hours** **n/a**

PATIENT GROUP	Frequency	Frequency	Frequency	Frequency
Ischaemic Stroke (without thrombolysis or thrombectomy) **n/a**	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)
Haemorrhagic Stroke (ICH) (with blood pressure alteration) **n/a**	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)

	0-24 hours **n/a**	**24-48 hours** **n/a**	**48-72 hours** **n/a**	**Beyond 72 hours** **n/a**
PATIENT GROUP	**Frequency**	**Frequency**	**Frequency**	**Frequency**
Haemorrhagic Stroke (ICH) (without blood pressure alteration) **n/a**	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)
Potential Hemi-craniectomy **n/a**	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)
	0-24 hours **n/a**	**24-48 hours** **n/a**	**48-72 hours** **n/a**	**Beyond 72 hours** **n/a**
PATIENT GROUP	**Frequency**	**Frequency**	**Frequency**	**Frequency**

Other Patient Group (please specify) **n/a**	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)	15 mins 30 mins Hourly Two hourly Four Hourly Six Hourly (QDS) Eight Hourly (TDS) Twelve hourly (BD)

16. Are there points of a patient pathway or times of day when neurological assessment/monitoring is more likely to be missed? (e.g. during ward rounds, overnight e.t.c)

Section 3 Neurological Deterioration

17. **What is that you observe in a patient that makes you aware that they have deteriorated? (i.e. is it a exact score change, specific change in condition, or another marker?)**

_____

_____

_____

18. **If neurological deterioration is noted on an assessment what actions are taken? (please tick all that apply and add detail/comments if appropriate)**

Action	If taken (tick)	Comments
None (continue routine observations)		
Additional Observations (if frequency increased please specify to what)		
Inform Senior Nurse		
Medical Review		
Additional Scan		
Treatment to alter blood pressure		
Glycaemic control		
Neuro-surgical review		
Other (please specify)		

Section 4 Experience

19. How important is it to neurologically monitor patients who have any of the following characteristics or co-morbidities? Please circle one score for each patient group.

	Not important	Slightly important	Important	Fairly Important	Very Important
Post thrombectomy	1	2	3	4	5
Post thrombolysis	1	2	3	4	5
Severe ischaemic stroke (without thrombolysis or thrombectomy)	1	2	3	4	5
Minor ischaemic stroke (without thrombolysis or thrombectomy)	1	2	3	4	5
Haemorrhagic Stroke (ICH) (with blood pressure alteration)	1	2	3	4	5
Haemorrhagic Stroke (ICH) (without blood pressure alteration)	1	2	3	4	5
Brainstem stroke	1	2	3	4	5
Over 75yrs of age	1	2	3	4	5
Diabetes	1	2	3	4	5
Cardiac Arrhythmias (inc. AF)	1	2	3	4	5
High blood pressure	1	2	3	4	5
Low blood pressure	1	2	3	4	5
Heart Failure	1	2	3	4	5
Infection	1	2	3	4	5

20. **Consider the statements below and indicate your level of agreement for each statement.**

	Strongly disagree	**Disagree**	**Unsure**	**Agree**	**Strongly agree**
Neurological assessment after stroke is a neglected area of practice	1	2	3	4	5
I am satisfied with the level of neurological assessment education provided on my ward/unit	1	2	3	4	5
I am satisfied with the neurological assessment guidelines provided on my ward/unit	1	2	3	4	5
I am satisfied with the level of neurological assessment in my ward/unit	1	2	3	4	5
I am satisfied with the guidelines for the management of neurological deterioration provided on my ward/unit	1	2	3	4	5
I am satisfied with the level of response to neurological deterioration provided to patients in my ward/unit	1	2	3	4	5

21. **For each statement please indicate your level of agreement.**

Neurological assessment/monitoring after stroke is important to:

	Strongly disagree	**Disagree**	**Unsure**	**Agree**	**Strongly agree**
Identify baseline neurological status	1	2	3	4	5
Identify change in neurological status	1	2	3	4	5
Quantify change in neurological status	1	2	3	4	5
Assist in care planning and delivery	1	2	3	4	5
For national audit data	1	2	3	4	5
Aid communication between staff in handover or emergency situations	1	2	3	4	5
Identify complications	1	2	3	4	5
Identify patients who require new or further intervention	1	2	3	4	5
Help predict patient outcome	1	2	3	4	5

22. **For each statement please indicate your level of agreement**

Staff who complete neurological assessment/monitoring are:

	Strongly disagree	**Disagree**	**Unsure**	**Agree**	**Strongly agree**
Adequately educated and trained in it	1	2	3	4	5
Assessed as competent in using scales	1	2	3	4	5
Confident the scales/tools they use are suitable for all stroke patients	1	2	3	4	5
Confident in assessing patients neurologically	1	2	3	4	5
Confident in quantifying and recording a neurological assessment	1	2	3	4	5
Confident in identifying neurological deterioration	1	2	3	4	5
Confident in knowing what to do if neurological deterioration is noted	1	2	3	4	5
Confident in delivering treatment/intervention to address neurological deterioration	1	2	3	4	5
Confident in reporting change to other colleagues	1	2	3	4	5
Confident in reporting change to patients and their families/carers	1	2	3	4	5

23. **Barriers and Facilitators to providing neurological assessment/monitoring. For each statement please indicate your level of agreement.**

Neurological assessment/monitoring after stroke is:

