

The accuracy of administrative coding: a population-based validation study of aneurysmal subarachnoid haemorrhages, lessons for neuroscience nurses

Dr, Linda Nichols¹

 [0000-0003-3711-993X](https://orcid.org/0000-0003-3711-993X)

¹ Academic University of Tasmania

Abstract

Background: Large-scale administrative data collection is more efficient than single centre, prospective designs for rare events such as aneurysmal subarachnoid haemorrhage (aSAH). However, there are limited studies reporting data validation of International Classification of Diseases (ICD) search algorithms for aSAH. Equally there are limited examples of utilising coding for neuroscience nursing research. Despite this neuroscience nurses play a critical and functional role in contributing to ICD coding. Nurses are the primary source of clinical data entry, they are at the point of care and complete critical pieces of documentation including admission, transfer and discharge forms as well as daily progress notes and forms. ICD coding is critical in the provision of a standardised and uniform language for classifying diseases, injuries and symptoms. Coding enables accurate health data tracking, comparative research between hospitals, institutions and regions and surveillance of diseases and mortality.

Methods: A population-based search of non-traumatic subarachnoid haemorrhages cases was undertaken from 2010 to 2014. The research was undertaken to fulfil the data collection for a PhD thesis. Using sequential combinations of the subarachnoid haemorrhage ICD-10 codes I60.0-I60.9, sensitivity, specificity and positive predictive values (PPV) were calculated on a three-stage basis. Individual access to discharge summary documentation, digital medical records (Gold standard) and data linkage with death registry records and ambulance records was used to validate aSAH cases using the three ICD-10 code algorithms.

Results: A total of 414 events were identified, including 282 non-traumatic subarachnoid haemorrhage admissions of which 172 (60.99%) were subsequently confirmed as aSAH. Medical record review of cases within ICD codes I60.0-I60.9 resulted in sensitivity and specificity (95% CI) of 0.90 (0.86-0.94) and 0.23 (0.16-0.31) and a PPV of 65.1%. When analysed on a population basis the sensitivity and specificity were 0.74 (0.68-0.79) and 0.25 (0.18-0.33) and a PPV of 67.4% when using 'I60.0-I60.9'. When analysing specific aneurysm morphology, the sensitivity and specificity was 0.52 (0.45-0.58) and 0.91 (0.85-0.95) with a PPV of 90.5%.

Conclusion: Applying broad administrative database search algorithms to uncommon diseases such as aSAH results in a number of inaccuracies, particularly an overestimation of results when considering the results as representative of a population or when attempting to describe specific aneurysm morphology. For nurses this highlights the importance of ensuring accuracy and completeness when documenting patient files. Whilst nurses do not generally assign ICD-10 codes, they play a vital role in ensuring that there is the necessary clinical information including (but not limited to) flowsheets, transcription of diagnostic details, narratives and care plans. This process of documentation is not only a legal requirement and part of nursing daily practice, but it plays an important role in billing, insurance claims, health service research, resource allocation, policy making and quality development as well as coding.

Keywords: Administrative data, ICD coding, aneurysmal subarachnoid haemorrhage, neuroscience nursing,

Questions or comments about this article should be directed to Dr. Linda Nichols

Email address: linda.nichols@utas.edu.au
DOI: 10.2478/ajon-2026-0008

Copyright © 2026 ANNA

Introduction

Neuroscience nurses work in some of the most intricate, highly integrated, specialised and information intense settings. Practice across neuroscience settings is constantly evolving with ongoing changes and developments related to diagnostics, treatments and pharmaceuticals. Of the neuroscience settings, neurosurgical units and practices are perhaps one of the most challenging areas to work and to document one's work. One specific disease process that requires specific, precise, accurate, unambiguous and particularly descriptive documentation to capture the complexity of the diagnosis and disease process is Aneurysmal subarachnoid haemorrhage (aSAH). aSAH is not well studied (Nichols et al., 2018; Rehman et al., 2021). Classified as haemorrhagic strokes and representing an estimated 5-11% of stroke events (Sarti et al., 1991), the age-standardised rate for aSAH is reported to be 9.99 cases per 100,000 person years (95% confidence interval [CI], 8.69-11.29) (Nichols et al., 2018). aSAH affect a significantly younger population than ischaemic stroke and has a high early fatality rate resulting in a significant loss of productive years (Feigin et al., 2003; Nichols et al., 2016). The most common cause of death associated with aSAH is the initial bleed (Rehman et al., 2021) with an estimated risk of death at or close to the time of ictus ranging from 12-16% (Huang & Van Gelder, 2002; Macpherson et al., 2011; Nichols et al., 2018). The result is that a significant proportion of individuals will die out of hospital or at smaller regional hospitals and thus will not be captured in tertiary hospital records. Prospectively studying rare diseases such as aSAH can be costly and time consuming (English(a) et al., 2016). As a result, many studies of

aSAH have been retrospective with the use of administrative databases, however the validity of these methods are rarely reported (Thigpen et al., 2015).

Prospectively studying rare diseases such as aSAH can be costly and time consuming (English(a) et al., 2016; García-Pérez et al., 2021). As a result, many studies of aSAH are undertaken retrospectively using searches of administrative databases. Case identification is achieved using searches of diagnostic coding, including the International Classification of Diseases (ICD) version 10 coding (Langelaan et al., 2017; Schwarz et al., 2019). This disease classification process involves the scientific classification of diseases through a coding process to provide a basis for clinical diagnosis and treatment (Wang et al 2023). The validity of ICD-10 is dependent on the accuracy of documentation as the key source of information influencing coding. However, the quality of documentation particularly discharge summaries is variable (Schwarz et al., 2019), with 60% of essential information missing in one study of more than 2000 discharge summaries (Langelaan et al., 2017). This can lead to inconsistencies and omissions in ICD-10 coding that could contribute to misleading findings. Validation is often restricted to small and undefined non-traumatic subarachnoid haemorrhage cohorts reported within general stroke studies resulting in questionable sensitivity and positive predictive values (PPV's) (Haesebaert et al., 2013; Tirschwell & Longstreth Jr, 2002). Applying aSAH validation from broad stroke studies is problematic given the distinct epidemiological differences between aSAH other non-traumatic subarachnoid haemorrhages and ischaemic strokes. Fundamental differences include cohort size; aSAH are generally researched in smaller

cohort sizes when compared to ischaemic stroke studies. This not only limits risk factor profiling and subsequent risk analysis (Korja et al., 2013), but also influences the validity, reliability, and generalisability of the findings. It is not to say that aSAH research is less scientifically rigorous or unethical but given the smaller cohort sizes consideration needs to be given to low statistical power, the high risk of random errors and bias, and the methodological, analytical and ethical challenges. This is further evidence for the importance of capturing a complete cohort when studying aSAH. Furthermore, reporting of results should always be reflective of the cohort structure, including the limitations.

