Fidelity of Administrative Data When Researching Down Syndrome.
Kristin M. Jensen, M.D., M.Sc.\textsuperscript{1,A}
Colin R. Cooke, M.D., M.Sc.\textsuperscript{2,A}
Matthew M. Davis, M.D., M.A.P.P.\textsuperscript{1,3,B,C}

1. University of Michigan, Departments of Internal Medicine and Pediatrics, Ann Arbor, MI
2. University of Michigan, Division of Pulmonary and Critical Care Medicine and Center for Healthcare Outcomes & Policy, Ann Arbor, MI
3. Ford School of Public Policy, University of Michigan, Ann Arbor, MI
A. Graduate of Robert Wood Johnson Foundation Clinical Scholars Program, University of Michigan, 2009-2001
B. Graduate of Robert Wood Johnson Foundation Clinical Scholars Program, University of Chicago, 1998-2000
C. Co-Director, University of Michigan, Robert Wood Johnson Foundation Clinical Scholars Program

Corresponding Author (Current Contact Information):
Kristin M. Jensen, M.D., M.S.
Linda Crnic Institute for Down Syndrome
Departments of Pediatrics and Internal Medicine
University of Colorado – Denver
Research Complex 2, Mail Stop 8608
12700 E. 19th Ave
Aurora, CO 80045
Phone: 303-724-6481
Fax: 303-724-5741
Email: kristin.jensen@ucdenver.edu

Keywords/Possible MeSH terms: Down syndrome; Medical Record Linkage; Databases, Factual/Standards; Registries/Standards; Epidemiologic Methods; International Classification of Diseases; Humans; Research

Definitions of Abbreviations & Acronyms:
Down Syndrome = DS
ICD-9 = International Classification of Diseases, Ninth Revision
**Acknowledgements:**

**Author Contributions:**
Dr. Jensen had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Study concept and design:* Jensen and Davis  
*Acquisition of data:* Jensen and Davis  
*Analysis and Interpretation of Data:* Jensen, Cooke, and Davis  
*Drafting of the manuscript:* Jensen  
*Critical revision of the manuscript for important intellectual content:* Jensen, Cooke, and Davis  
*Obtained funding:* Jensen  
*Administrative, technical, and material support:* Jensen and Davis  
*Study supervision:* Jensen and Davis

**Funding/Support:** This work was funded by the Robert Wood Johnson Foundation Clinical Scholars Program. During this project, Dr. Jensen and Dr. Cooke were Clinical Scholars and Dr. Davis served as the Co-Director of the Clinical Scholars program at the University of Michigan. None of the authors received any form of payment for production of this manuscript.

**Financial Disclosure:** none reported

**Role of the Sponsor:** The sponsor had no role in the design or conduct of the study, in the collection, analysis or interpretation of the data, or in the preparation, review, or approval of the manuscript.

**Conflicts of Interest:** The authors have no conflicts of interest to report.

**Additional Contributions:** The authors would like to thank Sonya M. DeMonner, M.P.H. for her assistance with the dataset.
What’s known on this subject?

The use of administrative data to study the quality of pediatric care is affected by how well diagnostic codes can identify patients with the diagnoses of interest.

What this study adds:

Investigators can improve the fidelity of their data by using clinical records to validate diagnoses identified within administrative databases. This is particularly important for less common diagnoses like Down syndrome, in which the presence of miscoded patients can significantly skew results.
Abstract:
Objective: To compare the fidelity of administrative data with clinical data when researching Down syndrome.

Patients and Methods: From outpatient, inpatient, and emergency department administrative claims within our institution, we identified 252 patients ages 18-45 years with encounters coded for Down syndrome (DS) by ICD9=758.0 from 2000-2008. We evaluated these cases for false positive errors – cases in which DS was not actually present in clinical descriptions. Subsequently, we identified false negative errors (cases in which DS was present without encounters coded as such) through examination of the medical record for all patients within our study frame who had one of several common DS comorbidities, including congenital heart disease, hypothyroidism, and atlanto-axial instability.

Results: Among 252 persons with an administrative code for DS, 53 (21%) did not have DS documented in their medical record. While searching for false negative errors, 29 additional persons were discovered with DS documented in the medical record who had not been previously identified. This led to a final cohort of 228 persons with DS. The presence of a billing code for DS had moderate sensitivity (87%) and positive predictive value (79%), but high specificity (99.9%).

