Title

Variations in Electronic Health Record-Based Definitions of Diabetic Retinopathy Cohorts A Literature Review and Quantitative Analysis

Journal

Ophthalmology Science, 4(4)

ISSN

2666-9145

Authors

Chen, Jimmy S
Copado, Ivan A
Vallejos, Cecilia
et al.

Publication Date

2024-07-01

DOI

10.1016/j.xops.2024.100468

Peer reviewed

AMERICAN ACADEMY OF OPHTHALMOLOGY®

Variations in Electronic Health Record-Based Definitions of Diabetic Retinopathy Cohorts

A Literature Review and Quantitative Analysis

Jimmy S. Chen, MD,[1,2] Ivan A. Copado, MPH,[1,2] Cecilia Vallejos, BS,[1,2] Fritz Gerald P. Kalaw, MD,[1,2] Priyanka Soe,[1,2] Cindy X. Cai, MD,[3] Brian C. Toy, MD,[4] Durga Borkar, MD, MMCi,[5] Catherine Q. Sun, MD,[6,7] Jessica G. Shantha, MD, MSc,[6,7] Sally L. Baxter, MD, MSc[1,2]

Purpose: Use of the electronic health record (EHR) has motivated the need for data standardization. A gap in knowledge exists regarding variations in existing terminologies for defining diabetic retinopathy (DR) cohorts. This study aimed to review the literature and analyze variations regarding codified definitions of DR.

Design: Literature review and quantitative analysis.

Subjects: Published manuscripts.

Methods: Four graders reviewed PubMed and Google Scholar for peer-reviewed studies. Studies were included if they used codified definitions of DR (e.g., billing codes). Data elements such as author names, publication year, purpose, data set type, and DR definitions were manually extracted. Each study was reviewed by $\geq$ 2 authors to validate inclusion eligibility. Quantitative analyses of the codified definitions were then performed to characterize the variation between DR cohort definitions.

Main Outcome Measures: Number of studies included and numeric counts of billing codes used to define codified cohorts.

Results: In total, 43 studies met the inclusion criteria. Half of the included studies used datasets based on structured EHR data (i.e., data registries, institutional EHR review), and half used claims data. All but 1 of the studies used billing codes such as the International Classification of Diseases 9th or 10th edition (ICD-9 or ICD-10), either alone or in addition to another terminology for defining disease. Of the 27 included studies that used ICD-9 and the 20 studies that used ICD-10 codes, the most common codes used pertained to the full spectrum of DR severity. Diabetic retinopathy complications (e.g., vitreous hemorrhage) were also used to define some DR cohorts.

Conclusions: Substantial variations exist among codified definitions for DR cohorts within retrospective studies. Variable definitions may limit generalizability and reproducibility of retrospective studies. More work is needed to standardize disease cohorts.

Financial Disclosure(s): Proprietary or commercial disclosure may be found in the Footnotes and Disclosures at the end of this article. *Ophthalmology Science 2024;4:100468*

Supplemental material available at www.ophthalmologyscience.org.

Diabetic retinopathy (DR), a common microvascular complication of diabetes mellitus, is one of the leading causes of preventable blindness in the adult working population.[1,2] It is estimated that by 2045, DR will affect 160.5 million adults in the world, disproportionately affecting patients in the Middle East, North Africa, and the Western Pacific.[1] Despite advances in DR treatments through clinical trials[3] and advances in DR telehealth screening,[4,5] there remain significant disparities in screening and treatment outcomes.[6,7] To address this gap in knowledge, retrospective studies have been crucial in advancing our understanding of epidemiologic differences,[8–10] treatment outcomes,[11,12] and health disparities.[13,14]

The adoption of the electronic health record (EHR) has resulted in significantly increased data availability in clinical practice and secondary use such as research and billing. Within ophthalmology, large data sets such as the American Academy of Ophthalmology Intelligent Research In Sight (IRIS) Registry[15] and the National Institutes of Health *All of US* Research Program[16] have integrated structured EHR data for retrospective big data studies. Additionally, EHR data has facilitated the development of artificial intelligence models, which have been particularly useful in automated DR diagnosis and screening.[17,18] However, developing big data studies and artificial intelligence models with generalizable outcomes is highly dependent

https://doi.org/10.1016/j.xops.2024.100468
ISSN 2666-9145/24

on the utilization of diverse data sets, which is limited by data sharing.[19,20] Standardized representation of data, including structured data and images, is needed to facilitate data sharing and has the potential to increase the quality of clinical care and research.[20–22]

With increasing interest in data standards within the ophthalmic community,[20,23] there remains a secondary need for standardized terminologies and definitions such as disease (and associated severity) and visual acuity. Because these concepts are often used to define cohorts or outcomes, it is critical these clinical concepts are standardized to ensure robust and generalizable analyses and artificial intelligence models. Currently, terminologies such as the International Classification of Diseases (ICD) and Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT) are often used to define disease-based cohorts. Tavakoli et al[24] recently showed that coverage of billing codes defined by SNOMED and ICD for ophthalmic infections and trauma demonstrated discrepant alignment, meaning that depending on the use of disease terminologies, different disease-based cohorts could be extracted. This is problematic as it may result in differing demographics across varying disease cohorts, resulting in ungeneralizable and potentially inaccurate study conclusions, and has profound implications in understanding health disparities.[25] Using DR as a use case, our study had the following aims: (1) to perform a literature review of retrospective studies using codified DR definitions and (2) to broadly analyze variations between codified DR definitions.

Methods

Literature Review

We reviewed articles identified using PubMed and Google Scholar using a combination of the following keywords: "Electronic Health Record," "Diabetic Retinopathy," "cohort definitions," "billing codes," and synonyms and acronyms when appropriate. Cited articles in reviewed references were used to expand our search. This study adhered to the Declaration of Helsinki. The requirement for informed consent was waived by the University of California San Diego Institutional Review Board for this study as no patient data were used.

We performed a literature review for articles published before March 2023, with the most recent search date on 3/12/2023. No date range limitation before this date was imposed. Four individual reviewers (J.S.C., I.A.C., C.V., and P.S.) manually reviewed each article for inclusion if the study was (1) a retrospective study design (e.g., retrospective cohort, cross sectional, or case-control); (2) used and listed codified definitions such as International Classification of Diseases, Ninth Revision (ICD-9) or Tenth Revision (ICD-10) to define study cohorts based on incident DR and its associated complications (i.e., diabetic macular edema [DME] or proliferative diabetic retinopathy [PDR]); and (3) contained a complete list of codes used to define DR and DR subtypes within the text or in Table S1 (available at www.ophthalmologyscience.org). International Classification of Diseases codes are managed by the World Health Organization, and in the United States, there are additions to the base coding system called Clinical Modifications.[26] International Classification of Diseases-10-Clinical Modifications was included in the ICD-10 category in this study. Articles were excluded if other criteria such as clinical assessment by fundoscopic examination or imaging were used to classify DR diagnosis,[27–29] a DR cohort could not be identified,[30–32] or the study did not follow a retrospective study design. For each study reviewed, a second reviewer verified the inclusion of articles, using the same inclusion criteria in the primary review. With consensus between both reviewers, articles were considered eligible for review within this study. The full search query, inclusion, and exclusion criteria are articulated in Table 2.

For each study, the following data were extracted: authors and year published, data set source (i.e., EHR or claims database), sample size of DR cohorts, location (single or multicenter), DR cohort type (new diagnosis or known diagnosis), study purpose and aims, and data terminology (ICD codes, Current Procedural Terminology [CPT], Healthcare Common Procedure Coding System, etc.). We additionally extracted whether the DR definitions were used to create the study's inclusion cohort or used as an outcome measure, labeled as "inclusion" and "outcome" respectively. We then generated descriptive statistics of all variables to identify variations among the included studies.

