CURRENT OPINION

Big data in ophthalmology: comparative databases and research applications

Mohammad Ayoubi[a,*], *Mohamed S. Sayed*[b,*], *Mohamed M. Khodeiry*[a,c], *Abdelrahman M. Elhusseiny*[d] and *Richard K. Lee*[a]

Purpose of review

This review aims to highlight the expanding role of big data in ophthalmology, provide a comparison of the most prominent databases, and their use in glaucoma-specific research. Understanding the strengths and limitations of each database allows researchers to tailor their research questions appropriately.

Recent findings

Several large-scale databases have emerged in ophthalmology research. Some databases offer detailed ocular exam findings and imaging, supporting artificial intelligence-driven diagnostics and treatment evaluation. Others are broader in scope, providing real-world population data to study trends and outcomes. Some databases even integrate genomic and systemic data, enabling novel explorations of disease risk and health disparities. Collectively, big data has enabled large-scale studies on a wide range of topics, advancing the field of ophthalmology in every aspect.

Summary

Big data platforms are transforming ophthalmology research, from uncovering systemic-ocular relationships to enabling artificial intelligence applications. Researchers can select platforms based on the availability of imaging, systemic data, or genomic information to better address specific research questions. Doing so can enhance precision medicine, address care disparities, and drive innovation in disease pathophysiology discovery, detection, and management.

Keywords

artificial intelligence, big data, glaucoma, ophthalmology

INTRODUCTION

Big data refers to large sets of data that typically require specialized tools and techniques to capture, store, manage, and analyze. In healthcare, the digitization of health records and medical imaging has led to the creation of large databases that support developing clinical insights and aiding research studies [1]. Ophthalmology has quickly adopted big data due to its high reliance on numerical biological data points [visual acuity, intraocular pressure (IOP), etc.] and high-volume/low-cost imaging to track diseases over time [2]. One of the most significant uses of big data in ophthalmology is for population-level analysis and disease surveillance. Researchers can track patterns of disease such as risk factors, associations, treatment outcomes, and practice variation across wide populations [3].

Several large-scale databases are being implemented in ophthalmology research, each offering unique strengths depending on the research question. These include general-purpose health data networks such as TriNetX, Cosmos, and All of Us, as well as more ophthalmology-focused resources like the Intelligent Research in Sight (IRIS) Registry, the Sight Outcomes Research Collaborative (SOURCE), and the United Kingdom (UK) Biobank. Although this review focuses on a few major databases, many additional databases exist that may support ophthalmic research, and future studies may benefit from evaluating their utility based on specific research goals or subspecialty needs. Nonetheless, this review provides a comparative analysis of

[a]Department of Ophthalmology, Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, Florida, USA, [b]Eye Institute, Cleveland Clinic Abu Dhabi, Abu Dhabi, UAE, [c]Department of Ophthalmology and Visual Sciences, University of Louisville, Louisville, Kentucky and [d]Harvey and Bernice Jones Eye Institute, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

Correspondence to Richard K. Lee, MD, PhD, Department of Ophthalmology, Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, 900 NW 17th Street, Miami, FL 33136, USA. Tel: +1 305 326 6000; e-mail: rlee@med.miami.edu

[*]M.A. and M.S.S. contributed equally as co-first authors.

Curr Opin Ophthalmol 2026, 37:81–88

DOI:10.1097/ICU.0000000000001190

KEY POINTS

- Big data platforms in ophthalmology vary widely in the types of data they provide, ranging from clinical exam findings to imaging and genomics.
- Ophthalmology-specific databases like IRIS and SOURCE offer detailed eye exam data that can evaluate ocular disease severity and progression.
- Broader datasets like TriNetX and Cosmos offer the opportunity to evaluate associations between systemic health conditions, systemic medications, and ocular conditions.
- Each database has distinct access requirements, with some limited to affiliated institutions and others open to approved external researchers.
- Combining complementary datasets enhances the potential to address complex questions in glaucoma research and supports more equitable, artificial intelligence-driven care.

major ophthalmic databases, their contributions in artificial intelligence-driven and glaucoma research, and the opportunities and challenges they present for future studies.