Conclusions: Administrative claims misclassify a sizeable proportion of persons with DS. Judgments about quality of care based on samples identified using administrative claims may not accurately reflect the experience of patients with the conditions in question. When using administrative databases to study the quality of care for patients with DS, diagnostic verification within the clinical record is advisable whenever possible.
**Background:**

Administrative data are frequently used by researchers to evaluate trends in health care. Their appeal lies in the relative ease in which investigators can access large volumes of information with detail. It is nearly impossible to replicate the convenience and cost-savings of using already-gathered information with the volume of patient encounters available in many of these datasets.\(^1\)-\(^3\) Secondary data analyses have led to valuable findings within epidemiology, clinical effectiveness, quality improvement, risk assessment, and health care utilization and expenditure research.\(^1\)-\(^3\)-\(^14\)

When evaluating uncommon pediatric conditions, many investigators face a substantial challenge to reach large enough sample sizes to be able to draw meaningful conclusions. Utilization of administrative data proves particularly helpful in these circumstances. Pooling together the experience of a large cohort of patients with rare diagnoses allows clinicians to look for trends and patterns in care utilization that can inform diagnostic, therapeutic, and policy decisions not otherwise possible.

Nevertheless, it is important for investigators to recognize that administrative data were not gathered prospectively or for research purposes.\(^15\) These data can be extremely difficult to validate and often do not contain all domains of interest.\(^1\)-\(^7\) Additionally, reimbursement incentives linked to specific diagnostic codes can profoundly affect the integrity of billing data, leading to systematic errors in coding.\(^2\) Studies correlating the accuracy of billing codes with clinical records reveal widely varying levels of agreement. For example, Quan et al (2004) found that major procedures performed in the operating room (and therefore likely the primary reason for admission) were coded much more reliably than minor procedures routinely performed on inpatient wards.\(^16\) Analyses of diagnoses such as diabetes mellitus, thrombocytopenia, and psoriasis show concordance levels with the clinical record of 94%, 83%, and 57% respectively.\(^5\)-\(^7\)-\(^9\) Questions remain about the usefulness of administrative data when evaluating conditions that function more often as comorbidities than the primary reasons for clinical encounters, especially since rare diagnoses and vague diagnostic criteria can increase the potential for miscoding.\(^1\)-\(^6\),\(^9\),\(^10\)-\(^17\)-\(^19\) These factors directly impact the quality of pediatric research and conclusions drawn from administrative data. It is important for investigators to recognize that cohorts identified strictly through billing codes may not accurately represent populations with the diagnoses of interest and to utilize methods to improve the quality of their data whenever possible.
Very little is known about the ability of the billing codes associated with Down syndrome to correctly identify patients. This is important, as Down syndrome occurs in nearly 1 in 700 live births\textsuperscript{20} and is the most commonly identifiable cause of intellectual disability.\textsuperscript{21} Given that individuals with Down syndrome now survive increasingly into adulthood with a corresponding increase in overall prevalence in the United States,\textsuperscript{22,23} the ability to access administrative data to evaluate trends in this population appeals to many investigators. The purpose of this study is to compare the fidelity of administrative data with clinical data when researching a cohort of patients with Down syndrome.

**Patients and Methods**

**Study Cohort**

We identified our original cohort from administrative data gathered from the Central Data Repository at the University of Michigan, which captures all outpatient, inpatient, and emergency care within our health system. Patients ages 18-45 years were included in the initial cohort if they were seen at our institution at any time between January 1, 2000, and June 30, 2008, with any diagnosis for “chronic and severe conditions originating in childhood”\textsuperscript{24} within any of the 15 diagnostic fields available for each encounter (n=108,216).

We then refined our cohort to only persons with Down syndrome listed within any of the 15 diagnostic fields available for each encounter (\textit{International Classification of Diseases, Ninth Revision, Clinical Modification} (ICD-9-CM) code 758.0).\textsuperscript{25} Diagnosis of Down syndrome was confirmed through manual chart review utilizing a search engine for free-text documents within the electronic medical record, known as the Electronic Medical Record Search Engine (EMERSE).\textsuperscript{26,27} Search terms included the following: "Down syndrome", "Down's syndrome", "Downs syndrome", "Trisomy 21", "Tri21", and "Tri 21". This resulted in 53 cases (21\%) of persons without Down syndrome being excluded from our cohort.