Quantitative Analysis of Terminologies Used

All data analyses were performed in R (version 4.0.2).[33] An analysis examining the distribution of terminologies defining DR versus data set source was performed. The broad categories of cohort definitions for each study were labeled as: "ICD Only," "Other Definitions," or "ICD + Other Definitions." A study was labeled in the "ICD Only" category if it defined DR using ICD-9 or ICD-10 codes, as well as associated codes and modifications such as Clinical Modifications codes. Conversely, a study was labeled in the "Other Definitions" category if it defined DR using any other non-ICD codes such as CPT, Anatomical Therapeutic Chemical codes, and SNOMED. A study was labeled as "ICD + Other Definitions" if it incorporated a combination of ICD-based definitions and non-ICD terminologies.

Studies were additionally labeled as "database" or "retrospective EHR" studies. "Database" studies were defined as studies using data from aggregated, multiinstitution data from a third party, and "retrospective EHR" studies were defined as studies whose data came from secondary use of their institutions' EHR. The sum of studies using the combinations of "ICD Only" ± "Other Definitions" as well as "database" or "retrospective EHR studies" was calculated. A Sankey Diagram was then created to visualize the combination of these data.

Subgroup Analysis of ICD Terminologies Used in DR Cohorts

A subgroup analysis of specific ICD terminologies was performed for all included studies. As part of the study review process, all unique ICD-9 or ICD-10 codes were extracted for each study. All codes were truncated to 2 decimal places as this level of precision captured all relevant unique diagnoses with mostly preserved disease severity/staging for simplicity of analysis; however, this did collapse diagnoses with or without DME for a given severity into the same category for ICD-10. If ".x" was included as part of the included ICD code (i.e., ICD-9 361.x), then all possible ICD values containing the given code to 2 decimal places were considered included in that study. The sum of each specific ICD code used across all included studies was then calculated.

Table 2. Search Query, Inclusion, and Exclusion Criteria for This Study

Literature Review Elements	Description
Databases included	PubMed Google Scholar
Search keywords	"Diabetic retinopathy ± electronic health records," "diabetic retinopathy ± EHR, "diabetic retinopathy ± big data," "diabetic retinopathy ± "International Classification of Diseases," "diabetic retinopathy ± ICD codes," "diabetic retinopathy" ± "ICD" ± "billing codes" ± "codes" ± retrospective
Inclusion criteria	1. The disease(s) of interest included diabetic retinopathy (DR) or a specified stage of severity associated with the condition (i.e., nonproliferative DR, vision-threatening DR, or diabetic macular edema). 2. Was peer-reviewed. 3. Communicated original research. 4. Was published before March 2023. 5. Identified participants using electronic health record (EHR) data or a clinical repository. 6. Used a retrospective study design (e.g., retrospective cohort, cross sectional, or case-control). 7. Used EHR-based definitions and not clinical definitions (i.e., fundoscopic imaging) to define DR. 8. Used the International Classification of Diseases, Ninth Revision (ICD-9) or International Classification of Diseases, Tenth Revision (ICD-10) codes. 9. A complete list of codes used to define DR and DR subtypes was found within the text or supplementary materials to ensure reproducibility.
Exclusion criteria	1. Clinical assessment by fundoscopic examination or imaging was used to classify DR diagnosis. 2. A DR cohort could not be identified. 3. Did not follow a retrospective study design.

Results

Characteristics of Included Studies

Our initial literature search yielded 207 articles. After eliminating duplicates, 184 articles were screened for relevance and eligibility for inclusion. After screening for lack of relevance (N = 28 articles) and applying our exclusion criteria, (N = 109), 42 articles were included in this study (Fig 1). Table 3[8–14,34–68] lists the included studies (N = 42) in this literature review. The articles included published studies from 2008 to 2022. Terminologies for defining DR cohorts included: ICD-9, ICD-10, Anatomical Therapeutic Chemical, CPT, Healthcare Common Procedure Coding System, Office of Population Censuses and Surveys Classification of Procedures, National Drug Codes, and SNOMED-CT. The majority of studies focused on associations between other clinical factors (i.e., dementia, ptosis) and DR (N = 17) and progression (N = 10). The remainder of the studies focused on factors such as the incidence of DR in a population (N = 5), accuracy of billing codes documented (N = 4), expenditures for DR (N = 2), defining DR phenotypes using EHR data (N = 2), and treatment trends (N = 2). Further details are available in Table S4 (available at www.ophthalmologyscience.org).

Table 5 summarizes the characteristics of the studies included in this study. The most common study design

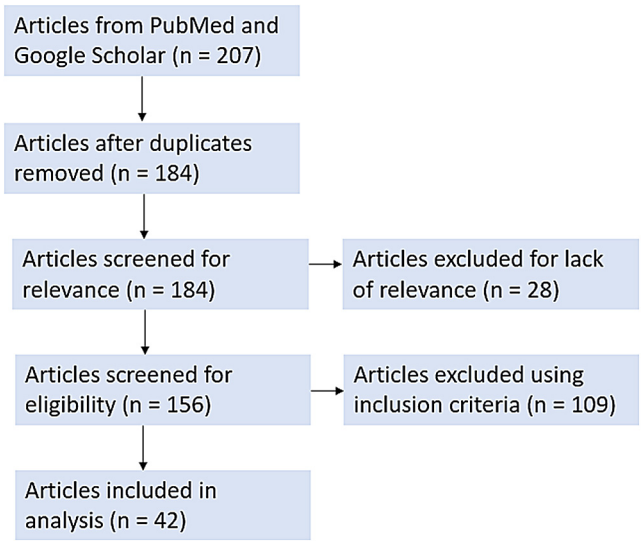

Fig 1. Flowchart of included and excluded literature in this literature review.

Table 3. Included Retrospective Studies in This Literature Review of Cohort Definitions of DR in Electronic Health Record Data

