Limited Accuracy of Administrative Data for the Identification and Classification of Adult Congenital Heart Disease

Abigail Khan, MD; Katrina Ramsey, MPH; Cody Ballard, PA-C; Emily Armstrong, PA-C; Luke J Burchill, MBBS, PhD; Victor Menashe, MD; George Pantely, MD; Craig S Broberg, MD, MCR

Background—Administrative data sets utilize billing codes for research and quality assessment. Previous data suggest that such codes can accurately identify adults with congenital heart disease (CHD) in the cardiology clinic, but their use has yet to be validated in a larger population.

Methods and Results—All administrative codes from an entire health system were queried for a single year. Adults with a CHD diagnosis code (*International Classification of Diseases, Ninth Revision*, (*ICD-9*) codes 745–747) defined the cohort. A previously validated hierarchical algorithm was used to identify diagnoses and classify patients. All charts were reviewed to determine a gold standard diagnosis, and comparisons were made to determine accuracy. Of 2399 individuals identified, 206 had no CHD by the algorithm or were deemed to have an uncertain diagnosis after provider review. Of the remaining 2193, only 1069 had a confirmed CHD diagnosis, yielding overall accuracy of 48.7% (95% confidence interval, 47–51%). When limited to those with moderate or complex disease (n=484), accuracy was 77% (95% confidence interval, 74–81%). Among those with CHD, misclassification occurred in 23%. The discriminative ability of the hierarchical algorithm (C statistic: 0.79; 95% confidence interval, 0.77–0.80) improved further with the addition of age, encounter type, and provider (C statistic: 0.89; 95% confidence interval, 0.88–0.90).

Conclusions—*ICD* codes from an entire healthcare system were frequently erroneous in detecting and classifying CHD patients. Accuracy was higher for those with moderate or complex disease or when coupled with other data. These findings should be taken into account in future studies utilizing administrative data sets in CHD. (*J Am Heart Assoc*. 2018;7:e007378. DOI: 10.1161/JAHA.117.007378.)

Key Words: administrative data • congenital heart disease • diagnosis • diagnosis code

Adults with congenital heart disease (CHD) represent a growing subset of the population and are currently estimated at 1.4 million in the United States.[1] In response, research on relevant outcomes, quality, and cost of care is now being pursued on a larger scale. In addition, there is an interest in developing a better understanding of the adult CHD (ACHD) population on a systems level, prompted by a strong focus on the high rate of resource utilization and interest in quality improvement.[2,3] Research in this field has been hampered by the heterogeneity of CHD diagnoses and

From the Adult Congenital Heart Disease Program, Knight Cardiovascular Institute (A.K., C.B., E.A., L.J.B., G.P., C.S.B.), Divisions of Biostatistics and Epidemiology (K.R.), Pediatric Cardiology (V.M.), Oregon Health & Science University, Portland, OR.

Correspondence to: Abigail Khan, MD, UHN 62, Knight Cardiovascular Institute, 3181 SW Sam Jackson Park Rd., Portland, OR 97239. E-mail: khaab@ohsu.edu

intervention and by the challenges of assembling large enough cohorts to accurately measure outcomes.

To address the wide knowledge gaps regarding CHD care, researchers have utilized administrative data sets with increasing frequency. These allow for the aggregation of large numbers of patients within or across health systems.[4] These data sets, which, in the United States, typically rely on *International Classification of Disease (ICD)* codes, are naturally suited to this purpose. *ICD* codes identify patient diagnoses and link them to specific patient encounters or procedures. Although the strengths of *ICD* codes include their ubiquity and ease of use, they also have limitations, such as their lack of granularity in CHD specifically. *ICD* codes, for instance, indicate the underlying diagnosis but do not provide further information about prior repairs or anatomical features that may be associated with outcomes. Furthermore, *ICD* codes are often entered by individuals who lack specific knowledge about congenital heart defects, and no mechanisms exist to verify the accuracy of such data at the time of either entry or subsequent analysis. Consequently, there are significant concerns regarding any conclusions drawn from research using such data.

Clinical Perspective

What Is New?

- This study demonstrates that *International Classification of Diseases (ICD)* codes frequently fail to correctly identify patients with adult congenital heart disease in a hospital-wide database.
- Research using *ICD* codes for adult congenital heart disease may be hampered by significant inaccuracy unless analyses are limited to selected cohorts, such as those with moderate or complex disease.

What Are the Clinical Implications?

- In addition to their research applications, *ICD* codes may also be used for administrative purposes, such as patient tracking, and for quality initiatives.
- Based on the results of this study, these initiatives should proceed with caution because there is high potential for incorrect identification of adults with congenital heart disease.
- Strategies such as limiting analyses to those with moderate or complex disease may improve accuracy but are still likely to yield imperfect results.
- This shortcoming should be acknowledged in the interpretation of data gathered for these purposes.