Within the Central Data Repository, we then identified all patients (meeting the aforementioned age and date criteria) who had visits coded for the commonly-associated Down syndrome comorbidities of congenital heart disease (ICD-9-CM codes 745-745.9, 746-746.9, 747-747.49), hypothyroidism (ICD-9-CM codes 243, 244.3, 244.8, 244.9), and atlanto-axial instability (847.0).
EMERSE was utilized to identify patients with documented Down syndrome in the electronic medical record with one of the aforementioned comorbidities who had not been previously identified by ICD-9-CM code 758.0. Given our low-return of newly-identified patients with Down syndrome within these high-yield comorbidities, the remainder of the original cohort of persons with “chronic and severe conditions originating in childhood” was treated as not having Down syndrome for our analysis. Only patients with documented Down syndrome, verified by the authors upon review of the clinical record, were included in the final cohort (Figure 1).  

Data Analysis

Bivariate comparison of patient attributes were made between persons coded for Down syndrome without evidence of that diagnosis in the medical record (false positives) and persons with Down syndrome present in the medical record without visits coded as such (false negatives) using chi-square, Kruskal-Wallis, and Student’s t-tests as appropriate. We determined binomial exact confidence intervals for all calculated proportions. We used Stata version 11.0 (Stata Corp, College Station, TX) for all analyses. The Institutional Review Board at the University of Michigan approved the study protocol.

Results

We originally identified a total of 252 patients with encounters listing the diagnostic code for Down syndrome at any time during 2000-2008. Forty-three percent of our cohort was female with a mean age of 29.5 years (Table 1). Upon manual chart review, we identified 53 false positives (21%, 95% CI 16-27%) in which Down syndrome was not documented in the medical record. The majority of these cases consisted of women of reproductive age being seen in obstetric clinics, likely because the billing code for Down syndrome refers to diagnostic screening for aneuploidy in their fetuses (Table 2).

After excluding these false positives from our cohort, we sought to identify possible false negatives—patients with documented Down syndrome without encounters coded as such. To do this, we manually reviewed the medical records for all patients with congenital heart disease, hypothyroidism, and atlanto-axial instability. Of 14,744 cases reviewed, 29 additional persons were identified with documented Down syndrome who had not previously been identified by the administrative claims. We considered the remainder of the original cohort of persons with “chronic and severe conditions originating in childhood” (n=107,935) as not having our diagnosis of
interest, given that only 0.2% of these patients with comorbidities common to Down syndrome were newly-identified with the diagnosis. This led to a final cohort of 228 adults with Down syndrome seen at our institution from 2000-2008.

We observed similarities in patient age, gender, and race/ethnicity between our final cohort and the original cohort identified strictly through ICD-9-CM code 758.0 (data not shown). More remarkable is the comparison of the false positive and false negative patient populations (Table 3). Patients who were inappropriately identified with Down syndrome through billing codes (false positives) were significantly more likely to be female with unknown race/ethnicity than patients with documented Down syndrome who had not been identified by billing codes (false negatives). Although the net difference in total sample size between the final and original cohorts is only 24 patients, the exclusion of false positives and inclusion of false negatives impacted our sample size by -21% and +15% respectively.

Our experience with clinical verification of this dataset reveals that ICD-9-CM code 758.0 is only moderately sensitive at 87.3% but highly specific at 99.9%. Similarly, the positive predictive value of ICD-9-CM code 758.0 is low at 79.0 %, but the negative predictive value is quite high at 99.9% (Figure 2).

**Discussion**

Investigators are increasingly aware of the potential for coding inaccuracies within medical billing data. Possible explanations for this problem include differences between the primary reason for an encounter and a patient’s chronic diagnoses or comorbidities, the presence of risk factors for the diagnosis of interest, vague diagnostic criteria for a given billing code, variable reimbursement incentives, lack of understanding of the clinical implications when choosing a specific diagnostic code, and lack of standardized coding practices. The implications of these coding inaccuracies are difficult to measure without systematic evaluation, but do not appear to adversely affect real-time clinical care or reimbursement. However, the consequences of miscoding have a greater impact when billing data is utilized to draw conclusions about areas outside of their intended purpose, such as epidemiology, health care quality, and practice patterns.

