Hierarchical Algorithms for Querying Physiologically Distinct Groups in Adult Congenital Heart Disease Using OMOP CDM

<Seohu Lee>[1], <Jong Ko>[2], <Haeun Lee>[1], <Ari Cedars>[2,3]
[1]Biomedical Informatics and Data Science, School of Medicine, Johns Hopkins University, Baltimore, MD, USA
[2]Division of Pediatric Cardiology and [3]Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, USA

Background

Adult Congenital Heart Disease (ACHD) encompasses a spectrum of congenital heart defects that persist from birth into adulthood, affecting the structure of the heart or the blood vessels near the heart. These defects vary significantly in type and severity. The heterogeneity and rarity of ACHD require large and diverse datasets to derive meaningful conclusions, highlighting the necessity and unique utility of multicenter observational research in this population.[1] Engaging in collaborative efforts across multiple centers may enable a more comprehensive understanding of the clinical landscape of ACHD than has previously been achievable. This approach might not only enhance the generalizability of findings but also facilitate the identification of rare disease subtypes or patterns that might be overlooked in smaller, single-center studies.[2, 3]

Data sources mapped to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) offer the opportunity to make such collaborative efforts a reality. The OMOP CDM ensures interoperability and standardization, offering a robust platform for comparative analyses among healthcare systems.[4] By harmonizing data, the OMOP CDM permits us to systematically explore variations in ACHD outcomes, treatment responses, and healthcare utilization practices. This methodology transcends institutional and national boundaries, allowing for a holistic examination of ACHD on a broad scale.

Realizing the promise offered by the OMOP CDM for large-scale multicenter observational ACHD studies, however, requires preliminary groundwork. A first essential step is to determine if ACHD anatomical diagnoses and historical surgical repairs are clearly linked within the OMOP framework, and to define hierarchical algorithms that adequately pool ACHD groups based on physiology. The current study aims to convert ACHD-relevant ICD-10-CM codes to SNOMED CT codes, and then validate them by converting them back to ICD-10-CM codes using two different mapping sites. Additionally, we will define hierarchical algorithms for ACHD physiological groups and analyze patient counts to evaluate and refine these algorithms.

Methods

This study used OMOP CDM databases from Johns Hopkins School of Medicine (JHM). The JHM dataset included Electronic Health Record (EHR) data converted to OMOP CDM from December 2016 to June 2023. The number of patients in the JHM OMOP CDM database is about 2.1 million. To define the Congenital Heart Disease (CHD) concept set, a clinician specializing in ACHD selected 92 ICD-10-CM codes. The concept set includes only the ICD-10-CM codes related to CHD from the ranges Q20-Q26. These ICD-10-CM codes were then converted into SNOMED CT codes using Analytical Tools and Languages for Analysis and Sharing (ATLAS) and Annotations & Terminology Harmonization Ensuring Networked Access (ATHENA). The SNOMED CT codes were reconverted to ICD-10-CM codes using two mapping sites: the

National Library of Medicine (NLM) Interactive Map-Assisted Generation of ICD Codes (I-MAGIC)[5] and International SNOMED[6]. We calculated the percent agreement between the original ICD-10-CM codes and the reconverted ICD-10-CM codes. The process was as follows:

1. Selecting ICD-10-CM codes needed for the ACHD hierarchical algorithms by a clinician.

2. Mapping the selected ICD-10-CM codes to SNOMED CT codes using ATLAS and ATHENA.

3. Reconverting the SNOMED CT codes back to ICD-10-CM codes for validation by two sites.

4. Calculating the percent agreement between the original and reconverted ICD-10-CM codes.

Following an evaluation of the 92 selected ICD-10-CM codes, a clinician defined hierarchical algorithms for the ACHD physiological groups. Using a CHD concept set including these 92 codes, we defined cohorts based on the hierarchical algorithms and generated patient counts. With these patient counts, we discussed the algorithms' effectiveness and potential areas for improvement.

Results

We used the codes mapped as 'Non-standard to Standard (OMOP)' in the *relationship* field of ATLAS and ATHENA to convert 92 ICD-10-CM codes to SNOMED CT codes. Among the 92 codes, 40 codes were agreed upon by both I-MAGIC and SNOMED International after reconverting the codes to ICD-10-CM. 15 codes were only agreed upon by I-MAGIC, and 7 codes were only agreed upon by SNOMED International. The remaining 30 codes were disagreed upon by both I-MAGIC and SNOMED International. We compared the original codes and the reconverted codes that were discrepant with both I-MAGIC and SNOMED International (Table 1).