	Strongly disagree	**Disagree**	**Unsure**	**Agree**	**Strongly agree**
Clearly specified in national guidelines (they tell us exactly what to do)	1	2	3	4	5
Important for all stroke patients	1	2	3	4	5
Important to monitor for change in a patient's condition	1	2	3	4	5
Tailored to the individual patient	1	2	3	4	5
Is considered a priority by the stroke multidisciplinary team	1	2	3	4	5
Only for patients who are likely to or have received treatments e.g. thrombolysis, thrombectomy or neurosurgical intervention	1	2	3	4	5
Clearly understood by clinicians	1	2	3	4	5
Only needed for hemorrhagic stroke patients	1	2	3	4	5
Not always possible to complete due to other ward demands	1	2	3	4	5
Completed accurately and consistently	1	2	3	4	5
Time consuming	1	2	3	4	5
Is better left than done incorrectly	1	2	3	4	5
Hard in patients with communication or cognitive difficulties	1	2	3	4	5
More likely to be done with electronic observation systems	1	2	3	4	5
Well documented	1	2	3	4	5
More important than letting a patient sleep	1	2	3	4	5
Reassuring to patients and families	1	2	3	4	5
Clearly communicated between clinicians	1	2	3	4	5
Required to identify neurological deterioration	1	2	3	4	5
Important as we need to identify neurological deterioration quickly	1	2	3	4	5
Not as important as experience and intuition in identifying neurological deterioration	1	2	3	4	5
Supported by clear guidance on what the response should be if neurological deterioration is identified	1	2	3	4	5

Please specify other barriers and facilitators:

24. Do you think changes are needed in neurological assessment of patients after stroke?

☐ Yes (If yes – what would you change?)

☐ No (If no- why not?)

25. If change in neurological assessment was suggested what do you see as the potential barriers to those changes? (on any level from the individual completing them to organisational wide challenges)

Section 5 Training

26. What stroke-specific training do staff undertake/ receive relating to neurological assessment/monitoring. Please list all the courses in the table and complete which staff they are for, whether the courses are internal or external and the mode of delivery.
It also asks if the courses are registered on the Stroke Specific Education Framework (SSEF)?

Key: D= Doctors PA= Physician Associates Nur= Nurses SpecNur= Specialist Nurses/Nurse Practitioners
SALT= Speech & Language Therapists OT= Occupational Therapists PT= Physiotherapists HCA= Healthcare Assistants

COURSE NAME (please provide course name as accurately as possible)	**WHICH STAFF (circle all that apply- Key located above)**	**INTERNAL or EXTERNAL COURSE (circle one)**	**COURSE FORMAT (circle one)**	**SSEF REGISTERED (circle one)**
	D PA Nur SpecNur SALT OT PT HCA Other (please specify)	Internal External If external who provided training:	Face to Face Online Other (please specify)	Yes/ No/ Don't Know
	D PA Nur SpecNur SALT OT PT HCA Other (please specify)	Internal External If external who provided training:	Face to Face Online Other (please specify)	Yes/ No/ Don't Know

COURSE NAME (please provide course name as accurately as possible)	**WHICH STAFF (circle all that apply- Key located above)**	**INTERNAL or EXTERNAL COURSE (circle one)**	**COURSE FORMAT (circle one)**	**SSEF REGISTERED (circle one)**
	D PA Nur SpecNur SALT OT PT HCA Other (please specify)	Internal External If external who provided training:	Face to Face Online Other (please specify)	Yes/ No/ Don't Know
	D PA Nur SpecNur SALT OT PT HCA Other (please specify)	Internal External If external who provided training:	Face to Face Online Other (please specify)	Yes/ No/ Don't Know
	D PA Nur SpecNur SALT OT PT HCA Other (please specify)	Internal External If external who provided training:	Face to Face Online Other (please specify)	Yes/ No/ Don't Know
	D PA Nur SpecNur SALT OT PT HCA Other (please specify)	Internal External If external who provided training:	Face to Face Online Other (please specify)	Yes/ No/ Don't Know
	D PA Nur SpecNur SALT OT PT HCA Other (please specify)	Internal External If external who provided training:	Face to Face Online Other (please specify)	Yes/ No/ Don't Know

27. **Do you have any other informal training/ mentorship opportunities in relation to neurological assessment /monitoring?**

	Yes
	No

If Yes please state/describe what they are:

_____

_____

_____

28. **Do you have any competency assessments in relation to neurological assessment /monitoring?**

	Yes
	No

If Yes please state/describe what they are:

_____

_____

_____

If you have a competency assessment in relation to neurological assessment/monitoring, if you are willing and allowed, please could you return a copy with the survey or email to asrmcloughlin1@uclan.ac.uk

Thank you for taking the time to complete this survey.
Please return the survey and any of the following that the Trust are happy to share:

- policies/protocols or documents related to neurological assessment /monitoring
- dedicated neurological assessment/monitoring documentation you use
- policies/protocols or documents related to response if neurological deterioration is noted
- tools or documentation to support neurological assessment not listed in the survey
- competency assessment documents

Many of these documents may already be in the public domain but please check with the appropriate authorities that they are happy for you to share copies with the research team. These documents if not publicly available will not be shared further or referred to by the researchers without returning to the Trust for explicit written consent.
Please return the survey and any additional documents in the pre-paid provided or alternatively you can scan and email a copy to asrmcloughlin1@uclan.ac.uk
With your consent we would like to collect and securely hold your contact details for the following purposes (please tick if you consent for your data to be used for each purpose- you do not need to agree to any or can agree to a combination the choice is yours):
To be entered into the prize draw for the £50 Amazon voucher
(this data will be destroyed once the prize draw is completed) ☐

To be a named contact for receiving the results summary from the survey
(this data will be destroyed once the results summaries are circulated) ☐
If you have consented to any of the above please provide your contact details below:
Name:
Title:
Address:

e-mail:

If you have any questions, please contact Alison Mcloughlin (PhD student) directly on asrmcloughlin1@uclan.ac.uk or Tel: 01772 894950. You can withdraw consent at any time by contacting us and your details will be removed and permanently deleted. All data will be held securely in line with University procedures and all relevant data protection legislation and permanently destroyed after its intended purpose for retention is completed. It will not be used for any purpose other than those stipulated that you have consented to and will not be shared with anyone outside the study team.