Author	Year	Purpose of Study	Cohort of Interest	Data Terminology
Bearelly et al[14]	2008	Identification of DME from clinical records	DME	ICD-9 codes
Lee et al[34]	2008	Cost of employees with DR	DR of any severity	ICD-9 and CPT Codes
Schmier et al[35]	2009	Estimation of expenditures for DR by Medicare in the US	DR of any severity	ICD-9 codes
Mollazadegan et al[36]	2013	Assessment of risk of developing DR from DM1 and celiac pts	Incident DR of any severity	ICD-9 and ICD-10 codes
Nwanyanwu et al[37]	2013	Assessment of risk factors of progression from NPDR to PDR	DR of any severity	ICD-9 and CPT codes
Jeng et al[38]	2016	Association between DR progression in patients with and without DM nephropathy	DR of any severity	ICD-9 codes
Restrepo et al[39]	2016	Identification of DR cohort of DM2 African Americans using a developed algorithm	DR of any severity	ICD-9 and CPT codes
Chiu et al[40]	2017	Correlation of ICD-9 diabetic complications with problem lists	DR of any severity	ICD-9 codes
Ooley et al[41]	2017	Association between level of DR and diabetic neurosensory hearing loss	DR of any severity	ICD-9 codes
Wang et al[42]	2017	Identification of risk factors for DR progression in patients with DM1 and DM2	DR of any severity	ICD-9 and CPT codes
Douros et al[43]	2018	Association between GLP1 agonists and incident DR	DR of any severity	ICD-10 codes
Kawasaki et al[44]	2018	Association between lipid-lowering meds and development of DR	DR with and without DME	ICD-10 and ATC codes
Lau et al[45]	2018	Accuracy of DR billing codes	DR of any severity	ICD-9 codes for DR and CPT/HCPCS codes for treatment
O'Brien et al[46]	2018	Association between bariatric surgery and development of DM complications such as DR	DR of any severity	ICD-9 and CPT codes
Obeid et al[47]	2018	Determination of lost to follow-up rates in pts with PDR	PDR	ICD-9, ICD-10, and CPT codes
Benoit et al[48]	2019	Determination of rates of eye examination visits in DM pts and estimate cumulative incidence of DR	DR of any severity	ICD-9 codes
Chapman et al[49]	2019	Assessment of health care resource utilization related to managing DM2 and its complications	DR of any severity	ICD-10 or OPCS codes
Chung et al[50]	2019	Assessment of effects of SGLT2 and DPP4 inhibitors on DR progression	DR of any severity	ICD-10 codes
Lee et al[51]	2019	Association between 3 age-related eye diseases (including DR) and dementia-related neuropathology	DR of any severity	ICD-9 codes
Chung et al[9]	2020	Progression of DR in Korean pts	DR of any severity	ICD-10 codes
Gange et al[52]	2020	Determination of rates of eye examinations and diabetic eye disease in first 5 yrs of DM2 diagnosis	DR of any severity	ICD-9 and CPT codes
Kozioł et al[8]	2020	Assessment of prevalence of DR in Poland	DR of any severity	ICD-9 and ICD-10 codes
Kume and Kashiwagi[53]	2020	Factors associated with DME development in DR patients	DME	ICD-10 codes
Moshfeghi et al[54]	2020	5-yr patterns of DR progression in the US	DR of any severity	ICD-9 and ICD-10 codes
Suresh et al[55]	2020	Assessment of the proportion of PDR patients lost to follow-up	PDR	ICD-9 and ICD-10 codes
Yu et al[56]	2020	Association between kidney function and vision-threatening DR	PDR or DME	ICD-9 and ICD-10 codes
Bagdasarova et al[57]	2021	Assessment of protective effects of cataract surgery with risk of development of CRVO or BRVO	DR of any severity	ICD-9, ICD-10, and CM codes
Cai et al[58]	2021	Accuracy of DR and complications during transition from ICD-9 to ICD-10	DR of any severity	ICD-9 and ICD-10 codes
Gange et al[13]	2021	Progression to PDR from initial DM diagnosis	PDR	ICD-9 and CPT codes
Gong et al[59]	2021	Assessment of the change in treatment patterns for PDR over time	PDR	ICD-9, ICD-10, and CPT codes
Hwang et al[60]	2021	Association between ophthalmic conditions and dementia	DR of any severity	ICD-9 codes
Wang et al[61]	2021	Early detection of DR through a predictive model using the EHR model	DR of any severity	ICD-9 and ICD-10 codes
Wittenborn et al[62]	2021	Identification of the prevalence of major eye disorders in Medicare recipients	DR of any severity	ICD 10 codes, CPT, HCPCS, and NDC
Wykoff et al[11]	2021	Association between newly diagnosed DR and risk of blindness compared with patients with good vision (20/40 or better)	DR of any severity	ICD-9, ICD-10, and CM codes
Bathelt et al[63]	2022	Validating the use of OMOP data on the effect of COVID on DR diagnosis	DR of any severity	ICD-10-GM
Chan et al[12]	2022	Association between socioeconomic factors and DR	DR of any severity	SNOMED-CT and CPT codes

Table 3. (Continued.)

Author	Year	Purpose of Study	Cohort of Interest	Data Terminology
Lee et al[64]	2022	Effectiveness of oral hypoglycemic medications on glycemic control and chronic complications	DR of any severity	ICD-10 codes
Lundeen et al[65]	2022	10-yr trends of DME or VTDR in Medicare beneficiaries	PDR, DME	ICD-9 and CPT codes
Mauricio et al[66]	2022	Associations between Alzheimer's disease and DR	DR of any severity	ICD-10 codes
Sugimoto et al[10]	2022	Prevalence/progression of DR associated with hyperglycemia conditions during pregnancy	DR of any severity	ICD 10 codes
Um et al[67]	2022	Association between DR and insomnia risk	DR of any severity	ICD 10 codes
Lin et al[68]	2023	Incidence of ptosis in DR patients	DR of any severity	ICD-9 codes

ATC = Anatomical Therapeutic Chemical code; BRVO = branch retinal vein occlusion; CM = Clinical Modifications; COVID = coronavirus disease; CPT = Current Procedural Terminology; CRVO = central retinal vein occlusion; DM = diabetes mellitus; DM1 = type I diabetes mellitus; DM2 = type II diabetes mellitus; DME = diabetic macular edema; DPP4 = dipeptidyl peptidase 4; DR = diabetic retinopathy; EHR = electronic health record; GLP1 = glucagon-like peptide 1; HCPCS = Healthcare Common Procedure Coding System; ICD = International Classification of Diseases; NDC = National Drug Code; NPDR = nonproliferative diabetic retinopathy; OMOP = Observational Medical Outcomes Partnership; OPCS = Office of Population Censuses and Surveys Classification of Procedures; PDR = proliferative diabetic retinopathy; SGLT2 = sodium glucose cotransporter 2; SNOMED-CT = Systematized Nomenclature of Medicine Clinical Terms; US = United States.

was a retrospective cohort study (N = 27) and less commonly cross sectional (N = 11) and case-control (N = 4). The most common data sources were EHR or claims-based data sets only (N = 20) respectively, whereas only a couple of studies used multimodal data (N = 2). The majority of studies included in this review were from multicenter institutions (N = 39). The majority of DR cohorts focused on known DR (N = 29) and were used as the inclusion data set (N = 24).

Distribution of Codified Definitions of DR

The majority of studies included some components of billing codes designated by the Centers for Medicare & Medicaid Services to define DR. Of the 42 studies included, 27 studies used a third-party database as a data set, and 15 used a retrospective EHR-based institutional data set. Of the studies conducted on a third-party database, 15 used ICD-based definitions only, 11 used ICD with other definitions, and 1 study used non-ICD-based codes to define their cohort; in this study, SNOMED was used to define DR.[12] Of the studies conducted on a retrospective EHR data set, 10 used ICD-based definitions only, and 5 used ICD with other definitions. The distribution of codified definitions utilized in the included studies is shown in Figure 2.

Table 5. Characteristics of Included Studies in this Literature Review

Characteristic of Included Studies	Count, n (%)
Study design	
Case-control	4 (9.5%)
Retrospective cohort	27 (64.3%)
Cross sectional	11 (26.2%)
Data source	
Claims database only	20 (47.6%)
EHR only	20 (47.6%)
Multimodal (EHR + claims)	2 (4.8%)
Geographic scope	
Single center	4 (9.5%)
Multicenter	38 (90.5%)
DR cohort	
Known DR	29 (69.0%)
New DR	12 (28.6%)
New and known DR	1 (2.4%)
DR identifier	
Inclusion	24 (57.1%)
Outcome	10 (23.8%)
Inclusion and outcome	8 (19.1%)

DR = diabetic retinopathy; EHR = electronic health record.

