Journal of Clinical Epidemiology 70 (2016) 146–154

Enriched administrative data can be used to retrospectively identify all known cases of primary subarachnoid hemorrhage

Shane W. English[a,b,*], Lauralyn McIntyre[a,b], Dean Fergusson[b], Alexis Turgeon[c], Cathy Sun[d], Marlise P. dos Santos[b,e], Cheemun Lum[e], John Sinclair[f], Alan Forster[d], Carl van Walraven[b,d]

[a]*Department of Medicine (Critical Care), The Ottawa Hospital, Ottawa, Canada*
[b]*Ottawa Hospital Research Institute, Clinical Epidemiology Program, University of Ottawa, Ottawa, Canada*
[c]*Department of Anesthesia (Critical Care), Hôpital de L'Enfant-Jésus, Québec City, Canada*
[d]*Department of Medicine, The Ottawa Hospital, Ottawa, Canada*
[e]*Department of Medical Imaging (Neuro-Interventional Section), The Ottawa Hospital, Ottawa, Canada*
[f]*Department of Surgery (Neuro-Surgery), The Ottawa Hospital, Ottawa, Canada*

Accepted 20 August 2015; Published online 21 September 2015

Abstract

Objective: We derived and validated a method to screen all hospital admissions for 1° subarachnoid hemorrhage (SAH) by retrospectively implementing recognized diagnostic criteria.

Study Design and Setting: A screen for 1° SAH was developed using two previously created registries. Screen-positive cases underwent diagnosis confirmation with primary record review. A review of all patient hospital encounters with the diagnostic code for 1° SAH, and cross-referencing with an existing SAH registry was undertaken to identify missed cases.

Results: Three subscreens were combined to form the 1° SAH screen (sensitivity: 98.4% [95% CI: 91.7–99.7%], specificity: 93.4% [95% CI: 90.4–95.4%], $n = 455$ patients in validation sample). From 1,699 screen-positive admissions between July 1, 2002 and June 30, 2011, we identified 831 true cases of SAH of which 632 patients had 1° SAH from ruptured aneurysm/arteriovenous malformation (sensitivity: 96.5% [95% CI: 94.8–97.8%], specificity: 40.3% [95% CI: 38.1–42.6%]). A review of all encounters with a diagnostic code for 1° SAH yielded additional 22 true cases.

Conclusion: When positive, our 1° SAH screen significantly increases the probability of this diagnosis in a particular hospitalization. The addition of patient hospitalizations encoded with the diagnostic code for 1° SAH improved sensitivity. Together, these methods represent the best way to retrospectively identify all cases of 1° SAH within an extensive sampling frame.

Keywords: Subarachnoid hemorrhage; Screen; Administrative data; Cohort; Diagnostic criteria; Retrospective disease registry

1. Introduction

Cohorts of patients with particular diseases are needed to study the epidemiology of those conditions. In the absence of prospectively collected registries, researchers can try to create patient cohorts by screening health administrative databases or other routinely collected data using diagnostic codes. Primary data of patients having these codes need to be reviewed to confirm the diagnosis. This process is especially problematic with uncommon diseases. Consider a disease which occurs in only 1 per 10,000 hospitalizations with a relatively accurate code (sensitivity and specificity of 90% and a positive likelihood ratio of 9). In this situation, the probability that a hospitalization with that code actually has the disease is only 0.09% (Appendix at www.jclinepi.com). As a result, approximately 1,000 charts coded with the disease would need to be reviewed manually to identify one true case. Therefore, alternative methods are needed to retrospectively create patient cohorts for uncommon diseases.

One example of an uncommon disease is subarachnoid hemorrhage (SAH) due to ruptured aneurysm or

Conflicts of interest: None.

Funding: Dr. English is the recipient of a Canadian Blood Services Fellowship Award.

* Corresponding author. Centre for Practice Changing Research, Ottawa Hospital Research Institute, 501 Smyth Road, Box 201B, Ottawa, Ontario K1H 8L6, Canada. Tel.: 613-737-8899x72818.

E-mail address: senglish@ohri.ca (S.W. English).

http://dx.doi.org/10.1016/j.jclinepi.2015.08.024
0895-4356/

What is new?