UCLan's Data Protection Officer is the Information Governance Manager and they are contactable on DPFOIA@uclan.ac.uk or Tel: 01772 892561.

If you wished to complain to the supervisory authority in the UK you would need to contact the ICO (Information Commissioner's Office) all their contact details are available via https://ico.org.uk/global/contact-us/ or their helpline number is 03031231113 (local rate call).

Appendix 5.2 Letter of ethical approval from the Science. Technology, Engineering, Medicine, and Health (STEMH) ethics committee at the University of Central Lancashire

31 May 2019
Liz Lightbody/ Alison McLoughlin
School of Nursing
University of Central Lancashire

Dear Liz and Alison

Re: STEMH Ethics Committee Application
Unique Reference Number: STEMH 1018

The STEMH ethics committee has granted approval of your proposal application 'Standardised Neurological Observation Schedule for Stroke (SNOBSS)'. Approval is granted up to the end of project date*.

It is your responsibility to ensure that

- the project is carried out in line with the information provided in the forms you have submitted
- you regularly re-consider the ethical issues that may be raised in generating and analysing your data
- any proposed amendments/changes to the project are raised with, and approved, by Committee
- you notify EthicsInfo@uclan.ac.uk if the end date changes or the project does not start
- serious adverse events that occur from the project are reported to Committee
- a closure report is submitted to complete the ethics governance procedures (Existing paperwork can be used for this purposes e.g. funder's end of grant report; abstract for student award or NRES final report. If none of these are available use e-Ethics Closure Report Proforma).

Yours sincerely

pp
St John Crean
Chair
STEMH Ethics Committee

* for research degree students this will be the final lapse date

NB - Ethical approval is contingent on any health and safety checklists having been completed, and necessary approvals gained.

Appendix 5.3 Letter of ethical approval from the Health Research Authority (HRA)

Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

Dr Catherine Elizabeth Lightbody
University of Central Lancashire
Brook Building 415,
Victoria Street, Preston
PR1 7QR

Email: hra.approval@nhs.net
HCRW.approvals@wales.nhs.uk

23 October 2019

Dear Dr. Lightbody,

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**Neurological Assessment Practices after StrokePart of: Standardised Neurological OBservation Schedule for Stroke (SSNOBS)**
IRAS project ID:	**261850**
REC reference:	**19/HRA/4113**
Sponsor	**University of Central Lancashire**

I am pleased to confirm that **HRA and Health and Care Research Wales (HCRW) Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

Please now work with participating NHS organisations to confirm capacity and capability, in line with the instructions provided in the "Information to support study set up" section towards the end of this letter.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) have been sent to the coordinating centre of each participating nation. The relevant national coordinating function/s will contact you as appropriate.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?
HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?

The *"After HRA Approval – guidance for sponsors and investigators"* document on the HRA website gives detailed guidance on reporting expectations for studies with HRA and HCRW Approval, including:

- Registration of Research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics and is updated in the light of changes in reporting expectations or procedures.

Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **261850**. Please quote this on all correspondence.

Yours sincerely,
Laura Greenfield

Approvals Specialist

Email: hra.approval@nhs.net

Copy to: *Professor StJohn Crean*

List of Documents

The final document set assessed and approved by HRA and HCRW Approval is listed below.

Document	*Version*	*Date*
Confirmation of any other Regulatory Approvals (e.g. CAG) and all correspondence [University Ethics Approval]	1	31 May 2019
Confirmation of any other Regulatory Approvals (e.g. CAG) and all correspondence [University Data Protection Checklist]	1	01 September 2018
Contract/Study Agreement template [Neurological Assessment PIC Agreement WalesProtocol]	1	22 August 2019
Contract/Study Agreement template [Neurological Assessment Practices after Stroke PIC Agreement Scotland]	1	22 August 2019
Contract/Study Agreement template [Neurological Assessment Practices after Stroke PIC Agreement Northern Ireland]	1	22 August 2019
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Insurance Certificate]	1	01 August 2018
Interview schedules or topic guides for participants [Neurological Assessment Practices after Stroke Interview Schedule]	1	20 May 2019
IRAS Application Form [IRAS_Form_02072019]		02 July 2019
Letter from funder [NIHR Successful Fellowship Letter]	1	24 July 2018
Letter from sponsor [Letter of Sponsor Approval]	1	31 May 2019
Letters of invitation to participant [E-mail to confirm survey approach]	1	20 May 2019
Letters of invitation to participant [Covering letter sent with survey]	1	20 May 2019
Letters of invitation to participant [E-Mail to named contact to request distribution of Interview PISs]	1	20 May 2019
Letters of invitation to participant [Covering Letter being sent with Interview PISs]	1	20 May 2019
Non-validated questionnaire [Neurological Assessment Practices after Stroke Survey]	1	20 May 2019
Other [Statement regarding IP]	1	22 August 2019
Participant consent form [Interview Consent Form]	1	20 May 2019
Participant information sheet (PIS) [Survey PIS]	Version 3	10 October 2019
Participant information sheet (PIS) [Interview PIS]	Version 3	10 October 2019
Referee's report or other scientific critique report [Reserach programme Approval Letter]	1	21 January 2019
Research protocol or project proposal [Neurological Assessment Practices after Stroke Protocol]	1	20 May 2019
Schedule of Events or SoECAT	1	18 September 2019
Summary CV for Chief Investigator (CI) [Dr Lightbody CV]	1	11 June 2019
Summary CV for student [Alison McLoughlin CV]	1	01 June 2019
Summary CV for supervisor (student research) [Dr Lightbody CV]	1	11 June 2019
Summary CV for supervisor (student research) [Professor Dame Caroline Watkins CV]	1	07 June 2019
Summary CV for supervisor (student research) [Dr Chris Price CV]	1	01 May 2019
Summary CV for supervisor (student research) [Dr Philippa Olive CV]	1	18 June 2019
Summary of any applicable exclusions to sponsor insurance (non-NHS sponsors only) [Insurance Detail]	1	01 July 2019

IRAS project ID	261850

Information to support study set up

The below provides all parties with information to support the arranging and confirming of capacity and capability with participating NHS organisations in England and Wales. This is intended to be an accurate reflection of the study at the time of issue of this letter.

