AMERICAN ACADEMY OF OPHTHALMOLOGY®

Implementing a Common Data Model in Ophthalmology: Mapping Structured Electronic Health Record Ophthalmic Examination Data to Standard Vocabularies

Justin C. Quon, MD,[1] Christopher P. Long, MD,[1] William Halfpenny, MBBS, MEng,[2] Amy Chuang, MS,[3] Cindy X. Cai, MD, MS,[4] Sally L. Baxter, MD, MSc,[2] Vamsi Daketi, MS,[5] Amanda Schmitz, BS,[5] Neil Bahroos, MS,[3] Benjamin Y. Xu, MD, PhD,[1] Brian C. Toy, MD[1]

Objective: To identify and characterize concept coverage gaps of ophthalmology examination data elements within the Cerner Millennium electronic health record (EHR) implementations by the Observational Health Data Sciences and Informatics Observational Medical Outcomes Partnership (OMOP) common data model (CDM).

Design: Analysis of data elements in EHRs.

Subjects: Not applicable.

Methods: Source eye examination data elements from the default Cerner Model Experience EHR and a local implementation of the Cerner Millennium EHR were extracted, classified into one of 8 subject categories, and mapped to the semantically closest standard concept in the OMOP CDM. Mappings were categorized as exact, if the data element and OMOP concept represented equivalent information, wider, if the OMOP concept was missing conceptual granularity, narrower, if the OMOP concept introduced excess information, and unmatched, if no standard concept adequately represented the data element. Descriptive statistics and qualitative analysis were used to describe the concept coverage for each subject category.

Main Outcome Measures: Concept coverage gaps in 8 ophthalmology subject categories of data elements by the OMOP CDM.

Results: There were 409 and 947 ophthalmology data elements in the default and local Cerner modules, respectively. Of the 409 mappings in the default Cerner module, 25% (n = 102) were exact, 53% (n = 217) were wider, 3% (n = 11) were narrower, and 19% (n = 79) were unmatched. In the local Cerner module, 18% (n = 173) of mappings were exact, 54% (n = 514) were wider, 1% (n = 10) were narrower, and 26% (n = 250) were *unmatched*. The largest coverage gaps were seen in the local Cerner module under the visual acuity, sensorimotor testing, and refraction categories, with 95%, 95%, and 81% of data elements in each respective category having mappings that were not exact. Concept coverage gaps spanned all 8 categories in both EHR implementations.

Conclusions: Considerable coverage gaps by the OMOP CDM exist in all areas of the ophthalmology examination, which should be addressed to improve the OMOP CDM's effectiveness in ophthalmic research. We identify specific subject categories that may benefit from increased granularity in the OMOP CDM and provide suggestions for facilitating consistency of standard concepts, with the goal of improving data standards in ophthalmology.

Financial Disclosure(s): Proprietary or commercial disclosure may be found in the Footnotes and Disclosures at the end of this article. *Ophthalmology Science 2025;5:100666 © 2024 by the American Academy of Ophthalmology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).*

Supplemental material available at www.ophthalmologyscience.org.

The clinical adoption of electronic health records (EHRs) by ophthalmologists in the United States has become highly prevalent, following federal incentives and mandates for meaningful use from 2009 to 2014, rising to an estimated 72% nationwide in 2018.[1−3] Disparate EHR implementations have resulted in limited interoperability among different institutions because of data heterogeneity, arising from differences in terminology, data organization, and database infrastructure.[4,5] Harmonizing and aggregating patient data from different sources is often a manual and immensely time-consuming effort. This may pose barriers to appropriate data access for patient care, reimbursement, research, and quality improvement needs. This issue may be exacerbated in the field of ophthalmology, which has

https://doi.org/10.1016/j.xops.2024.100666
ISSN 2666-9145/24

specialized clinical workflow and documentation, particularly for the physical examination and diagnostic testing. For example, some EHRs accept only text-based input and do not support drawing or annotation features.[6]

Common data models (CDMs) present a potential solution to reconcile these issues by promoting adoption of standardized data structures and vocabularies in both clinical settings and research.[7−13] Vocabulary standards facilitate interoperability between different EHR datasets for clinical quality measurement, benchmarking, and reimbursement, which can be leveraged to perform algorithmic interpretations of computable value sets for decision support, health information exchange, and public health surveillance. Additionally, employing standardized data structures may promote efficient data collection and reduce the impact of significant time expenditures in EHRs.[14−16] One of the most widely used CDMs is the Observational Medical Outcomes Partnership (OMOP) CDM, originally developed through the OMOP project and currently maintained by the Observational Health Data Sciences and Informatics collaborative workforce.[17,18] A central component of the OMOP CDM is the inclusion of open-source standard vocabulary concepts, which allow for the transformation of source data into standardized OMOP CDM structures.[19−24] The OMOP standard concepts are developed from multiple existing terminologies, such as the Systematized Nomenclature of Medicine Clinical Terms, Logical Observation Identifiers, Names, and Codes (LOINC), and RxNorm.[25,26] Despite the wide range of vocabularies included in the OMOP CDM, the standard concepts may not necessarily represent all clinically relevant data found in EHR ophthalmology modules. For instance, Cai et al demonstrated that coverage gaps by the OMOP CDM were found in all areas of the general eye examination in the Epic EHR system.[27] Gaps in OMOP concept coverage for ophthalmology examination findings may exist in other source EHR systems and potentially limit the utility of CDMs in clinical practice and data analysis.