Specific ICD-9 and ICD-10 codes were extracted for all included in this study. For studies utilizing ICD-9 data (N = 27), the majority of studies utilized the 362.01 (N = 24), 362.02 (N = 21), 362.03 (N = 20), 362.04 (N = 20), 362.05 (N = 21), and 362.06 (N = 19) codes, which define mild to severe or unspecified nonproliferative DR and PDR without specification of the presence or absence of macular edema. Additionally, 19 studies used ICD-9 codes, including 362.07 for DME. Some studies used the 250.5x code for diabetes with ocular manifestations (N = 14), and far fewer studies used the ICD-9 249.5 code for secondary diabetes mellitus with ophthalmic manifestations (N = 2). Some studies also used complications of DR such as tractional retinal detachment or other retinal complications in defining their DR cohorts (N = 10). For studies utilizing ICD-10 data (N = 20), the majority of studies utilized E08.3x (N = 8), E09.3x (N = 8), E10.3x (N = 15), E11.3x (N = 17), E12.3x (N = 4), and E13.3x (N = 11) codes, corresponding to the full spectrum of nonproliferative DR and PDR severity. Some studies additionally used H36.0, or DR (N = 11). A few studies also included complications of DR in their ICD-10 inclusion criteria, including neovascular glaucoma (H40.x) and vitreous hemorrhage (H43.1).[10,44,58] One study also chose to include various degrees of vision impairment (H54.x) in patients with known diabetes to identify their DR cohort.[49] The distribution of ICD-9 and ICD-10 codes is shown in Figure 3A and 3B respectively. International Classification of Diseases codes and associated diagnoses are referenced in Table S1.

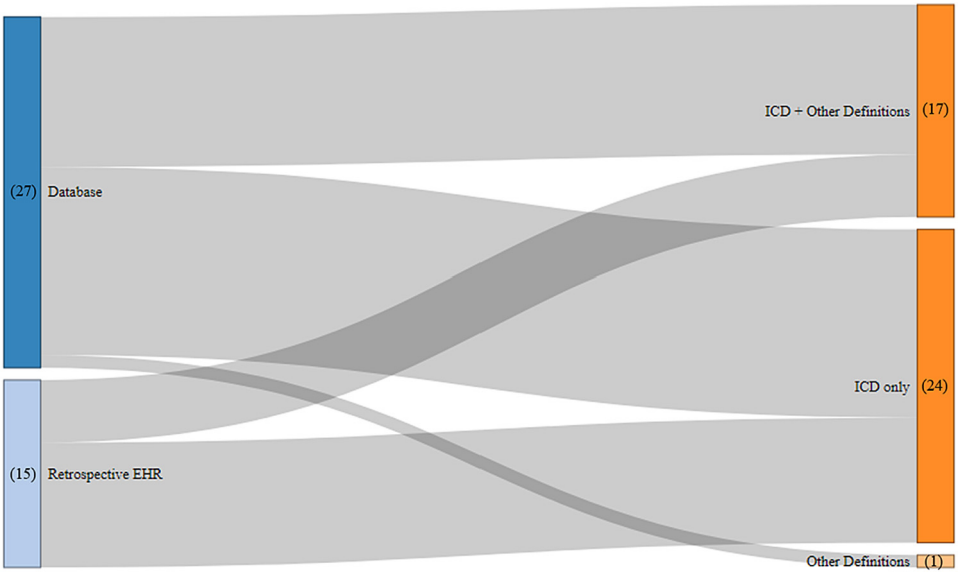

Fig 2. Sankey Diagram of studies utilizing Database vs. Retrospective Electronic Health Record (EHR) data vs. definitions used in their cohorts for diabetic retinopathy.

Discussion

In this study, we reviewed manuscripts using codified cohorts for DR and analyzed variations in coding and characteristics between these studies. This study has 3 key findings: (1) there was substantial variation in coding systems used to define DR cohorts; (2) even with studies using the same coding systems, different definitions were used; and (3) the vast majority of studies used provider-defined billing diagnosis codes, with only a small minority using procedure codes or other clinical data elements.

The first key finding is that there was substantial variation in coding systems used to define DR cohorts. Although the majority of included studies used ICD-9 and ICD-10 to define DR, a significant number of studies included other terminology systems such as SNOMED, CPT codes (i.e., treatments or procedures), or other disease qualifiers (i.e., complications), etc. across a variety of data sets (claims, IRIS registry, etc.). Of note, there was a transition in the usage of ICD-9 to ICD-10 codes, likely due to the formal transition in 2015,[69] although this is problematic as the diagnosis codes did not necessarily map 1 to 1.[58] A key difference with ICD-10 codes is that a single code can specify DR severity and DME status, whereas in ICD-9, only the presence or absence of DME is codified (362.07). Although these ICD codes were generally accurate in staging DR severity[70] and identifying treatment,[45]

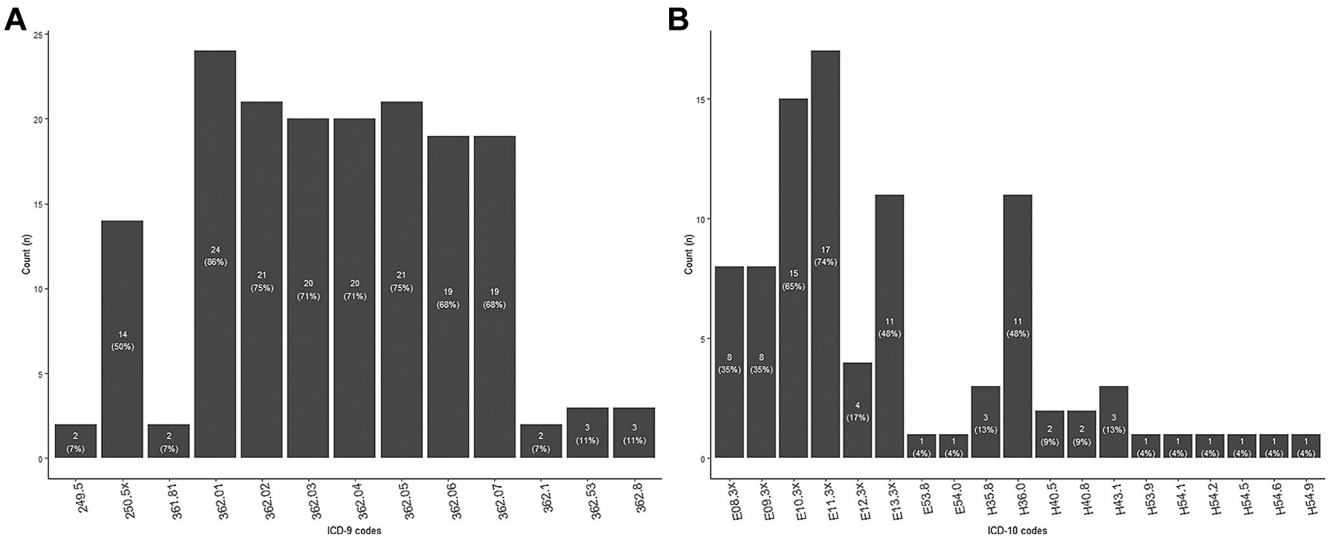

Fig 3. The distribution of (A) ICD-9 and (B) ICD-10 codes.

comparisons between DR cohorts defined by other nonbilling coding systems such as SNOMED may prove problematic and limit our ability to compare outcomes between studies using different systems to define DR cohorts. Attempts to standardize and aggregate large population-based data sets such as IRIS and *All of Us* may help address this variation and availability of data for analyses, although they are limited by the types of data collected. Harmonization of cohort definitions across different vocabularies may be achieved by standardized mapping of concepts, such as the Observational Medical Outcomes Partnership model.[23,71,72] This case for standardization for large data analyses extends beyond coding systems used and has important implications for imaging standards for ophthalmology.[73] These large data sets will not only need to address standardized methodologies used to define diseases but also ensure accurate and complete data and accessibility.