Previous research from our group demonstrated that *ICD, Ninth Revision* (*ICD-9*) codes can be used to identify ACHD patients within the cardiology clinic, with high sensitivity and specificity.[5] However, the use of *ICD* codes for CHD patients outside the specialty clinic has not yet been validated. For several reasons, *ICD* codes may be considerably less accurate when utilized outside a cardiology setting. This has important implications for the accuracy and validity of work using administrative data sets across large healthcare systems without reference to other factors such as type of provider. We sought to examine the validity of *ICD* used across a large tertiary referral center for both identifying and classifying patients with CHD and to examine major sources of error. Our secondary objective was to determine whether the addition of commonly available variables within an algorithmic analysis could be used to improve the accuracy of the data for CHD determination.

Methods

The data, analytic methods, and study materials will not be made available to other researchers for purposes of reproducing the results or replicating the procedure. Study approval did not include the use of such repositories, and there is an abundance of personal identifiers in the data set. The study was approved by the institutional review board at the Oregon Health and Science University. The informed consent requirement was waived.

Identification of Patients

The electronic health record (EHR) system for Oregon Health and Science University Hospital was queried for all adults receiving care in 2010. Patients with any *ICD-9* code of 745 to 747 related to CHD were identified as the initial cohort of interest. To better ensure representation of all applicable patients, relevant codes from any year were included as long as the patient received care during the target calendar year. For example, a patient coded in 2009 with "coarctation" and then seen in 2010 for an ophthalmology examination would be included.

Using the available codes, we applied a hierarchical algorithm (1) to determine whether CHD was truly present and (2) to classify patients into 1 of 14 major defect subgroups based on the codes (Table 1). This algorithm was previously tested, revised, and validated in a limited clinical setting in which CHD patients are common.[5] The hierarchical algorithm ranks codes by diagnostic complexity. For example, code 745.2 ("tetralogy of Fallot") would take precedence over code 745.4 ("ventricular septal defect") in a given patient, and code 745.1 ("transposition of the great arteries") would be superseded by code 745.3 ("common ventricle"). The algorithm included limited physiologic codes ("cyanosis," "pulmonary hypertension"), syndromes ("Shone syndrome," "hypoplastic left heart syndrome"), and modifiers for past procedures ("s/p Fontan"), given their association with CHD, whenever available.

Separately, health records including clinic notes, progress reports, and imaging tests for each patient in the cohort were reviewed by a cardiology provider to determine the accuracy of the algorithmic coding. Demographics, codes, encounters, dates, provider names, and clinic notes were exported from the EHR to a customized database for ease of viewing and assigning categories. To improve accuracy, we reviewed all available notes from any year, not just those from the target calendar year. Based on review, each patient was dichotomously determined to be either CHD or non-CHD. Next, providers selected the most appropriate defect subgroup for those with CHD. If designations were not clear from initial review, the patient was flagged for secondary review within the EHR by a second provider.

Conditions such as hypertrophic cardiomyopathy, mitral valve prolapse, familial aortopathies (including Marfan syndrome), inheritable cardiomyopathies, and channelopathies were excluded, as were iatrogenic complications such as a ventricular septal defect or pulmonary vein stenosis, consistent with our objective to identify only patients who would be

Table 1. Overview of Identification and Classification of Adults With CHD

Subgroup	Total by Codes,* n	Total by Review,* n	Correct by Codes,* n	Missed, n (%)	Incorrect by Codes,* n	Wrongly Included, n (%)
Eis/cyanosis/PAH	19	20	19	1 (5)	0	0
Fontan	96	46	37	9 (20)	23	59 (61)
TGA-DORV	118	90	86	4 (4)	18	32 (27)
TOF-PA-truncus	156	142	131	11 (8)	14	25 (16)
Coarctation	134	117	107	10 (9)	15	27 (20)
AVSD	4	58	4	54 (93)	0	0
Ebstein anomaly	23	20	19	1 (5)	5	4 (17)
Pulmonary vein anomaly	20	17	13	4 (24)	2	7 (35)
Subaortic stenosis	13	9	5	4 (44)	2	8 (62)
Anomalous coronary	44	17	15	2 (12)	29	29 (66)
Pulmonary stenosis	56	36	29	7 (19)	12	27 (48)
Shunts	647	193	171	22 (11)	420	476 (74)
Bicuspid aortic valve	182	268	154	114 (43)	26	28 (15)
Other	681	36	30	6 (17)	558	651 (96)
Total	2193	1069	820	249 (23)	1124	1373 (63)

AVSD indicates atrioventricular septal defect; CHD, congenital heart disease; Eis, Eisenmenger syndrome; PAH, pulmonary arterial hypertension; TGA-DORV, transposition of the great arteries or double-outlet right ventricle; TOF-PA-truncus, tetralogy of Fallot, pulmonary atresia, or truncus arteriosus.
*The total number of patients by manual review, by algorithmic interpretation of coded data, and the numbers of patients either missed or wrongly included.

commonly evaluated in a CHD clinic. Certain diagnoses commonly associated with other anatomic heart defects, namely, isolated dextrocardia, heterotaxy, or left superior vena cava (SVC), were considered congenital because such diagnoses may warrant a visit to a referral center for consideration of other potential congenital findings. Syndromes such as Down syndrome or Turner syndrome were, similarly, not considered congenital unless specific cardiac defects were identified. Vascular conditions such as an enlarged aorta in the absence of CHD such as patent ductus arteriosus, coarctation, bicuspid valve, or family history were not considered congenital.