Neuroscience nurses contributed data for this study through forms, progress notes and assessments. Nursing documentation can hold the key to transforming research, patient care, resource allocation, policy making, quality development and the operations of health care systems. It is important to note that documentation can occur either during a patient encounter or after the event (Silfen, 2006). Whilst it is best practice to document during an event, there can be delays related to workload constraints, for example progress notes are often not written until the end of a shift. Neuroscience nurses are the primary source of clinical data entry related to aSAH, data entry is also supported by emergency nurses and doctors, ambulance staff, neurosurgical and radiological doctors and clinicians. However, it is nurses who are often the first to triage or provide care and the last to document and provide details of episodes of care before handing over to the next staff, or transitioning a patient to the next phase of care. Nursing documentation plays a fundamental role in what is an accelerated shift towards data driv-

en decision making (Darach et al., 2025). Research involving nurses is integral to not only improving patient outcomes, but also healthcare quality as a whole. This research and its fundamental basis of the ICD coding is reliant on coding staff being able to discern diagnoses retrospectively from nursing and medical notes and forms (Wang et al., 2023). There is also an emphasis on coder experience and familiarity of diseases (Hernandez et al., 2018; Ma et al., 2018) and in the case of aSAH the intricacies and implications of correctly coding aneurysm location, vessel, and/or region of the brain (anterior/posterior)

From a clinical perspective the diagnosis of a subarachnoid haemorrhage (SAH) is fairly unambiguous, however identifying an exact cause and the underlying pathology together with differentiating aneurysmal from non-aneurysmal causes can be challenging. While ruptured saccular aneurysms account for approximately 85% of spontaneous SAH cases, up to 15% are non-aneurysmal (NASAH), arising from other vascular issues, venous bleeding, or mimic conditions (Hasan et al., 2018; Roman-Filip et al., 2023). Obtaining a correct diagnosis with any rare disease can be challenging, and a lengthy process and can vary between settings (Boycott et al., 2017). This is reflected in the reported incidence of aSAH, which varies significantly from 3-25 per 100,000 (de Rooij et al., 2007). It is advocated to utilise multiple overlapping sources to enable an accurate estimate of aSAH incidence (Nichols et al., 2018). Several validation studies of subarachnoid haemorrhage have been conducted (English(a) et al., 2016; English(b) et al., 2016; Kirkman et al., 2009; Washington et al., 2014). However, these studies have varied greatly in regard to their study populations and cohort inclusion,

with a general focus on combining aneurysmal with non-traumatic subarachnoid haemorrhages. Whilst aSAH represent up to 85% of non-traumatic subarachnoid haemorrhage hospital admissions (Risselada et al., 2011) previous studies have not undertaken specific validation for aSAH as a single entity.

The catastrophic, haemorrhagic and life threatening nature of aSAH results in the need for endovascular or surgical care that is only available at tertiary hospitals (Nguyen et al., 2022; Nichols et al., 2020). Inter-hospital transfer rates for aSAH are significantly higher for aSAH than ischaemic stroke, they are influenced by a multiplicity of delays in reaching a tertiary neurosurgical hospital that can treat them (Doukas et al., 2019; Goertz et al., 2021; Nichols et al., 2020). Optimal treatment time post aSAH is 12 hours post ictus, noting that there is a decreasing evidence benefit when treatment occurs 12-24 hours post ictus (Buscot et al., 2022). aSAH patients suffer significantly higher early fatality rates when compared to other stroke events. A high early fatality rate poses significant challenges when researching aSAH, with Nichols et al. (2021) reporting 54% of their cohort dying within 24 hours of ictus or initial bleed. This is significant given the large proportion of individuals who die out of hospital or possibly at a regional centre and are not accounted for in tertiary hospital-based data such ICD coding. Of those who reach hospital, it is estimated that a further 16% of individuals will die (Buscot et al., 2022). In-hospital death significantly impacts administrative coding by shifting the focus toward mortality reporting, affecting quality metrics, and increasing the need for precise documentation of comorbidities. In-hospital deaths are highly sensitive to documentation errors, which can affect the accuracy

of the underlying cause-of-death coding (Ben-Tovim et al., 2010; Peng et al., 2017).

Understanding the accuracy of data collection methods and search algorithms is fundamental to establishing effective health care policy as well as leveraging towards the inclusion of aSAH in long-term surveillance programs. This includes exploring differences in the quality and quantity of data included in discharge summaries, incorporating an approach described by McCormick et al. (2015) as a gold-standard chart review and modern data linkage with death registry data. Validation of aSAH search algorithms across both regional and neurosurgical admissions, in respect to both causative aneurysm morphology and admission type, is essential in developing a valid and accurate case definition for future research. As we usher in a new era of digital healthcare it is fundamental that this occurs with sufficient training and education that optimises nursing roles and their contributions to data collection and use in research. This study aims to address the absence of ICD-10 coding validation pertaining specifically to aSAH. The object of this study is to validate aSAH search algorithms following a non-traumatic subarachnoid haemorrhage administrative data search using ICD-10 coding.

Methods

Study Setting

This study was set in Tasmania, an island state in the southeast of Australia. Tasmania has an overall low population density (7.24 people / km²), with the population of 511,200 (2011) relatively decentralised with 65% of individuals residing in inner regional areas, 33% in outer regional areas and 2% residing

in remote and very remote areas. Tasmania has one single public neurosurgical unit with all regional and private aSAH admissions referred to this unit, limiting the potential to miss cases and reducing referral bias.

Data Sources

This study includes a Tasmanian non-traumatic subarachnoid hemorrhage cohort identified over a five-year period from 2010-2014 (Nichols et al., 2018). Multiple overlapping data sources were used including a statewide hospital administrative data search of regional and neurosurgical discharge and emergency databases using primary and secondary ICD 10 codes I60.0-I60.9 (Table 1). Events were linked using state-wide individual healthcare identifying numbers as well as correlating the date and time of admissions. Early out of hospital deaths were identified through data linkage with the Registry of Births, Deaths and Marriages, to identify and then verify all aSAH deaths that were not admitted to hospital and missed through initial administrative data searches.

Identification of aSAH Cases

Individual digital medical records were accessed through a digital medical records system. Following guidance and training from an experienced coder and removal of any links or reference to the ICD_10 coding, records were

independently and blindly reviewed by a nurse who has clinical neurosurgical expertise. Training included both education in ICD coding as well as training in its application. Previous experience working on a neurosurgical unit also enabled the author to evaluate for the presence or absence of necessary documentation and the completeness of the clinical record. Once the diagnosis of a subarachnoid haemorrhage was established, the causative factors including, where applicable, the exact aneurysm morphology were noted and coded. A Neurosurgical Registrar contributed to the auditing, with discrepancies examined and resolved by discussion with experienced coding staff. The review of records and coding also took place on a regular basis. A percentage of all expert coder's records were reviewed by a physician-peer evaluator, in keeping with standard international practice (Silfen, 2006). When completed, the data were combined with the original ICD-10 coding undertaken by the Health Department coding staff.

Records were assessed for completeness, discrepancies, accuracy and reliability. A valid aSAH was defined as a non-traumatic subarachnoid haemorrhage identified as having a definite or probable rupture of a cerebral aneurysm demonstrated on imaging (computed tomography or digital subtraction angiography) and not associated with major trauma or surgical complications. Subarachnoid

TABLE 1: ICD-10 CODES AND DESCRIPTIONS.

I60.0 – I60.5	Non-traumatic Subarachnoid haemorrhage from specified anatomical locations including carotid siphon and bifurcation, middle cerebral, anterior communicating, posterior communicating, basilar and vertebral arteries.
I60.6	Subarachnoid haemorrhage from other intracranial arteries.
I60.7	Subarachnoid haemorrhage from unspecified intracranial artery.
I60.8	Other subarachnoid haemorrhage.
I60.9	Subarachnoid haemorrhage, unspecified.

haemorrhages caused by arteriovenous malformations, extensions of intracranial haemorrhages or with a likely perimesencephalic haemorrhage pattern were excluded. Additional admissions were identified through a ward-based search that helped to identify missing and regional admissions. Additional cases captured through the Registry of Births Deaths and Marriages that were never admitted to hospital were considered as missed cases, as they are generally not considered in hospital-based validation studies. These early and sudden deaths were identified through coding linked to autopsy reports. A list of known cases was provided to the data linkage unit; this list was compared to all causes of death from non-traumatic subarachnoid haemorrhage for the same time period and a de-identified list of people who died from aSAH was created. All deaths linked to significant co-morbidities or with suggestions of a traumatic event were excluded.

