Published in final edited form as:

Neurosurgery. 2015 August ; 77(2): 269–273. doi:10.1227/NEU.0000000000000778.

Validation of an ICD-9 Code Algorithm for Identifying Chiari Malformation Type 1 Surgery in Adults

Jacob K. Greenberg, BA¹, Travis R. Ladner, BA⁶, Margaret A. Olsen, PhD, MPH^{4,5}, Chevis N. Shannon, MBA, MPH, DrPH⁶, Jingxia Liu, PhD³, Chester K. Yarbrough, MD¹, Jay F. Piccirillo, MD², John C. Wellons III, MD, MSPH⁶, Matthew D. Smyth, MD¹, Tae Sung Park, MD¹, and David D. Limbrick, MD, PhD¹

¹Department of Neurological Surgery, Washington University School of Medicine in St. Louis, St. Louis, Missouri

²Department of Otolaryngology, Washington University School of Medicine in St. Louis, St. Louis, Missouri

³Division of Biostatistics, Washington University School of Medicine in St. Louis, St. Louis, Missouri

⁴Division of Infectious Diseases, Washington University School of Medicine in St. Louis, St. Louis, Missouri

⁵Division of Public Health Sciences, Washington University School of Medicine in St. Louis, St. Louis, Missouri

⁶Department of Neurological Surgery, Vanderbilt University Medical Center, Nashville, Tennessee

Abstract

Background—The use of administrative billing data may enable large-scale assessments of treatment outcomes for Chiari Malformation Type I (CM-1). However, to utilize such datasets, validated International Classification of Diseases, Ninth Revision (ICD-9-CM) code algorithms for identifying CM-1 surgery are needed.

Objective—To validate two ICD-9-CM code algorithms identifying patients undergoing CM-1 decompression surgery.

Methods—We retrospectively analyzed the validity of two ICD-9-CM code algorithms for identifying adult CM-1 decompression surgery performed at two academic medical centers between 2001 and 2013. Algorithm 1 included any discharge diagnosis code of 348.4 (CM-1), as well as a procedure code of 01.24 (cranial decompression) or 03.09 (spinal decompression, or laminectomy). Algorithm 2 restricted this group to patients with a primary diagnosis of 348.4. The positive predictive value (PPV) and sensitivity of each algorithm were calculated.

Correspondence: Jacob K. Greenberg, BA, St. Louis Children's Hospital, One Children's Way, 4S20, St. Louis, MO 63110, greenbergj@wusm.wustl.edu.

Disclosure: The authors have no personal, financial, or institutional interest in any of the drugs, materials, or devices described in this article.

Results—Among 340 first-time admissions identified by Algorithm 1, the overall PPV for CM-1 decompression was 65%. Among the 214 admissions identified by Algorithm 2, the overall PPV was 99.5%. The PPV for Algorithm 1 was lower in the Vanderbilt (59%) cohort, males (40%), and patients treated between 2009 and 2013 (57%), whereas the PPV of Algorithm 2 remained high (99%) across subgroups. The sensitivity of Algorithms 1 (86%) and 2 (83%) were above 75% in all subgroups.

Conclusion—ICD-9-CM code Algorithm 2 has excellent PPV and good sensitivity to identify adult CM-1 decompression surgery. These results lay the foundation for studying CM-1 treatment outcomes using large administrative databases.

Keywords

Chiari Malformation Type 1; Health Services Research; Neurosurgery; Administrative Data Research; Validation Studies

INTRODUCTION

Administrative billing databases are increasingly popular tools used to study a wide range of neurosurgical conditions.¹⁻⁴ These large datasets, which are produced by a variety of public and private organizations, offer a time- and cost-effective means of studying questions related to clinical outcomes and resource utilization.⁵ The information they contain about various diagnoses and procedures is typically based on International Classification of Diseases, Ninth Revision (ICD-9-CM) coding, rather than data elements recorded for research purposes.^{5, 6} Consequently, the clinical information conveyed by ICD-9-CM codes may be incomplete in some instances, making ICD-9-CM codes unreliable for identifying certain neurological conditions.⁷ Therefore, one important aspect of reliably utilizing administrative billing data for research purposes is establishing the validity of ICD-9-CM codes for identifying relevant diagnoses and procedures.