In this study, we derived and validated a systematic, comprehensive, and reproducible method to screen enriched administrative data in a novel way by implementing actual diagnostic criteria, usually applied prospectively, to accurately identify retrospectively every patient with 1° subarachnoid hemorrhage regardless of treatment outcome.

arteriovenous malformation (AVM). This is the major cause of primary SAH (1° SAH) [1−4] with an incidence that ranges between 2 and 16 cases per 1,00,000 person-years [5]. Nine studies have found that the positive predictive value (PPV) of several diagnostic codes for 1° SAH ranges from 33% to 100% [6−14]. However, SAH prevalence in these studies was very high, ranging from 1% to 28%. These values vastly exceed the prevalence of SAH in all hospitalizations. Because PPVs decrease as prevalence decreases [15,16], the accuracy of SAH cohorts that are identified using administrative diagnostic codes in unselected populations is questionable [17]. Simply using such diagnostic codes to identify rare diseases for epidemiologic study, where prevalence is low, may lead to important inaccuracies, both due to false positives (as demonstrated in the example previously) or from missed cases. Some have attempted to compensate for this shortfall by using prospectively collected data including diagnostic codes and therapeutic or outcome registries to identify cohorts with subsequent validation of the diagnosis using diagnostic criteria [18]. However, this may lead to missed patients, with a selection favoring those undergoing certain treatments or with certain outcomes, and bias subsequent epidemiologic and disease outcome studies.

We presently have no method to screen routinely collected hospital data for 1° SAH. In this study, we derived and validated a systematic, comprehensive, and reproducible method to screen enriched administrative data in a novel way by implementing actual diagnostic criteria, usually applied prospectively, to accurately identify retrospectively every patient with 1° SAH, regardless of treatment or outcome.

2. Methods

2.1. Study setting

The Ottawa Hospital (TOH) is a 1,000-bed tertiary care teaching hospital that supplies all neurosurgical services to Ottawa and the Champlain Local Health Integrated Network (approximate population in 2011 of 1.2 million [19]). The Ottawa Hospital Data Warehouse (TOHDW) is a collection of health data sets containing clinical and administrative data for all inpatient encounters at TOH. The primary holdings of TOHDW, available from July 1, 2002, include the Discharge Abstract Database (which contains diagnostic, procedural, demographic, and administrative information for each hospitalization), all laboratory tests (including results), all pharmacy orders, radiology information (including all imaging reports), and every autopsy report [20]. Appendix at www.jclinepi.com lists all terms and abbreviations used in this study.

2.2. Patient cohorts used to derive and validate a screening algorithm for 1° SAH

When deriving a screening algorithm for a particular disease within a database, a patient cohort known to have that disease (commonly referred to as the ''Gold-standard reference'') is necessary. In this study, we used two SAH patient registries at our hospital: (1) the TOH Interventional Neuroradiology Database, a retrospectively gathered data set of consecutive primary subarachnoid (1° SAH) patients who underwent endovascular coiling since January 2003 and (2) the TOH Neurosurgical Database, a prospective collection of consecutive 1° SAH patients who underwent surgical clipping of cerebrovascular aneurysms or AVMs from January 2002. For this study, we identified from these registries all cases of 1° SAH that occurred between January 1, 2007 and December 31, 2008. All cases in this ''Known 1° SAH Group'' met criteria for 1° SAH including blood in the subarachnoid space that resulted from an aneurysmal rupture; an AVM rupture; a perimesencephalic bleed (defined as a nonaneurysmal 1° SAH with a distinct peribrainstem hemorrhagic pattern with no discernible cause on imaging [21]); or an idiopathic cause. In particular, the diagnosis of 1° SAH required the exclusion of secondary causes including trauma and neoplasm. 1° SAH due to ruptured aneurysm or AVM was classified as aneurysmal SAH.

To derive and validate our screening methods, we created a random sample of other TOH patients without 1° SAH. This random sample of controls (named ''Random Sample Control Group'') was included to measure the accuracy of our algorithms to screen for 1° SAH. The absence of 1° SAH was confirmed by primary chart review.

2.3. Part I: derivation and validation of a 1° SAH screen

To identify all hospitalized patients older than 17 years with a possible 1° SAH, we created three subscreens to search for all patients that met any of the following three criteria:

1. SAH or ruptured aneurysm on computed tomography (CT) head, CT angiogram, or catheter angiogram;
2. Xanthochromia or more than 5×10^6/L red blood cells (RBCs) in cerebrospinal fluid (CSF);
3. Postmortem examination indicating SAH.

This diagnostic approach is supported by the most recent American Heart Association guidelines [22,23] with a