Our findings highlight what is known about administrative data – billing codes can be inappropriately assigned to describe a patient encounter. We observed moderate sensitivity (87%) and positive predictive value
(79%) of the billing code specific to Down syndrome. These values mirror the experiences of Cooke et al (2011) and Galdarossa et al (2012) with observed sensitivities of 76% and 83% between administrative and clinical records and are slightly higher than the experience by Icen et al., which demonstrated 57% concordance, for their respective diagnoses of interest.⁵,⁶,⁹

In our cohort, the net difference between our original and final cohorts is only 24 cases. However, the differences between these cohorts are substantive. One-fifth of cases in the original cohort were inappropriately assigned the diagnosis of Down syndrome. The vast majority of these individuals were women of reproductive age at obstetric visits, suggesting a significantly healthier population than our population of interest. Inclusion of these cases without diagnostic verification would significantly skew most outcomes of interest.

Compounding the ramifications of high false positive rates for a given diagnosis of interest are the implications of the loss of information from patients with that diagnosis who were not identified by the associated billing code (false negatives). Thirteen percent of our final cohort would not have been identified without our systematic search for false negatives in comorbidities commonly associated with Down syndrome. While perhaps less of an issue with large cohorts of common diagnoses, the loss of 13% of cases for a diagnosis with as low of prevalence as Down syndrome can mean the difference between enough power to observe meaningful differences and to draw statistically significant conclusions or not. One must also consider whether the reason for omission of a diagnostic code results from clinical case differences (e.g., whether patients are less complicated at baseline), as this could lead to non-random omission and skew of results.²

In its 2011 report “Child and Adolescent Health and Health Care Quality”, the Institute of Medicine highlights the fact that no single data source exists that can provide “valid and reliable indicators about health and health care quality in children and adolescents.”³¹ Therefore, researchers must continue to optimize the information available in current administrative datasets. To this end, several investigators have proposed alternative methods to validate administrative coding. Herbert et al (2004) recommends combining clear definitions of the outcomes of interest with regular audits of data registries.¹⁹ To do this successfully, however, investigators must approach administrative data with prospective questions and build in short windows of time for the retrospective audits. Cooke et al (2011) noted improved specificity of administrative coding when combining pharmacy data with ≥1 outpatient encounter for the ICD-9 code in question.⁶ Similarly, Parker et al (2003)
observed significant improvements in predictions of hospital readmissions when combining pharmacy-based disease markers with administrative data.\textsuperscript{8} Icen et al (2008) determined that the Positive Predictive Value of a given code representing the diagnosis of interest increased with repeated visits for the same diagnostic code.\textsuperscript{5} In our study, we observed lack of fidelity in administrative data when evaluating a cohort of persons with Down syndrome. False positive errors were slightly more common than false negative errors. However, the populations of patients covered by both errors are substantially different from each other, raising concerns about the validity of conclusions drawn from administrative data without diagnostic verification in the medical record. To remedy this situation, we strongly recommend validation of diagnoses of interest with the medical record as the gold standard. When this is not possible, implementation of coding algorithms, such as those employed in the studies described above, and linkage to confirmatory data sources (such as pharmacy data or the electronic medical record) can serve to greatly improve the fidelity and quality of research involving administrative data.

This study has several limitations. First, our work reflects evaluation of a single diagnosis in administrative data. Although we anticipate that rates of diagnostic accuracy will vary between diagnoses, our findings lie within the ranges described in the medical literature for a wide range of clinical conditions. Second, our analysis is restricted to encounters at a single academic medical center. While our experience may not reflect national trends in fidelity of administrative data, we are encouraged that our findings are consistent with what has already been published in the medical literature and do not anticipate that national trends will be significantly different. Third, our cohort was identified from a larger cohort of adults with “chronic and serious illnesses originating in childhood.” It is possible that several more false negative cases are present in our health system that were not included in that cohort, although we anticipate the yield of such analysis to be quite low based on our experience of identifying only 0.2% additional patients through review of comorbidities. These limitations notwithstanding, we expect that our experience of identifying inappropriately coded encounters within administrative data provides a reasonable approximation of similar databases nationally.

Conclusion

Administrative datasets are used by investigators with increasing frequency to evaluate trends in areas such as epidemiology, clinical effectiveness, risk assessment and health care utilization, as well as to inform
decisions regarding clinical care and health policy. Our study highlights potential pitfalls of such analyses.