This study examines the concept coverage of the OMOP CDM for 2 implementations of the Cerner Millennium EHR (Oracle Health) by mapping Cerner source data to their closest OMOP standard concepts. Identification of these coverage gaps will enable future iterations of the OMOP CDM to improve representation of ophthalmology source data.

Methods

Source Data

Source data elements for the eye examination were obtained from 2 implementations of the Cerner Millennium EHR: the default Cerner Model Experience and a localized Cerner ophthalmology module developed at the University of Southern California. No human subjects or human medical records were involved in this study but only the database organization for each implementation. Institutional review board approval was exempted by the University of Southern California institutional review board. All research adhered to the tenets of the Declaration of Helsinki. After the development and release of the default Cerner Model Experience in 2015, several institutions independently developed local models that expanded on the default model. The local University of Southern California module was developed in 2020 as a collaborative grassroots effort with input from ophthalmologists at multiple institutions nationwide. In the Cerner EHR, data are organized based on different clinical workflows and tasks. Data captured during these tasks are stored in discrete task assays. Discrete task assays are unique groupings of data elements in the EHR used to define parameters for each individual test or query. In this study, data elements were extracted from these discrete task assays and mapped to the OMOP CDM. Compared with the default Cerner eye examination, the localized Cerner ophthalmology module was developed to have more granularity for certain general eye examination findings and included more clinical findings for ophthalmology subspecialties (Fig 1A). All data elements were categorized into one of 8 subject categories: core eye examination, visual acuity (VA), refraction, sensorimotor testing, diagnostic testing, orbital examination, anterior segment examination, or posterior segment examination. The core eye examination category included data elements related to intraocular pressure (IOP), pupil exam, extraocular movements, and confrontational visual fields. Data elements unrelated to the eye examination (e.g., chief complaint, review of systems, etc) were omitted from analysis.

Mapping Source Data to OMOP CDM

A standardized protocol for mapping EHR source data to OMOP standard concepts was adapted from Cai et al.[27] The first mapping iteration was performed with the Automated Terminology Harmonization, Extraction, and Normalization for Analytics (Athena) web application (v1.13.0.20.230130.1236 with OMOP vocabulary v5.0) and the USAGI software tool (v1.4.3).[28,29] The Athena application provides a comprehensive list of standard concepts in the OMOP CDM, and USAGI is a software program that maps source data to OMOP standard concepts. USAGI proposes multiple possible matches based on similarity of terms between source data concepts and OMOP standard concepts, and the semantically closest mapping was manually selected for each source data element based on the mapping protocol. Concepts in the OMOP CDM that were designated as “nonstandard” or “invalid” in Athena were excluded as potential targets for source data mappings. For all OMOP standard concepts used as target mappings, the associated OMOP concept identification number, source vocabulary, and source vocabulary identification code were documented.

All mappings were then categorized based on semantic equivalence of the source data element and the OMOP standard concept. These categories were adapted from the Health Level 7 Fast Healthcare Interoperability Resources ConceptMap equivalencies.[30] A designation of exact was used when the OMOP concept represented the source data element with no loss or addition of information. Wider mappings indicated that the closest OMOP standard concept had loss of information compared with the source data. Each wider mapping was subcategorized based on the type of missing information, which included laterality (left or right eye), one or more general concepts (such as the testing method), or both. Narrower mappings were used when the closest OMOP concept did not fully represent the scope of the source data or included additional information that was not necessarily accurate. Finally, source data elements were designated as unmatched when there was no OMOP standard concept that adequately represented the source data.

After selection of the closest standard concept with USAGI and categorization of the mapping type, all mappings were verified or

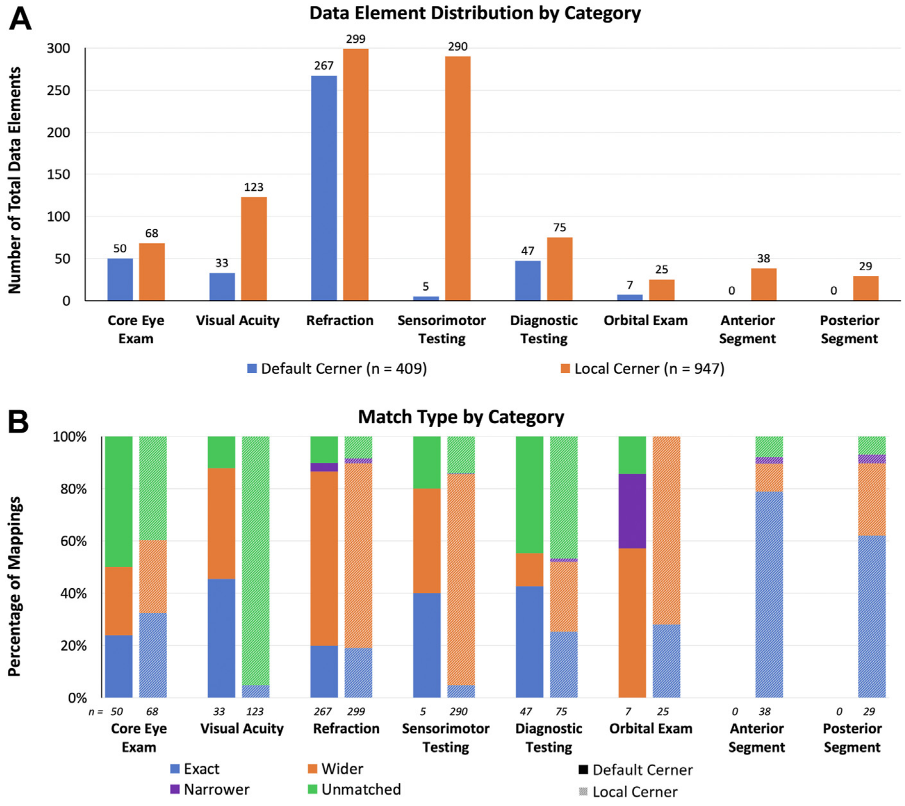

Figure 1. A, Distribution of data elements classified into 8 subject categories for the default and local Cerner modules. **B,** Match type breakdown of data elements in each category.