Reconciliation of certain codes warranted special consideration. The code for an atrial-level shunt was deemed CHD, as opposed to a patent foramen ovale (PFO) if there was progress note documentation of an atrial septal defect, specifically, right heart enlargement, coexisting pulmonary hypertension, or cyanosis, or a confirmatory imaging/procedure report stating atrial septal defect; therefore, a patient with a prior stroke and right-to-left shunt of agitated saline during echocardiography was not considered CHD. Codes for congenital aortic valve abnormality were considered CHD if a previous clinic note, imaging test, or operative note documented a specific congenital valve lesion.

After completion of the primary review, a secondary review was performed only for flagged records (n=188). For each record, providers determined whether a CHD diagnosis was indeed present, absent, uncertain, or incomplete by reviewing the entire EHR. Uncertain diagnoses were those in which there was clinical debate about a congenital abnormality, such as a bicuspid versus tricuspid aortic valve, or whether a small pulmonary valve gradient could be considered mild congenital stenosis. Incomplete records were those in which the EHR documentation was inadequate to conclusively arbitrate, such as a code for congenital heart anomaly in an echo referral but no clinic visits to review, a history of prior aortic valve surgery in which valve morphology was not described, or patients with suspected CHD referred for further workup but who did not complete the evaluation. All questionable, uncertain, or incomplete designations were thereafter reviewed by a third provider for a final designation.

For each patient in whom the algorithm either misidentified a congenital patient or wrongly classified them by diagnosis, a third review was performed to confirm the finding and determine the source of error if possible. Errors were classified into common types and summarized.

Finally, the source of codes was also considered. Codes associated with a clinic visit were identified by provider type, specifically, ACHD, pediatric cardiology, general cardiology, obstetrics, or other. Codes associated with either echocardiography or ECG were also noted, as were those associated with an inpatient encounter. Codes that came from general problem lists were not linked to a specific encounter or provider. More than 1 source of codes was potentially identified for each patient. The predictive value of the codes was compared by source.

Statistical Analyses

Algorithmic determinations of CHD were compared with provider review for each patient. Each patient's designation as CHD by codes was determined to be correct, missed, or wrongly included. The overall accuracy of administrative codes was determined with 95% confidence intervals (CIs). For each defect subgroup, we tabulated the number of individuals incorrectly included and the number missed, both expressed as a percentage of the total number of those within that subgroup found by provider review. Because of considerable disparity in the size of subgroups, we used the Jeffreys 95% CI.[6] For those with CHD by provider review, we quantified the number of patients misclassified into defect subgroups, both overall and by initial subgroup category.

To evaluate our ability to use additional EHR or claims data to identify individuals with true CHD, we compared the area under the curve (AUC) of predicted probabilities from logistic models that included candidate variables singly and in combination. The AUC reflects the probability that a randomly selected true case will have higher prediction than a randomly selected noncase and ranges from 0.5 for a poor classifier (equal chance of success or failure) to 1.0 for a perfect classifier. To avoid overfitting and small groups in our *ICD-9*–based variables, we generated 3 indicators to reflect 4 levels of defect subgroups determined by their previously estimated positive predictive values (grouped conceptually as 0–0.29, 0.3–0.59, 0.6–0.85, and 0.86–1.0). These constituted our initial model. Additional variables included age, sex, and encounters with specialists in ACHD or adult or pediatric cardiology, obstetrics, or echocardiogram or ECG testing. We used a forward stepwise procedure to add predictors, retaining those that resulted in a significant gain in AUC ($P<0.05$ or change in AUC >0.01) for the predicted probabilities from the logistic model. Although encounters with CHD specialists were significant predictors of a correct CHD diagnosis, we speculated that these encounters may not always be distinguishable from other office encounters and calculated our final model both with and without an indicator for this type of encounter. We also estimated the marginal effects of each variable in the final model to aid the interpretation of the variable's contribution to predicted probability of a true diagnosis. Variables with large marginal effects may not increase the AUC substantially if they affect only small numbers of patients or if they overlap with other effects in the model. As a final step, we performed 5-fold cross-validation. Analyses were performed using Stata version 14 (StataCorp).

Results

The initial query yielded 2399 potential adults with any CHD-related *ICD-9* code who were seen for outpatient or inpatient care or procedures during the 2010 calendar year. Of these, 146 did not have a CHD diagnosis by administrative codes, mainly those with only physiologic codes such as cyanosis without a structural heart defect. In addition, 36 were considered incomplete and 24 uncertain. The remaining 2193 were algorithmically found to have code-based CHD and constituted the main study cohort.

By provider review, 1069 patients were confirmed with a CHD diagnosis, whereas the remaining 1124 individuals were not; therefore, the overall accuracy was 48.7% (95% CI, 47–51%). CHD patients tended to be younger than non-CHD patients (43.9 ± 19.4 versus 58.2 ± 15.6 years, $P<0.001$) and more evenly matched by sex (51% versus 60% female in those without CHD, $P<0.001$). A breakdown of accuracy by defect subgroup is shown in Table 1. When only moderate or complex lesions were included, the administrative codes identified 627 patients, 484 of which had CHD by provider review (77.2%; 95% CI, 74–81%).