Individual cases were considered as correctly coded if the ICD-10 coding matched the documented diagnosis that had been recorded during each individual admission by either a neuroradiologist or neurosurgeon. Two methods were used; firstly, discharge summary documentation was reviewed and coded using the same method that is used by administrative coding staff. To assess the accuracy of the discharge summaries, documentation held in the individual digital medical record was accessed and coded as the gold standard. Three commonly used sequential ICD-10 code algorithms were tested including I60.0-I60.5 that focused on the range of code that specifically linked to aneurysmal locations and dichotomized as anterior and posterior haemorrhages. The second algorithm used the codes I60.0-I60.8, to include haemorrhages

coded from other intracranial arteries, unspecified intracranial arteries and other subarachnoid haemorrhages. The third algorithm I60.0-I60.9 undertook a broad range of codes including the previously described codes and subarachnoid haemorrhages unspecified. The three coding algorithms were validated using discharge summary data, gold-standard chart review and a combination of chart review and data linkage with death registry data.

Statistical Analysis

Data are displayed as counts and percentages for haemorrhage types noting that due to regional admissions and inter-hospital transfers some individuals experienced multiple admissions. Bivariate analyses were performed, and differences were assessed with chi-squared test for hospital admission type. The sensitivity, specificity and likelihood ratios and positive predictive values with associated 95% confidence intervals (CI) were calculated using the three specified algorithms. Analyses were stratified according to neurosurgical admission, a combination of regional and neurosurgical admission and on a population-basis using data linkage with the Registry of Births Deaths and Marriages. A final calculation focused on specific anterior or posterior haemorrhage location. All analyses were undertaken using the statistical software R (The R Foundation, 2013).

Ethics Approval

The Tasmanian Health and Medical Human Research Ethics Committee approved this study in December 2014 (H0014563).

Results

The initial Statewide administrative data search for non-traumatic subarachnoid haemorrhages identified 1470 events that were admitted to hospital. The removal of duplicate, rehabilitation and elective admissions as well as the exclusion of community-based deaths resulted in the final cohort including 383 admissions, representing 291 individuals, with community-based deaths included as early and sudden out of hospital deaths. Of these ICD-10 coding and full access to medical rec-

ords were available for 360 admissions representing 282 individuals (Figure 1).

A manual review of individual medical records confirmed 172 (60.99%) aSAH cases that were linked to the ICD-10 codes described (Table 2). A total of 44 regional cases were not transferred for neurosurgical care and if excluded would have resulted in 23.4% and 7.81% missing aSAH admissions using the ICD-10 code algorithms of (I60.0-I60.8) and (I60.0-I60.9) respectively (data not shown). Overall, there were no statistically significant differences in the percentage of valid aSAH

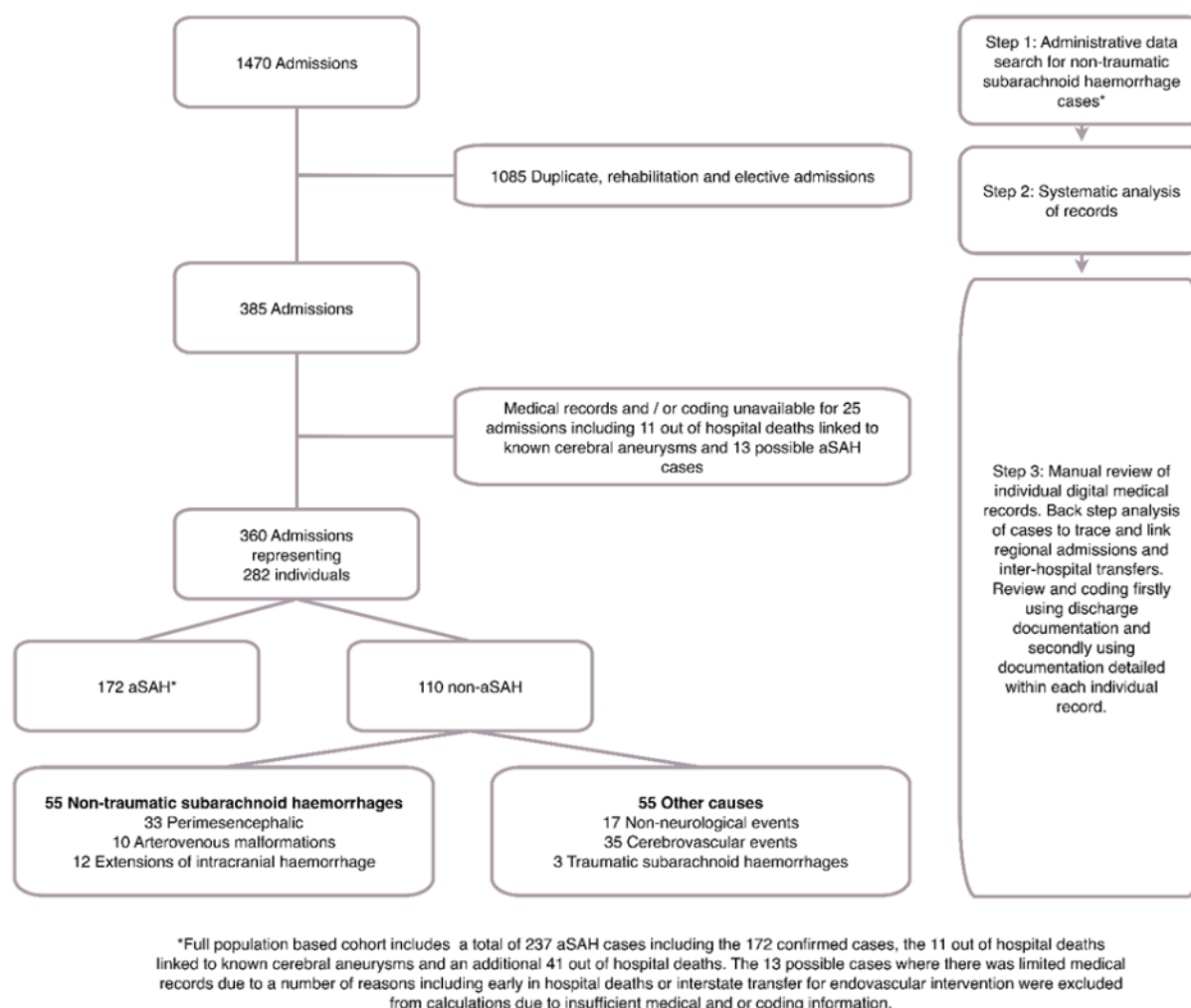


Figure 1: Flowchart describing the identification of aneurysmal subarachnoid haemorrhage cases from initial administrative data search using international classification of diseases-10 codes to final validation of cases.

Table 2: Percentage of valid aneurysmal subarachnoid haemorrhage events using International Classification of diseases-10th edition coding by hospital admissions, with ICD coding available for 360 admissions representing 282 individual cases.