Chiari Malformation Type 1 (CM-1) is a common neurosurgical condition diagnosed in 1% to 4% of brain and cervical spine magnetic resonance imaging studies.^{8, 9} However, despite the prevalence and the significant morbidity associated with CM-1, almost all evidence regarding clinical outcomes from CM-1 surgery comes from small, single-center studies. Administrative billing databases offer an efficient method of studying CM-1 outcomes in a multi-center, large-scale context. However, to our knowledge, no study has evaluated the validity of an ICD-9-CM code algorithm to identify CM-1 decompression surgery. Therefore, the objective of this study was to develop and validate two novel ICD-9-CM algorithms for identifying adult patients surgically treated for CM-1 in order to lay the foundation for future studies utilizing administrative datasets.

METHODS

To develop an ICD-9-CM code algorithm for identifying patients undergoing CM-1 decompression, we reviewed diagnosis and procedure codes defined by the National Center for Health Statistics.¹⁰ The diagnosis code for CM-1 is 348.4, though that code may also refer to brain compression from other conditions such as hemorrhage. As CM-1 surgery

typically spans the craniovertebral junction, there are two procedure codes potentially appropriate for specifying a CM-1 decompression: 01.24 (cranial decompression) and 03.09 (spinal decompression, or laminectomy).

To ensure the ICD-9-CM diagnosis and procedure codes we selected seemed broadly appropriate, we conducted a pilot review of the medical records of a subset of patients known to have received a CM-1 decompression at Washington University in St. Louis Medical Center (WU). After confirming the frequent use of these codes in this population, we sought to formally test two different algorithms for CM-1 decompression: Algorithm 1 included any discharge diagnosis code of 348.4, as well as a procedure code of 01.24 or 03.09; Algorithm 2 included a primary diagnosis code of 348.4, as well as a procedure code of 01.24 or 03.09.

We evaluated the performance of each algorithm by searching all adult (> 18 years) inpatient admissions between January 1, 2001 and May 20, 2013 at two academic medical centers, WU (including adult patients treated at Barnes Jewish Hospital and St. Louis Children's Hospital) and Vanderbilt University Medical Center (VU). We then performed medical record review to determine whether a CM-1 decompression had been performed during the hospitalization, excluding cases where a definitive diagnosis was not possible from available information. Because each patient may have had multiple admissions during the study period, we only analyzed first surgical admissions for each patient. In addition, we used a combination of the ICD-9-CM code search procedure and a medical record search for Current Procedural Terminology (CPT) code 61343 to generate a list of all known CM-1 surgeries at each center, which served as the basis for sensitivity calculations. The initial internal validation was done at WU and the subsequent external validation at VU was done independently and blinded to the results at WU.

Statistical analysis

Using medical record review as the gold standard, we calculated algorithm positive predictive value (PPV) as the number of CM-1 decompressions accurately identified by each ICD-9-CM algorithm (ie, confirmed by chart review) divided by the total number of eligible admissions resulting from each ICD-9-CM algorithm search. We calculated sensitivity as the number of CM-1 decompressions identified by each algorithm divided by all known CM-1 decompressions in the study period. Results were stratified by treating institution (WU and VU), gender, and year of surgery.

For PPV and sensitivity calculations, 95% confidence intervals (CI) were calculated with exact binomial methods. All analyses were performed using SAS version 9.3 (SAS Institute, Cary, North Carolina). The WU (IRB ID#: 201102012) and VU (IRB ID#: 131644) Institutional Review Boards approved all study procedures.

RESULTS

During the study period between 2001 and 2013, Algorithm 1 identified 340 unique patient admissions. Within that group, 214 admissions were also identified by the more-restrictive Algorithm 2 (Table 1). Among the 340 patients in the overall population, 149 were from the

WU cohort compared to 191 treated at VU. The median age was similar at WU (39.4 years; interquartile range [IQR] 32.8 – 47.6) and VU (39.9 years; IQR 30.3 – 52.3), as was the proportion of women (69.1% at WU vs. 66.0% at VU). The number of admissions identified increased in recent years, with approximately half the study population (55.7% of the WU cohort and 44.5% of the VU cohort) admitted between 2009 and 2013.