The second key finding is that different definitions were used even within the same coding system. For example, for both ICD-9 and ICD-10 based cohorts, studies demonstrated numeric variability among the codes used to bill for nonproliferative DR (362.0x and E08-E13.x respectively) (Fig 3). Although we analyzed broadly by ICD codes for all types of DR cohorts, the specific composition of cohorts was likely affected by the authors' choices to incorporate factors beyond diagnosis codes such as treatments given and other diabetic complications. For example, only 1 study[49] used ICD10 codes for low vision and visual impairment (H54.x) as part of their cohort definition for DR. Additionally, some studies[11,13,14] only used the diagnoses codes to define their DME cohorts, whereas others incorporated treatments such as anti-VEGF or steroids.[38,44,59] There also existed differential usage of CPT codes, for example, Lee et al[34] and Hwang et al[60] used different sets of CPT codes (67036, 67038, 67039, 67040, 67107, 67108 vs. 67113, 67042, respectively) to define vitrectomy. Although these variations may be in part due to a given study's cohort of interest, these findings emphasize the concern regarding the lack of standardization in defining DR cohorts. Similar issues regarding the standardization of data elements have been raised by Hribar et al[74] regarding definitions of visual acuity after cataract surgery and intraocular pressure after glaucoma surgery for assessing ophthalmic quality measures. Thus, there exists a secondary need for standardizing disease definitions within a coding system for various DR entities (i.e., PDR, center-involving DME, DR with VH, and DR with tractional retinal detachment) among other ophthalmic diseases. Although there have been attempts to create standardized definitions of chronic diseases[75] and DR,[76] the overall lack of universally accepted standardized definitions for DR affects the ability to compare outcomes between studies, assess the generalizability of multistudy findings in reviews, and adversely affects the reproducibility of studies in other settings.

The third key finding is that the vast majority of studies used provider-defined billing diagnosis codes, with only a small minority using procedure codes or other clinical data elements. Specifically, diagnosis billing codes such as ICD-9/10 are defined by the provider to bill for the visit. Although some studies used other clinical data (i.e., medications, and treatments as defined by CPT) to supplement their ICD-based cohorts of DR (Fig 2), some of these codified definitions require manual input from the investigators to define DR, although some vocabularies such as CPT and RxNorm may be employed to identify diagnoses programmatically. Creating codified phenotypes for diseases such as DR may be a potential solution to accurately and potentially automatically define DR severity and complications without physician input, which may be inaccurate and incomplete (i.e., unspecified).[77] Although there exist tools to aggregate disease phenotypes and other features to create cohort definitions and atlases, these tools are not used widely outside of the research sphere.[78] Although codified phenotypes to define type II diabetes mellitus[75,79,80] and coronavirus disease[81,82] exist, there is a need for more specific definitions in ophthalmology, for which many diseases are defined based on examination-based and image-based findings documented in the EHR, which were excluded from our study but would likely introduce even more variation. However, these data may be challenging to include due to variability in documentation and incompleteness. Specific documentation standards and guidelines and physician buy-in will be needed to promote standardized data entry conducive to codified cohort phenotyping without increasing the documentation burden of the EHR.[83,84] Inclusion of other data such as visual acuity may have important implications in outcomes studies for these diabetes patients and would enable improved public health reporting to potentially address disparities in diabetes and DR.[25] Standardized vocabularies such as the Observational Medical Outcomes Partnership model may be a promising means to codify clinical findings such as visual acuity. Other solutions may include high-throughput phenotyping using EHR data, which has been employed extensively in genomic analyses, and may also have a role in reproducibly selecting patient cohorts.[85,86]

This study has additional limitations. First, our study only examined codified definitions of DR. Future work will be needed to understand how we can incorporate fundoscopic definitions for DR, which were used in many seminal DR trials,[87,88] but are difficult to incorporate into discrete DR diagnoses in retrospective studies. Second, we did not quantify the impact of various coding schemas and specific coding definitions used for DR cohort composition. Future research would be invaluable in elucidating the consequences of such variation in cohorts based on various studies' DR definitions. Third, our study did not analyze how DR defined using other codified definitions, such as SNOMED, varied compared with other billing code-based criteria. This was in part due to the low number of studies that used non-ICD codes in defining DR. More work is needed to understand the impact of DR definitions using various definitions of DR and how these cohorts compare to each other.

Overall, there exists significant variation in which coding systems are used and how they are used to define DR cohorts. This has important implications for the generalizability and reproducibility of retrospective studies in ophthalmology. Ongoing efforts in standardization are needed to create reproducible and interoperable codified definitions and cohorts for DR and other ophthalmic diseases.

Footnotes and Disclosures

Originally received: July 11, 2023.
Final revision: January 4, 2024.
Accepted: January 11, 2024.
Available online: January 24, 2024.Manuscript no. XOPS-D-23-00167R2.

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

[2] UCSD Health Department of Biomedical Informatics, University of California San Diego, La Jolla, California.

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

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

[5] Department of Ophthalmology, Duke Eye Center, Duke University, Durham, North Carolina.

[6] F.I. Proctor Foundation, University of California San Francisco, San Francisco, California.

[7] Department of Ophthalmology, University of California San Francisco, San Francisco, California.

Disclosure(s):

All authors have completed and submitted the ICMJE disclosures form.

The author(s) have made the following disclosure(s):

B.C.T.: Payment or honoraria for lectures, presentations, speakers bureaus — Alimera, Eyepoint, Alcon.

D.B.: Consulting fees — AbbVie/Allergan, Genentech, Glaukos, Iveric Bio, Verana Health; Speaker fees — Iveric Bio.

C.Q.S.: Grants or contracts — National Eye Institute, All May See Foundation, UCSF Senate Grant.

J.G.S.: Other financial or non-financial interests — General payment 229.60, Bausch and Lomb 122.53 and EyePoint 107.07.

S.L.B.: Equipment — Optomed and Topcon; Consulting fees — voxelcloud.io; Speaker fees — iVista Medical Education, outside the submitted work.

The other authors have no proprietary or commercial interest in any materials discussed in this article.

S.L.B. is supported by NIH Grants DP5OD29610, P30EY022589, R01EY034146, OT2OD032644, T35EY033704, R01MD014850 (Bethesda, Maryland, United States) and an unrestricted departmental grant from Research to Prevent Blindness (New York, New York, United States).

HUMAN SUBJECTS: No human subjects were included in this study. This study adhered to the Declaration of Helsinki. The requirement for informed consent was waived by the University of California San Diego Institutional Review Board for this study as no patient data were used.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Chen, Copado, Baxter

Data collection: Chen, Copado, Vallejos, Kalaw, Soe, Baxter

Analysis and interpretation: Chen, Copado, Cai, Toy, Borkar, Sun, Shantha, Baxter

Obtained funding: Baxter

Overall responsibility: Chen, Baxter

Abbreviations and Acronyms:
CPT = Current Procedural Terminology; **DME** = diabetic macular edema; **DR** = diabetic retinopathy; **EHR** = electronic health record; **ICD** = International Classification of Diseases; **PDR** = proliferative diabetic retinopathy; **SNOMED-CT** = Systematized Nomenclature of Medicine Clinical Terms.

Keywords:
Big data, Data standardization, Diabetic retinopathy, Electronic health records, Informatics.

Correspondence:
Sally L. Baxter, MD, MSc, Division of Ophthalmology Informatics and Data Science, Viterbi Family Department of Ophthalmology and Shiley Eye Institute, UCSD Health Department of Biomedical Informatics, 9415 Campus Point Dr MC0946, La Jolla, CA 92093. E-mail: s1baxter@health.ucsd.edu.