	Valid (%)	Not Valid (%)	P value
Individual Admissions			
Total, n= 282	172 (60.99)	110 (39.01)	
Regional admission, n= 122 *	85 (69.67)	37 (30.33)	
Neurosurgical admission, n= 238	184 (77.31)	54 (22.69)	0.15
*78 admissions experienced an inter-hospital transfer for neurological evaluation			

events between regional and neurosurgical admissions (Table 2).

When validating the coding using the discharge summaries, the sensitivity for combined results of both neurological admissions and regional admissions ranged from 0.71 (0.65-0.77) for the codes of I60.0-I60.8 to 0.90 (0.86-0.94) when the code I60.9 was included

medical records. The combined sensitivity ranged from 0.67 (0.60-0.73) for the codes of I60.0-I60.8 to 0.90 (0.86-0.94) when using the broad range of codes including I60.0-I60.9. The PPV increased to 80.5% for the codes of I60.0-I60.8 and 67.4% for the broad range of codes (Table 4). The specificity of the broad algorithm was less than the I60.0-I60.8 algorithm using a search of both neurosurgical

TABLE 3: VALIDATION CODING UTILISING INFORMATION GATHERED FROM DISCHARGE SUMMARIES.

	Sensitivity (95% CI)	Specificity (95% CI)	Positive Predictive Value (95% CI)
Neurosurgical: I60.0-I60.8, n= 238	0.83 (0.76-0.88)	0.70 (0.59-0.80)	85.2% (78.7-90.4)
Combined: I60.0-I60.8, n= 360	0.71 (0.65-0.77)	0.77 (0.69-0.84)	83.2% (77.1-88.2)
Neurosurgical: I60.0-I60.9, n= 238	0.91 (0.86-0.95)	0.27 (0.18-0.39)	72.4% (65.7-78.4)
Combined: I60.0-I60.9, n= 360	0.90 (0.86-0.94)	0.23 (0.16-0.31)	65.1% (59.5-70.4)

in a broader algorithm. The PPV was 83.2% for the code range of I60.0-I60.8 and dropped to 65.1% for the broad range of codes, with a noticeably lower specificity associated with the broad range of codes when compared to the I60.0-I60.8 algorithm (Table 3). A gold standard manual review of the individual medical records identified decreased sensitivity and increased positive predictive values when compared to the validation using the digital

records and a combined search using regional records. In addition to the 172 aSAH events identified, an additional 54 out of hospital deaths were identified of which 52 were confirmed as aSAH after the exclusion of cases linked to multiple comorbidities and traumatic events. When considering aSAH on a population-level, the sensitivity for the broad range of codes was 0.74 (0.68-0.79) with a PPV of 67.4%.

TABLE 4: VALIDATION CODING UTILISING GOLD STANDARD INFORMATION GATHERED FROM INDIVIDUAL MEDICAL RECORDS.

	Sensitivity (95% CI)	Specificity (95% CI)	Positive Value (95% CI)	Predictive
Neurosurgical: I60.0-I60.8, n= 238	0.82 (0.75-0.87)	0.65 (0.54-0.76)	82.1% (75.1-87.7)	
Combined: I60.0-I60.8, n= 360	0.67 (0.60-0.73)	0.72 (0.63-0.79)	80.5% (74.2-85.9)	
Neurosurgical: I60.0-I60.9, n= 238	0.91 (0.85-0.95)	0.26 (0.17-0.37)	70.4% (63.7-76.6)	
Combined: I60.0-I60.9, n= 360	0.90 (0.86-0.94)	0.24 (0.17-0.32)	67.4% (61.9-72.6)	
Population: I60.0-I60.8, n= 414	0.54 (0.48-0.60)	0.72 (0.64-0.80)	80.5% (74.2-85.9)	
Population: I60.0-I60.9, n= 414	0.74 (0.68-0.79)	0.25 (0.18-0.33)	67.4% (61.9-72.6)	

The validation of specific aneurysm morphology using the definition of anterior and posterior resulted in a sensitivity of 0.68 (0.59-0.75) when analysing neurosurgical admissions and 0.52 (0.45-0.58) when analysing combined neurosurgical and regional admissions, with the positive predictive values of 91.9% and 90.5% respectively (Table 5).

Discussion:

Administrative data is a valuable research resource for population-based research (Cavallaro et al., 2023). In this study we found less than optimal accuracy of ICD-10 coded aSAH when validated using both the discharge summaries and a gold standard review of individual medical records. Only one algorithm of I60.0-I60.8 exceeded the 80% PPV cut-off used for suitability in scientific research

(Borkar et al., 2019). Applying broad level stroke results to smaller subgroups has been previously identified as unreliable (Haesebaert et al., 2013; Thigpen et al., 2015; R. Woodfield et al., 2015). It has been recommended to consider both the stroke subtype and research question when selecting a search algorithm and reporting results (Tirschwell & Longstreth Jr, 2002). Our data demonstrates variable results when applied specifically to aSAH, including the poor capture of specific aneurysm morphology. This study also demonstrated that whilst a validated search strategy can improve the probability of identifying cases, reviewing individual records is required to confirm the diagnosis (English(a) et al., 2016), as demonstrated by the improved results when the full medical records were analysed compared to the discharge summaries alone.

TABLE 5: VALIDATION CODING OF SPECIFIC HAEMORRHAGE LOCATION UTILISING GOLD STANDARD INFORMATION GATHERED FROM INDIVIDUAL MEDICAL RECORDS.

	Sensitivity (95% CI)	Specificity (95% CI)	Positive Value (95% CI)	Predictive
Neurosurgical, n= 238	0.68 (0.59-0.75)	0.90 (0.81-0.95)	91.9% (85.2-96.2)	
Combined, n= 360	0.52 (0.45-0.58)	0.91 (0.85-0.95)	90.5% (84.0-95.0)	

This study demonstrated that in both validation sets the algorithm of ICD-10 codes I60.0-I60.8 had a lower sensitivity, but higher specificity. When selecting algorithms, it is imperative to limit false positive results through using a highly specific search (Maxim et al., 2014). In this study the risk of the reduced specificity associated with the broad algorithm I60.0-I60.9 is evidenced by an over-estimation of results. From an epidemiological point of view, the use of data sets obtained through ICD-10 search algorithms present a number of challenges. A general consensus is that results are more likely to be over-estimated than under-estimated (Kirkman et al., 2009; Krarup et al., 2007). The application of stroke, or more specifically subarachnoid haemorrhage, validation to an aSAH cohort is far from precise. Many studies have inadvertently included a multitude of subarachnoid haemorrhage events, related to the coding selected through a failure to validate search algorithms that rely on unspecified coding (English(a) et al., 2016; Lai & Morgan, 2012). The implicit assumption that the ICD-10 code I60.9 is representative of aSAH cases is fraught with potential inaccuracy, given the PPV of 67.4% (61.9-72.6) and low specificity. When our validation was extended to include the code I60.9, there was evidence of a potentially gross over-estimate of results.