Table 1. Discrepancy Table of Original and Reconverted Codes of ICD-10-CM

No.	Original Codes	Reconverted Codes	
	ICD-10-CM	I-MAGIC	SNOMED International
1	Q20 Congenital malformations of cardiac chambers and connections	Q24.9 Congenital malformation of heart, unspecified	Q24.9 Congenital malformation of heart, unspecified
9	Q20.8 Other congenital malformations of cardiac chambers and connections	Q24.9 Congenital malformation of heart, unspecified	Q24.9 Congenital malformation of heart, unspecified
10	Q20.9 Congenital malformation of cardiac chambers and connections, unspecified	Q24.9 Congenital malformation of heart, unspecified	Q24.9 Congenital malformation of heart, unspecified
11	Q21 Congenital malformations of cardiac septa	Q21.9 Congenital malformation of cardiac septum, unspecified	Q21.9 Congenital malformation of cardiac septum, unspecified
18	Q21.14 Superior sinus venosus atrial septal defect	Q21.16 Sinus venosus atrial septal defect, unspecified	Q21.1 Atrial septal defect
19	Q21.15 Inferior sinus venosus atrial septal defect	Q21.16 Sinus venosus atrial septal defect, unspecified	Q21.1 Atrial septal defect
21	Q21.19 Other specified atrial septal defect	Q21.10 Atrial septal defect, unspecified	Q21.1 Atrial septal defect
23	Q21.20 Atrioventricular septal defect, unspecified as to partial or complete	Q21.23 Complete atrioventricular septal defect	Q21.2 Atrioventricular septal defect
24	Q21.21 Partial atrioventricular septal defect	Q21.23 Complete atrioventricular septal defect	Q21.2 Atrioventricular septal defect
25	Q21.22 Transitional atrioventricular septal defect	Q21.23 Complete atrioventricular septal defect	Q21.2 Atrioventricular septal defect
29	Q21.8 Other congenital malformations of cardiac septa	Q21.9 Congenital malformation of cardiac septum, unspecified	Q21.9 Congenital malformation of cardiac septum, unspecified

31	Q22 Congenital malformations of pulmonary and tricuspid valves	Q24.8 Other specified congenital malformations of heart	Q24.8 Other specified congenital malformations of heart
39	Q22.8 Other congenital malformations of tricuspid valve	Q22.9 Congenital malformation of tricuspid valve, unspecified	Q22.9 Congenital malformation of tricuspid valve, unspecified
41	Q23 Congenital malformations of aortic and mitral valves	Q24.8 Other specified congenital malformations of heart	Q24.8 Other specified congenital malformations of heart
47	Q23.8 Other congenital malformations of aortic and mitral valves	Q24.8 Other specified congenital malformations of heart	Q24.8 Other specified congenital malformations of heart
48	Q23.9 Congenital malformation of aortic and mitral valves, unspecified	Q24.8 Other specified congenital malformations of heart	Q24.8 Other specified congenital malformations of heart
49	Q24 Other congenital malformations of heart	Q24.9 Congenital malformation of heart, unspecified	Q24.9 Congenital malformation of heart, unspecified
57	Q24.8 Other specified congenital malformations of heart	Q24.9 Congenital malformation of heart, unspecified	Q24.9 Congenital malformation of heart, unspecified
59	Q25 Congenital malformations of great arteries	Q27.9 Congenital malformation of peripheral vascular system, unspecified	Q27.9 Congenital malformation of peripheral vascular system, unspecified
63	Q25.21 Interruption of aortic arch	Q25.49 Other congenital malformations of aorta	Q25.4 Other congenital malformations of aorta
67	Q25.40 Congenital malformation of aorta unspecified	Q25.49 Other congenital malformations of aorta	Q25.4 Other congenital malformations of aorta
70	Q25.43 Congenital aneurysm of aorta	Q25.49 Other congenital malformations of aorta	Q25.4 Other congenital malformations of aorta
71	Q25.44 Congenital dilation of aorta	Q25.49 Other congenital malformations of aorta	Q25.4 Other congenital malformations of aorta
75	Q25.48 Anomalous origin of subclavian artery	Q25.49 Other congenital malformations of aorta	Q25.4 Other congenital malformations of aorta
81	Q25.72 Congenital pulmonary arteriovenous malformation	I28.8 Other diseases of pulmonary vessels	Q25.7 Other congenital malformations of pulmonary artery
83	Q25.8 Other congenital malformations of other great arteries	Q27.9 Congenital malformation of peripheral vascular system, unspecified	Q27.9 Congenital malformation of peripheral vascular system, unspecified
84	Q25.9 Congenital malformation of great arteries, unspecified	Q27.9 Congenital malformation of peripheral vascular system, unspecified	Q27.9 Congenital malformation of peripheral vascular system, unspecified
85	Q26 Congenital malformations of great veins	Q26.9 Congenital malformation of great vein, unspecified	Q26.9 Congenital malformation of great vein, unspecified
90	Q26.4 Anomalous pulmonary venous connection, unspecified	Q26.9 Congenital malformation of great vein, unspecified	Q26.8 Other congenital malformations of great veins
91	Q26.8 Other congenital malformations of great veins	Q26.9 Congenital malformation of great vein, unspecified	Q26.9 Congenital malformation of great vein, unspecified

We calculated the number and percentage of agreements between the original and reconverted ICD-10-CM codes for both I-MAGIC and SNOMED International (Table 2). Of the 92 modified codes, I-MAGIC agreed with 55 codes, and SNOMED International agreed with 47 codes, resulting in agreement rates of 59.78% and 51.09%, respectively. I-MAGIC showed eight more codes with agreement than SNOMED International.