Positive Predictive Value

The performance of both algorithms is summarized in Table 1. In the overall population, Algorithm 1 had a PPV of 65% compared to 99.5% for Algorithm 2. While Algorithm 1 performed moderately better in the WU Cohort (PPV = 72%) than the VU Cohort (PPV = 59%), the PPV for Algorithm 2 was similar in both populations (99% for both). Likewise, the PPV for Algorithm 1 was lower in men (40%) and in patients treated between 2009 and 2013 (57%), whereas the PPV of Algorithm 2 remained high (99%) across demographic groups and treatment periods.

False Positives

Among 120 patients misclassified as a CM-1 decompression by Algorithm 1, the most common true reason for admission was surgery for intracranial hemorrhage/hematoma (43.3%), including cases associated with cerebral edema and traumatic brain injury. Other common reasons included surgery for cerebral infarction (28%), tumor (8%), and gunshot wound (3%). Notably, VU had substantially more false positives resulting from decompression for hemorrhage/hematoma (38/78, 49%) than WU (14/42, 33%), whereas the proportion of false positives resulting from decompression for infarction was higher at WU (16/42, 38%) than VU (17/78, 22%). Only one case was miscoded by Algorithm 2, a Chiari Malformation Type 2 (CM-2) surgery incorrectly classified as a CM-1 decompression. A complete list of the true reasons for admission for the false positive results is shown in Table 2.

Sensitivity

Of the 257 patients with known CM-1 decompression during the study period, Algorithm 1 identified 220 (sensitivity = 86%), while Algorithm 2 identified 213 (sensitivity = 83%). Although the sensitivity of Algorithm 2 was slightly lower in the VU cohort (81%), men (78%), and patients treated between 2005 and 2008 (77%), the magnitude of the observed differences was relatively small (differences of 10% or less across groups) (Table 1).

False Negatives

There were 44 false negative CM-1 decompressions not captured by Algorithm 2. Within the 44 false negatives, 37 were also false negative with Algorithm 1, and seven were identified by Algorithm 1 but not by the more-restrictive Algorithm 2 requiring CM-1 as the primary diagnosis code. Among the 44 false negative results for Algorithm 2 (including the subset that were also false negatives for Algorithm 1), 21 (48%) lacked 348.4 as the primary diagnosis code, 12 (27%) lacked any diagnosis of 348.4, and 30 (68%) lacked a procedure code of 01.24 or 03.09. Notably, the number of false negatives lacking 348.4 as the primary diagnosis code was higher in the WU (12/19, 63%) than VU (9/25, 36%) cohort, whereas

the number of patients lacking one of the specified procedure codes was higher in the VU (21/25, 84%) than WU (9/19, 47%) group.

Other than a diagnosis of CM-1, the most common primary diagnoses among the false negative results were CM-2 (21%), syringomyelia (5%), and cerebral cysts (5%). The most common procedures, other than a cranial or spinal decompression, were meninges repair (39%) and craniectomy (23%). A list of the most common diagnosis and procedure codes among false negative patients is shown in Table 3.

DISCUSSION

Key Results

In this study, we evaluated the performance of two different ICD-9-CM code algorithms to identify CM-1 decompression surgery, comparing the impact of listing CM-1 as any diagnosis (Algorithm 1) vs. only including admissions listing CM-1 as the primary diagnosis (Algorithm 2). We found that the PPV for Algorithm 2 was considerably higher than that for Algorithm 1, with only a small corresponding decrease in sensitivity. These effects were robust across treating centers, patient demographic groups, and treatment years, demonstrating that Algorithm 2 is the preferred method for studying CM-1 outcomes using administrative data.