The identified prevalence of aSAH cases can vary significantly depending on the primary focus of the search strategy and is often not considered when reporting results. It should be noted that sensitivity and specificity can still be relatively high with a low PPV if the prevalence is low (Ranganathan & Aggarwal, 2018). In their broad stroke validation study [Hall et al. \(2016\)](#) failed to explore the prevalence of subarachnoid haemorrhages when

reporting that subarachnoid haemorrhages demonstrated the highest sensitivity yet the lowest PPV in their study. In contrast our study was limited to a smaller representative non-traumatic subarachnoid haemorrhage search strategy. Validating a broader search strategy would be impracticable as our reported five-year results represented 72 combined aSAH emergency presentations per year, equating to 1.4 per 10,000 emergency department presentations, calculated using published data for the same time period and location (Morley et al., 2018). Routine screening of all admissions, or even all stroke admissions, to confirm aSAH cases would be especially problematic. This is illustrated by a scenario where relatively accurate coding with a sensitivity and specificity of 80%, and a positive likelihood ratio of 8, yields a probability that admission with the correct coding for aSAH would only be 0.08% and unrealistic in regard to validating cases. However, as PPV increases with increasing prevalence (Ranganathan & Aggarwal, 2018), there is a risk of bias associated with an over-estimation of PPV's due to the high prevalence within subarachnoid haemorrhages or stroke samples compared to the general hospital populations (English(a) et al., 2016; Rebecca Woodfield et al., 2015). Herein lies the importance of having clear and consistent descriptions of both the original search strategy and search algorithm when validating rare diseases such as aSAH as well as reporting sensitivity, specificity and PPV.

The specialised neurosurgical care and interventions required following an aSAH often require inter-hospital transfers from regional hospitals (Yiannakoulis et al., 2003) resulting in the potential for duplication. Previous studies have also acknowledged a significant over-estimation due to the inability to distinguish

inter-hospital transfers from new admissions (Spolaore et al., 2005). The use of gold-standard access to individual medical records enabled what Brennan et al. (2012) described as inevitable data entry errors to be assessed, as well as facilitating the tracking of individual cases through multiple admissions. Regional admissions are also at risk of being underestimated, with 44 (36.07%) of regional admissions in this cohort not being transferred for neurosurgical evaluation. The sensitivity and PPV associated with a combined search of regional and neurosurgical admissions were also less than that of the neurosurgical hospital alone, which indicated a more generalized coding approach within the regional hospitals. Despite this the homogenous nature of regional coding is an important consideration when validating cases. (Thigpen et al., 2015) surmised that it is reasonable to accept that the accuracy of coding is considerably less in non-specialised centres. Coding staff in regional settings often work alone and undertake a broad range of coding, resulting in more general coding results (Yiannakoulis et al., 2003) Documentation precision, limited neurological expertise and clinical interest in coding have also been identified as contributing factors to inaccurate coding (Abo-Elboulé et al., 2013; Haesebaert et al., 2013; Mazzali & Duca, 2015; McCormick et al., 2015; Williams & Mann, 2002). However, admission outcomes also need to be considered. The focus within a regional hospital following an aSAH often concerns diagnosis and transfer, compared to a neurosurgical centre where there is an imperative to refer specifically to an aneurysm location in reference to surgical and endovascular interventions. Admission outcome is a significant factor influencing the care and discharge documentation and subsequent ICD-10 coding.

Administrative datasets are commonly used to provide population characteristics, including the reporting of incidence, prevalence and temporal trends of specific diseases (Cook & Collins, 2015; Hall et al., 2016). This is fraught with bias depending on the data set utilised, as cohorts selected by treatment, interventions or outcomes are not representative of the entire population (English(a) et al., 2016; van Walraven et al., 2016) and such evaluation is only possible when data sets represent the entire population (Mazzali & Duca, 2015). Previous studies have used a variety of data sets including hospital morbidity databases (Lai & Morgan, 2012), imaging and operative data sets (Pobereskin, 2001). Reported population characteristics linked to aneurysmal location are also problematic given the low sensitivity identified in this study. Lai and Morgan (2012) reported similar anterior/posterior rupture results to this study from the administrative data set, however the aneurysm location was only specified in 45.7% of their cases and the coding was not validated. A third factor that significantly impacts population characteristics is the failure to capture individuals who were not admitted to hospital following an aSAH. Compared to other stroke subtypes aSAH have a high early fatal rate, with 23.2% of individuals experiencing an aSAH in this cohort dying before reaching medical attention. Hospital based cohort estimates may therefore be inherently inaccurate. Previous research has validated the use of autopsy reports in the validation of subarachnoid haemorrhages with markedly high sensitivity of 100% and PPV of 100% (Tolonen et al., 2007). However, a systematic review of stroke validation identified suboptimal sensitivity in detecting fatal stroke, providing further evidence of the distinctly different presentation and likely outcome following an aSAH com-

pared with insidious events such as ischemic stroke that do not have a high early mortality rate.

Strengths and Limitations:

The retrospective nature of any study in identifying cases with rare diseases will always have inherent limitations (Rehman et al., 2021) that can only be overcome with rigorous case attainment strategies. The strength of this study is the comprehensiveness of the case attainment using individual digital medical records and the data linkage across different facilities, including regional admissions and death registry data. Unlike previous studies that have acknowledged a significant over-estimation due to the inability to distinguish inter-hospital transfers from new admissions (Spolaore et al., 2005), The use of unique statewide medical identification numbers in this study enabled individual cases to be tracked through multiple admissions. Overlapping search strategies (English(b) et al., 2016) and the manual gold standard search strategy using individual medical records over whole calendar years were also vital to maximizing case attainment (Cook & Collins, 2015) and ensuring the comparability and quality of studies (Sudlow & Warlow, 1996). This study provides validation beyond the reported text from discharge summaries where errant transcription and terminology errors occur (English(a) et al., 2016), including a second level of validation based on full access to each individual medical record. A third level of validation was undertaken using data linkage with death registry data. One limitation is that the death registry data was de-identified and unable to be validated, however given that aSAH results in significant rates of sudden and unexpected death (Nichols et al.,

2021), it would be expected that all coding would be based on autopsy results. Secondly all cases of sudden out of hospital deaths where there were significant co-morbidities coded or suggestions of a traumatic event were excluded.

A second strength of this study is the exhaustive approach taken in validating the cases, including the use of trained neurosurgical staff and methods that ensure validation was undertaken blinded to the original coding. One limitation is that the validation exercise was undertaken by a single team member, however this replicates the method of the original coding that is only undertaken by single coding staff and all cases in question were discussed and verified, as would happen in normal coding practice (Lawthers et al., 2000). Validity of administrative data is dependent on the training of staff to locate and interpret information (Hall et al., 2016), with St. Germaine-Smith et al. (2012) going as far to comment that coding is an art rather than a science. The author went to lengths to ensure that she was familiar with the main diagnosis selection and the classification of aSAH codes before undertaking the exercise. The clinical expertise of neurosurgical staff enabled the identification of the exact aneurysm morphology and location when compared to the original coding. The regular and constructive communication between the author and the clinical staff was a significant strength. It is acknowledged that data collected is from one public hospital system, however this did include four separate hospitals (three regional referral and one tertiary hospital) and a number of trained coders. This study was limited in the reporting of meaningful specificities given our estimates of non-aSAH cases were restricted to non-traumatic subarachnoid related codes. In an

attempt to limit bias, multiple predictive variables were analysed across both neurosurgical admissions and combined regional and neurosurgical admissions as well as on a population basis using differing algorithms. It is recommended that all studies use a combination of specific validation data linked to the disease process being studied as well as a sample validation to ensure the accuracy for the specific health care setting being studied.