Sources of Error

Errors resulting in misidentification as CHD are given in Table 2. The most common was the erroneous inclusion of PFO patients as having an atrial shunt (n=274). Similarly, several patients who were seen in a hereditary hemorrhagic telangiectasia clinic with intrapulmonary shunts were also miscoded as having an intracardiac shunt. The second most common source of error (n=123) was the inclusion of pregnant women in whom the fetus had CHD, either known or suspected. Tricuspid aortic valve was frequently miscoded as bicuspid valve (n=70). Other types of acquired valve disease (usually mitral stenosis or regurgitation) were often coded as congenital (n=37). Interestingly, patients with multiple sclerosis ("MS") were at times miscoded as having congenital mitral stenosis (n=16). Some were flagrant errors: Codes for specific conditions including tetralogy of Fallot, coarctation, Ebstein anomaly, and hypoplastic left heart syndrome were assigned to patients with any suspected rationale (n=13). Such codes were in some cases erroneously carried forward from one clinical encounter to another, meaning that the number of instances that the code was encountered had no bearing on accuracy. Another error was the inclusion of patients with prior CHD who had undergone heart transplant (n=13). The bulk of the remaining miscellaneous errors were either uncertain or too infrequent to classify but included screening visits for CHD when none was present, thoracic anatomic problems (pericardial cyst, hypoplastic lung), muscular dystrophy, iatrogenic ventricular septal defect or pulmonary vein stenosis, and a healthy control in a CHD research study, among others.

Table 2. Sources of Error in the Identification of CHD Patients Using Administrative Codes

Error Type	n	Incorrectly Identified Patients (%)
PFO misclassified as ASD	274	24
Fetal CHD mistakenly applied to mother	123	11
TAV mistaken for BAV	70	6
Noncongenital valve problems (MS, MR)	37	3
Intrapulmonary shunt mistaken for ASD	20	2
Congenital arrhythmias misclassified as other congenital	18	2
Multiple sclerosis coded as mitral stenosis	16	1
Vascular problems mistaken for congenital aortic disease	15	1
Inexplicable coding error	13	1
Posttransplant patients with prior CHD	13	1
Dilated aortas misclassified as congenital aortopathy	12	1
Myocardial bridge misclassified as anomalous coronary	12	1
HCM misclassified as subaortic stenosis	11	1
Patient history that was erroneous	6	1
Unclassified/miscellaneous	484	43
Total	1124	100

ASD indicates atrial septal defect; BAV, bicuspid aortic valve; CHD, congenital heart disease; HCM, hypertrophic cardiomyopathy; MR, mitral regurgitation; MS, mitral stenosis; PFO, patent foramen ovale; TAV, tricuspid aortic valve

Categorization

Of the 1069 patients with confirmed CHD, 820 (77%) were correctly classified into 1 of the 14 defect subgroups. The remaining 249 (23%) were wrongly classified, although correctly diagnosed as having CHD. A breakdown of accuracy by defect type is shown in Table 3. Codes for the category of Eisenmenger syndrome, cyanosis, and pulmonary arterial hypertension and for Ebstein anomaly were correct in 19 of 20 patients each. Conversely, only 4 of 58 patients with atrioventricular septal defect were correctly categorized; the remaining patients were either misclassified as having a simple shunt or a single ventricle, largely due to the overlapping code for common atrioventricular connection with single ventricle. A wide range of other values were found across other defect categories, none of which identified lesions perfectly (Figure).

Source of Codes

For patients within the final cohort, 1753 different codes in 800 individual patients could be identified in at least 1 of our target sources by provider (ACHD, pediatric cardiology, general cardiology, or obstetrics) or procedure (echocardiogram or ECG) or by service location (inpatient). For the remaining 1393 patients, the codes originated from other types of providers or from general problem lists. Comparatively, the latter patients were older (55.8±19.3 versus 43.2±15.6 years, $P<0.001$) and more often female (60% versus 48%, $P<0.001$) than those with an identifiable code source. For those with an identifiable source, the number of patients and the accuracy of codes for detecting CHD are shown in Table 4. The codes for patients who had encounters with an ACHD provider were the most accurate (93%; 95% CI, 91–95%), followed by inpatient codes (80%; 95% CI, 74–84%). The accuracy of codes from obstetric providers was far lower (16%; 95% CI, 5–27%), often reflecting fetal CHD diagnosis rather than a maternal diagnosis. The accuracy of classification into defect subgroups was generally similar between sources, with inpatient codes being the most reliable.

Algorithm Optimization

The performance of the hierarchical algorithm for the identification of individuals with CHD was measured using ROC curves. The C statistic for the algorithm alone was 0.79 (95% CI, 0.77–0.80). By comparison, the C statistic based on an ACHD specialist encounter was 0.72 (95% CI, 0.71–0.74). Demographic characteristics alone, such as age (C statistic: 0.71; 95% CI, 0.69–0.73) and sex (C statistic: 0.54; 95% CI, 0.52–0.56) showed only fair or poor discriminative ability. The addition of other factors to the model (hierarchy plus age, fetal echocardiogram, echocardiogram or ECG, and/or ACHD encounter) improved discriminative ability (C statistic: 0.89; 95% CI, 0.88–0.90; Table 5).