Assessing the accuracy of ICD-9-CM codes is an important first step in using administrative billing data to study various diseases. Indeed, the importance of validating these algorithms is underscored by prior experience with conditions such as hemorrhagic transformation of ischemic stroke, where early studies relying on ICD-9-CM codes were later demonstrated to be unreliable.¹¹

Among neurological conditions, there is substantial variability in the accuracy of ICD-9-CM codes, with algorithms tested for certain conditions such as mild traumatic brain injury having PPVs and sensitivities below 50%.^{7, 12} By comparison, ICD-9-CM codes are often more accurate for major surgeries,^{13, 14} and algorithms identifying certain spine surgeries have high sensitivity (90%) and specificity (> 99%).¹⁵ Nonetheless, the detail provided by ICD-9-CM procedure codes for many neurosurgical procedures remains limited, suggesting administrative databases may not be appropriate for studying all interventions.¹⁶

In this multicenter validation, we examined the performance of algorithms combining ICD-9-CM diagnosis and procedure codes. Like previous studies using this integrated approach,^{15, 17} we found excellent algorithm performance, particularly when restricting to admissions listing CM-1 as the primary discharge diagnosis (Algorithm 2). Even with a PPV greater than 99%, Algorithm 2 had good sensitivity (83%) that was only marginally lower than that for Algorithm 1 (86%). Maximizing PPV at a small cost to sensitivity is desirable to ensure the population studied truly represents the condition of interest.

Although the overall results were similar in the WU and VU cohorts, there were some notable differences. For instance, while Algorithm 2 had only one false positive result overall, application of Algorithm 1 at VU led to a relatively high number of false positive

results related to decompression for hemorrhage/hematoma. This distinction from the WU cohort may explain the lower PPV of Algorithm 1 at VU and could result from variability in both case composition and also coding practices. Likewise, although the sensitivities of Algorithms 1 and 2 were similar, the disparate causes of false negative cases at WU (typically failure to code diagnosis 348.4) and VU (most often lack of procedure code 01.24 or 03.09) provided further evidence of differing billing practices across institutions. Thus, while the overall high performance of Algorithm 2 at both centers indicates its utility in administrative data research, the distinctions noted suggest that variable coding practices and case mix across centers may be one potential limitation when applying this Algorithm to multicenter administrative datasets.

Notably, although these results suggest coding practices may vary across institutions, the general approach to coding is the same at WU and VU. At both institutions, professional coders review operative notes and other clinician notes from inpatient admissions and subsequently assign ICD-9-CM diagnosis and procedure codes. In cases of ambiguity, surgeons may be consulted for clarification, but that is not routine practice. Given the diverse clinical indications and surgical techniques for CM-1 decompression, this overall similarity in the billing methods used at these two centers emphasizes the potentially important role of individual coder training and instruction in understanding inter-hospital differences in ICD-9-CM coding. A focused investigation into such practices, beyond the scope of this study, may be indicated.

Generalizability

While ICD-9-CM code Algorithm 2 had good to excellent performance internally (WU) and in a limited assessment of external validity (VU), there are limitations in the extent to which these results generalize to the broad range of academic and community centers included in administrative datasets. Specifically, while generalizability occurs on a spectrum rather than being defined based on a precise number of participating centers, assessment at an additional academic hospital, as well as one or more community hospitals, would provide further support for the external validity of these results. Moreover, while this validation included high-volume centers from two states, further external validation should evaluate hospitals in one or more additional states as well as hospitals with differing case volumes to ensure the durability of these findings.

Limitations

Beyond threats to its generalizability, this study has additional limitations. First, because independent lists were not available for either institution, we used a combination of ICD-9-CM and CPT code searches to generate a reference list of all known CM-1 surgeries. While the number of surgeries identified is consistent with the general expectations of participating surgeons, relying on administrative queries—albeit using broad ICD-9-CM and CPT codes—may have caused us to miss some cases, potentially biasing our sensitivity calculations. Second, although the PPV of Algorithm 2 was excellent, the sensitivity was relatively lower, indicating that some CM-1 surgeries may be missed when utilizing this algorithm for administrative data research. Finally, the ICD code algorithms described herein will need to be updated with the coming implementation of ICD-10.

CONCLUSION

Based on results from two academic hospitals, ICD-9-CM diagnosis and procedure codes can reliably identify patients receiving decompression surgery for CM-1. Restricted to patients with a primary discharge diagnosis of CM-1, Algorithm 2 has excellent PPV while maintaining good sensitivity, and therefore is the preferred method for identifying adult CM-1 surgeries. While additional external validation should be pursued, these results lay the foundation for our study of CM -1 treatment outcomes using large administrative billing databases, published elsewhere in this issue.¹⁸

Acknowledgments

We thank Ms Mary Amann, Ms Mary Montani, and Ms Jane Zbinden for their assistance with data acquisition at Washington University. We thank Dr Jesse M. Ehrenfeld, Director, Vanderbilt Anesthesiology & Perioperative Informatics Research Division, and his team for their assistance with data acquisition at Vanderbilt University.