The rapid development of information technology platforms has resulted in medical records increasingly being transferred to digitised versions (Gonzalez & Chiodo, 2015). Changes include but are not limited to; typed record keeping, the use of transcription and key-stroke-driven software, voice recognition and template driven documentation. This study was undertaken in a period prior to keystroke-driven electronic medical records (EMR) and whilst EMR may improve linking with, and coding of ICD codes. The implementation of technology including EMR is complex and does impact the depth and details recorded related to clinical nursing events (Jedwab et al., 2022). It is estimated that 40% of individuals do not receive evidenced based processes of care (Rehman et al., 2021) and there is evidence to support that prompted EMR may play a role in addressing this improving both documentation and outcomes following aSAH. Documentation standards are also influenced by limited medical and nursing training in regard to the importance of clinical structured documentation and terminology that is in alignment with coding algorithms. Overall, the transfer and digitisation of medical records have significantly improved the efficiency of utilising medical records for research purposes (Wang et al., 2023) This move has also improved overall documentation standards,

organisational performance and recognition of research importance (Janerka et al., 2024).

The importance of moving towards EMR will be evidenced in future research outcomes with the hope of more certainty in reporting results, than currently achievable. Rehman et al. (2020) and alike have currently been limited to reporting results with only a level of certainty for example terms such as 'a marginally greater risk'. Artificial intelligence is a significant changing factor that will influence the way we collect data. Studies such as this work need to be replicated in the near future using datasets that have been collected using the technology of embedded EMR with the benefits of artificial intelligence algorithms.

Conclusion

ICD-10 coding has the potential to be a powerful epidemiological tool and our findings support the notion that administrative data searches can be used to identify aSAH from a non-traumatic subarachnoid haemorrhage cohort, but with caution. The identification of the sensitivity and specificity related to aSAH are not arbitrary and will enable these results to be incorporated into subsequent statistical analysis to account for misclassification and coding discrepancies. The results of this study will substantiate and provide an important part of the epidemiological armamentarium used to reduce bias within ICD searches of administrative databases. This study demonstrates that complete cohorts of rare diseases such as aSAH can be identified, by combining administrative search strategies and data linkage to facilitate an accurate and cost-effective rare disease cohort for epidemiological research. Noting that the identification of biases are study specific and can be complex, and

are disease specific. An analysis of one's cohort and coding should be undertaken before applying results. This study highlights the importance of documentation standards for neuroscience nurses. Nurses are responsible and are the main contributors to health records, and consistent, accurate and standardised documentation is vital for the integrity of record keeping. Clinical judgement regarding documentation is a learned process and skills sets are often reflective of years of experience and or knowledge of a specialty. Education in this area is lacking, nurses generally do not think on a daily basis that their contributions are influencing staffing, resource distribution, research and the improvement of patient outcomes. Data analytics are an evolving practice, and technological advances hold considerable promise for research, and clinical care fields alike. The key to addressing documentation void is to ensure that future research is integrated on a system wide basis and that the results of studies or analyses using coded data are shared with nursing staff on all levels.

Acknowledgements:

The author would like to thank all those involved in the study including the coding staff who assisted and trained the author in the process. Thanks also to Stankovich, J, Stirling, C, and Gall S who supervised the overall aSAH PhD thesis including the data collection. I would like to thank all three of my supervisors for reviewing drafts of this publication and making comment. A special thanks to Nichols, D who assisted with the programming and the conceptualising of the statistical analysis for this paper. Nichols, D also reviewed and edited numerous iterations of this manuscript, proof reading and reviewing the work.

Conflict of interest:

None to declare.

References

- Aboa-Eboulé, C., Mengue, D., Benzenine, E., Hommel, M., Giroud, M., Béjot, Y., & Quantin, C. (2013). How accurate is the reporting of stroke in hospital discharge data? A pilot validation study using a population-based stroke registry as control. *J Neurol*, 260(2), 605–613. <https://doi.org/10.1007/s00415-012-6686-0>
- Ben-Tovim, D. I., Pointer, S. C., Woodman, R., Hakendorf, P. H., & Harrison, J. E. (2010). Routine use of administrative data for safety and quality purposes—hospital mortality. *Medical journal of Australia*, 193, S100–S103.
- Borkar, D. S., Sobrin, L., Hubbard, R. A., Kempen, J. H., & VanderBeek, B. L. (2019). Techniques for improving ophthalmic studies performed on administrative databases. *Ophthalmic Epidemiol*, 26(3), 147–149. <https://doi.org/10.1080/09286586.2018.1554160>
- Boycott, K. M., Rath, A., Chong, J. X., Hartley, T., Alkuraya, F. S., Baynam, G., Brookes, A. J., Brudno, M., Carracedo, A., & den Dunnen, J. T. (2017). International cooperation to enable the diagnosis of all rare genetic diseases. *The American Journal of Human Genetics*, 100(5), 695–705.
- Brennan, L., Watson, M., Klaber, R., & Charles, T. (2012). The importance of knowing context of hospital episode statistics when reconfiguring the NHS. *BMJ : British Medical Journal*, 344, e2432. <https://doi.org/10.1136/bmj.e2432>
- Buscot, M.-J., Chandra, R. V., Maingard, J., Nichols, L., Blizzard, L., Stirling, C., Smith, K., Lai, L., Asadi, H., Froelich, J., Reeves, M. J., Thani, N., Thrift, A., & Gall, S. (2022). Association of Onset-to-Treatment Time With Discharge Destination, Mortality, and Complications Among Patients With Aneurysmal Subarachnoid Hemorrhage. *JAMA network open*, 5(1), e2144039–e2144039. <https://doi.org/10.1001/jamanetworkopen.2021.44039>
- Cavallaro, F. L., Cannings-John, R., Lugg-Widger, F., Gilbert, R., Kennedy, E., Kendall, S., Robling, M., & Harron, K. L. (2023). Lessons learned from using linked administrative data to evaluate the Family Nurse Partnership in England and Scotland. *International journal of population data science*, 8(1), 2113.

