ICD-10 to SNOMED-CT and OMOP Mapping: Performance, Limitations, and Clinical Implications of Automated Athena Mappings

Kirsten Colpaert[1, 2,3] **(**Kirsten.Colpaert@uzgent.be),
Hadewig Claeys[4] (hadewigclaeys@hotmail.com),
Athiana Ryland[4] (athiana.ryland@hotmail.com),
Maartje Deschacht[5] (Martha.Deschacht@uzgent.be),
Mieke Deschepper[1] (Mieke.Deschepper@uzgent.be)

[1]Data Science Institute, Ghent University Hospital, Ghent, Belgium
[2]Department of Intensive Care, Ghent University Hospital, Ghent, Belgium,
[3]Department of Internal Medicine and Pediatrics, Ghent University, Ghent, Belgium
[3]Ghent University, Ghent, Belgium
[5]Strategic Policy Cell, Ghent University Hospital, Ghent, Belgium

Corresponding author: Mieke.Deschepper@uzgent.be

Background

Interoperability between clinical coding systems is essential for leveraging high-quality health data across clinical, administrative, and research contexts. ICD-10 is widely used for diagnostic reporting and health monitoring [1,2,3], whereas SNOMED CT provides greater semantic detail and supports advanced clinical documentation. Due to their structural differences—ICD-10's mono-hierarchical format versus SNOMED CT's poly-hierarchical organization—mapping between them is complex and often sensitive to granularity, terminology variation, and semantic ambiguity [4, 5].
Automated tools such as the Athena platform (OHDSI) have emerged to streamline large-scale mapping tasks. Athena applies OMOP vocabulary standards and promises efficiency by rapidly converting ICD-10 codes into corresponding SNOMED CT concepts. However, the degree to which automated mappings preserve clinical specificity and reliability remains uncertain [6]. This study therefore focuses exclusively on evaluating Athena's automated mappings for muscular dystrophy-related ICD-10 codes (G71 group) and quantifying their accuracy compared to manual expert mapping [7].

Methods

A structured evaluation was performed using ICD-10 codes within the G71 group, covering muscular dystrophies and their subtypes. Codes were retrieved from the ICD-10-CM browser [8] and entered individually into the Athena platform under the "Non-standard to Standard map (OMOP → SNOMED)" option. For each ICD-10 code, Athena returned a single SNOMED CT concept [7].
In parallel, an independent manual mapping process was conducted using the international SNOMED CT browser [9]. Two researchers reviewed available SNOMED CT concepts per ICD-10 term, considering synonyms, parent/child concepts, and clinical context. Ambiguities were resolved through group consensus.
Athena outputs were then compared to manual mappings in a blinded assessment. Each mapping was categorized as:

1. Exact match – fully identical to the manual selection.
2. Insufficiently specific – technically correct but lacking clinical precision.
3. Incorrect – semantically mismatched or clinically invalid.

Results were summarized as absolute counts and percentages. While AI-based mappings (LLMs) were first evaluated in a preliminary step, their inconsistent and unreliable performance led to their exclusion from the focused analysis

Results

Athena produced one SNOMED CT concept for each of the 34 ICD-10 codes assessed. Its performance varied depending on diagnostic complexity and the specificity of the ICD-10 entry.

Accuracy Overview (Table 1):

- 22 of 34 mappings (64.7%) were exact matches with the manually validated SNOMED CT concept.
- 12 mappings (35.3%) differed from manual results:
    - 10 (29.4%) insufficiently specific, usually broader than clinically appropriate.
    - 2 (5.9%) incorrect, returning concepts misaligned with clinical meaning.

Comparison: Athena vs Manual Mapping		
Coding with exact match (“Athena”/manual)	22	64,7%
Coding with errors or lacking specificity	12	35,3%
Coding with unsifficient specificity	*10*	*29,4%*
Coding with errors	*2*	*5,9%*
Total	**34**	**100%**

Table 1 Comparison: Athena vs Manual Mapping. Wrong coding are split up between true errors and lack of specificity

Patterns in Athena’s Mapping Behavior

1. *Tendency toward broader concepts*
   Athena frequently selected higher-level SNOMED CT concepts for ICD-10 entries with detailed clinical implications.
   Example: Limb-girdle muscular dystrophy subtypes (G71.031–G71.039) were mapped to generic parent concepts, despite SNOMED CT offering multiple, more precise alternatives.
2. *Challenges with “other specified” and “unspecified” ICD-10 terms*
   ICD-10 categories such as G71.09 or G71.19 inherently lack specificity. SNOMED CT avoids such vague groupings, forcing Athena to choose broad or non-ideal matches.
3. *Single-output consistency but lack of contextual nuance*
   Athena’s design ensures one result per code, which enhances reproducibility and efficiency. However, this rigidity prevents Athena from proposing alternative valid concepts, particularly when ICD-10 terms correspond to multiple SNOMED CT expressions.
4. *Low rate of fully incorrect mappings*
   Only 5.9% of mappings were categorically incorrect. Although encouraging, even small percentages may significantly influence clinical or research workflows when dealing with large datasets.