Discussion

We found a large number of patients within a large academic healthcare institution to be inaccurately identified as having CHD based on administrative codes. The algorithmic interpretation of such codes when applied to a university-wide data set was far worse than when limited to patients seen by ACHD providers.[5] These findings have important implications for projects that have or will seek to use similar administrative data sets to perform research or quality improvement efforts in ACHD. In addition, the findings provide better understanding of the types of errors inherent in such data sets and how to potentially avoid them.

The lower accuracy of such codes in this cohort than in a cardiology practice cohort reflects both the lower prevalence of such patients in the general population and the higher prevalence of overlapping conditions, especially acquired valve disease and PFO.[5]

Table 3. Comparison Between the Algorithm-Based Diagnosis and the Final Diagnosis After Manual Provider Review

Code-Based Diagnosis	Final Diagnosis After Provider Review															Total
	Eis/Cyanosis/PAH	Fontan	TGA-DORV	TOF-PA-Truncus	Coarctation	AVSD	Subaortic Stenosis	Ebstein Anomaly	Anomalous PV	Anomalous Coronary	Pulmonary Stenosis	Shunts	Bicuspid Aortic Valve	Other	Not ACHD	
Eis/cyanosis/PAH	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19
Fontan	0	37	3	5	3	23	0	0	0	0	1	0	1	0	23	96
TGA-DORV	1	4	86	3	2	0	0	0	0	0	1	2	1	0	18	118
TOF-PA-truncus	0	0	0	131	0	0	1	1	0	0	4	4	1	0	14	156
Coarctation	0	0	0	0	107	1	0	0	0	0	0	2	7	2	15	134
AVSD	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	4
Subaortic stenosis	0	0	0	0	0	0	5	0	0	0	0	0	3	0	5	13
Ebstein anomaly	0	0	0	0	0	0	0	19	0	0	0	1	1	0	2	23
Anomalous PV	0	0	0	0	0	1	0	0	13	0	0	1	3	0	2	20
Anomalous coronary	0	0	0	0	0	0	0	0	0	15	0	0	0	0	29	44
Pulmonary stenosis	0	2	0	0	0	1	0	0	0	0	29	2	7	3	12	56
Shunts	0	1	0	2	1	26	3	0	2	1	0	171	19	1	420	647
Bicuspid aortic valve	0	0	0	0	0	0	0	0	0	0	0	2	154	0	26	182
Other	0	2	1	1	4	2	0	0	2	1	1	8	71	30	558	681
Total	20	46	90	142	117	58	9	20	17	17	36	193	268	36	1124	2193
Correctly classified	19	37	86	131	107	54	5	19	13	15	29	171	154	30	···	820
Misclassified	1	9	4	11	10	4	4	1	4	2	7	22	114	6	···	249
Correct (%)	95	80	96	92	91	93	56	95	76	88	81	89	57	83	···	

ACHD indicates adult congenital heart disease; AVSD, atrioventricular septal defect; Eis, Eisenmenger syndrome; PAH, pulmonary arterial hypertension; PV, pulmonary vein; TGA-DORV, transposition of the great arteries or double-outlet right ventricle; TOF-PA-truncus, tetralogy of Fallot, pulmonary atresia or truncus arteriosus.
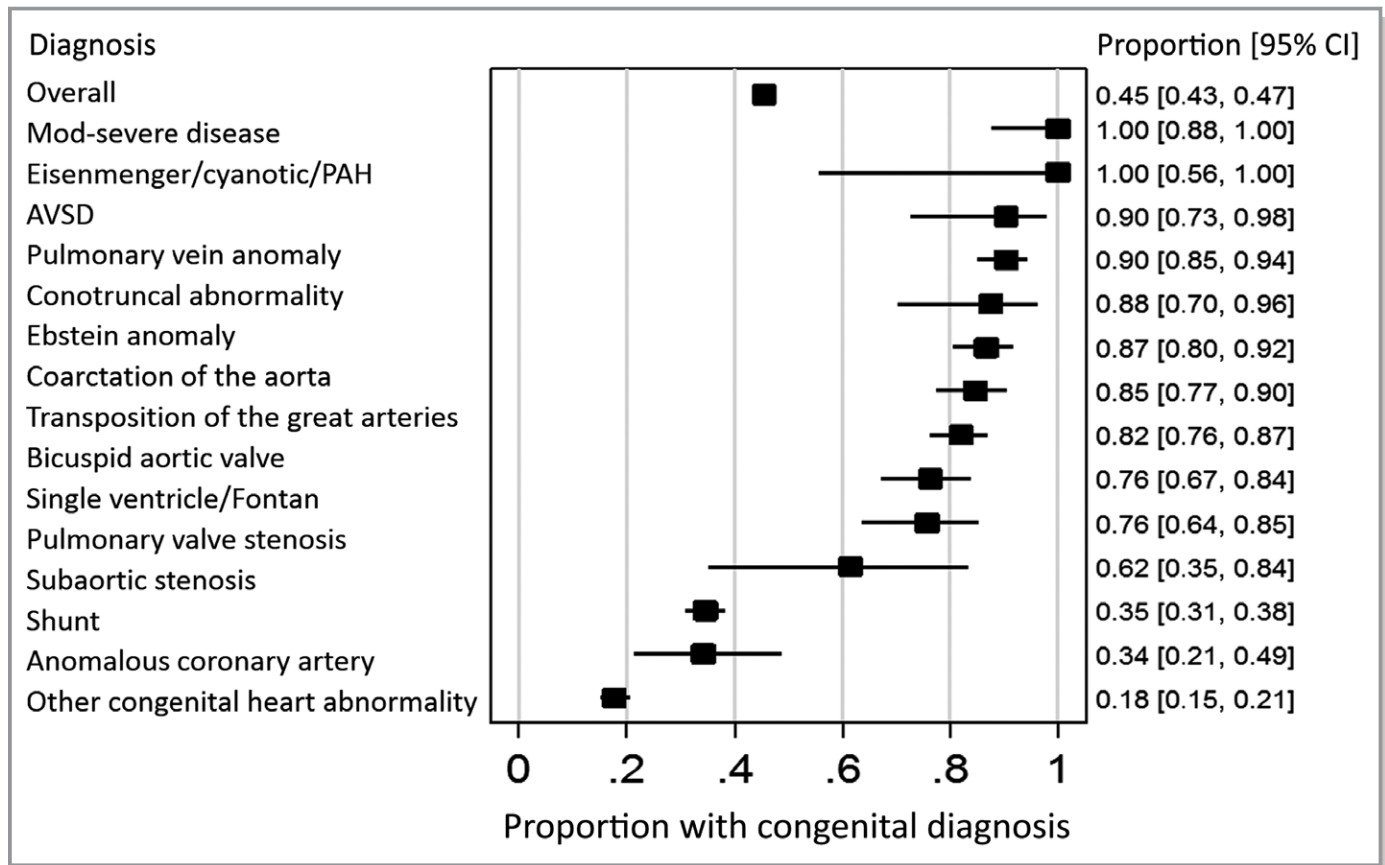
Figure. Proportion of diagnoses initially identified using codes that were found to be congenital after chart review, by defect subgroup. AVSD indicates atrioventricular septal defect; Mod, moderate; PAH, pulmonary arterial hypertension.