updated by manually browsing data hierarchies in Athena. The initial mapping was performed by a medical student (J.Q.), with independent verification or revision of all mappings by an ophthalmology resident (C.L.). The final consensus mapping was reached through adjudication in discussion with a board-certified ophthalmologist (B.T.). Mappings after the first iteration included semantically equivalent synonyms in search queries to find OMOP standard concepts that may not have been proposed by USAGI, which instead uses a term-similarity approach. Mappings with USAGI were then updated if a closer match was found with manual browsing. The total number of mappings and the number of each mapping type were then recorded for descriptive statistical analysis. Quantitative analysis was then performed on mappings that were not exact to determine which concepts were not represented and categorize the reasons behind imprecise mappings to the OMOP CDM.

Results

In total, there were 409 source elements for the ophthalmology examination in the default Cerner Model experience and 947 elements for the localized Cerner ophthalmology module.

Table 1 demonstrates the overall breakdown of mappings for both EHR implementations by match type, vocabulary of the target standard concept, and subject category. Mappings classified as exact were in the minority of mapping types for both source datasets, with only 25% (102/409) and 18% (173/947) of source data elements having exact OMOP representations for the default Cerner and localized Cerner module, respectively. Most mappings were classified as wider, accounting for 53% (217/409) of default Cerner and 54% (514/947) of local Cerner mappings. No mappings were found for 19% (79/409) of default Cerner and 26% (250/947) of local Cerner source data elements.

Table 1. Distribution of Mappings for the Default Cerner Millennium Eye Examination and a Localized Cerner Ophthalmology Module by Mapping Type and Vocabulary

	Default Cerner EHR (n = 409)	Local Cerner EHR (n = 947)
Match type		
Exact	102 (24.9%)	173 (18.3%)
Wider	217 (53.1%)	514 (54.3%)
Narrower	11 (2.7%)	10 (1.0%)
Unmatched	79 (19.3%)	250 (26.4%)
Vocabulary		
SNOMED	120 (29.3%)	415 (43.8%)
LOINC	204 (49.9%)	276 (29.1%)
Other	6 (1.5%)	6 (0.6%)
Unmatched	79 (19.3%)	250 (26.4%)

EHR = electronic health record; LOINC = Logical Observation Identifiers, Names, and Codes; SNOMED = Systematized Nomenclature of Medicine.

The distribution of data elements in each category is demonstrated in Figure 1A. In the default Cerner module, data elements related to refractive measurements accounted for the majority of the dataset with 267 out of 409 terms. Notably, the default Cerner module did not include data elements related to the anterior segment or posterior segment examinations, and it had few terms dedicated to sensorimotor testing (n = 5) and the orbital examination (n = 7). Compared with the default Cerner implementation, the local Cerner module expanded upon all 8 categories with additional data elements (Fig 1A). Figure 1B depicts the match type distribution of each data element category for the default and local Cerner modules. With the exceptions of the anterior and posterior segment categories for the local Cerner module, most data elements in each category for both modules had imprecise mappings (wider, narrower, or unmatched) to the closest standard concept in the OMOP CDM. All analyzed data elements with their corresponding OMOP standard concepts and match type are provided in Tables S2 and S3 (available at www.ophthalmologyscience.org).

In the local Cerner module, 40% (27/68) of data elements in the core eye exam category were unmatched due to insufficient clinical granularity of OMOP standard concepts (Fig 1B). Examples include the following: pupil diameter measurements in different lighting conditions (Diameter Dark oculus dexter [OD] and Diameter Light OD had the closest OMOP concept of Right Eye Pupil Diameter Manual), the method used to measure IOP, and confrontational visual fields separated into specific quadrants for each eye (OD Confrontational Testing Inner Lower was mapped to the OMOP concept Right Eye Visual Field Defects by Confrontation). Rather than representing raw measurement or observation data from each eye, many of the concepts in the OMOP CDM for visual field findings were instead interpretations of data from both eyes (e.g., left homonymous inferior quadrantanopia).

In the VA category for the local Cerner module, 95% (117/123) of data elements were unmatched. In addition to converting Snellen VA measurements to logarithm of the minimum angle of resolution (logMAR), the local Cerner module included distinct terms for the numerator and denominator when recording VA, to preserve information for the distance at which VA measurements were obtained. The local Cerner module also included additional data elements for the number of "missed" or "incorrect" numbers or letters during the VA examination. For example, the local Cerner module included distinct data elements for Near VA oculus sinister (OS) Numerator, Near VA OS Denominator, and Near VA OS Missed. However, these distinctions were not found in the default Cerner module or the OMOP CDM, thus accounting for the drastically different match type distribution between the 2 modules, as well as the high unmatched proportion in the local Cerner module.