Funding: This work was supported by the Clinical and Translational Science Award (CTSA) program of the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health (NIH) under Award Numbers UL1 TR000448 and TL1 TR000449 and by funding from the Agency for Healthcare Research and Quality (R24 HS19455). This work was also supported in part through grants from the American Syringomyelia & Chiari Alliance Project (ASAP) and Washington University School of Medicine, as well as philanthropic funding provided by the Park-Reeves Syringomyelia Research Consortium, the O'Keefe family, and Mateo Dalla Fontana. T.R.L. was also supported by a CSNS/CNS Medical Student Fellowship.

References

1. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. *JAMA*. Apr 7; 2010 303(13):1259–1265. [PubMed: 20371784]
2. Simon TD, Riva-Cambrin J, Srivastava R, Bratton SL, Dean JM, Kestle JR. Hospital care for children with hydrocephalus in the United States: utilization, charges, comorbidities, and deaths. *J Neurosurg Pediatr*. Feb; 2008 1(2):131–137. [PubMed: 18352782]
3. Wen T, Attenello F, He S, et al. Racial and Socioeconomic Disparities in Incidence of Hospital Acquired Complications Following Cerebrovascular Procedures. *Neurosurgery*. Mar 21; 2014
4. Zacharia BE, Deibert C, Gupta G, et al. Incidence, cost, and mortality associated with hospital-acquired conditions after resection of cranial neoplasms. *Neurosurgery*. Jun; 2014 74(6):638–647. [PubMed: 24618799]
5. Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med*. Oct 15; 1997 127(8 Pt 2):666–674. [PubMed: 9382378]
6. Jette N, Quan H, Hemmelgarn B, et al. The development, evolution, and modifications of ICD-10: challenges to the international comparability of morbidity data. *Med Care*. Dec; 2010 48(12):1105–1110. [PubMed: 20978452]
7. St Germaine-Smith C, Metcalfe A, Pringsheim T, et al. Recommendations for optimal ICD codes to study neurologic conditions: a systematic review. *Neurology*. Sep 4; 2012 79(10):1049–1055. [PubMed: 22914826]
8. Meadows J, Kraut M, Guarnieri M, Haroun RI, Carson BS. Asymptomatic Chiari Type I malformations identified on magnetic resonance imaging. *J Neurosurg*. Jun; 2000 92(6):920–926. [PubMed: 10839250]
9. Strahle J, Muraszko KM, Kapurch J, Bapuraj JR, Garton HJ, Maher CO. Chiari malformation Type I and syrinx in children undergoing magnetic resonance imaging. *J Neurosurg Pediatr*. Aug; 2011 8(2):205–213. [PubMed: 21806364]
10. ICD-9 CM Index to Diseases and Procedures. National Center for Health Statistics, Centers for Disease Control and Prevention; Atlanta, GA: 2011.