COMPARATIVE ANALYSIS OF KEY DATABASES

In recent years, several large-scale data platforms have emerged, offering valuable resources for clinical and epidemiological research. Some are designed specifically for eye care, while others are broader health databases that also include some ophthalmic information among a wide range of specialties. In this section, we outline the key features of six prominent big data platforms commonly used in ophthalmology research: TriNetX, UK Biobank, All of Us, IRIS Registry, Cosmos, and SOURCE. We summarize their population sizes, specific relevance to ophthalmology research, and highlight their respective strengths and limitations (Table 1).

IRIS REGISTRY

The IRIS Registry was developed by the American Academy of Ophthalmology. It is the world's largest ophthalmic clinical data registry, with de-identified data from more than 83 million patients and 2800 ophthalmology practices from around the United States [4]. IRIS includes data regarding patient demographics, diagnoses, and procedures. One of IRIS Registry's main strengths is the inclusion of ocular examination findings such as visual acuity, IOP, optic nerve appearance, and fundus exam observables [5]. Over 150 original articles have been published in prominent ophthalmology journals using the IRIS Registry. One major theme has focused on clinical outcomes and treatment effectiveness [6,7]. Another theme is the exploration of demographic disparities and health inequities [8,9]. Other studies have analyzed postoperative complications and surgical outcomes [10–12].

The IRIS Registry has several important limitations to keep in mind. It lacks raw imaging data and associated quantitative metrics from modalities like visual fields and optical coherence tomography (OCT), limiting its utility for certain artificial intelligence or imaging-based analyses [13]. Furthermore, it does not capture data from all ophthalmology practices and may disproportionately represent private practices or those serving Medicare patients, as automated patient quality reporting has been a key incentive for participating locations. In addition, academic medical centers are greatly underrepresented in the database, making up only 1% of the registry [5]. Second, clinics often lack comprehensive systemic health information, which limits their ability to study how systemic factors may influence eye health. Despite these limitations, IRIS has been instrumental in numerous large-scale studies and continues to grow as more practices contribute data. Its sheer size and specialty-specific focus make it a great resource for big data research, especially for examining trends in care and outcomes over time.

SOURCE

SOURCE is a multicenter ophthalmology data repository that integrates detailed electronic health record (EHR) data from more than 20 major academic medical centers. It includes records for roughly 6 million patients, 36 million visual acuity assessments, and over 600 million eye exam findings across the participating institutions [14▪]. Each academic center shares de-identified clinical and ocular imaging data, which is then pooled together for research purposes. SOURCE contains patient demographics, ophthalmic exam findings (visual acuity, IOP), ocular and systemic diagnoses, procedures, medications, and ocular imaging (OCT, visual fields) [15]. SOURCE also contains information on social determinant health metrics, personal income, net worth, and the highest education level completed [14▪]. One of the main strengths of SOURCE is its inclusion of systemic healthcare received at SOURCE sites. This allows researchers to study ocular diseases in the context of systemic diseases and medications.

A key study utilized SOURCE data to facilitate the development of a high-precision algorithm for

Table 1. Comparative summary of major big data resources

Database name	Population size	Ocular exam findings	Ocular imaging	Genomic data	Advantages	Limitations
IRIS Registry	>80 million	Yes	No	No	Broad coverage across the whole United States, provides information on laterality, large population size	Lack of systemic health data, surgical records may be inconsistent, majority of data comes from community-based practices, with limited representation of the academic centers
SOURCE	~6 million	Yes	Yes (fundus, OCT)	No	Inclusion of imaging has enabled artificial intelligence (AI)-driven studies, includes systemic health data	Smaller size compared to other databases, limited access can lead to bias in research, only operates through Epic EHR
Epic Cosmos	>290 million	Yes	No	No	Large population size, potential to become a powerful database in ophthalmology	Lacks ocular imaging, only operates through Epic EHR
All of Us	~850 000	No	No	Yes	Special emphasis on underrepresented groups, comprehensive general health data for patients	Small population size, less generalizable, limited ophthalmic data
UK Biobank	~500 000	Yes	Yes (fundus, OCT)	Yes	Comprehensive systemic and ocular health data supported many AI-driven studies with its imaging dataset.	Self-reporting of ocular histories may lead to recall bias and misclassification, not representative of the general population
TriNetX	>250 million	No	No	No	Large population size with global contributions	Lacks ophthalmic data, lacks ocular imaging

EHR, electronic health record; IRIS, Intelligent Research in Sight; SOURCE, Sight Outcomes Research Collaborative.

identifying patients with ocular conditions in EHR data. By combining structured data (e.g. billing codes) with unstructured data from clinical notes, the researchers developed a probability-based algorithm that significantly improved diagnostic accuracy. It was developed in response to the limitations of using billing codes alone to identify study populations, which can be prone to inaccuracies and bias due to miscoding, lack of clinical context, and variability in coding practices [16]. The model created in the study achieved a 95% positive-predictive value and 100% negative-predictive value in identifying patients with exfoliation syndrome, substantially outperforming code-based identification methods [17]. As other ophthalmic databases continue to evolve, integrating similar natural language processing techniques could greatly improve cohort identification and the overall reliability of real-world evidence.