Most data elements in the refraction category for both default and local Cerner modules had wider mappings, with respective percentages of 67% (178/267) and 71% (211/299). In this category, the closest OMOP standard concepts for most of these mappings had multiple levels of missing clinical granularity, including both laterality and details about the type of refraction performed. For example, the data element OD Dry Retinoscopy Refraction Horizontal Prism had the closest OMOP mapping of Prism Strength.

The sensorimotor examination was poorly represented in the default Cerner module (n = 5) and was greatly expanded in the local Cerner module, with 290 terms accounting for 31% of the local module dataset. However, the sensorimotor examination was poorly represented in the OMOP CDM, with less than 5% of data elements (14/290) in the local Cerner module having an exact mapping. This was primarily due to poor concept coverage of measurements related to gaze deviations, such as the data elements Down Gaze Horizontal Max and Down Right Gaze Vertical Max both having the closest mapping to the OMOP concept Gaze down and right.

For the diagnostic testing category, 45% (21/47) and 48% (35/75) of data elements in the default and local Cerner modules were unmatched, respectively. The unmatched concepts were largely related to parameters or interpretations of various diagnostic tests. Specifically, the OMOP CDM did not have adequate concepts for the lipid layer thickness of each eye, settings used in brightness acuity testing for glare disability, the type of Schirmer test, and measurements of matrix metalloproteinase-9 in tears. There were no adequate concepts for interpretation of corneal sensation testing, corneal pachymetry, potential acuity, and retinal acuity meter testing. Measurements related to keratometry and optical coherence tomography scans were also not fully represented in the OMOP CDM. Finally, the Cerner modules had separate data elements for the number of correct and total plates for color vision testing, although this distinction was not found in the OMOP CDM.

The orbital exam had minimal representation in the default Cerner module with 7 terms in total, which was expanded to 25 data elements in the local module. Of the imprecise mappings for these 25 terms, the gaps in concept coverage included not specifying laterality for a variety of measurements (such as Upper Lid Crease OD or Superior Scleral Show OD) and not differentiating whether margin reflex distance was measured before or after phenylephrine administration, which is clinically relevant information that could be used to guide diagnosis and treatment options.

The default Cerner module did not have any terms dedicated to the anterior or posterior segment examinations. The local Cerner module included 38 terms for the anterior segment examination, with 79% (n = 30) having an exact mapping. The posterior segment examination had 29 terms, of which 62% (n = 18) had an exact mapping. The OMOP CDM did not have concepts for the method of anterior or posterior segment testing and whether a drawing was included in either examination. Unrepresented anterior segment concepts included physical examination findings for the adnexa and orbits. In the posterior segment examination, there were no distinction in concepts pertaining to examination findings of the peripheral versus central retina.

Overall, the vast majority of possible matches (exact, wider, or narrower) for both EHR implementations were mapped to either Systematized Nomenclature of Medicine (SNOMED) or LOINC vocabularies (Fig 2A). Mapping

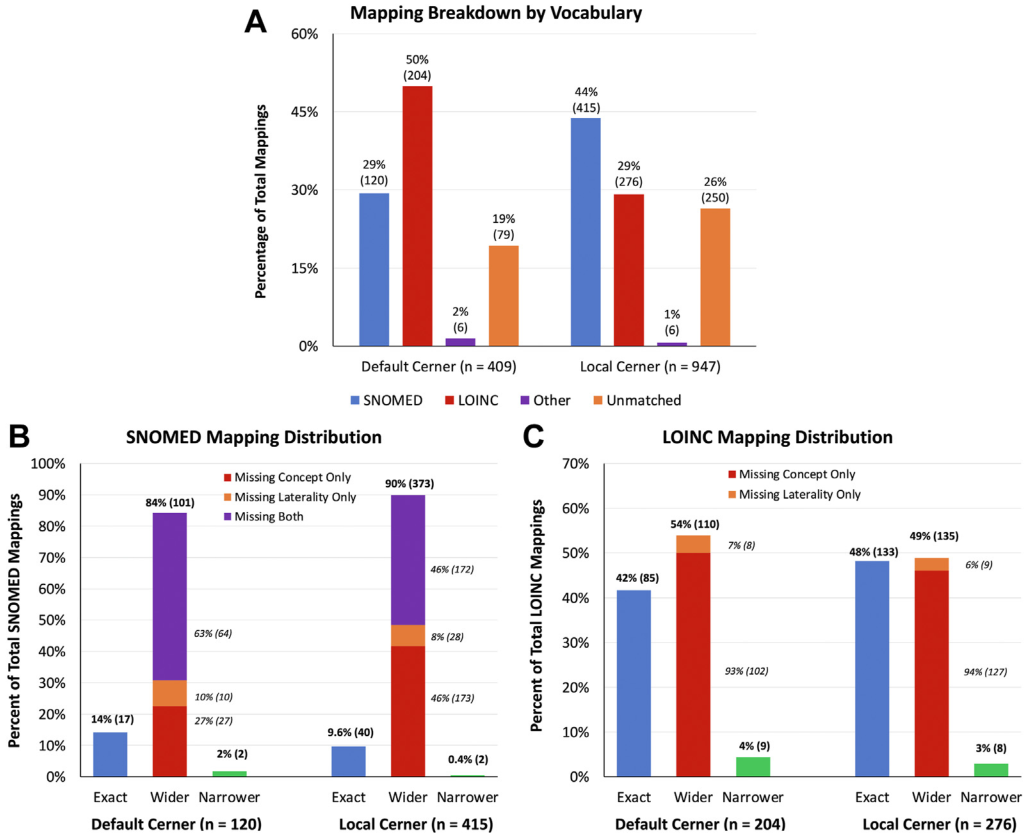

Figure 2. **A,** Distribution of all default Cerner and local Cerner module mappings by vocabulary of the target standard concept. Match type distribution for mappings with target standard concepts in SNOMED (**B**) or LOINC (**C**) vocabularies. LOINC = Logical Observation Identifiers, Names, and Codes; SNOMED = Systematized Nomenclature of Medicine.

