Chiari malformation Type I surgery in pediatric patients. Part 1: validation of an ICD-9-CM code search algorithm

Travis R. Ladner, BA,1 Jacob K. Greenberg, MD, MSCI,2 Nicole Guerrero, BA,3 Margaret A. Olsen, PhD, MPH,4,5 Chevis N. Shannon, MBA, MPH, DrPH,1 Chester K. Yarbrough, MD, MPH,2 Jay F. Piccirillo, MD,6 Richard C. E. Anderson, MD,3 Neil A. Feldstein, MD,3 John C. Wellons III, MD, MSPH,1 Matthew D. Smyth, MD,2 Tae Sung Park, MD,2 and David D. Limbrick Jr., MD, PhD2

1Department of Neurological Surgery, Vanderbilt University Medical Center, Nashville, Tennessee; Departments of 2Neurological Surgery, Medicine, 3Surgery, and 4Otolaryngology–Head and Neck Surgery, Washington University School of Medicine in St. Louis, Missouri; and 5Department of Neurosurgery, Columbia University Medical Center, New York, New York

OBJECTIVE Administrative billing data may facilitate large-scale assessments of treatment outcomes for pediatric Chiari malformation Type I (CM-I). Validated International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) code algorithms for identifying CM-I surgery are critical prerequisites for such studies but are currently only available for adults. The objective of this study was to validate two ICD-9-CM code algorithms using hospital billing data to identify pediatric patients undergoing CM-I decompression surgery.

METHODS The authors retrospectively analyzed the validity of two ICD-9-CM code algorithms for identifying pediatric CM-I decompression surgery performed at 3 academic medical centers between 2001 and 2013. Algorithm 1 included any discharge diagnosis code of 348.4 (CM-I), as well as a procedure code of 01.24 (cranial decompression) or 03.09 (spinal decompression or laminectomy). Algorithm 2 restricted this group to the subset of patients with a primary discharge diagnosis of 348.4. The positive predictive value (PPV) and sensitivity of each algorithm were calculated.

RESULTS Among 625 first-time admissions identified by Algorithm 1, the overall PPV for CM-I decompression was 92%. Among the 581 admissions identified by Algorithm 2, the PPV was 97%. The PPV for Algorithm 1 was lower in one center (84%) compared with the other centers (93%–94%), whereas the PPV of Algorithm 2 remained high (96%–98%) across all subgroups. The sensitivity of Algorithms 1 (91%) and 2 (89%) was very good and remained so across subgroups (82%–97%).

CONCLUSIONS An ICD-9-CM algorithm requiring a primary diagnosis of CM-I has excellent PPV and very good sensitivity for identifying CM-I decompression surgery in pediatric patients. These results establish a basis for utilizing administrative billing data to assess pediatric CM-I treatment outcomes.

http://thejns.org/doi/abs/10.3171/2015.10.PEDS15370

KEY WORDS Chiari malformation Type I; health services research; neurosurgery; administrative data research; validation studies

The field’s knowledge of pediatric Chiari malformation Type I (CM-I) decompression surgery outcomes is largely derived from single-center studies conducted at academic hospitals.1,5,8,13,14,20 Given variations in demographic and clinical factors across patient populations, as well as differences in patient selection criteria and interventions across centers, large, multicenter studies are warranted. One efficient means for large-scale assessment of CM-I outcomes is through review of data contained in administrative billing databases, which offer valuable information regarding outcomes and resource utilization.5,10 These tools have become more popular in neurosurgery outcomes research in recent years.7,18,22,23

Hospitals use International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) coding,10 although the clinical information conveyed by
these codes may be incomplete or inaccurate in some instances.19 Therefore, ensuring valid search criteria for a particular disease/intervention is a key prerequisite for using these data sets for research purposes. Nonvalidated algorithms have been previously used in pediatric CM-I; however, the criteria used might have identified patients with spina bifida.17 Unfortunately, there is no discrete ICD-9-CM coding for CM-I decompression. Although two ICD-9-CM algorithms for identifying patients undergoing CM-I decompression have previously been validated in adults,2 the performance of these particular algorithms in children remains uncertain. Given differences in clinical and epidemiological features between children and adults with CM-I, these algorithms must be independently validated in the pediatric population. In this multicenter study, we validate two ICD-9-CM algorithms for identifying pediatric patients surgically treated for CM-I to facilitate future studies of pediatric CM-I using administrative data sets.