References

1. Teo ZL, Tham YC, Yu M, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. *Ophthalmology*. 2021;128:1580−1591. https://doi.org/10.1016/j.ophtha.2021.04.027.
2. Ting DSW, Cheung GCM, Wong TY. Diabetic retinopathy: global prevalence, major risk factors, screening practices and public health challenges: a review. *Clin Exp Ophthalmol*. 2016;44:260−277. https://doi.org/10.1111/ceo.12696.
3. Sun JK, Jampol LM. The Diabetic Retinopathy Clinical Research Network (DRCR.net) and Its Contributions to the Treatment of Diabetic Retinopathy. *Ophthal Res*. Diabetic Retinopathy Clinical Research Network. *DRCR.net*. 2019;62:225−230. https://doi.org/10.1159/000502779.
4. Ellis MP, Bacorn C, Luu KY, et al. Cost analysis of teleophthalmology screening for diabetic retinopathy using teleophthalmology billing codes. *Ophthalmic Surg Lasers Imaging Retina*. 2020;51:S26−S34. https://doi.org/10.3928/23258160-20200108-04.
5. Baxter SL, Quackenbush Q, Cerda J, Gregg C, Millen M, Thorne C. Implementing clinical informatics tools for primary care-based diabetic retinopathy screening. *Am J Manag Care*. 2022;28:e355−e362. https://doi.org/10.37765/ajmc.2022.89253.
6. Coney JM, Scott AW. Racial disparities in the screening and treatment of diabetic retinopathy. *J Natl Med Assoc*. 2022;114:171−181. https://doi.org/10.1016/j.jnma.2021.12.011.
7. Hemmerich C, Jones G, Staggs J, Anderson RM, Bacani R, Vassar M. Inequities and research gaps in ophthalmology: A scoping review. *JAMA Ophthalmol*. 2023;141:63−70. https://doi.org/10.1001/jamaophthalmol.2022.5237.
8. Kozioł M, Nowak MS, Udziela M, Piątkiewicz P, Grabska-Liberek I, Szaflik JP. First nation-wide study of diabetic retinopathy in Poland in the years 2013-2017. *Acta Diabetol*. 2020;57:1255−1264. https://doi.org/10.1007/s00592-020-01540-6.
9. Chung YR, Ha KH, Lee K, Kim DJ. Diabetic retinopathy and related clinical practice for people with diabetes in Korea: A 10-year trend analysis. *Diabetes Metab J*. 2020;44:928−932. https://doi.org/10.4093/dmj.2020.0096.
10. Sugimoto M, Sampa K, Tsukitome H, et al. Trends in the prevalence and progression of diabetic retinopathy associated with hyperglycemic disorders during pregnancy in Japan. *J Clin Med*. 2021;11:165. https://doi.org/10.3390/jcm11010165.

11. Wykoff CC, Khurana RN, Nguyen QD, et al. Risk of blindness among patients with diabetes and newly diagnosed diabetic retinopathy. *Diabetes Care*. 2021;44:748−756. https://doi.org/10.2337/dc20-0413.
12. Chan AX, McDermott JJ, Lee TC, et al. Associations between healthcare utilization and access and diabetic retinopathy complications using All of Us nationwide survey data. *PLoS One*. 2022;17:e0269231. https://doi.org/10.1371/journal.pone.0269231.
13. Gange WS, Lopez J, Xu BY, Lung K, Seabury SA, Toy BC. Incidence of proliferative diabetic retinopathy and other neovascular sequelae at 5 years following diagnosis of type 2 diabetes. *Diabetes Care*. 2021;44:2518−2526. https://doi.org/10.2337/dc21-0228.
14. Bearelly S, Mruthyunjaya P, Tzeng JP, et al. Identification of patients with diabetic macular edema from claims data: A validation study. *Arch Ophthalmol*. 2008;126:986−989. https://doi.org/10.1001/archopht.126.7.986.
15. Chiang MF, Sommer A, Rich WL, Lum F, Parke DW; II. The 2016 American Academy of Ophthalmology IRIS® Registry (Intelligent Research in Sight) database: characteristics and methods. *Ophthalmology*. 2018;125:1143−1148. https://doi.org/10.1016/j.ophtha.2017.12.001.
16. All of US Research Program Investigators, Denny JC, Rutter JL, et al. The "all of Us" research program. *N Engl J Med*. 2019;381:668−676. https://doi.org/10.1056/NEJMsr1809937.
17. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316:2402−2410. https://doi.org/10.1001/jama.2016.17216.
18. Ting DSW, Cheung CYL, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA*. 2017;318:2211−2223. https://doi.org/10.1001/jama.2017.18152.
19. Lee CS, Brandt JD, Lee AY. Big data and artificial intelligence in ophthalmology: where are we now? *Ophthalmol Sci*. 2021;1:100036. https://doi.org/10.1016/j.xops.2021.100036.
20. Shweikh Y, Sekimitsu S, Boland MV, Zebardast N. The growing need for ophthalmic data standardization. *Ophthalmol Sci*. 2023;3:100262. https://doi.org/10.1016/j.xops.2022.100262.
21. Baxter SL, Lee AY. Gaps in standards for integrating artificial intelligence technologies into ophthalmic practice. *Curr Opin Ophthalmol*. 2021;32:431−438. https://doi.org/10.1097/ICU.0000000000000781.
22. Halfpenny W, Baxter SL. Towards effective data sharing in ophthalmology: data standardization and data privacy. *Curr Opin Ophthalmol*. 2022;33:418−424. https://doi.org/10.1097/ICU.0000000000000878.
23. 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 (USCDI). *Ophthalmol Sci*. 2022;2:100210. https://doi.org/10.1016/j.xops.2022.100210.
24. Tavakoli K, Kalaw FGP, Bhanvadia S, Hogarth M, Baxter SL. Concept coverage analysis of ophthalmic infections and trauma among the standardized medical terminologies SNOMED-CT, ICD-10-CM, and ICD-11. *Ophthalmol Sci*. 2023;3:100337. https://doi.org/10.1016/j.xops.2023.100337.
25. Baxter SL, Nwanyanwu K, Legault G, Lee AY. Data sources for evaluating health disparities in ophthalmology: where we are and where we need to go. *Ophthalmology*. 2022;129:e146−e149. https://doi.org/10.1016/j.ophtha.2022.06.035.
26. National Center for Health Statistics. International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM). October 17, 2022 https://www.cdc.gov/nchs/icd/icd-10-cm.htm. Accessed June 5, 2023.
27. Lee CS, Lee AY, Baughman D, et al. The United Kingdom Diabetic Retinopathy Electronic Medical Record Users Group: Report 3: Baseline Retinopathy and Clinical Features Predict Progression of Diabetic Retinopathy: Report 3. *Am J Ophthalmol*. 2017;180:64−71. https://doi.org/10.1016/j.ajo.2017.05.020.
28. Semeraro F, Parrinello G, Cancarini A, et al. Predicting the risk of diabetic retinopathy in type 2 diabetic patients. *J Diabetes Complications*. 2011;25:292−297. https://doi.org/10.1016/j.jdiacomp.2010.12.002.
29. Newton KM, Peissig PL, Kho AN, et al. Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. *J Am Med Inform Assoc*. 2013;20:e147−e154. https://doi.org/10.1136/amiajnl-2012-000896.
30. Nguyen CTN, Yosef M, Khalatbari S, Shah AR. Sociodemographic variables associated with risk for diabetic retinopathy. *Clin Diabetes Endocrinol*. 2022;8:7. https://doi.org/10.1186/s40842-022-00144-z.
31. Malhotra NA, Greenlee TE, Iyer AI, Conti TF, Chen AX, Singh RP. Racial, ethnic, and insurance-based disparities upon initiation of anti-vascular endothelial growth factor therapy for diabetic macular edema in the US. *Ophthalmology*. 2021;128:1438−1447. https://doi.org/10.1016/j.ophtha.2021.03.010.
32. Obeid A, Su D, Patel SN, et al. Outcomes of eyes lost to follow-up with proliferative diabetic retinopathy that received panretinal photocoagulation versus intravitreal anti-vascular endothelial growth factor. *Ophthalmology*. 2019;126:407−413. https://doi.org/10.1016/j.ophtha.2018.07.027.
33. RStudio Team. RStudio: Integrated Development Environment for R. http://www.rstudio.com/; 2015.
34. Lee LJ, Yu AP, Cahill KE, et al. Direct and indirect costs among employees with diabetic retinopathy in the United States. *Curr Med Res Opin*. 2008;24:1549−1559. https://doi.org/10.1185/030079908X297303.
35. Schmier JK, Covert DW, Lau EC, Matthews GP. Medicare expenditures associated with diabetes and diabetic retinopathy. *Retina*. 2009;29:199−206. https://doi.org/10.1097/IAE.0b013e3181884f2d.
36. Mollazadegan K, Kugelberg M, Montgomery SM, Sanders DS, Ludvigsson J, Ludvigsson JF. A population-based study of the risk of diabetic retinopathy in patients with type 1 diabetes and celiac disease. *Diabetes Care*. 2013;36:316−321. https://doi.org/10.2337/dc12-0766.
37. Harris Nwanyanwu K, Talwar N, Gardner TW, Wrobel JS, Herman WH, Stein JD. Predicting development of proliferative diabetic retinopathy. *Diabetes Care*. 2013;36:1562−1568. https://doi.org/10.2337/dc12-0790.
38. Jeng CJ, Hsieh YT, Yang CM, Yang CH, Lin CL, Wang IJ. Diabetic retinopathy in patients with diabetic nephropathy: development and progression. *PLoS One*. 2016;11:e0161897. https://doi.org/10.1371/journal.pone.0161897.
39. Restrepo NA, Farber-Eger E, Crawford DC. Searching in the dark: phenotyping diabetic retinopathy in a de-identified electronic medical record sample of African Americans. *AMIA Jt Summits Transl Sci Proc*. 2016;2016:221−230.
40. Chiu S, Davis J, Giaconi J, et al. Variable validity of computer extracted problem lists for complications of diabetes mellitus within the VA Greater Los Angeles Health System. *Diabetes*