distributions after stratifying by vocabulary are shown in Figure 2B and 2C. Less than 15% of source data elements mapped to a SNOMED target concept had exact matches for both the default Cerner (17/120) and local Cerner (40/415) modules. Systematized Nomenclature of Medicine mappings classified as wider were missing both laterality and concepts for 63% (64/101) of default Cerner mappings and 46% (172/373) for the local Cerner module.

Compared with mappings with SNOMED target concepts, the LOINC vocabulary had a greater proportion of exact matches (42% with 85/204 matches for default Cerner and 48% with 133/276 matches for the local Cerner module). Furthermore, none of the wider LOINC mappings were missing both laterality and concepts. Instead, 93% to 94% (102/110 for default Cerner and 127/135 for local Cerner) of wider LOINC mappings for both EHR datasets were missing only a concept, and 6% to 7% (8/110 for default Cerner and 9/135 for local Cerner) were only missing laterality in the target LOINC concept.

Discussion

The primary objective of this study was to describe the representation of eye examination source data elements implemented in the Cerner EHR by standard concepts in the OMOP CDM. There were considerable gaps in concept coverage for 2 implementations of the Cerner ophthalmology module, with only 25% of default Cerner and 18% of localized Cerner source data elements having an exact representation by OMOP standard concepts. Gaps were found in all 8 categories of the eye examination, comprising either imprecise mappings or inability to map because there were no OMOP standard concepts that adequately reflected the semantic content of the original source data.

The most common cause of imprecise mappings for Cerner ophthalmology data elements was loss of clinically relevant information when mapping to OMOP standard concepts. Lack of laterality specification in OMOP standard concepts comprised a large proportion (38% of wider mappings in standard Cerner; 41% of wider mappings in localized Cerner), which was particularly notable when the target concept was derived from the SNOMED vocabulary. Addressing this gap may be beneficial in widespread standardization of ophthalmology data, given that the American Academy of Ophthalmology has selected SNOMED as the preferred medical ontology for ophthalmology terms.[26] Systematized Nomenclature of Medicine by convention employs postcoordination to describe diagnoses and procedures. Because not specifying laterality will detract from efficiency of data analysis in research studies, where clear criteria including laterality specification are often used for defining cohorts and analyzing study outcomes, some efforts have been made to improve this workflow, including prespecifying postcoordinated descriptors.[31,32] In contrast, the LOINC ontology precoordinates laterality information by convention. Additional work within OMOP may focus on harmonizing and linking semantically equivalent concepts within the various ontologies that comprise the OMOP CDM. Some of this work is being undertaken by the Observational Health Data Sciences and Informatics Eye Care and Vision Workgroup.

Other imprecise mappings also involved loss of information such as the method or condition used for a clinical measurement, observation, or diagnostic test. Some examples of information loss included room brightness conditions when measuring pupil diameter (important for the nuanced evaluation of anisocoria), whether current contact lenses were soft contact lenses or rigid gas-permeable lenses (necessary to track complications or outcomes with a particular lens type), and the degree of gaze deviation (relevant to strabismus surgery outcomes). Thus, these details could all potentially be critical for developing a computable definition for clinically relevant cohorts. Figure 1B demonstrates that data elements for concepts in the VA, refraction, and sensorimotor testing categories were especially poorly represented in the OMOP CDM, with <20% of each category in the local Cerner module having exact matches. Therefore, adding clinical granularity to OMOP standard concepts in future versions of the CDM may help reduce these coverage gaps and further improve representation of EHR source data. This is especially true for VA, which is a cardinal "vital sign" in ophthalmology.

One of the key strengths in this study is the analysis of 2 Cerner ophthalmology modules: the default Cerner Model Experience, as well as a localized implementation that vastly expanded on the data elements for all 8 subject categories. The local Cerner module was developed with feedback from practicing ophthalmologists nationwide, including the United States Veterans Health Administration, United States Department of Defense, Children's Hospital Los Angeles, University of Pittsburgh Medical Center, University of New Mexico, and Indiana University, to ensure that the additional information was useful for clinical decision making and had an appropriate level of granularity. For example, the local Cerner module separated the numerator and denominator for VA measurements in the general eye examination, thus allowing for the retention of measurement distance information that could otherwise be lost upon conversion to logMAR or Snellen equivalent. Another important difference is that the default Cerner module did not have structured fields or data elements for the anterior and posterior segment examinations. Instead, it utilized a single free-text box for the physical examination, which was shared with all other medical specialties, and this unstructured text format was not conducive to extract-transform-load (ETL) processes needed for storing concepts in a structured manner. In contrast, the localized Cerner module accommodated structured data for the anterior segment and posterior segment examinations, as well as additional fields for the orbital exam and sensorimotor testing.