Types of participating NHS organisation	Expectations related to confirmation of capacity and capability	Agreement to be used	Funding arrangements	Oversight expectations	HR Good Practice Resource Pack expectations
1 site type	PIC activities should not commence until a PIC Agreement is in place. HRA and HCRW recommend use of the standard Participating NHS Organisation to PIC agreement available here.	HRA and HCRW recommend use of the standard Participating NHS Organisation to PIC agreement, available here.	Participating NHS organisations will not receive funding to undertake this study	Principal Investigators or Local Collaborators are not expected for the Participant Identification Centres	It is expected that patient identification at the Participant Identification Centres will be carried out by the authorised local employees and therefore, Honorary Research Contracts or Letters of Access are not expected for these sites.

Other information to aid study set-up and delivery

This details any other information that may be helpful to sponsors and participating NHS organisations in England and Wales in study set-up.
The applicant has indicated that they intend to apply for inclusion on the NIHR CRN Portfolio.

Appendix 6.1 Early Interview Schedule

Early Interview Schedule

COHERENCE- sense making

What do you think are the aims of neurological assessment and monitoring?
(area, trust, personal)

Do you think neurological assessment and monitoring is important? What do you see as the values/benefits of neurological assessment and monitoring?

Do you think that neurological assessment and monitoring improves care? How/ Why not?

COGNITIVE PARTICIPATION- relational work

Who leads on neurological assessment and monitoring practice?

Is there a shared agreement within and between different staff groups and about what neurological assessment and monitoring is for, and how it is to be used? What makes you say that?

What are the local/national guidelines/ protocols around neurological assessment and monitoring? Are they helpful? Do you use them? How/why?

How acceptable is neurological assessment and monitoring? What factors effect this?

Do you have any reservations about its use?

Do you think everyone understands and is happy what is required of them to make neurological assessment and monitoring workable?

Is there an agreed terminology/effective communication around neurological assessment and monitoring? How does this work?

How confident are you and are in neurological assessment and monitoring? Do you feel supported enough in neurological assessment and monitoring, by management, resources, training?

COLLECTIVE ACTION-operational work

What is your experience of neurological assessment and monitoring?

Where is neurological assessment and monitoring being done? Who does it? Is it consistently done? Done enough/too much?

Are there factors that help or hinder its use in clinical practice?

Are there any conflicts in neurological assessment and monitoring? Is everyone enrolled into neurological assessment and monitoring, or do some people not use it as much as others?

How do you record neurological assessment and monitoring and treatment recommendation? Is it communicated between staff and if so how?

Is neurological assessment and monitoring the same for all patients? Are there patient characteristics, co-morbidities etc that affect what we do or should do?
What training have you received? Has it been useful? Are there avenues of informal training, mentoring and skills improvement?

How is the decision to use neurological assessment and monitoring usually made? What is the procedure if deterioration is noted in a patient?

Do you have set protocols in terms of what constitutes a deterioration and what you should do?

REFLEXIVE MONITORING- appraisal work

Overall, do you think/feel neurological assessment and monitoring is effective?

Do you thinks neurological assessment and monitoring needs changing on your unit or within stroke care generally?

How would you do that? What might help or hinder that?
If No why not?

Have you identified any gaps in your training? Is there anything you would like to add to the existing training?

How acceptable do you think neurological assessment and monitoring is to patients/carers?

Appendix 6.2 Final Interview Schedule

(Colour coded to show NPT constructs through the questions. Italicised text represents prompts)

Thank you (name) for agreeing to take part in this interview. As per the information and consent form this interview will ask you about your professional experiences of using neurological assessment and monitoring in acute stroke. It will also ask you about some of the factors that could influence its use in practice. The interview is expected to last approximately 40 minutes. If at any point you need a break or want to stop just let me know. You may hear me making notes. It will just be prompts for things that I might require for further clarification later.
Before we begin, can I just check you are happy to proceed with the interview (Response) and that you are aware the interview is being audio-recorded for later transcription (Response).
Can I please ask:

- Your job title?
- How long have you worked in the stroke service?

Views/Opinions/Importance:
Can you tell me about the neurological assessment and monitoring patients on your stroke unit receive?
[Just describe what generally happens
What would you say are the main aims of neurological assessment and monitoring?
How is the decision to monitor usually made, including frequency and duration?
What is it that would reassure you that you are doing the right thing?]
How do you think neurological assessment and monitoring impacts on care?
[individual patients, unit, organisation, culture]

What factors impact on neurological monitoring in clinical practice?
[What helps or hinders its use? Individual, Ward, Trust levels]

Are there specific patient groups or characteristics that influence your assessment and monitoring practices?
What are the local/national guidelines/ protocols around neurological assessment and monitoring? *(Do you use them? How/why? Are they sufficient?)*

Common language across the stroke pathway
How do you record neurological assessment and monitoring and treatment recommendations?
[How is neurological assessment communicated between staff. Is there set paperwork? Is a patient's neurological assessment included in handover? If so how? Do you think staff understand the terminology? How would you describe the quality of communication around neurological assessment and monitoring?]

How does the multi-disciplinary team approach neurological assessment and monitoring?
[How acceptable is neurological assessment and monitoring to clinical staff? What factors effect this? What is the ethos in the team/organisation around neurological assessment and monitoring? Is there common agreement on it?]

Deterioration:
What would you see in a patient that would make you instigate further assessment and/or intervention?