Among the limitations when it comes to SOURCE, almost all of SOURCE data comes from large academic medical centers. These tertiary centers often manage patients with more complex and/or rare diseases compared to private or community-based practices. Moreover, many patients may receive only specialized care at academic centers (such as for glaucoma or retinal disease) while obtaining routine or longitudinal eye care elsewhere. Another limitation involves access to its data. SOURCE is currently restricted to members of the consortium. This can lead to bias in research representation, as studies may disproportionately reflect the patient populations, practice patterns, and regional characteristics of participating institutions, thereby limiting their generalizability [18]. SOURCE is also limited to centers that use Epic as their EHR. Thus, many private practices, Veterans Affairs Medical Centers, and many other large institutions that do not use Epic may be excluded from future expansion [18]. Despite these limitations, SOURCE represents a powerful model for ophthalmic data sharing. By standardizing EHR and test data across centers, it enables high-resolution research that bridges clinical data and imaging.

EPIC COSMOS

Cosmos is a large EHR-based data network developed by Epic Systems that collects records from hundreds of participating healthcare organizations that use the Epic EHR platform [19]. As of 2025, Cosmos contains de-identified data from over 290 million patients from more than 1600 hospitals and 37 000 clinics across the United States. This makes Cosmos one of the largest aggregated clinical datasets in the world. It contains data from both inpatient and outpatient encounters, as well as surgical and emergency department visits, combining all of them into longitudinal patient profiles. These profiles include patient demographics, diagnoses, medications, procedures, and other structured data from routine care [19]. It has enabled broad epidemiological studies and public health surveillance for all specialties in medicine by providing real-world data on diverse patient populations [20–22].

Despite its growing role in clinical research across various medical specialties such as internal medicine and cardiology, Cosmos has not been as widely adopted in ophthalmology, partly because Kaleidoscope in Epic needs to be implemented for ophthalmology data collection. Currently, very few ophthalmic studies have been published using this database [23,24]. While these studies show that Cosmos can be used in ophthalmology research, more work is needed to make it a widely used tool in the field.

One limitation is the lack of ophthalmic-specific data elements that are often crucial for meaningful analysis. Its data previously did not include ocular exam findings such as visual acuity and IOP, although recent updates (i.e. Kaleidoscope) are working to incorporate these measures. Furthermore, free-text clinical notes, which often contain descriptions of ocular findings, are not accessible in Cosmos, further limiting ophthalmic data. Lastly, raw ophthalmic imaging is not incorporated into Cosmos, although there are plans to incorporate them in the future. This restricts the utility of Cosmos for questions that rely on disease staging, progression, or response to therapy.

Despite its limitations, Cosmos holds significant potential for future research. Its database can be helpful in studying systemic conditions that affect eye health and identifying systemic risk factors in ocular diseases. Given the sheer quantity of patients in its database and the number of institutions moving towards Epic as their main EHR, Cosmos has strong potential to become the most comprehensive real-world data source for ophthalmology research, setting a new standard for big data research.

ALL OF US RESEARCH PROGRAM

The All of Us Research Program is a longitudinal cohort study by the National Institutes of Health that aims to build a diverse health database [25]. As of 2024, All of Us has enrolled over 849 000 participants from across the country, placing special emphasis on individuals from historically underrepresented communities. A key strength is the comprehensiveness of the database. It contains surveys (sociodemographic characteristics, lifestyle, barriers to healthcare access), EHR information (diagnoses, procedures, medications, and other structured information), genomics, physical measurements, and digital health technologies (wearable health tracking devices) [25,26▪]. This richness makes All of Us a valuable resource for studying genetic associations, lifestyle factors, and a variety of health outcomes in an underrepresented cohort.