Methods

The development of the ICD-9-CM algorithms for CM-I surgery in adults has been previously reported.2 Two algorithms were used. Algorithm 1 included any discharge diagnosis code of 348.4 (compression of brain), as well as a procedure code of 01.24 (cranial decompression) or 03.09 (spinal decompression or laminectomy); Algorithm 2 included a primary diagnosis code of 348.4, as well as a procedure code of 01.24 or 03.09. For the current study, the initial development and internal validation was done at Washington University Medical Center (WU)—including pediatric patients treated at St. Louis Children’s Hospital or Barnes Jewish Hospital. The subsequent external validations at Vanderbilt University Medical Center (VU) and Columbia University Medical Center (CU) were done independently and blinded to the results at WU and also to each other.

Positive Predictive Value of Algorithm Search

Each algorithm was evaluated by searching all inpatient admissions for patients 18 years of age or younger for the period from January 1, 2001 to May 20, 2013 at the 3 academic medical centers using billing records. Medical records were then reviewed to determine whether a CM-I decompression was performed during the respective hospitalization, excluding cases where there was insufficient documentation of information in the medical record for confident determination. Only the first surgical admission was analyzed in cases involving patients with multiple surgical admissions (e.g., for revision or because of infection). There were 5 additional CM-I decompression cases (< 1% of the study cohort) identified outside of the algorithmic billing search, which, while not found on this query, still contained the appropriate ICD-9-CM codes; these were considered true positives as well.

Sensitivity of Algorithm Search

In parallel, a list of all known CM-I surgeries performed during the study period at each center was generated using a combination of the ICD-9-CM code search procedure and a medical record search for Current Procedural Terminology (CPT) Code 61343. This served as the basis for sensitivity calculations. An independent surgery case list maintained at CU was cross-referenced with this CPT/ICD-9 billing code search criterion, and 100% concordance was found.

Statistical Analysis

Using medical record review as the gold standard, algorithm positive predictive value (PPV) was calculated as the number of CM-I decompressions accurately identified by each ICD-9-CM algorithm divided by the total number of eligible admissions resulting from each ICD-9-CM algorithm search. Sensitivity was calculated as the number of CM-I decompressions correctly identified by each algorithm divided by all known CM-I decompressions in the study period. Results were stratified by treating institution, patient sex, and year of surgery (2001–2004, 2005–2008, 2009–2013).

For PPV and sensitivity calculations, we calculated 95% confidence intervals with exact binomial methods. All analyses were performed using IBM SPSS Statistics version 21 (IBM Corp.) and Microsoft Excel 2011 (Microsoft Corp.). The WU, VU, and CU institutional review boards approved all study procedures.

Results

Algorithm 1 identified 626 unique patient admissions across all institutions; 1 patient had insufficient clinical data and was excluded. Within the group of 625 patients, 581 admissions (93%) were also identified by the more-restrictive Algorithm 2 (Table 1). Of the entire group of 625 patients, 330 (53%) were from WU, 129 (21%) were from VU, and 166 (27%) were from CU. The median age was similar at WU (10.5 years, interquartile range [IQR] 5.6–14.7 years), VU (9.8 years, IQR 5.9–14.2 years), and CU (10.2 years, IQR 5.6–14.5 years). The proportion of female patients was also similar (WU 54%, VU 50%, CU 46%). The number of admissions identified by the ICD-9-CM algorithms increased in recent years, with almost half the study population (WU 43%, VU 50%, CU 44%) admitted between 2009 and 2013.

Positive Predictive Value

The performance of both ICD-9-CM algorithms is summarized in Table 1. In the overall population, Algorithm 1 had a PPV of 92% (572/625) compared with 97% (561/581) for Algorithm 2. While Algorithm 1 performed moderately better in the WU and CU cohorts (PPV 93%–94%) than the VU cohort (PPV 84%), the PPV for Algorithm 2 was similar in all 3 populations (96%–98%). The PPV of Algorithm 2 remained high (96%–98%) across all demographic groups and treatment periods as well.

False Positives

Among 53 patients misclassified as having undergone CM-I decompression by Algorithm 1, the most common true reason for admission was surgery for intracranial hemorrhage/hematoma or hematoma associated with
traumatic brain injury and cerebral edema (n = 14, 26%), followed by foramen magnum stenosis (n = 9, 17%), and Chiari malformation Type II (CM-II, n = 9, 17%). The majority of foramen magnum stenosis (n = 6, 71%) and CM-II (n = 6, 67%) cases occurred at WU, whereas the majority of hemorrhage/hematoma cases (n = 11, 79%) were at VU.

With the more-restrictive Algorithm 2, only 20 false-positive cases were identified, and there were no cases of misattributed hemorrhage/hematoma. The most common misclassified diagnoses for Algorithm 2 were foramen magnum stenosis (n = 7, 35%) and CM-II (n = 5, 25%). A complete list of the reasons for hospital admission from medical record review for the false-positive search results is shown in Table 2.