The prevalence of standard OMOP concept coverage gaps was not equal among all ophthalmic subspecialty content. In the local Cerner module, deficiencies were particularly apparent for the sensorimotor exam with 14 exact mappings out of 290 terms (including details regarding eye alignment, stereo vision, and nystagmus that are important to neuro-ophthalmologists, strabismus surgeons, and pediatric ophthalmologists), refraction with 57 exact mappings out of 299 terms (including specialized details of refractive error measurements important for optometry and pediatric ophthalmology), and point-of-care diagnostic testing with 20 exact mappings out of 47 terms.

A prior study examined concept representation of the general ophthalmology module in another widely used EHR, the Epic Foundation System. Cai et al reported that coverage gaps existed in all areas of the Epic EHR general eye examination, with only 25% of all mappings classified as exact.[27] Our findings closely corroborate the reported coverage gaps, with 25% of default Cerner and 18% of localized Cerner data elements having exact OMOP representations. Of note, 90% of source data elements (excluding unmatched terms) for the Epic eye examination were matched to a SNOMED target concept. In our study, SNOMED mappings accounted for only 36%−60% of Cerner source data, after excluding unmatched terms. Excluding 6 terms matched to the UK Biobank ontology, the remaining 40% to 64% of data elements were successfully matched to concepts in the LOINC vocabulary. One reason for this discrepancy may be the present study included both eyes for Cerner source data elements with specified laterality, and LOINC terms tended to precoordinate laterality information into the standard concept itself, compared with SNOMED standard concepts where laterality specification would require a postcoordination modifier (Fig 2B and C). In contrast, Cai et al mapped only right eye concepts. This discrepancy highlights the need to develop standardized practices for mapping disparate datasets to the OMOP CDM. Harmonizing disparate datasets to the OMOP CDM through standardized mapping practices would enhance data interoperability and accessibility, thereby facilitating reproducibility of clinical studies and enabling more efficient health information exchange. One approach taken by the OMOP Eye Care workgroup has been to employ an ETL dashboard to coordinate among multiple institutions performing OMOP mapping.[33]

In addition to gaps in concept coverage for both clinical granularity and laterality, the OMOP CDM also contains inconsistencies and redundant entries that may benefit from standardization. In prior versions of the OMOP CDM, the term left eye IOP had redundant entries in the same measurement domain from both SNOMED and LOINC vocabularies, namely Intraocular pressure of left eye [SNOMED, OMOP concept ID 44805433] and Left eye Intraocular pressure [LOINC; OMOP concept ID 21491757]. These semantically equivalent concepts were harmonized in the v20240229 release of the OMOP CDM by designating the LOINC term as the preferred standard concept and creating a nonstandard to standard link from the SNOMED concept, thus preventing further ambiguous mappings and providing backwards compatibility to databases which had previously mapped to the SNOMED concept (now nonstandard). There are also inconsistencies in the inclusion of related terms. For example, Right eye Prism base distance Phoropter *[LOINC]* is included as a standard concept, but the equivalent Left eye Prism base distance Phoropter is not an existing entry in the OMOP CDM.

Future efforts for improving data standardization by the OMOP CDM may include developing best practices for the storage location of data elements. Currently, the domains in which some eye examination data are stored are not consistent. For instance, the observation table includes a term for ogMAR VA left eye, but the measurement table includes another equivalent term in Visual acuity logMAR Eye—left. Additionally, there are equivalent concepts of OS Intraocular pressure in the measurement domain and Intraocular pressure left eye in the condition domain, indicating that harmonizing the storage location for eye examination data would help improve consistency in data standards by the OMOP CDM.

Future work may define "minimum data sets" for the ophthalmic examination. This can be driven by concepts determined to be clinically relevant based on published clinical literature. Another approach may be to perform a utilization analysis of the data concepts to determine those that are populated in real world use from multiple databases among early ETL adopters. This would not only guide the removal of redundant concepts for parsimony, but also enable the merging or consolidation of unnecessarily specific concepts for defining a "minimum data set." Additionally, these efforts would also align with ongoing initiatives to include eye-specific data elements in the United States Core Data for Interoperability to support data standardization and interoperability.[8]

One limitation of this study is the possibility that some source data elements may not be mapped to the semantically closest standard concept. This may result from human error or the inherent inability of the USAGI application's term-similarity approach to browse for synonymous terms. For example, the closest standard concept for the source data element OD Retinoscopy Cylinder was determined to be Right eye Cylinder—W cycloplegia by Objective refraction, where retinoscopy was deemed to imply objective refraction with cycloplegia. However, this standard concept was not included in any of the mappings suggested by USAGI and was instead found by manual browsing through data hierarchies in the Athena web application. The possibility of overlooked standard concepts was minimized in this study through multiple mapping methods (e.g. utilization of semiautomated USAGI suggestions and extensive manual browsing using Athena) as well as a hierarchal iterative verification process, where mappings performed by a medical student were independently verified or refined by an ophthalmology resident and then a board-certified ophthalmologist.

The current study also did not address whether unmatched data elements would be better matched to SNOMED, LOINC, or another vocabulary standard. The general convention in the OMOP CDM is that concepts in the "measurement" domain are mapped to LOINC terms, and concepts in the "observation" domain are mapped to SNOMED terms. Regardless, one of the key advantages of OMOP is the ability to harmonize semantically equivalent terms from different vocabularies into the same OMOP concept identifier.

Because the primary focus of this manuscript was to identify data elements in Cerner-based EHR implementations that were not found in OMOP standard concepts, another limitation of this work is that it did not comprehensively address what concepts were included in the OMOP CDM but not in the Cerner modules.