Administrative codes have been found to have variable accuracy in other studies. A recent meta-analysis of the validity of heart failure diagnoses in administrative data sets suggests that the specificity and positive predictive value of these codes are high but with low sensitivity.[7] Other studies have shown that the positive predictive value of *ICD* codes is highly variable[8–11] but also that the performance of these codes may be improved by the application of algorithms or the addition of multiple associated billing codes or further clinical information, such as we have tried to include.[8,12] Notably, the current investigation examined the specificity of ACHD codes but not the sensitivity, given the lack of ability to identify the total number of ACHD patients across the hospital system.

As expected, the less accurate codes tended to be those associated with less complex types of ACHD, and codes were more accurate when simple defects were excluded. Complex lesions such as Eisenmenger syndrome or cyanotic heart disease were accurate, whereas codes for simple lesions such as shunts were less so, with significant heterogeneity between these extremes.

The use of administrative codes has several limitations, including (1) the inherent limitations of the codes themselves, (2) provider misuse, and (3) medical uncertainty. Examples of all these were found in this analysis. Inherent limitations include the lack of discriminatory detail in the *ICD-9* classification scheme, which, despite its comprehensive nature, was never designed to extensively differentiate the CHD spectrum. The 2 often-cited examples are the inability to distinguish between atrial septal defect and PFO or bicuspid aortic valve from a dysfunctional tricuspid valve. Text

Table 4. Analysis of Code Accuracy by the Source of Codes

Source	n	CHD, n (%)	95% CI, %	Misclassified as CHD, n (%)
Adult CHD	563	524 (93)	91–95	114 (22)
Pediatric cardiology	273	159 (58)	52–64	33 (21)
Adult general cardiology	519	283 (55)	50–59	62 (22)
Echocardiography	610	434 (71)	68–75	103 (24)
ECG	502	292 (58)	54–63	72 (25)
Fetal/obstetric visit	43	7 (16)	5–27	2 (29)
Inpatient	251	200 (80)	75–85	36 (18)

CHD indicates congenital heart disease; CI, confidence interval.

Table 5. Success of Variables From Administrative Database in Identifying CHD, as Measured by AUC and Marginal Effect on Predicted Probabilities From the Logistic Model