**Sensitivity**

Of 628 patients with known CM-I decompression during the study period, Algorithm 1 successfully identified 572 (sensitivity = 91%), while Algorithm 2 identified 561 (sensitivity = 89%, Table 1). The sensitivity of Algorithm 2 increased over time (83% in 2001–2004 vs 94% in 2009–2013, Table 1). Algorithm 2 was also more sensitive in the WU cohort (97%) compared with the VU and CU cohorts (each 82%).

**False Negatives**

There were 67 false-negative CM-I decompressions not captured by Algorithm 2. Of those, 56 false-negative results were also not captured by Algorithm 1. Among the patients accounting for false-negative results for Algorithm 2, 47/67 (70%) had 348.4 as the primary diagnosis code but lacked 01.24 or 03.09 as a procedure code; 15/67 (22%) lacked 348.4 as the primary diagnosis code but had 01.24 or 03.09 as a procedure code; and 5/67 (7%) lacked the 348.4 primary diagnosis code and both 01.24 and 03.09 as a procedure code. The number of false negatives lacking 348.4 as the primary diagnosis code was lower in the CU cohort (8/33, 24%) compared with the VU (8/23, 35%) and WU (4/11, 36%) cohorts; however, the number of patients lacking one of the specified procedure codes was somewhat lower in the WU cohort (7/11, 64%) than the other centers (CU 28/33, 85%; VU: 17/23, 74%).

Other than a diagnosis of CM-I, the most commonly coded primary diagnoses among the false-negative results were CM-II (n = 5, 7%), syringomyelia (n = 4, 6%), and cerebral palsy (n = 2, 3%). For the patients with a primary diagnosis of syringomyelia, CM-I (348.4) was coded in 3/4 (75%) cases, but not as the primary diagnosis. The most common procedures, other than a cranial or spinal decompression, were craniectomy (40%) and meninges repair (12%). A list of the most common diagnosis and procedure codes among patients with false-negative results is shown in Table 3.

**Discussion**

We evaluated the performance of two different ICD-9-CM code algorithms for identifying CM-I decompression surgery in pediatric patients, comparing the utility of

---

**TABLE 1. Validation of two ICD-9-CM coding algorithms for patients treated for CM-I**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>PPV, % (95% CI)</th>
<th>Sensitivity, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Algorithm 1: CM-I = Any Dx (n = 625)</td>
<td>Algorithm 2: CM-I = Primary Dx (n = 581)</td>
</tr>
<tr>
<td>No. of Pts</td>
<td>Algorithm 1: CM-I = Any Dx (n = 628)</td>
<td>Algorithm 2: CM-I = Primary Dx (n = 628)</td>
</tr>
<tr>
<td>Total</td>
<td>625</td>
<td>92 (89–94)</td>
</tr>
<tr>
<td>Cohort</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WU</td>
<td>330</td>
<td>94 (91–96)</td>
</tr>
<tr>
<td>VU</td>
<td>129</td>
<td>84 (77–90)</td>
</tr>
<tr>
<td>CU</td>
<td>166</td>
<td>93 (90–97)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>319</td>
<td>93 (90–96)</td>
</tr>
<tr>
<td>Male</td>
<td>306</td>
<td>90 (86–93)</td>
</tr>
<tr>
<td>Year of surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001–2004</td>
<td>115</td>
<td>93 (88–98)</td>
</tr>
<tr>
<td>2005–2008</td>
<td>231</td>
<td>93 (90–96)</td>
</tr>
<tr>
<td>2009–2013</td>
<td>279</td>
<td>90 (86–93)</td>
</tr>
</tbody>
</table>

CU = Columbia University; Dx = diagnosis; pts = patients; VU = Vanderbilt University; WU = Washington University.

---

**TABLE 2. True reasons for admission and surgery for cases misclassified as a CM-I decompression by Algorithms 1 and 2**

<table>
<thead>
<tr>
<th>True Reason for Surgery</th>
<th>Algorithm 1</th>
<th>Algorithm 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgery: intracranial hemorrhage/hematoma†</td>
<td>14 (26%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Surgery: foramen magnum stenosis</td>
<td>9 (17%)</td>
<td>7 (35%)</td>
</tr>
<tr>
<td>Surgery: CM-II</td>
<td>9 (17%)</td>
<td>5 (25%)</td>
</tr>
<tr>
<td>Surgery: tumor</td>
<td>4 (8%)</td>
<td>1 (5%)</td>
</tr>
<tr>
<td>Surgery: basilar invagination</td>
<td>3 (6%)</td>
<td>1 (5%)</td>
</tr>
<tr>
<td>Surgery: arachnoid cyst</td>
<td>2 (4%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Other</td>
<td>12 (23%)</td>
<td>6 (30%)</td>
</tr>
<tr>
<td>Total</td>
<td>53</td>
<td>20</td>
</tr>
</tbody>
</table>