In conclusion, this study demonstrates considerable gaps in representation of Cerner-based ophthalmology data elements by standard concepts in the OMOP CDM. There are opportunities to improve concept coverage in all aspects of the default general eye examination, as well as subspecialty-specific diagnostics and clinical findings. Future efforts to improve source data representation may involve adding new concepts or modifying existing standard concepts to include more granularity. Consistency may also be improved by standardizing storage location for eye examination data, as well as ensuring there are not any semantically equivalent and redundant entries to minimize ambiguous mappings. Advancements in concept coverage and data consistency by the OMOP CDM would enable more efficient access to clinically relevant information, resulting in wide-ranging benefits for research studies, quality improvement, and patient care.

Data Availability

The datasets analyzed in this study are derived from proprietary source data elements. We have provided descriptors, and the specific Cerner DTA (discrete task assay) and Concept CKI (Cerner Knowledge Index) details are available upon reasonable request to the corresponding author by other authorized users of the Cerner electronic health record.

Footnotes and Disclosures

Originally received: June 15, 2024.
Final revision: October 28, 2024.
Accepted: November 25, 2024.
Available online: November 28, 2024.Manuscript no. XOPS-D-24-00190.

[1] Department of Ophthalmology, Roski Eye Institute, Keck School of Medicine, University of Southern California, Los Angeles, California.

[2] Division of Ophthalmology Informatics and Data Science, Viterbi Family Department of Ophthalmology and Shiley Eye Institute, University of California San Diego, La Jolla, California.

[3] Southern California Clinical and Translational Science Institute, Los Angeles, California.

[4] Wilmer Eye Institute, Johns Hopkins School of Medicine, Baltimore, Maryland.

[5] Department of Information Technology, Keck Medicine of USC, Los Angeles, California.

Disclosure(s):

All authors have completed and submitted the ICMJE disclosures form.

The authors have made the following disclosures:

A.C.: Financial support − CTSI NCATS.

C.X.C.: Grants − Regeneron Pharmaceuticals, Inc.

S.L.B.: Financial support − Research to Prevent Blindnesss; Consultant − Topcon; Honoraria − Topcon; Travel expenses − Topcon; Others − Optomed.

B.C.T.: Consultant − Alimera Sciences, Eyepoint Pharmaceuticals, Regeneron, Bausch and Lomb.

This work was supported by grant UL1TR001855 from the National Center for Advancing Translational Science (NCATS) and P30EY029220 from the National Eye Institute (NEI) of the US National Institutes of Health, as well as Research to Prevent Blindness, New York, NY (Unrestricted Grant to the USC Roski Eye Institute). B.C.T. is supported by K23EY032985 from the National Eye Institute of the National Institutes of Health. C.X.C. is supported by K23EY033440 from the National Eye Institute of the National Institutes of Health. C.X.C. is the Jonathan and Marcia Javitt Rising Professor of Ophthalmology. W.H. is supported by a grant from NIH (DP5OD029610). S.L.B. is supported by NIH grants P30EY022589, R01EY034146, and OT2OD032644. The sponsor or funding organization had no role in the design or conduct of this research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

HUMAN SUBJECTS: No human subjects were included in this study. Institutional review board approval was exempted by the USC Institutional Review Board. All research adhered to the tenets of the Declaration of Helsinki.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Quon, Halfpenny, Cai, Baxter, Toy

Data collection: Quon, Long, Chuang, Daketi, Schmitz, Bahroos, Toy

Analysis and interpretation: Quon, Xu, Toy

Obtained funding: Toy

Overall responsibility: Quon, Long, Halfpenny, Chuang, Cai, Baxter, Daketi, Schmitz, Bahroos, Xu, Toy

Abbreviations and Acronyms:
CDM = common data model; **EHR** = electronic health record; **ETL** = extract-transform-load; **IOP** = intraocular pressure; **logMAR** = logarithm of the minimum angle of resolution; **LOINC** = Logical Observation Identifiers, Names, and Codes; **OD** = oculus dexter; **OMOP** = Observational Medical Outcomes Partnership; **OS** = oculus sinister; **OU** = oculus uterque; **SNOMED** = Systematized Nomenclature of Medicine; **VA** = visual acuity.

Keywords:
Common data model, Data standards, Electronic health record, OHDSI, OMOP.

Correspondence:
Brian C. Toy, MD, USC Roski Eye Institute, Keck School of Medicine, 1450 San Pablo St, Los Angeles, CA 94539. E-mail: Brian.Toy@med.usc.edu.