- Cook, J. A., & Collins, G. S. (2015). The rise of big clinical databases. *Br J Surg*, 102(2), e93–e101. <https://doi.org/10.1002/bjs.9723>
- Darach, N., Kim, M. S., Wisersith, W., & Collins, E. G. (2025). Data Analytics and Administrative Decision Making in Nursing Management: A Systematic Review. *Journal of nursing management*, 2025(1), 4344–47.
- de Rooij, N. K., Linn, F. H., van der Plas, J. A., Algra, A., & Rinkel, G. J. (2007). Incidence of subarachnoid haemorrhage: a systematic review with emphasis on region, age, gender and time trends. *J Neurol Neurosurg Psychiatry*, 78(12), 1365–1372. <https://doi.org/10.1136/jnnp.2007.117655>
- Doukas, A., Barth, H., Petridis, K. A., Mehdorn, M., & von der Bröle, C. (2019). Misdiagnosis of acute subarachnoid hemorrhage in the era of multimodal diagnostic options. *The American Journal of Emergency Medicine*, 37(11), 2079–2083. <https://doi.org/https://doi.org/10.1016/j.ajem.2019.03.001>
- English(a), S. W., McIntyre, L., Ferguson, D., Turgeon, A., Dos Santos, M. P., Lum, C., Chassé, M., Sinclair, J., Forster, A., & van Walraven, C. (2016). Subarachnoid hemorrhage admissions retrospectively identified using a prediction model. *Neurology*, 87(15), 1557–1564. <https://doi.org/10.1212/wnl.0000000000003204>
- English(b), S. W., McIntyre, L., Ferguson, D., Turgeon, A., Sun, C., dos Santos, M. P., Lum, C., Sinclair, J., Forster, A., & van Walraven, C. (2016). Enriched administrative data can be used to retrospectively identify all known cases of primary subarachnoid hemorrhage. *Journal of Clinical Epidemiology*, 70, 146–154.
- Feigin, V. L., Lawes, C. M., Bennett, D. A., & Anderson, C. S. (2003). Stroke epidemiology: a review of population-based studies of incidence, prevalence, and case-fatality in the late 20th century. *The Lancet Neurology*, 2(1), 43–53.
- García-Pérez, L., Linertová, R., Valcárcel-Nazco, C., Posada, M., Gorostiza, I., & Serrano-Aguilar, P. (2021). Cost-of-illness studies in rare diseases: a scoping review. *Orphanet Journal of Rare Diseases*, 16(1), 178.
- Goertz, L., Pflaeging, M., Hamisch, C., Kabbasch, C., Pennig, L., von Spreckelsen, N., Laukamp, K., Timmer, M., Goldbrunner, R., Brinker, G., & Krischek, B. (2021). Delayed hospital admission of patients with aneurysmal subarachnoid hemorrhage: clinical presentation, treatment strategies, and outcome. *Journal of Neurosurgery JNS*, 134(4), 1182–1189. <https://doi.org/https://doi.org/10.3171/2020.2.JNS20148>
- Gonzalez, T., & Chiodo, C. (2015). ICD 10: a primer for the orthopedic foot and ankle surgeon. *Foot & Ankle International*, 36(9), 1110–1116.
- Haesebaert, J., Termoz, A., Polazzi, S., Mouchoux, C., Mechtouff, L., Derex, L., Nighoghossian, N., & Schott, A.-M. (2013). Can hospital discharge databases be used to follow ischemic stroke incidence? *Stroke*, 44(7), 1770–1774.
- Hall, R., Mondor, L., Porter, J., Fang, J., & Kapral, M. K. (2016). Accuracy of administrative data for the coding of acute stroke and TIAs. *Canadian Journal of Neurological Sciences*, 43(6), 765–773.
- Hasan, T. F., Duarte, W., Akinduro, O. O., Goldstein, E. D., Hurst, R., Haranhalli, N., Miller, D. A., Wharen, R. E., Tawk, R. G., & Freeman, W. D. (2018). Nonaneurysmal “Pseudo-Subarachnoid Hemorrhage” Computed Tomography Patterns: Challenges in an Acute Decision-Making Heuristics. *Journal of Stroke and Cerebrovascular Diseases*, 27(9), 2319–2326. <https://doi.org/https://doi.org/10.1016/j.jstrokecerebrovasdis.2018.04.016>
- Hernandez, R. K., Wade, S. W., Reich, A., Pirolli, M., Liede, A., & Lyman, G. H. (2018). Incidence of bone metastases in patients with solid tumors: analysis of oncology electronic medical records in the United States. *BMC cancer*, 18(1), 44.
- Huang, J., & Van Gelder, J. M. (2002). The probability of sudden death from rupture of intracranial aneurysms: a meta-analysis. *Neurosurgery*, 51(5), 1101–1107.
- Janerka, C., Leslie, G. D., Gallagher, O., Mellan, M., Lane, M., & Gill, F. J. (2024). Understanding research capacity and culture of nurses and midwives in two health services in Western Australia. *Collegian*, 31(3), 137–143.
- Jedwab, R. M., Manias, E., Hutchinson, A. M., Dobroff, N., & Redley, B. (2022). Nurses’ experiences after implementation of an organization-wide electronic medical record: qualitative descriptive study. *JMIR nursing*, 5(1), e39596.
- Kirkman, M. A., Mahattanakul, W., Gregson, B. A., & Mendelow, A. D. (2009). The accuracy of hospital discharge coding for hemorrhagic stroke. *Acta Neurol Belg*, 109(2), 114–119.
- Korja, M., Silventoinen, K., Laatikainen, T., Jousilahti, P., Salomaa, V., Hernesniemi, J., & Kaprio, J. (2013). Risk Factors and Their

Combined Effects on the Incidence Rate of Subarachnoid Hemorrhage – A Population-Based Cohort Study. *Plos one*, 8(9), e73760. <https://doi.org/10.1371/journal.pone.0073760>

Krarup, L.-H., Boysen, G., Janjua, H., Prescott, E., & Truelsen, T. (2007). Validity of stroke diagnoses in a National Register of Patients. *Neuroepidemiology*, 28(3), 150–154.

Lai, L., & Morgan, M. K. (2012). Incidence of subarachnoid haemorrhage: an Australian national hospital morbidity database analysis. *Journal of Clinical Neuroscience*, 19(5), 733–739.

Langelaan, M., Baines, R. J., de Bruijne, M. C., & Wagner, C. (2017). Association of admission and patient characteristics with quality of discharge letters: posthoc analysis of a retrospective study. *BMC health services research*, 17(1), 225.

Lawthers, A. G., McCarthy, E. P., Davis, R. B., Peterson, L. E., Palmer, R. H., & Iezzoni, L. I. (2000). Identification of in-hospital complications from claims data: is it valid? *Medical care*, 38(8), 785–795.

Ma, X.-L., Zhai, X., Liu, J.-W., Xue, X.-X., Guo, S.-Z., Xie, H., Chen, J.-X., Zhao, H.-H., & Wang, W. (2018). Study on the biological basis of hypertension and syndrome with liver-fire hyperactivity based on data mining technology. *World Journal of Traditional Chinese Medicine*, 4(4), 176–180.

Macpherson, K. J., Lewsey, J. D., Jhund, P. S., Gillies, M., Chalmers, J. W., Redpath, A., Briggs, A., Walters, M., Langhorne, P., & Capewell, S. (2011). Trends in incidence and in short term survival following a subarachnoid haemorrhage in Scotland, 1986-2005: a retrospective cohort study. *BMC neurology*, 11(1), 38.

Maxim, L. D., Niebo, R., & Utell, M. J. (2014). Screening tests: a review with examples. *Inhalation toxicology*, 26(13), 811–828.

Mazzali, C., & Duca, P. (2015). Use of administrative data in healthcare research. *Internal and emergency medicine*, 10(4), 517–524.

McCormick, N., Bhole, V., Lacaille, D., & Avina-Zubieta, J. A. (2015). Validity of Diagnostic Codes for Acute Stroke in Administrative Databases: A Systematic Review. *PloS One*, 10(8), e0135834. <https://doi.org/10.1371/journal.pone.0135834>

Morley, C., Stankovich, J., Peterson, G., & Kinsman, L. (2018). Planning for the future: emergency department presentation patterns in Tasmania, Australia. *International emergency nursing*, 38, 34–40.

Nguyen, T. P., Stirling, C., Kitsos, G., Nichols, L., Chandra, R. V., Rehman, S., Smith, K., Mosley, I., Lai, L., Asadi, H., Dubey, A., Froelich, J., Thrift, A. G., & Gall, S. (2022). Barriers and facilitators to more timely treatment of aneurysmal subarachnoid haemorrhage across two tertiary referral centres in Australia: A thematic analysis. *Australasian Emergency Care*, 25(3), 267–272. <https://doi.org/https://doi.org/10.1016/j.auec.2022.01.003>

Nichols, L., Gall, S., Stankovich, J., & Stirling, C. (2021). Associations between socioeconomic status and place of residence with survival after aneurysmal subarachnoid haemorrhage. *Internal Medicine Journal*, 51(12), 2095–2103.