11. Mullen MT, Moomaw CJ, Alwell K, et al. ICD9 codes cannot reliably identify hemorrhagic transformation of ischemic stroke. *Circ Cardiovasc Qual Outcomes*. Jul; 2013 6(4):505–506. [PubMed: 23819956]
12. Bazarian JJ, Veazie P, Mookerjee S, Lerner EB. Accuracy of mild traumatic brain injury case ascertainment using ICD-9 codes. *Acad Emerg Med*. Jan; 2006 13(1):31–38. [PubMed: 16365331]
13. De Coster C, Li B, Quan H. Comparison and validity of procedures coded With ICD-9-CM and ICD-10-CA/CCI. *Med Care*. Jun; 2008 46(6):627–634. [PubMed: 18520318]
14. Quan H, Parsons GA, Ghali WA. Validity of procedure codes in International Classification of Diseases, 9th revision, clinical modification administrative data. *Med Care*. Aug; 2004 42(8):801–809. [PubMed: 15258482]
15. Wang MC, Laud PW, Macias M, Nattinger AB. Strengths and limitations of International Classification of Disease Ninth Revision Clinical Modification codes in defining cervical spine surgery. *Spine (Phila Pa 1976)*. Jan 1; 2011 36(1):E38–44. [PubMed: 20975624]
16. Faciszewski T, Jensen R, Berg RL. Procedural coding of spinal surgeries (CPT-4 versus ICD-9-CM) and decisions regarding standards: a multicenter study. *Spine (Phila Pa 1976)*. Mar 1; 2003 28(5):502–507. [PubMed: 12616165]
17. Ramirez LF, Thisted R. Using a national health care data base to determine surgical complications in community hospitals: lumbar discectomy as an example. *Neurosurgery*. Aug; 1989 25(2):218–225. [PubMed: 2770986]
18. Greenberg JK, Ladner TR, Olsen MA, et al. Complications and resource use associated with surgery for Chiari malformation type 1 in adults: a population perspective. *Neurosurgery*. 2015

Table 1

Validation of two ICD-9-CM coding algorithms for patients treated for CM-1 at two medical centers.

	No. Patients	Positive Predictive Value, % (95% CI)		Sensitivity, % (95% CI)	
		Algorithm 1: CM-1 = Any Diagnosis (n=340)	Algorithm 2: CM-1 = Primary Diagnosis (n=214)	Algorithm 1: CM-1 = Any Diagnosis (n=257)	Algorithm 2: CM-1 = Primary Diagnosis (n=257)
Total	340	65 (59-70)	99.5 (97-100)	86 (81-90)	83 (78-87)
Cohort					
WU	149	72 (64-79)	100 (97-100)	86 (78-91)	85 (77-91)
VU	191	59 (52-66)	99 (95-100)	86 (78-91)	81 (73-87)
Gender					
Female	229	77 (71-82)	99 (97-100)	87 (81-91)	84 (78-89)
Male	111	40 (30-49)	100 (92-100)	82 (69-91)	78 (65-88)
Year of Surgery					
2001-2004	76	72 (61-82)	100 (93-100)	89 (78-95)	84 (72-92)
2005-2008	96	72 (62-81)	99 (92-100)	79 (69-87)	77 (67-85)
2009-2013	168	57 (49-65)	100 (96-100)	89 (81-94)	87 (79-93)

WU=Washington University. VU=Vanderbilt University.

Table 2

True reasons for admission and surgery for cases misclassified as a CM-1 decompression by Algorithms 1 and 2. Results are pooled for the Washington University and Vanderbilt Cohorts.

	Algorithm 1, n (%)	Algorithm 2, n (%)
Surgery: Intracranial Hemorrhage/hematoma ^a	52 (43.3)	0 (0)
Surgery: Cerebral infarction	33 (27.5)	0 (0)
Surgery: Tumor	9 (7.5)	0 (0)
Surgery: Gunshot wound	3 (2.5)	0 (0)
Surgery: Abscess	2 (1.7)	0 (0)
Surgery: Chiari 2 Malformation	1 (0.83)	1 (100)
Other	20 (16.7)	0 (0)
Total	120	1

^aThis category included decompressive surgeries for hemorrhage or hematoma associated with traumatic brain injury and cerebral edema.

Table 3

A list of the common procedure codes and primary diagnosis codes among the 44 false negative CM-1 surgeries not captured by ICD-9-CM Algorithm 2.

Primary Diagnoses	n (%)*	Procedure Codes	n (%)*
348.4 (CM-1)	23 (52.3)	01.24/03.09 (cranial/spinal decompression)	14 (31.8)
741.0/741.01 (CM-2)	9 (20.5)	02.12 (meninges repair)	17 (38.6)
336.0 (syringomyelia)	2 (4.5)	01.25 (craniectomy)	10 (22.7)
348.0 (cerebral cyst)	2 (4.5)	04.42 (cranial nerve decompression)	4 (9.1)
Other	8 (18.2)	81.01 (Atlantoaxial fusion)	3 (6.8)

CM-2 = Chiari Malformation Type 2.

* % refers to the percentage of the 44 false negative patients.