	Probability of CHD	Effect of variables on predicted probability of CHD+ from logistic model (marginal effects)			
	Defect subgroups alone		Including age and encounter types		
Defect subgroups from *ICD-9* codes					
Group 1 [base] (n=356): Eis/cyanosis/PAH, conotruncal abnormality, coarctation of the aorta, AVSD, Ebstein anomaly, PV anomaly	0.91 (0.88–0.94)	Ref	Ref	Ref	Ref
Group 2 (n=465): single ventricle/Fontan, TGA, pulmonary valve stenosis, bicuspid aortic valve, subaortic stenosis	0.82 (0.78–0.85)	−0.09 (−0.13 to −0.04)	−0.09 (−0.15 to −0.04)	−0.08 (−0.13 to −0.02)	−0.03 (−0.10 to 0.04)
Group 3 (n=691): anomalous coronary artery, shunt	0.35 (0.31–0.39)	−0.56 (−0.60 to −0.51)	−0.50 (−0.55 to −0.44)	−0.47 (−0.52 to −0.41)	−0.36 (−0.43 to −0.30)
Group 4 (n=681): other congenital heart abnormalities	0.18 (0.15–0.21)	−0.73 (−0.77 to −0.68)	−0.67 (−0.72 to −0.62)	−0.62 (−0.67 to −0.56)	−0.48 (−0.54 to −0.41)
Age: each 10-y increase after age 18			−0.05 (−0.06 to −0.04)	−0.06 (−0.06 to −0.05)	−0.04 (−0.05 to −0.03)
Encounter types					
Obstetrics				−0.39 (−0.46 to −0.32)	−0.36 (−0.43 to −0.29)
ECG or echocardiography				+0.15 (0.11–0.18)	+0.05 (0.02–0.09)
Specialist in ACHD					+0.36 (0.31–0.42)
C statistic/AUC		0.82 (0.80–0.83)	0.85 (0.83–0.87)	0.87 (0.86–0.89)	0.89 (0.88–0.90)

The 95% confidence intervals are shown in parentheses. C statistic/AUC, the probability that a randomly selected true congenital case has higher prediction from the logistic model than a randomly selected noncongenital patient, reflects performance of the combination of variables as classifiers. Variables were excluded because of poor performance, sex, adult and pediatric cardiology encounters. ACHD indicates adult congenital heart disease; AUC, area under the curve; AVSD, atrioventricular septal defect; CHD, congenital heart disease; Eis, Eisenmenger syndrome; *ICD-9, International Classification of Diseases, Ninth Revision*; PAH, pulmonary arterial hypertension; PV, pulmonary vein; TGA, transposition of the great arteries.

descriptors that accompany the codes at the time of selection are often more comprehensive but map to the same numeric code, meaning that differentiation is lost when the descriptor disappears. In the future, text-based analysis of the EHR may be helpful in this regard, although it was not possible for this study.

Furthermore, we identified several important operational errors. Although uncommon, some specific CHD codes seemed to have found their way into the chart of a non-CHD patient and stayed there uncorrected and replicated across multiple encounters. An unexpected source of error was fetal abnormalities attributed to the mother. Another was the inclusion of diagnoses with similar acronyms such as "MS," which, when typed into a search field, will identify 2 distinct diagnoses. Such errors tended to be rare, and a certain amount of miscoding would be expected in any large data set. Misclassification can occur because of lack of understanding on the part of those entering the codes. For example, support staff without specific expertise in CHD may

theoretically be more apt to select a generic code such as "congenital heart anomaly" in place of a more specific code. In our study, however, we found only 14 patients with generic codes that had a moderate or complex defect for which a more descriptive code could have been used. Most often, such codes were not indicative of true CHD. Inpatient codes are often reviewed by professional coders, which may explain the relatively better performance we found from such codes.

Moreover, poor discrimination may be inherent in the disease itself. There may be diagnostic uncertainty based on questionable or inconsistent imaging findings, such as cases in which a valve was assumed to be bicuspid as an etiology of valve dysfunction without clear visualization of inherent valve morphology. There may be questions about what is truly congenital. Conditions such as mitral valve prolapse, hypertrophic cardiomyopathy, or long QT syndrome can all be said to be existing at or before birth to a degree, yet they are not generally thought of as congenital CHD. We note that providers are rarely, if ever, trained in the utility of administrative coding.

Our data also highlight the importance of further refining CHD case definitions and improving the breadth and accuracy of *ICD* codes to better reflect the ACHD population.

A major goal of our study was not only to test the accuracy of administrative data but to fine-tune an algorithm for the identification and classification of patients with CHD across a healthcare system by inclusion of other information that may likely be available in administrative data sets. The overall discriminative ability of our algorithm, as measured by the C statistic, was improved by the addition of other factors that, we postulated, could affect the likelihood of CHD—namely, age, an encounter with an ACHD provider, and echocardiography and ECG. Interestingly, age alone had fair predictive value for the diagnosis of ACHD, in part because it forced the exclusion of some of the more common acquired conditions in older patients. Age may be less discriminatory in the future as the CHD population ages. Surprisingly, a visit with either an adult or pediatric cardiologist did not provide additive value, suggesting that perhaps the cardiologists did not modify inappropriate codes when they were already provided. This highlights the importance of appropriate training in administrative coding for physicians.

The *ICD,10th Revision* (*ICD-10*) codes significantly expand the number of available codes and, theoretically, will increase their specificity in ACHD patients. Nevertheless, it is important to note that many of the limitations we describe will still be relevant to *ICD-10*—based data sets. First, *ICD-10* granularity for CHD diagnoses is still imperfect. For example, the *ICD-10* codes still fail to differentiate between atrial septal defect and PFO. Second, because this type of research is typically done in a retrospective manner, researchers are likely to rely on *ICD-9*—based databases for the foreseeable future.