Metab Syndr. 2017;11:S611−S615. https://doi.org/10.1016/j.dsx.2017.04.013.
41. Ooley C, Jun W, Le K, et al. Correlational study of diabetic retinopathy and hearing loss. *Optom Vis Sci*. 2017;94:339−344. https://doi.org/10.1097/OPX.0000000000001025.
42. Wang SY, Andrews CA, Herman WH, Gardner TW, Stein JD. Incidence and risk factors for developing diabetic retinopathy among youths with type 1 or type 2 diabetes throughout the United States. *Ophthalmology*. 2017;124:424−430. https://doi.org/10.1016/j.ophtha.2016.10.031.
43. Douros A, Filion KB, Yin H, et al. Glucagon-like peptide 1 receptor agonists and the risk of incident diabetic retinopathy. *Diabetes Care*. 2018;41:2330−2338. https://doi.org/10.2337/dc17-2280.
44. Kawasaki R, Konta T, Nishida K. Lipid-lowering medication is associated with decreased risk of diabetic retinopathy and the need for treatment in patients with type 2 diabetes: a real-world observational analysis of a health claims database. *Diabetes Obes Metab*. 2018;20:2351−2360. https://doi.org/10.1111/dom.13372.
45. Lau M, Prenner JL, Brucker AJ, VanderBeek BL. Accuracy of billing codes used in the therapeutic care of diabetic retinopathy. *JAMA Ophthalmol*. 2017;135:791−794. https://doi.org/10.1001/jamaophthalmol.2017.1595.
46. O'Brien R, Johnson E, Haneuse S, et al. Microvascular outcomes in patients with diabetes after bariatric surgery versus usual care: a matched cohort study. *Ann Intern Med*. 2018;169:300−310. https://doi.org/10.7326/M17-2383.
47. Obeid A, Gao X, Ali FS, et al. Loss to follow-up in patients with proliferative diabetic retinopathy after panretinal photocoagulation or intravitreal anti-VEGF injections. *Ophthalmology*. 2018;125:1386−1392. https://doi.org/10.1016/j.ophtha.2018.02.034.
48. Benoit SR, Swenor B, Geiss LS, Gregg EW, Saaddine JB. Eye care utilization among insured people with diabetes in the U.S., 2010-2014. *Diabetes Care*. 2019;42:427−433. https://doi.org/10.2337/dc18-0828.
49. Chapman D, Foxcroft R, Dale-Harris L, Ronte H, Bidgoli F, Bellary S. Insights for care: the healthcare utilisation and cost impact of managing type 2 diabetes-associated microvascular complications. *Diabetes Ther*. 2019;10:575−585. https://doi.org/10.1007/s13300-018-0548-4.
50. Chung YR, Ha KH, Lee K, Kim DJ. Effects of sodium-glucose cotransporter-2 inhibitors and dipeptidyl peptidase-4 inhibitors on diabetic retinopathy and its progression: A real-world Korean study. *PLoS One*. 2019;14:e0224549. https://doi.org/10.1371/journal.pone.0224549.
51. Lee CS, Larson EB, Gibbons LE, et al. Ophthalmology-based neuropathology risk factors: diabetic retinopathy is associated with deep microinfarcts in a community-based autopsy study. *J Alzheimers Dis JAD*. 2019;68:647−655. https://doi.org/10.3233/JAD-181087.
52. Gange WS, Xu BY, Lung K, Toy BC, Seabury SA. Rates of eye care and diabetic eye disease among insured patients with newly diagnosed type 2 diabetes. *Ophthalmol Retina*. 2021;5:160−168. https://doi.org/10.1016/j.oret.2020.07.004.
53. Kume A, Kashiwagi K. Systemic and ocular diseases associated with the development of diabetic macular edema among Japanese patients with diabetes mellitus. *BMC Ophthalmol*. 2020;20:309. https://doi.org/10.1186/s12886-020-01578-8.
54. Moshfeghi A, Garmo V, Sheinson D, Ghanekar A, Abbass I. Five-year patterns of diabetic retinopathy progression in US clinical practice. *Clin Ophthalmol*. 2020;14:3651−3659. https://doi.org/10.2147/OPTH.S275968.
55. Suresh R, Yu HJ, Thoveson A, et al. Loss to follow-up among patients with proliferative diabetic retinopathy in clinical practice. *Am J Ophthalmol*. 2020;215:66−71. https://doi.org/10.1016/j.ajo.2020.03.011.
56. Yu Y, Ying GS, Maguire MG, VanderBeek BL. Association of diagnosis code-based and laboratory results-based kidney function with development of vision threatening diabetic retinopathy. *Ophthal Epidemiol*. 2020;27:498−503. https://doi.org/10.1080/09286586.2020.1773869.
57. Bagdasarova Y, Lee AY, Maring M, et al. Cataract surgery is not associated with decreased risk of retinal vein occlusion. *Ophthalmol Sci*. 2021;1:100041. https://doi.org/10.1016/j.xops.2021.100041.
58. Cai CX, Michalak SM, Stinnett SS, Muir KW, Fekrat S, Borkar DS. Effect of ICD-9 to ICD-10 transition on accuracy of codes for stage of diabetic retinopathy and related complications: results from the CODER study. *Ophthalmol Retina*. 2021;5:374−380. https://doi.org/10.1016/j.oret.2020.08.004.
59. Gong D, Hall N, Elze T, et al. Temporal trends in the treatment of proliferative diabetic retinopathy: an AAO IRIS® Registry analysis. *Ophthalmol Sci*. 2021;1:100037. https://doi.org/10.1016/j.xops.2021.100037.
60. Hwang PH, Longstreth WT, Thielke SM, et al. Ophthalmic conditions associated with dementia risk: the cardiovascular Health Study. *Alzheimers Dement*. 2021;17:1442−1451. https://doi.org/10.1002/alz.12313.
61. Wang R, Miao Z, Liu T, et al. Derivation and validation of essential predictors and risk index for early detection of diabetic retinopathy using electronic health records. *J Clin Med*. 2021;10:1473. https://doi.org/10.3390/jcm10071473.
62. Wittenborn JS, Gu Q, Erdem E, et al. The prevalence of diagnosis of major eye diseases and their associated payments in the Medicare fee-for-service program. *Ophthal Epidemiol*. 2023;30:129−141. https://doi.org/10.1080/09286586.2021.1968006.
63. Bathelt F, Reinecke I, Peng Y, et al. Opportunities of digital infrastructures for disease management—exemplified on COVID-19-related change in diagnosis counts for diabetes-related eye diseases. *Nutrients*. 2022;14:2016. https://doi.org/10.3390/nu14102016.
64. Lee KA, Jin HY, Kim YJ, Kim SS, Cho EH, Park TS. Real-world comparison of mono and dual combination therapies of metformin, sulfonylurea, and dipeptidyl peptidase-4 inhibitors using a common data model: a retrospective observational study. *Med (Baltim)*. 2022;101:e28823. https://doi.org/10.1097/MD.0000000000028823.
65. Lundeen EA, Andes LJ, Rein DB, et al. Trends in prevalence and treatment of diabetic macular edema and vision-threatening diabetic retinopathy among medicare part B fee-for-service beneficiaries. *JAMA Ophthalmol*. 2022;140:345−353. https://doi.org/10.1001/jamaophthalmol.2022.0052.
66. Mauricio D, Vlacho B, Barrot de la Puente J, et al. Associations between diabetic retinopathy and Parkinson's disease: results from the Catalonian primary care cohort study. *Front Med (Lausanne)*. 2021;8:800973. https://doi.org/10.3389/fmed.2021.800973.
67. Um YH, Kim TW, Jeong JH, Hong SC, Seo HJ, Han KD. Association between diabetic retinopathy and insomnia risk: a nationwide population-based study. *Front Endocrinol*. 2022;13:939251. https://doi.org/10.3389/fendo.2022.939251.
68. Lin CJ, Hsu AY, Tien PT, et al. Diabetic retinopathy as a potential risk factor for ptosis: A 13-year nationwide population-based cohort study in Taiwan. *Front Epidemiol*. 2023;3:1093064. https://doi.org/10.3389/fepid.2023.1093064.