If you spot a change in a patient's condition what guides your actions?
[Do you have set protocols in terms of what constitutes a deterioration and what you should do? What is the procedure if deterioration is noted in a patient? Can you make these decisions, or do you have to go to others? Who decides? How is it reviewed? What stops you?]

Support:
Do you feel supported to undertake neurological assessment and monitoring in your Trust?
[by management, resources, training]

Training:
What can you tell me about the training you have received around neurological assessment and monitoring?
[Has it been useful? Are there any gaps? Is there any further training you would like?]

Patients/Carers:
What do you think the patients/carers experience is in terms of neurological assessment and monitoring?
[Is there anything you would introduce/do differently? Are there patient or care factors that impact on assessment and monitoring e.g. tolerance, disturbance, anxiety…..?]

Overarching:
If you were going to tell someone why you do neurological assessment and monitoring what would you say?

What would you change, if anything, in relation to neurological assessment and monitoring?
[How would you change it? Why has that change not previously occurred?]

Key
Coherence
Cognitive Participation
Collective Action
Reflexive Monitoring

Appendix 7.1 Session plans for the expert group meetings.

Overview of session plans for expert group meetings to develop the SNOBSS (Standardised Neurological OBservation Schedule for Stroke). Items in ***Bold*** indicate activities for expert group members. Although actual content of sessions changed as per chapter 7 the planned content was covered in different formats with the expert group completing item ranking and other exercises outside of the group sessions.

Session 1:

Time	Activity	Methods & Resources
14.00-14.10	Brief introduction to the group and each other. Obtain consent to record sessions	PowerPoint, Teams
14.10-14.15	Project and work completed overview	PowerPoint
14.15-14.25	Outline of group focus, aims and agenda	PowerPoint, discussion for clarification
14.25-14.45	Key findings overview: • Is change warranted? • Variation in current practice outline • Current recognition of deterioration • Scales in use • Clinimetric properties overview and key points	PowerPoint, survey data, reviews data
	Comfort Break	
14.50-15.20	**What items are best to identify neurological deterioration?** Exercise to rate items previously used as essential, desirable or should be omitted from assessment to detect END in an acute setting within 72 hours of stroke.	Review data, Jamboard
15.20-15.40	**Item Hierarchy** Ranking of items selected	Jamboard, Discussion
15.40-15.55	**Is it possible to identify a single set of items for all patient groups?**	PowerPoint, Discussion, Jamboard excercise
15.55-16.00	Summary and agreement of next steps	Discussion

Session 2:

Time	Activity	Methods & Resources
14.00-14.10	Repeat of housekeeping: introductions if needed, reminder the session is being recorded and that this is an open forum for discussion.	PowerPoint, Teams
14.10-14.15	Repeat of the aims tweaked to ensure they accommodate the construct agreement from session one	PowerPoint, agreement check
14.15-15.00	**Agreement of items to be included in SNOBSS and the way that should be assessed**	PowerPoint, item ranking data, discussion

15.00-15.20	**Agreement of whether the same items should be used for all stroke patients?**	PowerPoint, survey data, clinical scenarios, discussion
15.20-15.40	**Agreement of schedule and frequency of assessment and monitoring**	PowerPoint, survey data, discussion
15.45-15.55	**Agreement of what response should be if deterioration noted**	PowerPoint, survey data, discussion
15.55-16.00	Summary and agreement of next steps	Discussion

Appendix 7.2 SNOBSS Evaluation Questions

These questions were devised to guide the evaluation of the SNOBSS with clinical teams. A PowerPoint presentation was then devised to run sessions on Microsoft Teams to cover the content:

If you were going to tell someone about your current practice in relation to neurological and physiological monitoring, what would you say? If you were going to tell someone why you do neurological monitoring what would you say?

INTRODUCE SNOBSS

If someone asked you to complete this schedule on all patients coming into your unit what would your first thought be?
What would you need to ask?
Who would complete this schedule in your service?
What would you do to ensure successful implementation of the schedule in practice?
What would be the challenges to implementing it? (freq)
Do you think this schedule will capture deterioration across the whole stroke population? Are there things that you think are important to assess that are not included? (If so, what?)
Escalation policy and frequency of monitoring will be decided locally on a per patient basis.
Who do you think should make that decision?
Do you think this schedule could be adopted into outlying areas of the stroke pathway, such as Emergency Departments?
How would you ensure consistency in assessment?
Any other comments about any aspect of neurological monitoring and the prototype schedul

Appendix 7.3 Useful Items Table

This table was sent to all members of the expert group after session one for them to decide on which items they felt were essential, desirable or should be omitted from an assessment to identify deterioration. They were also asked to provide justification for their decisions.

Which items are useful in detecting change in a patient?

Items	Essential	Desirable	Omit	Comments
Alertness				
Level of Consciousness (LOC)				
LOC Questions				
LOC Commands				
Orientation				
Vision & Sensory				
Visual Fields				
Diplopia				
Sensation				
Involuntary				
Gaze/ Conjugate Eye Deviation/ Extraocular Eye Movements				
Pupillary Abnormality				
Extinction/Neglect				
Muscle Tone				
Upper Limb Tone				
Lower Limb Tone				
Upper and lower limb asymmetry				
Plantar Reflexes				
Deep Tendon Reflexes				
Pathologic Reflexes				

Respiration				
Voluntary				
Facial Palsy				
Motor Power- Affected Arm				
Motor Power- Unaffected Arm				
Proximal Arm				
Distal Arm				
Motor-Power Affected Leg				
Motor-Power Unaffected Leg				
Proximal Leg				
Distal Leg				
Foot Dorsiflexion				
Shoulder Function				
Hand (movement, power)				
Wrist Extension				
Finger Strength				
Ataxia				
Gait/Walking				
Performance/Disability Status				
Speech/Best Language				
Dysarthria				
Dysphagia				
Dementia				
Higher cortical function (frontal, parietal)				
Other- What else could/should be added?				
Seizure Activity				

Appendix 7.4 Document to agree item order, method of assessment, and frequency of monitoring.