As noted, our study was unable to measure true sensitivity because of an inability to identify patients with ACHD who were not coded as such. Instead, it examined the accuracy of ACHD codes in patients who were given a diagnosis of ACHD. This would require reviewing records of all patients seen across the university, which was not feasible. We ensured that all patients seen in the ACHD clinic identified in our previous study were correctly detected through the coded data and that patients excluded by the algorithm, in fact, did not have CHD; however, this did not include patients seen outside of the ACHD clinic setting, some of whom may have had ACHD without being properly coded as such.

Importantly, the 14 defect subgroups, although convenient, do not account for the significant variation between patients within those groups. The type of surgical repair and/or intervention, for example, can affect care pathways and outcomes. This information is relevant to researchers but is not well differentiated by administrative data sets. We were also unable to identify the type of provider entering codes or whether codes were entered by a provider or by support staff.

Ours was a single-site study performed at the only ACHD center for the state of Oregon. Although we postulate that these findings are likely to be generalizable to other academic hospital populations, physician behaviors across 1 healthcare system may not be the same, and we cannot be sure that our findings are freely generalizable to other multi-institutional administrative data sets. Further validation would be required to determine whether our algorithm is applicable in health systems other than our own.

Healthcare data based on *ICD-9* codes for the identification and classification of CHD may have significant errors, particularly as related to simple CHD. Those seeking to use administrative data for research or other purposes should take specific codes into account and consider limiting their analyses to those with moderate or complex disease.

Disclosures

None.

References

1. Gilboa SM, Devine OJ, Kucik JE, Oster ME, Riehle-Colarusso T, Nembhard WN, Xu P, Correa A, Jenkins K, Marelli AJ. Congenital heart defects in the United States: estimating the magnitude of the affected population in 2010. *Circulation*. 2016;134:101–109.

2. Mackie AS, Pilote L, Ionescu-Ittu R, Rahme E, Marelli AJ. Health care resource utilization in adults with congenital heart disease. *Am J Cardiol*. 2007;99:839–843.

3. Gurvitz M, Marelli A, Mangione-Smith R, Jenkins K. Building quality indicators to improve care for adults with congenital heart disease. *J Am Coll Cardiol*. 2013;62:2244–2253.

4. Riehle-Colarusso TJ, Bergersen L, Broberg CS, Cassell CH, Gray DT, Grosse SD, Jacobs JP, Jacobs ML, Kirby RS, Kochilas L, Krishnaswamy A, Marelli A, Pasquali SK, Wood T, Oster ME; Congenital Heart Public Health Consortium. Databases for congenital heart defect public health studies across the lifespan. *J Am Heart Assoc*. 2016;5:e004148. DOI: 10.1161/jaha.116.004148.

5. Broberg C, McLarry J, Mitchell J, Winter C, Doberne J, Woods P, Burchill L, Weiss J. Accuracy of administrative data for detection and categorization of adult congenital heart disease patients from an electronic medical record. *Pediatr Cardiol*. 2015;36:719–725.

6. Brown LD, Cai TT, DasGupta A. Interval estimation for a binomial proportion. *Stat Sci*. 2001;16:101–133.

7. McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of heart failure diagnoses in administrative databases: a systematic review and meta-analysis. *PLoS One*. 2014;9:e104519.

8. Chung CP, Rohan P, Krishnaswami S, McPheeters ML. A systematic review of validated methods for identifying patients with rheumatoid arthritis using administrative or claims data. *Vaccine*. 2013;31(suppl 10):K41–K61.

9. van Walraven C, Austin P. Administrative database research has unique characteristics that can risk biased results. *J Clin Epidemiol*. 2012;65:126–131.

10. Williams DJ, Shah SS, Myers A, Hall M, Auger K, Queen MA, Jerardi KE, McClain L, Wiggleton C, Tieder JS. Identifying pediatric community-acquired pneumonia hospitalizations: accuracy of administrative billing codes. *JAMA Pediatr*. 2013;167:851–858.

11. Taneja C, Berger A, Inglese GW, Lamerato L, Sloand JA, Wolff GG, Sheehan M, Oster G. Can dialysis patients be accurately identified using healthcare claims data? *Perit Dial Int*. 2014;34:643–651.

12. Fan J, Arruda-Olson AM, Leibson CL, Smith C, Liu G, Bailey KR, Kullo IJ. Billing code algorithms to identify cases of peripheral artery disease from administrative data. *J Am Med Inform Assoc*. 2013;20:e349–e354.

Author/s:
Khan, A; Ramsey, K; Ballard, C; Armstrong, E; Burchill, LJ; Menashe, V; Pantely, G; Broberg, CS

Title:
Limited Accuracy of Administrative Data for the Identification and Classification of Adult Congenital Heart Disease

Date:
2018-01-23

Citation:
Khan, A., Ramsey, K., Ballard, C., Armstrong, E., Burchill, L. J., Menashe, V., Pantely, G. & Broberg, C. S. (2018). Limited Accuracy of Administrative Data for the Identification and Classification of Adult Congenital Heart Disease. JOURNAL OF THE AMERICAN HEART ASSOCIATION, 7 (2), https://doi.org/10.1161/JAHA.117.007378.

Persistent Link:
http://hdl.handle.net/11343/255375

File Description:
Published version
License:
CC BY-NC