References

1. Boland MV, Chiang MF, Lim MC, et al. Adoption of electronic health records and preparations for demonstrating meaningful use: an American Academy of Ophthalmology survey. *Ophthalmology*. 2013;120:1702−1710.
2. Lim SB, Shahid H. Distribution and extent of electronic medical record utilisation in eye units across the United Kingdom: a cross-sectional study of the current landscape. *BMJ Open*. 2017;7:e012682.
3. Lim MC, Boland MV, McCannel CA, et al. Adoption of electronic health records and perceptions of financial and clinical outcomes among ophthalmologists in the United States. *JAMA Ophthalmol*. 2018;136:164−170.
4. Li E, Clarke J, Ashrafian H, et al. The impact of electronic health record interoperability on safety and quality of care in high-income countries: systematic review. *J Med Internet Res*. 2022;24:e38144.
5. Bernstam EV, Warner JL, Krauss JC, et al. Quantitating and assessing interoperability between electronic health records. *J Am Med Inform Assoc*. 2022;29:753−760.
6. Chiang MF, Boland MV, Brewer A, et al. Special requirements for electronic health record systems in ophthalmology. *Ophthalmology*. 2011;118:1681−1687.
7. Halfpenny W, Baxter SL. Towards effective data sharing in ophthalmology: data standardization and data privacy. *Curr Opin Ophthalmol*. 2022;33:418−424.
8. Baxter SL, Reed AA, Maa A, et al. Ocular health and national data standards: a case for including visual acuity in the United States Core Data for Interoperability. *Ophthalmol Sci*. 2022;2:100210.
9. Sheehan J, Hirschfeld S, Foster E, et al. Improving the value of clinical research through the use of Common Data Elements. *Clin Trials*. 2016;13:671−676.
10. Son N, Kim B, Chung S, Han S. Korean pharmacovigilance system based on EHR-CDM. *Stud Health Technol Inf*. 2019;264:1592−1593.
11. Ji H, Kim S, Yi S, et al. Converting clinical document architecture documents to the common data model for incorporating health information exchange data in observational health studies: CDA to CDM. *J Biomed Inf*. 2020;107:103459.
12. Ahmadi N, Peng Y, Wolfien M, et al. OMOP CDM can facilitate data-driven studies for cancer prediction: a systematic review. *Int J Mol Sci*. 2022;23:11834.
13. Belenkaya R, Gurley M, Dymshyts D, et al. Standardized observational cancer research using the OMOP CDM oncology module. *Stud Health Technol Inf*. 2019;264:1831−1832.
14. Long CP, Tai-Seale M, El-Kareh R, et al. Electronic health record use among ophthalmology residents while on call. *J Acad Ophthalmol (2017)*. 2020;12:e143−e150.
15. Read-Brown S, Hribar MR, Reznick LG, et al. Time requirements for electronic health record use in an academic ophthalmology center. *JAMA Ophthalmol*. 2017;135:1250−1257.
16. Gali HE, Baxter SL, Lander L, et al. Impact of electronic health record implementation on ophthalmology trainee time expenditures. *J Acad Ophthalmol (2017)*. 2019;11:e65−e72.
17. Hripcsak G, Duke JD, Shah NH, et al. Observational health data Sciences and Informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inf*. 2015;216:574−578.
18. OMOP common data model. GitHub. https://ohdsi.github.io/CommonDataModel/. Accessed February 15, 2024.
19. Quiroz JC, Chard T, Sa Z, et al. Extract, transform, load framework for the conversion of health databases to OMOP. *PLoS One*. 2022;17:e0266911.

20. FitzHenry F, Resnic FS, Robbins SL, et al. Creating a common data model for comparative effectiveness with the observational medical outcomes partnership. *Appl Clin Inf*. 2015;6:536−547.
21. Peng Y, Henke E, Sedlmayr M, Bathelt F. Towards ETL processes to OMOP CDM using metadata and modularization. *Stud Health Technol Inf*. 2023;302:751−752.
22. Lima DM, Rodrigues-Jr JF, Traina AJM, et al. Transforming two decades of ePR data to OMOP CDM for clinical research. *Stud Health Technol Inf*. 2019;264:233−237.
23. Biedermann P, Ong R, Davydov A, et al. Standardizing registry data to the OMOP Common Data Model: experience from three pulmonary hypertension databases. *BMC Med Res Methodol*. 2021;21:238.
24. Katsch F, Hussein R, Korntheuer R, Duftschmid G. Converting HL7 CDA based nationwide Austrian medication data to OMOP CDM. *Stud Health Technol Inf*. 2023;302:899−900.
25. Chiang MF, Casper DS, Cimino JJ, Starren J. Representation of ophthalmology concepts by electronic systems: adequacy of controlled medical terminologies. *Ophthalmology*. 2005;112:175−183.
26. Hoskins HD Jr, Hildebrand PL, Lum F. The American Academy of Ophthalmology adopts SNOMED CT as its official clinical terminology. *Ophthalmology*. 2008;115:225−226.
27. Cai CX, Halfpenny W, Boland MV, et al. Advancing toward a common data model in ophthalmology: gap analysis of general eye examination concepts to standard observational medical outcomes partnership (OMOP) concepts. *Ophthalmol Sci*. 2023;3:100391.
28. GitHub - OHDSI/Athena. Web application for distributing and browsing the Standardized Vocabularies for all instances of an OMOP CDM. GitHub. https://github.com/OHDSI/Athena. Accessed February 15, 2024.
29. USAGI for vocabulary mapping. OHDSI observational health data Sciences and Informatics. https://www.ohdsi.org/analytic-tools/usagi/. Accessed February 15, 2024.
30. HL7 FHIR. Resource concept map − content. https://www.hl7.org/fhir/conceptmap.html. Accessed May 2, 2024.
31. Miñarro-Giménez JA, Martínez-Costa C, López-García P, Schulz S. Building SNOMED CT post-coordinated expressions from annotation groups. *Stud Health Technol Inf*. 2017;235:446−450.
32. Karlsson D, Nyström M, Cornet R. Does SNOMED CT post-coordination scale? *Stud Health Technol Inf*. 2014;205:1048−1052.
33. Ophthalmology Content ETL Dashboard. OHDSI eye care and vision WG github. https://ohdsi.github.io/EyeCareVisionWG/Approach.html. Accessed May 2, 2024.