Document sent to the expert group after session two to agree on item, order, method of assessment and frequency of monitoring:

Items selected as essential for identifying change after stroke.
Six items were chosen as the most useful ones to identify change when repeatedly assessed at the bedside by a range of staff. This prototype of items will need further testing, but the overall aim is to develop a range of items that can be tested regularly to identify deterioration which will trigger a response.
If anyone missed the session and wants more information about why other items were excluded please contact me asrmcloughlin1@uclan.ac.uk

Suggested order of items
(please make comment or change ordering if you DO NOT AGREE) :
Level of Consciousness
Speech/ Best Language
Facial Paresis/ Palsy
Gaze
Motor Power- Arms
Motor Power- Legs

Below for each item there is a table showing a range of ways to assess each individual item. As you will see there are a wide variety in terms of the numbers of options for each item and the wording of assessments. Numbers of options is important as we want to be able to notice change. Two-point responses are generally more about presence or absence and will not detect change as well as greater options but then too many options might increase complexity and confusion and ultimately increase variability in assessment. Remember this is not about scoring items as we are used to in established scales this is a method of identifying deterioration in individual items. Please keep in mind the changes that you see in patients as you complete this exercise.

INSTRUCTIONS- What I need for each table

1. How many grading options you think there should be for that item - this can be written in or just highlighted on the table if the number of options you think appropriate is represented.
2. I want you to identify the way that you think that item should be assessed in terms of language used or style of assessment. I have purposively not included which scales these assessment options are from. I know you will recognise many but I just want you to consider the language and content across them all and highlight any that you think are useful. This does not mean that you must choose an already established assessment set you can pick

individual words or phrases from across several that you think might explain the item or process of assessment best. Feel free to add your own ideas if you would like.

3. In the final column labelled Examiner notes/training I have included relevant scale instructions please indicate by highlighting what you think are appropriate. I have also included comments from the discussion about what the group thought were important points to clarify or include in the training. Please feel free to add any additional comments/suggestions.
4. Specific questions will be bolded outside of the tables so please add comments/ delete or highlight as appropriate.

This information will allow further iteration and where appropriate guide future testing to check the best way of describing and assessing these items regularly by a range of staff.

Quick question before you start on the tables:

In terms of presentation should item grading options be listed with best be at top and worst at bottom? YES/NO

Level of Consciousness (LOC)

I have added in tables for LOC questions and commands as the assessment style and language used you might feel is more appropriate or needs testing alongside the more traditional LOC assessments.

Level of Consciousness Table

2 or 3 options	4 options	5 options	6 options	Examiner Notes/ Training
Alert **Drowsy** **fully conscious** **somnolent, can be awaked to full consciousness** **reacts to verbal command, but is not fully conscious**	**coma** **stupor** **drowsiness** **normal** **Alert** **Voice** **Pain** **Unresponsive** **Fully conscious, alert** **Sleepy, can be awakened to full consciousness**	**Fully conscious** **Lethargic but mentally intact** **Obtunded** **Stuperous** **Comatose** Eye Response: **Eyelids open or opened, tracking or clinking to command**	**alert, keenly responsive** **drowsy but can be aroused by minor stimulation to obey, answer or respond** **requires repeated stimulation to attend, or is lethargic or obtunded, requiring strong or painful stimulation to make movements** **cannot be roused by any stimulation, does react purposefully to painful stimuli**	Difference between sleeping and reduced level of consciousness very important.

	Reacts to voice / stimulus, cannot be fully conscious **Coma: no response to stimulus** **Fully conscious** **Somnolent** **Stupor** **Comatose** **Alert; keenly responsive** **Not alert; but arousable by minor stimulation to obey, answer, or respond. Not alert; requires repeated stimulation to attend, or is obtunded and requires strong or painful stimulation to make movements (not stereotyped).** **Responds only with reflex motor or autonomic effects or totally unresponsive, flaccid, and areflexic.** **Spontaneous--open with blinking at baseline**	**Eyelids open but not to tracking** **Eyelids closed but opens to loud voice** **Eyelids closed but opens to pain** **Eyelids remain closed with pain stimuli**	**cannot be roused by any stimulation, does react with decerebration to painful stimuli** **cannot be roused by any stimulation, does not react to painful stimuli**	

	To verbal stimuli, command, speech **To pain only (not applied to face)** **No response**			

Should painful stimulus application be included? YES/NO
If yes best way to apply?

Level of Consciousness- Questions Table

2 question format	Examiner Notes/ Training
Age, Month **Answers both questions correctly.** **Answers one question correctly.** **Answers neither question correctly**	We discussed using other questions in speech assessments so there could be an overlap.

Should this be included in any format? YES/NO

Level of Consciousness- Commands Table

2 tasks	6 options	Examiner Notes/ Training
Open and close eyes, grip and release non=paretic hand **Performs both tasks correctly.** **Performs one task correctly.** **Performs neither task correctly.**	**Obeys commands for movement** **Purposeful movement to painful stimulus** **Withdraws in response to pain** **Flexion in response to pain (decorticate posturing)** **Extension response in response to pain (decerebrate posturing)** **No response**	Credit for unequivocal attempt is made Only first attempt scored Comprehension of commands- link to language (e.g. stick out your tongue, close your eyes, point to the door, place hand on ear)