Table 2. Number and Percentage of Agreement with Original and Reconverted ICD-10-CM Codes for I-MAGIC and SNOMED International

	I-MAGIC	SNOMED International
Number of Agreement	55	47
Percentage (%)	59.78	51.09

Using the ACHD concept set of 92 codes, a clinician developed hierarchical algorithms for 10 physiologically distinct ACHD groups based on ICD-10-CM codes. We then converted these codes to SNOMED CT and generated patient counts in ATLAS. All ACHD physiological groups had over 10 patients, except for Fontan/Glenn/Single Ventricle and D-Transposition, which had 2 and 0 patients, respectively, in the JHM ATLAS (Table 3).

Table 3. Hierarchical Algorithms for ACHD Physiological Groups and Corresponding Patient Count from JHM ATLAS

ACHD Physiological Groups	Hierarchical Algorithms with SNOMED CT Code	Patient Count (n)
1. Eisenmenger Syndrome/Shunt with pulmonary hypertension	[(434462 and/or 4099995 and/or 4100152 and/or 4289309 and/or 315922 and/or 4061819) and (4322024 or 4339214)] or 40493243	785
2. Fontan/Glenn/Single Ventricle	4339962 and/or 4208834 and/or 4050559 and/or 2107269 and/or 4051948 and/or 40491942	2
3-1. D-Transposition of the great arteries with atrial switch	(432431 and/or 40456182 and/or 313867) and (4221982 and/or 4075541 and/or 2107361)	0
3-2. D-Transposition of the great arteries with arterial switch	(432431 and/or 40456182 and/or 313867) and (4019932 and/or 4286184 and/or 4077745 and/or 4122006)	0
4. L-Transposition of the great arteries	(432431 and/or 40456182 and/or 313867) and (4100733 and/or 4101005)	59
5. Tetralogy of Fallot/DORV TOF type	313867 and/or 4101618 and/or 320835 and/or 4109337 and/or 4101619	771
6. Truncus arteriosus	441950 and/or 45766266	46
7. AV Canal defects	4100152 and/or 4235784 and/or 435912 and/or 37164933	311
8. Ebstein's anomaly	4069182	100
9. Shone Complex	(313006 and/or 441108 and/or 40404007 and/or 4100869) and (4324704 and/or 4062247 and/or 321119 and/or 4253808 and/or 314457 and/or 259123 and/or 4147787)	18
10. Sinus Venosus	4316879	45

Conclusion

This study demonstrates the feasibility and effectiveness of using the OMOP CDM to harmonize and query data for certain ACHD physiological groups across different coding systems. By converting ICD-10-CM codes to SNOMED CT codes and validating them through I-MAGIC and SNOMED International, we achieved substantial agreement, with I-MAGIC showing slightly higher concordance than SNOMED International. However, SNOMED International does not extend to the second decimal place of the ICD-10-CM codes, which may contribute to its lower agreement rate. If we recalculated the agreement

percentage using only the first decimal place, SNOMED International's number of agreements would increase to 73, resulting in a 79.35% agreement rate, while I-MAGIC's agreement remains unchanged. This reconversion validation process revealed several challenges and disagreements in converting ICD-10-CM to SNOMED CT, which may present barriers for newcomers to OHDSI network studies and clinicians initiating OHDSI research.

The development and application of hierarchical algorithms for various ACHD physiological groups enabled patient identification for ACHD, underscoring the utility of this approach in multicenter observational research. However, the absence of patient counts for specific physiologic groups, such as Fontan/Glenn/Single Ventricle and D-Transposition, indicates a need for further refinement. Future research should validate these findings using real patient data from multiple institutions to ensure accurate categorization at the patient level and broaden the applicability of hierarchical algorithms for ACHD. This preliminary work begins to lay the foundation for large-scale studies, which we anticipate will enhance our understanding and management of ACHD through comprehensive data analysis and collaboration.

References

1. Marelli AJ, Mackie AS, Ionescu-Ittu R, Rahme E, Pilote L. Congenital Heart Disease in the General Population. Circulation. 2007 Jan 16;115(2):163–72.
2. Moons P, Bovijn L, Budts W, Belmans A, Gewillig M. Temporal Trends in Survival to Adulthood Among Patients Born With Congenital Heart Disease From 1970 to 1992 in Belgium. Circulation. 2010 Nov 30;122(22):2264–72.
3. Opotowsky AR, Siddiqi OK, Webb GD. Trends in Hospitalizations for Adults With Congenital Heart Disease in the U.S. Journal of the American College of Cardiology. 2009 Jul;54(5):460–7.
4. Observational Health Data Sciences and Informatics (OHDSI) [Internet]. The Book of OHDSI; [cited 2024 Jun 16]. Available from: http://book.ohdsi.org/
5. I-MAGIC demo tool [Internet]. NLM; [cited 2024 Jun 16]. Available from: https://imagic.nlm.nih.gov/imagic/code/map
6. SNOMED International apping tool [Internet]. SNOMED International; [cited 2024 Jun 16]. Available from: https://mapping.ihtsdotools.org