Nichols, L., Stirling, C., Otahal, P., Stankovich, J., & Gall, S. (2018). Socioeconomic Disadvantage Is Associated with a Higher Incidence of Aneurysmal Subarachnoid Hemorrhage. *J Stroke Cerebrovasc Dis*, 27(3), 660–668. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2017.09.055>

Nichols, L., Stirling, C., Stankovich, J., & Gall, S. (2020). Time to treatment following an aneurysmal subarachnoid hemorrhage, rural place of residence and inter-hospital transfers. *Australasian Emergency Care*, 23(4), 225–232.

Nichols, L. J., Gall, S., & Stirling, C. (2016). Determining rural risk for aneurysmal subarachnoid hemorrhages: A structural equation modelling approach. *Journal of Neurosciences in Rural Practice*, 7(4), 559.

Peng, M., Southern, D. A., Williamson, T., & Quan, H. (2017). Under-coding of secondary conditions in coded hospital health data: impact of co-existing conditions, death status and number of codes in a record. *Health informatics journal*, 23(4), 260–267.

Pobereskin, L. (2001). Incidence and outcome of subarachnoid haemorrhage: a retrospective population based study. *Journal of Neurology, Neurosurgery & Psychiatry*, 70(3), 340–343.

Ranganathan, P., & Aggarwal, R. (2018). Common pitfalls in statistical analysis: Understanding the properties of diagnostic tests—Part 1. *Perspectives in clinical research*, 9(1), 40–43.

Rehman, S., Chandra, R. V., Lai, L. T., Asadi, H., Dubey, A., Froelich, J., Thani, N., Nichols, L., Blizzard, L., & Smith, K. (2021). Adherence to evidence-based processes of care reduces one-year mortality after aneurysmal subarachnoid hemorrhage (aSAH). *Journal of the neurological sciences*, 428, 117613.

- Rehman, S., Chandra, R. V., Zhou, K., Tan, D., Lai, L., Asadi, H., Froelich, J., Thani, N., Nichols, L., & Blizzard, L. (2020). Sex differences in aneurysmal subarachnoid haemorrhage (aSAH): aneurysm characteristics, neurological complications, and outcome. *Acta neurochirurgica*, 162(9), 2271–2282.
- Risselada, R., De Vries, L., Dippel, D., van Kooten, F., van der Lugt, A., Niessen, W., Firouzan, A., Stricker, B. C., & Sturkenboom, M. (2011). Incidence, treatment, and case-fatality of non-traumatic subarachnoid haemorrhage in the Netherlands. *Clinical neurology and neurosurgery*, 113(6), 483–487.
- Roman-Filip, I., Morosanu, V., Bajko, Z., Roman-Filip, C., & Balasa, R. I. (2023). Non-Aneurysmal Perimesencephalic Subarachnoid Hemorrhage: A Literature Review. *Diagnostics (Basel)*, 13(6). <https://doi.org/10.3390/diagnostics13061195>
- Sarti, C., Tuomilehto, J., Salomaa, V., Sivenius, J., Kaarsalo, E., Narva, E. V., Salmi, K., & Torppa, J. (1991). Epidemiology of subarachnoid hemorrhage in Finland from 1983 to 1985. *Stroke*, 22(7), 848–853.
- Schwarz, C. M., Hoffmann, M., Schwarz, P., Kamolz, L.-P., Brunner, G., & Sendlhofer, G. (2019). A systematic literature review and narrative synthesis on the risks of medical discharge letters for patients' safety. *BMC health services research*, 19(1), 158.
- Silfen, E. (2006). Documentation and coding of ED patient encounters: an evaluation of the accuracy of an electronic medical record. *The American Journal of Emergency Medicine*, 24(6), 664–678.
- Spolaore, P., Brocco, S., Fedeli, U., Visentin, C., Schievano, E., Avossa, F., Milan, G., Toso, V., Vanuzzo, D., & Pilotto, L. (2005). Measuring accuracy of discharge diagnoses for a region-wide surveillance of hospitalized strokes. *Stroke*, 36(5), 1031–1034.
- St. Germaine-Smith, C., Metcalfe, A., Pringsheim, T., Roberts, J. I., Beck, C. A., Hemmelgarn, B. R., McChesney, J., Quan, H., & Jette, N. (2012). Recommendations for optimal ICD codes to study neurologic conditions: a systematic review. *Neurology*, 79(10), 1049–1055.
- Sudlow, C., & Warlow, C. (1996). Comparing stroke incidence worldwide: what makes studies comparable? *Stroke*, 27(3), 550–558.
- The R Foundation, (2013). A language and environment for statistical computing.
- Thigpen, J. L., Dillon, C., Forster, K. B., Henault, L., Quinn, E. K., Tripodis, Y., Berger, P. B., Hylek, E. M., & Limdi, N. A. (2015). Validity of international classification of disease codes to identify ischemic stroke and intracranial hemorrhage among individuals with associated diagnosis of atrial fibrillation. *Circulation: Cardiovascular Quality and Outcomes*, 8(1), 8–14.
- Tirschwell, D. L., & Longstreth Jr, W. (2002). Validating administrative data in stroke research. *Stroke*, 33(10), 2465–2470.
- Tolonen, H., Salomaa, V., Torppa, J., Sivenius, J., Immonen-Räihä, P., Lehtonen, A., & Finstroke, R. (2007). The validation of the Finnish Hospital Discharge Register and Causes of Death Register data on stroke diagnoses. *European Journal of Preventive Cardiology*, 14(3), 380–385.
- van Walraven, C., English, S., & Austin, P. C. (2016). Administrative database code accuracy did not vary notably with changes in disease prevalence. *Journal of Clinical Epidemiology*, 79, 86–89.
- Wang, C., Yao, C., Chen, P., Shi, J., Gu, Z., & Zhou, Z. (2023). Artificial intelligence algorithm with ICD coding technology guided by embedded electronic medical record system in medical record information management. *Microprocessors and Microsystems*, 104962.
- Washington, C. W., Derdeyn, C. P., Dacey, R. G., Dhar, R., & Zipfel, G. J. (2014). Analysis of subarachnoid hemorrhage using the nationwide inpatient sample: the NIS-SAH severity score and outcome measure. *Journal of neurosurgery*, 121(2), 482–489.
- Williams, J., & Mann, R. (2002). Hospital episode statistics: time for clinicians to get involved? *Clinical Medicine*, 2(1), 34–37.
- Woodfield, R., Grant, I., Group, U. B. S. O., Follow-Up, U. B., Group, O. W., & Sudlow, C. L. (2015). Accuracy of electronic health record data for identifying stroke cases in large-scale epidemiological studies: a systematic review from the UK Biobank Stroke Outcomes Group. *Plos one*, 10(10), e0140533.
- Woodfield, R., Grant, I., & Sudlow, C. L. (2015). Accuracy of Electronic Health Record Data for Identifying Stroke Cases in Large-Scale Epidemiological Studies: A Systematic Review from the UK Biobank Stroke Outcomes Group. *Plos one*, 10(10), e0140533. <https://doi.org/10.1371/journal.pone.0140533>
- Yiannakoulis, N., Svenson, L. W., Hill, M. D., Schopflocher, D. P., Rowe, B. H., James, R. C., Wielgosz, A. T., & Noseworthy, T. W. (2003). Incident cerebrovascular disease in rural and urban Alberta. *Cerebrovascular diseases*, 17(1), 72–78.