Conclusion

This study demonstrates that Athena, while efficient and generally reliable, exhibits systematic limitations when translating ICD-10 codes to SNOMED CT—especially for clinically nuanced diagnostic groups such as muscular dystrophies. The 64.7% exact-match rate aligns with expectations for automated tools but underscores that more than one-third of mappings are either overly broad or incorrect.

The predominant issue was loss of clinical specificity, particularly in codes representing multiple subtypes or ambiguous categories. SNOMED CT's fine-grained hierarchical structure requires contextual interpretation, which Athena cannot provide. Over-generalized mappings may lead to reduced semantic accuracy, potentially influencing downstream activities such as cohort identification, phenotype construction, or clinical decision support [10, 11].

Despite these limitations, Athena offers critical advantages: it provides rapid, standardized, and reproducible mappings and serves as a strong foundation for automated processing of large datasets. For high-level reporting or systems seeking standardization across heterogeneous sources, Athena's mappings may be sufficient. However, for tasks requiring diagnostic precision, manual expert review remains indispensable [2].

A hybrid approach—automated mapping followed by targeted manual validation—emerges as the most practical and reliable strategy [2, 11]. Refinements in automated terminology alignment, especially for complex clinical domains, could strengthen the performance of tools like Athena. Training and guidance for tool users are also essential to ensure correct interpretation of broad or ambiguous outputs.

In conclusion, Athena provides an efficient baseline for ICD-10 to SNOMED CT conversion but cannot replace human oversight. Its tendency toward non-specific mappings highlights the ongoing need for expert-led curation to maintain the semantic integrity of clinical data.

References

1. Alyami, M.A., et al. *Health Decision Support System based on Patient provided data for both Patients and Physicians Use*. in *12th International Conference on Ubiquitous Information Management and Communication (IMCOM)*. 2018. Sungkyunkwan Univ, Langkawi, MALAYSIA.
2. Hong, N., et al., *Cross-Language Terminology Mapping Between ICD-10-CN and SNOMED-CT.* Studies in health technology and informatics, 2022. **290**: p. 42-46.
3. Vikström, A., et al., *Mapping the categories of the Swedish primary health care version of ICD-10 to SNOMED CT concepts: rule development and intercoder reliability in a mapping trial.* BMC medical informatics and decision making, 2007. **7**: p. 9.
4 . Willett, D.L., et al., *SNOMED CT Concept Hierarchies for Sharing Definitions of Clinical Conditions Using Electronic Health Record Data.* Applied Clinical Informatics, 2018. **9**(3): p. 667-682.
5. Bodenreider, O., R. Cornet, and D.J. Vreeman, *Recent Developments in Clinical Terminologies — SNOMED CT, 6OINC, and RxNorm.* Yearb Med Inform, 2018. **27**(01): p. 129-139.
6. Burrows, E.K., et al., *Standardizing Clinical Diagnoses: Evaluating Alternate Terminology Selection.* AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science, 2020. **2020**: p. 71-79.
7. Observational Health Data Sciences and Informatics. *Athena – OHDSI Vocabularies Repository*. 2024; v20250227:[Available from: https://athena.ohdsi.org/search-terms/start.
8. Centers for Disease Control and Prevention. *National Center for Health Statistics: ICD-10-CM*. 2024; Available from: https://icd10cmtool.cdc.gov/?fy=FY2024.
9. Health Terminology Standards Development Organization. *SNOMED CT Browser*. 2024; Available from: https://browser.ihtsdotools.org/?perspective=full&conceptId1=404684003&edition=MAIN/2024-12-01&release=&languages=en,nl,fr.
10. Fung, K.W., et al., *Achieving Logical Equivalence between SNOMED CT and ICD-10-PCS Surgical Procedures.* AMIA ... Annual Symposium proceedings. AMIA Symposium, 2017. **2017**: p. 724-733.
11. Mamou, M., et al. *ICD-11 (JLMMS) and SCT Inter-Operation*. in *10th eHealth Conference on Predictive Modeling in Healthcare - From Prediction to Prevention*. 2016. Vienna, AUSTRIA.