* Values represent numbers of cases (%). Results are pooled for the WU, VU, and CU cohorts.
† This category included decompressive surgeries for hemorrhage or hematoma associated with traumatic brain injury and cerebral edema.
including CM-I in any diagnosis field (Algorithm 1, Table 4) or as the primary discharge diagnosis (Algorithm 2, Table 5). The PPV for Algorithm 2 (97%) was higher than that for Algorithm 1 (92%), with a minimal difference in sensitivity (89% vs 91%), irrespective of treatment centers, patient demographic characteristics, or treatment years. Thus, we suggest that Algorithm 2 should be the preferred method for studying CM-I outcomes using administrative billing data.

Ensuring valid ICD-9-CM code search criteria is a key prerequisite for using billing data to study neurological disease. Algorithms for conditions such as spinal cord injury and traumatic brain injury have shown PPVs less than 50%. Relying on such low-fidelity algorithms can introduce substantial misclassification bias and lead to unreliable results. While ICD-9-CM codes are accurate for some major surgeries, they have poor specificity for some other procedures, including spinal decompression surgeries.

Previous work studying pediatric CM-I decompression outcomes in administrative databases has been limited by use of nonvalidated search criteria that likely limit the reliability of the results. For instance, one study included a diagnosis code of 741.X, which would also select CM-II (i.e., spina bifida with or without hydrocephalus) patients, who represent a clinically distinct and often more severe population. In addition, that study did not restrict to a primary diagnosis of 348.4, also potentially lowering the specificity of their results. Consequently, the administrative data results obtained using such ICD-9-CM search methods may not accurately reflect outcomes in CM-I.

In this multicenter validation, we examined the performance of algorithms combining ICD-9-CM diagnosis and procedure codes to identify CM-I surgeries in pediatric patients. We found excellent performance, especially after restricting to admissions listing CM-I as the primary discharge diagnosis (Algorithm 2). Algorithm 2 maximized PPV at a very small cost to sensitivity, ensuring that the population analyzed truly represents the disease of interest. For every 100 true CM-I decompressions in a given database, this algorithm would correctly identify 89 of them (sensitivity), and of every 100 potential cases thus identified, 97 would be a true CM-I decompression (PPV). Less than 1% of such cases were CM-II. Therefore, researchers using this algorithm may confidently assume that study populations identified in this way are representative CM-I decompressions.

Although the overall results were similar across centers, there were some notable differences, which may be accounted for by case composition (i.e., lower proportion of CM-I decompressions as a fraction of total operations at a given center) and heterogeneous coding practices (e.g., failure to include the specified procedure code at VU and CU). For example, the sensitivity of Algorithm 2 was moderately greater in the WU cohort than the VU and CU cohorts. In addition, the PPV of Algorithm 1 was lowest at VU, likely because of a higher proportion of decompressions for hemorrhage/hematoma coded as 348.4; however, with the more-restrictive Algorithm 2, these cases were excluded, and PPV became highest at VU. The remaining misattributed cases (generally decompression for foramen magnum stenosis or CM-II) were highest in the WU population. Sensitivity increased over time, which may reflect changes in coding practices.

This study has 4 main limitations. As above, variable coding practices and case mix across centers may limit reproducibility. Second, these algorithms were only validated at academic hospitals with high-volume pediatric neurosurgical practices in 3 regions (Midwest, Southeast, and Northeast), and these institutions may not be entirely representative of administrative billing data sets, which contain a mix of academic and community hospitals across the country; therefore, further work to compare these algorithms at both community hospitals and in other centers is necessary.

### Table 3. A list of the common procedure codes and primary diagnosis codes among the 67 false-negative CM-I surgeries not captured by ICD-9-CM Algorithm 2*

<table>
<thead>
<tr>
<th>Procedure Codes</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01.25 (craniectomy)</td>
<td>27 (40%)</td>
</tr>
<tr>
<td>01.24/03.09 (cranial/spinal decompression)</td>
<td>15 (22%)</td>
</tr>
<tr>
<td>02.12 (meninges repair)</td>
<td>8 (12%)</td>
</tr>
<tr>
<td>01.59 (other brain lesion excision)</td>
<td>4 (6%)</td>
</tr>
<tr>
<td>02.12 (meninges repair)</td>
<td>8 (12%)</td>
</tr>
<tr>
<td>04.42 (cranial nerve decompression)</td>
<td>2 (3%)</td>
</tr>
<tr>
<td>Other</td>
<td>9 (13%)</td>
</tr>
</tbody>
</table>

* Results are pooled for the WU, VU, and CU cohorts.