Should this be included in any format? YES/NO

Speech/ Best Language

3 options	4 options	5 options	Examiner Notes/ Training
Normal **Expressive Deficit** **Receptive Deficit** **impossible** **difficult** **normal**	**Normal: no communication difficulty** **Mild communication difficulty** **Moderate difficulty, no proper sentences** **Severe difficulty, 1 or 2 words or less** **no aphasia** **limited vocabulary or incoherent speech** **more than yes/no, but not longer sentences** **only yes/no or less** **Essentially no verbal output Moderate loss** **Mild loss** **Normal** **Verbal command for the patient to make a fist on the healthy side** **have the patient name an object such as a "watch"** **have the patient repeat familiar words such as "cherry blossoms"** **have the patient give his/her address and name family members** **No aphasia;** normal.	**Oriented** **Confused conversation, but able to answer questions points** **Inappropriate words** **Incomprehensible speech** **No response** **General conversation** **normal speech** **slight word-finding difficulty, conversation is possible** **severe word-finding difficulties, conversation is difficult** **only yes or no** **mute**	Assessment of functional speech Specific questions such as how are you? Where are you? What time is it? (links to orientation) or Object naming (e.g. Belt, Watch or Index Finger, Ring finger)- need to pick items that are recognisable and available) or Picture explanation (not well used) Examiners need an awareness of assessing for slurring but not formal dysphasic assessment Not included pure repetition as not considered appropriate by group.

	Mild-to-moderate aphasia; some obvious loss of fluency or facility of comprehension, without significant limitation on ideas expressed or form of expression. 2**Severe aphasia;** all communication is through fragmentary expression; great need for inference, questioning, and guessing by the listener. **Mute, global aphasia;** no usable speech or auditory comprehension.		

Facial Paresis/ Palsy

1 option	2 options	3 options	4 options	6 options	Examiner Notes/ Training
One side does not move as well as the other	**Prescence** **Absence** or **None/dubious** **Present** or **Paralysis** **Normal** or **Symmetrical** **Asymmetrical**	**Normal** **Asymmetry on forced grimaced** **Asymmetry or drooping at rest** **Normal** **Paresis** **Paralysis** **Severe weakness; drooling** **Moderate loss; asymmetry at rest**	**Intact** **Mild** **Moderate** **Severe** **Normal** symmetrical movements. **Minor paralysis** (flattened nasolabial fold, asymmetry on smiling). **Partial paralysis** (total or near-total	**Normal** **minor paralysis** (flattened nasolabial fold, asymmetry on smiling) **partial paralysis** (total or near-total paralysis of the lower face) **complete paralysis** of 1 or both sides (absence of facial movement in the upper and lower parts of the face).	Need to be aware that most people are not purely symmetrical. Ask or use pantomime to encourage the patient to show teeth or raise eyebrows and close eyes. Show teeth or smile Patient's face is examined while talking and smiling. Only the muscles in the lower half of the face are assessed

		Mild weakness; asymmetry on smiling **Normal**	paralysis of lower face). **Complete paralysis** of one or both sides (absence of facial movement in the upper and lower face).	**deficit of IX nerve** (soft palate paralysis) **deficit of XII nerve**	Score symmetry of grimace in response to noxious stimuli in the poorly responsive or non-comprehending patient. If facial trauma/bandages, orotracheal tube, tape or other physical barriers obscure the face, these should be removed to the extent possible.

Gaze

1 option	3 options	4 options	Examiner Notes/ Training
Horizontal Gaze (side to side)	**Normal conjugate movement, eyes move L & R equally** **Difficulty looking to affected side (lateral paresis)** **Eyes deviated at rest (away from affected side)** **no Gaze Palsy** **Gaze palsy present** **conjugate eye deviation** **forced** **gaze failure** **none** **Normal** **Partial Gaze Palsy, including CN III, IV, VI, INO, skew deviation**	Horizontal and vertical eye movements **normal** **partial Gaze Palsy (gaze is abnormal in 1 or both eyes, but forced deviation or total gaze paresis is not present)** **nystagmus and/or Horner's syndrome** **forced deviation or total gaze (paresis not overcome by the oculocephalic manoeuvre)** **Normal.** **Partial Gaze Palsy; gaze is abnormal in one or both eyes, but forced deviation or total gaze paresis is not present.** **Forced deviation, or total gaze paresis not overcome by the oculocephalic maneuver.** **Conjugate deviation of eyes**	Steadies the head Only horizontal Asks the patient to follow the examiner's finger. The examiner observes the resting eye position and subsequently the full range of movements by moving the finger from the left to the right, then vice versa. Voluntary or reflexive activity Establishing eye contact and then moving about the patient from side to side will occasionally clarify the presence of a partial gaze palsy.

	Total Gaze palsy or forced deviation or ophthalmoplegia **Gaze palsay, or persistent deviation** **Gaze preference, or difficulty with far lateral gaze** **Normal** **Normal** **Gaze preference or difficulty with far lateral Gaze** **Gaze Palsy or persistent deviation**	**Intact** **Mild** **Moderate** **Severe** **Normal** **median eye position, deviation to one side impossible 4** **lateral eye position, return to midline possible** **lateral eye position, return to midline impossible**	

Motor Power Arms

1 or 2 options	3 or 4 options	5 options	6 options	7 options	Examiner Notes/ Training
Arm Drift (close eyes and hold out both arms) **Abnormal- arm can't move or drifts down** **Equal** **Unequal**	ARM RAISING **impossible** **incomplete** **possible** **None** **Mild** **Significant** **Total**	Arm Outstretched **arm maintains position for 5 seconds** **arm maintains position for 5 seconds but affected hand pronates** **arm drifts before 5 seconds** **pass and maintains lower position** **arm can't maintain position but attempts to oppose gravity** **arm falls**	**Normal strength** **Contracts against resistance** **Elevates against gravity** **Gravity eliminated** **Flicker** **No movements** **No drift; limb holds 90 (or 45) degrees for full 10 seconds.**	**Normal Power** **Mild Weakness** **Moderate Weakness** **Severe Weakness** **Flexion to Pain** **Extension to Pain** **No Response** **No movement** **Trace movement only**	Sound side first Begin with none-paretic arm Scored for Proximal and Distal Limb placed Palms down 90 degrees if sitting 45 degrees if supine Use of pantomime