Judgments about quality of care based on samples identified from administrative claims may not accurately reflect patients with the conditions in question. This is especially important when evaluating uncommon diagnoses. To improve the quality of pediatric research produced from administrative databases, diagnostic verification within the clinical record is advisable whenever possible.
References


Figure 1: Study Cohort Identification

Assessed for eligibility (n=108,216):
All patients ages 18-45 years seen in this health system from January 1, 2000- June 30, 2008 with chronic and serious illnesses originating in childhood.

Patients with visits coded for Down syndrome (n=252).

Excluded, False Positives (n=53):
No evidence of Down syndrome upon review of electronic medical record.

Review of medical records (n=14,774):
All patients within study frame with visits coded for congenital heart disease, hypothyroidism, or atlanto-axial instability.

Included, False Negatives (n=29):
Patients with Down syndrome documented in medical record without a visit coded during study frame.

Total number of adults with Down syndrome within our health system (n=228).
Figure 2: Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Values of Billing Codes Associated with Down Syndrome (ICD-9-CM=758.0) within our Cohort

<table>
<thead>
<tr>
<th>Billing Code for Down Syndrome</th>
<th>Diagnosis of Down Syndrome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present</td>
</tr>
<tr>
<td>Present</td>
<td>199</td>
</tr>
<tr>
<td>Absent</td>
<td>29</td>
</tr>
</tbody>
</table>

*Based on yield of 0.2% new cases of persons with Down syndrome after review of nearly 15,000 records of patients with comorbidities common to Down syndrome, the remainder of the original cohort of persons with "chronic and serious illnesses originating in childhood" was treated as negative for Down syndrome in this analysis.

<table>
<thead>
<tr>
<th></th>
<th>Point estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>87.28%</td>
<td>(82.25-91.31%)</td>
</tr>
<tr>
<td>Specificity</td>
<td>99.95%</td>
<td>(99.94% - 99.96%)</td>
</tr>
<tr>
<td>Positive Predictive Value</td>
<td>78.97%</td>
<td>(73.41-83.83%)</td>
</tr>
<tr>
<td>Negative Predictive Value</td>
<td>99.97%</td>
<td>(99.96%-99.98%)</td>
</tr>
</tbody>
</table>
Table 1: Demographics in Final Cohort, n (%)  

<table>
<thead>
<tr>
<th>n</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>97 (43%)</td>
</tr>
<tr>
<td>Age, mean (S.D.)</td>
<td>29.5 yr (9.3)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>190 (83%)</td>
</tr>
<tr>
<td>African American</td>
<td>18 (8%)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4 (2%)</td>
</tr>
<tr>
<td>Asian</td>
<td>1 (&lt;1%)</td>
</tr>
<tr>
<td>Unknown/Other</td>
<td>15 (7%)</td>
</tr>
<tr>
<td>Category</td>
<td>Count</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Obstetric Visit</td>
<td>36</td>
</tr>
<tr>
<td>High Risk Pregnancy for Down Syndrome</td>
<td>3</td>
</tr>
<tr>
<td>Other Genetic Condition or Developmental Delay</td>
<td>7</td>
</tr>
<tr>
<td>Other</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>53</td>
</tr>
</tbody>
</table>
Table 3: Comparison of False Positive and False Negative Patients for Billing Code Associated with Down Syndrome (ICD-9-CM=758.0)

<table>
<thead>
<tr>
<th></th>
<th>False Positive</th>
<th>False Negative</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>n (%)</td>
<td>53</td>
<td>29</td>
<td>--</td>
</tr>
<tr>
<td>Female</td>
<td>47 (89%)</td>
<td>12 (41%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mean Age (S.D.)</td>
<td>28.8 yr (7.3)</td>
<td>28.5 yr (9.6)</td>
<td>0.85</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>20 (38%)</td>
<td>23 (79%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>African American</td>
<td>1 (2%)</td>
<td>2 (7%)</td>
<td>0.25</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1 (2%)</td>
<td>0</td>
<td>0.46</td>
</tr>
<tr>
<td>Asian</td>
<td>1 (2%)</td>
<td>0</td>
<td>0.46</td>
</tr>
<tr>
<td>Unknown/Other</td>
<td>30 (57%)</td>
<td>4 (14%)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>