### Table 4. Summary 2 × 2 contingency table for Algorithm 1*

<table>
<thead>
<tr>
<th>Identification</th>
<th>CM-I</th>
<th>No CM-I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identified</td>
<td>TP = 572</td>
<td>FP = 53</td>
</tr>
<tr>
<td>Not identified</td>
<td>FN = 56</td>
<td>NA</td>
</tr>
</tbody>
</table>

FN = false negative; FP = false positive; NA = not applicable; TP = true positive.

* Results are pooled for the WU, VU, and CU cohorts. The PPV of Algorithm 1 was 92% (572/625). The sensitivity was 91% (572/629).

### Table 5. Summary 2 × 2 contingency table for Algorithm 2*

<table>
<thead>
<tr>
<th>Identification</th>
<th>CM-I</th>
<th>No CM-I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identified</td>
<td>TP = 561</td>
<td>FP = 20</td>
</tr>
<tr>
<td>Not identified</td>
<td>FN = 67</td>
<td>NA</td>
</tr>
</tbody>
</table>

* Results are pooled for the WU, VU, and CU cohorts. The PPV of Algorithm 2 was 97% (561/581). The sensitivity was 89% (561/628).
United States regions is warranted. Potential variability in coding practices between academic and community-based practices is not likely to impact the performance of the algorithm in large administrative database studies for CM-I, as we have observed that 96% of CM-I decompressions are performed at teaching hospitals. From the data presented, centers that perform a higher proportion of decompressions for either foramen stenosis or CM-II than our centers would likely have a lower PPV, as these indications were a frequent reason for false positives.

Third, because independently prepared lists of all known CM-I decompressions were not available, reference lists were created using a combination of ICD-9-CM and CPT code searches; while the number of surgeries identified was consistent with our general expectations, some cases may have been missed, potentially resulting in an overestimate of the sensitivity to identify CM-I decompression surgery. However, an independent case list maintained at CU was cross-referenced with this CPT/ICD-9 billing code search criteria, and 100% concordance was found. Finally, the ICD-9-CM code algorithms used here will need to be updated once ICD-10 has been fully implemented with sufficient case volume and follow-up for study, and we hope this study will serve as a practical, reproducible model for that work. Even as ICD-10 is broadly implemented, the existing ICD-9 billing data set will serve a critical role in CM-I research for years to come, permitting large-scale studies into pediatric neurosurgical practices with an emphasis on geographic, socioeconomic, and demographic factors, and medical and surgical complications, among other important areas.

Conclusions
ICD-9-CM diagnosis and procedure codes can reliably and reproducibly be used to identify pediatric patients undergoing decompression surgery for CM-I. An algorithm restricted to patients with CM-I as the primary discharge diagnosis has excellent PPV (97%) and very good sensitivity (89%) and therefore is the preferred method for identifying pediatric CM-I surgeries. These results establish a basis for utilizing large administrative billing databases to assess pediatric CM-I treatment outcomes.

Acknowledgments
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 HS194555). 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 Council of State Neurosurgeons Medical Student Fellowship.

References
15. Quan H, Parsons GA, Ghali WA: Validity of procedure codes in International Classification of Diseases, 9th revision, clinical modification administrative data. Med Care 42:801–809, 2004
29. Quan H, Parsons GA, Ghali WA: Validity of procedure codes in International Classification of Diseases, 9th revision, clinical modification administrative data. Med Care 42:801–809, 2004


Disclosures
The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author Contributions
Conception and design: Ladner, Greenberg, Limbrick. Acquisition of data: Ladner, Greenberg, Guerrero. Analysis and interpretation of data: Ladner, Greenberg, Guerrero, Olsen, Shannon, Yarbrough, Limbrick. Drafting the article: Ladner, Greenberg. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Ladner. Statistical analysis: Ladner. Administrative/technical/material support: Shannon, Limbrick. Study supervision: Olsen, Shannon, Piccirillo, Anderson, Feldstein, Wellons, Smyth, Park, Limbrick.

Supplemental Information
Companion Papers

Correspondence
Travis R. Ladner, 2200 Children’s Way, 9222 Doctors’ Office Tower, Nashville, TN 37232. email: travis.r.ladner@vanderbilt.edu.