		Arm raising: **Normal** **straight arm, movement not full** **flexed arm** **trace movements** **no movement** **Normal** **Can raise a straight arm** **Can raise arm with flexion at the elbow** **Can move, but not against gravity** **No movements** **raises arm with normal strength** **raises arm with reduced strength (elbow straight)** **raises arm with flexion in elbow** **can move, but not against gravity** **paralysis, no movement**	**Drift; limb holds 90 (or 45) degrees, but drifts down before full 10 seconds; does not hit bed or other support.** **Some effort against gravity; limb cannot get to or maintain (if cued) 90 (or 45) degrees, drifts down to bed, but has some effort against gravity.** **No effort against gravity; limb falls.** **No movement.** **UN = Amputation or joint fusion, explain:** **No drift, limb holds for 10s** **Drift before 10s does not hit bed** **Drifts down to bed, but has some efforts against gravity** **No effort against gravity, limb falls or no movement**	**Motion without gravity only** **Moves against gravity but not against resistance** **Moderate weakness** **Mild weakness** **Positive drift of arm** **Normal**	

			Amputation, joint fusion, explain		

Motor Power Legs

1 or 2 options	3 or 4 options	5 options	6 options	7 options	Examiner Notes/ Training
Leg Drift (open eyes and lift each leg separately) **Abnormal- leg can't move or drifts down** **Equal** **Unequal**	LEG RAISING **Impossible** **gravity** **resistance** **normal** **None** **Mild** **Significant** **Total**	Leg maintained in position **leg maintains position for 5 seconds** **leg drifts to intermediate position by the end of 5 seconds** **leg drifts to bed within 5 seconds but not immediately** **leg falls to bed immediately** Leg flexing: **normal** **movement against resistance, reduced strength** **movement against gravity** **trace movements** **no movement** **Raises leg with normal strength** **[Raises straight leg with reduced strength]** **Raises leg against gravity but with bent knee** **Can move leg but not against gravity**	**Normal strength** **Contracts against resistance** **Elevates against gravity** **Gravity eliminated** **Flicker** **No movements** **No drift; leg holds 30-degree position for full 5 seconds.** **Drift; leg falls by the end of the 5-second period but does not hit bed.** **Some effort against gravity; leg falls to bed by 5 seconds, but has some effort against gravity.** **No effort against gravity; leg falls to bed immediately.** **No movement.**	**Normal Power** **Mild Weakness** **Moderate Weakness** **Severe Weakness** **Flexion to Pain** **Extension to Pain** **No Response** **No movement** **Trace movement only** **Motion without gravity only** **Moves against gravity but not against resistance** **Moderate weakness** **Mild weakness** **Positive drift of arm** **Normal**	Sound side first Begin with none-paretic leg Scored for Proximal and Distal Limb placed 30 degrees (always tested supine) Use of pantomime Close the eyes

		Paralysed, no movement **Normal** **Can raise a straight leg** **Can raise leg with flexion at the knee** **Can move, but not against gravity** **No movements**	**UN = Amputation or joint fusion, explain**: **:** **No drift, limb holds for 10s** **Drift before 10s does not hit bed** **Drifts down to bed, but has some efforts against gravity** **No effort against gravity, limb falls or no movement** **Amputation, joint fusion, explain**		

Frequency of Assessment

We did not get a lot of time to discuss this as a group and it was highlighted that we would need to explore through formal and organised testing what frequency would work best. However, it was agreed that I would share basic data on the most common frequency of assessment reported as being completed at set time points for specific patient groups so that people could consider potential frequencies.

The following table shows the most common frequencies reported after thrombolysis and thrombectomy over six specific time periods. The numbers in brackets reports the number of respondents that reported that frequency as the most common. These may seem low as the total respondent numbers were 125 but this just indicates the level of variation reported. The numbers post thrombectomy are considerably lower however as this was not applicable for lots of hospitals.

	Time period (hours)					
Patient Group	0-8	8-16	16-24	24-48	48-72	Beyond 72
Post thrombolysis	30 minutes (43)	Hourly (69)	Hourly (57)	Four Hourly (80)	Four Hourly (74)	Four Hourly (37) Six Hourly (36)
Post thrombectomy	30 minutes (10)	Hourly (16)	Hourly (13)	Four Hourly (25)	Four Hourly (27)	Four Hourly (18)

						Six Hourly (17)

The table below shows the common frequencies reported for other specific patient groups over four time periods. As above the numbers in brackets report the number of respondents that reported that frequency as the most common. Potential hemicraniectomy was not applicable at many hospitals.

Patient Group	**Time period (hours)**			
	0-24	24-48	48-72	Beyond 72
Ischaemic Stroke (without thrombolysis or thrombectomy	Four Hourly (61)	Four Hourly (76)	Four Hourly (65)	Four Hourly (37) Six Hourly (36)
Haemorrhagic Stroke (ICH) (with blood pressure alteration)	Hourly (45)	Hourly (38)	Four Hourly (58)	Four Hourly (46)
Haemorrhagic Stroke (ICH) (without blood pressure alteration)	Hourly (50)	Four Hourly (64)	Four Hourly (66)	Four Hourly (47)
Potential Hemicraniectomy	Hourly (30)	Hourly (26)	Hourly (14) Four Hourly (14)	Four Hourly (14)
Other	Four Hourly (21)	Four Hourly (20)	Four Hourly (37) Six Hourly (36)	Four Hourly (13) Six Hourly (16)

Please provide your thoughts on frequency of assessment: (This can include ideas of variations that need to be tested)
How often should it be completed?
Same for all patients or different for specific patient groups?
Reducing frequency over 72 hrs?
When would you discontinue neurological monitoring? Would this vary dependent upon patient group?
Links to physiological monitoring and NEWS scores.