69. Cartwright DJ. ICD-9-CM to ICD-10-CM codes: what? Why? How? *Adv Wound Care*. 2013;2:588−592. https://doi.org/10.1089/wound.2013.0478.
70. Muir KW, Gupta C, Gill P, Stein JD. Accuracy of international classification of disease (ICD-9-CM) billing codes for common ophthalmologic conditions. *JAMA Ophthalmol*. 2013;131:119−120. https://doi.org/10.1001/jamaophthalmol.2013.577.
71. Belenkaya R, Gurley MJ, Golozar A, et al. Extending the OMOP common data model and standardized vocabularies to support observational cancer research. *JCO Clin Cancer Inform*. 2021;5:12−20. https://doi.org/10.1200/CCI.20.00079.
72. Reinecke I, Zoch M, Reich C, Sedlmayr M, Bathelt F. The usage of OHDSI OMOP − a scoping review. *Stud Health Technol Inform*. 2021;283:95−103. https://doi.org/10.3233/SHTI210546.
73. Lee AY, Campbell JP, Hwang TS, Lum F, Chew EY; American Academy of Ophthalmology. Recommendations for standardization of images in ophthalmology. *Ophthalmology*. 2021;128:969−970. https://doi.org/10.1016/j.ophtha.2021.03.003.
74. Hribar M, Chen JS, Chen A, Chiang MF. Can EHR data automatically calculate ophthalmic quality measures? The impact of baselines. *Invest Ophthalmol Vis Sci*. 2022;63:3827.
75. CCW. Chronic Conditions Data Warehouse. https://www2.ccwdata.org. Accessed June 25, 2023.
76. PheKB. Diabetic retinopathy. https://phekb.org/phenotype/diabetic-retinopathy. Accessed June 25, 2023.
77. Verana Health. Verana Health to Present Findings of Diabetic Retinopathy Study Using American Academy of Ophthalmology IRIS® Registry Data at AMIA Annual Symposium. https://www.veranahealth.com/verana-health-to-present-findings-of-diabetic-retinopathy-study-using-american-academy-of-ophthalmology-iris-registry-data-at-amia-annual-symposium/. Accessed February 19, 2023.
78. Observational Health Data Sciences and Informatics. Defining cohorts. In: The Book of OHDSI. https://ohdsi.github.io/TheBookOfOhdsi/. Accessed October 9, 2023.
79. Pendergrass SA, Crawford DC. Using electronic health records to generate phenotypes for research. *Curr Protoc Hum Genet*. 2019;100:e80. https://doi.org/10.1002/cphg.80.
80. Wei WQ, Denny JC. Extracting research-quality phenotypes from electronic health records to support precision medicine. *Genome Med*. 2015;7:41. https://doi.org/10.1186/s13073-015-0166-y.
81. Kim J, Neumann L, Paul P, et al. Privacy-protecting, reliable response data discovery using COVID-19 patient observations. *J Am Med Inform Assoc*. 2021;28:1765−1776. https://doi.org/10.1093/jamia/ocab054.
82. Kuo TT, Pham A, Edelson ME, et al. Blockchain-enabled immutable, distributed, and highly available clinical research activity logging system for federated COVID-19 data analysis from multiple institutions. *J Am Med Inform Assoc*. 2023;30:1167−1178. https://doi.org/10.1093/jamia/ocad049.
83. Goldstein IH, Hwang T, Gowrisankaran S, Bales R, Chiang MF, Hribar MR. Changes in electronic health record use time and documentation over the course of a decade. *Ophthalmology*. 2019;126:783−791. https://doi.org/10.1016/j.ophtha.2019.01.011.
84. Chen JS, Hribar MR, Goldstein IH, et al. Electronic health record note review in an outpatient specialty clinic: who is looking? *JAMIA Open*. 2021;4:ooab044. https://doi.org/10.1093/jamiaopen/ooab044.
85. Son JH, Xie G, Yuan C, et al. Deep phenotyping on electronic health records facilitates genetic diagnosis by clinical exomes. *Am J Hum Genet*. 2018;103:58−73. https://doi.org/10.1016/j.ajhg.2018.05.010.
86. Pathak J, Bailey KR, Beebe CE, et al. Normalization and standardization of electronic health records for high-throughput phenotyping: the SHARPn consortium. *J Am Med Inform Assoc*. 2013;20:e341−e348. https://doi.org/10.1136/amiajnl-2013-001939.
87. Bressler SB, Odia I, Glassman AR, et al. Changes in diabetic retinopathy severity when treating diabetic macular edema with ranibizumab: DRCR.net Protocol I 5-Year Report. DRCR.net. *Retina*. 2018;38:1896−1904. https://doi.org/10.1097/IAE.0000000000002302.
88. Maturi RK, Glassman AR, Josic K, et al. Effect of intravitreous anti-vascular endothelial growth factor vs sham treatment for prevention of vision-threatening complications of diabetic retinopathy: The Protocol W Randomized Clinical Trial. *JAMA Ophthalmol*. 2021;139:701−712. https://doi.org/10.1001/jamaophthalmol.2021.0606.