All of Us offers unique opportunities to study connections between systemic health and eye disease. Its genomic data and extensive survey information (on diet, physical activity, and other lifestyle variables) could be leveraged to explore genetic risk factors, the impact of lifestyle on ocular diseases, and socioeconomic and racial disparities in ophthalmology [27–32]. Such studies highlight the ability of All of Us to shed light on the interplay between systemic health, social factors, and vision, especially as the dataset continues to grow in diversity and depth.

A limitation in All of Us is that it was designed as a general biomedical resource rather than an eye-specific dataset, and thus it contains only the eye health information that appears in participants' routine medical records or self-reports. This means diagnoses of common eye diseases (e.g. diabetic retinopathy and cataract) may be present through billing codes, but detailed ophthalmologic exam findings or specialized tests are not routinely collected. This is illustrated by the fact that currently only around 3000 participants, out of 849 000+, have recorded IOP measurements. As a result, All of Us is less suited for in-depth ophthalmology research compared to dedicated eye databases like IRIS or SOURCE. While All of Us aims for diversity, the number of participants with specific eye diseases is often limited, reducing statistical power for certain analyses.

UK BIOBANK

The UK Biobank is a large, UK-based prospective population cohort consisting of individuals registered with the National Health Service (NHS). It consists of about 500 000 participants aged 40–69 at the time of recruitment (2006–2010) [33]. Its database includes sociodemographic information, psychosocial factors, lifestyle and health status questionnaires, biochemical assays, genetic data (whole-genome and exome sequencing), and multimodal imaging [33]. In terms of ophthalmology, it includes visual acuity, refractive error, IOP, corneal hysteresis, and self-reported eye disease history from participants. Additionally, ocular imaging was also collected. Approximately 67 000 participants had gradable color fundus photographs and spectral-domain OCT scans of the retina [34,35]. The combination of comprehensive health profiles with detailed eye exams makes the UK Biobank a uniquely powerful resource for ophthalmic research.

The UK Biobank has supported a broad range of studies. Studies have investigated associations between lifestyle factors, systemic conditions, and genetic loci with ocular conditions [36–39]. The UK Biobank has also played a key role in advancing artificial intelligence research in ophthalmology. Its large, standardized retinal image dataset has been used to train artificial intelligence models capable of predicting not only eye diseases but also systemic conditions [40]. Collectively, these findings illustrate the unique value of UK Biobank in enabling detailed investigations into how behavioral, pharmacologic, and genetic factors interact with and influence ocular health.

One limitation of UK Biobank is that the ophthalmic exams were performed on a subset of patients (those who visited certain assessment centers during specific time frames), which may introduce some selection bias. Another limitation is its reliance on self-reporting for ocular histories. This can introduce recall bias and misclassification. Lastly, given that the UK Biobank cohort study was volunteer-based and predominantly composed of middle-aged White individuals (~95%), it may not be representative of the general population [34,38]. Nonetheless, UK Biobank remains one of the most comprehensive resources for eye epidemiology, bridging ocular phenotypes with genetic and systemic variables. It exemplifies how big data can be applied to uncover links between eye health and overall health.

TriNetX

TriNetX is a federated clinical data network containing de-identified records of over 250 million patients globally. It aggregates EHR data from more than 220 healthcare organizations (hospitals, clinics, and health systems) in over 30 countries [41]. Users can access the network via a web interface to identify cohorts meeting specific criteria (diagnoses, medications, lab results, etc.) and to analyze aggregate outcomes. The data are mapped to standard ontologies and updated regularly, enabling near real-time access to clinical insights across a broad patient population [42]. It contains structured data regarding patient demographics, diagnoses, procedures, medications, and laboratory values.

In the context of ophthalmology, TriNetX offers the advantage of large-scale and real-world diversity [43▪]. Studies primarily examine associations between systemic diseases, medications, and ocular conditions, as well as trends in ophthalmic procedures [44–50]. These studies demonstrate how TriNetX can support statistically robust ophthalmic research by enabling the analysis of large, matched patient cohorts across diverse real-world populations. Another key advantage is the composition of its contributing institutions, which include both academic medical centers and large private health systems. This distinguishes TriNetX from databases like IRIS, which draws predominantly from private practices, and SOURCE, which is composed almost entirely of academic centers.

One major limitation of TriNetX is the lack of key ocular exam metrics such as visual acuity and IOP in the database, making it difficult to account for baseline disease severity and potentially introducing confounding variables [18]. The database also lacks ocular imaging, further restricting the types of research that can be conducted. Additionally, the number of patients with documented ophthalmic visits is relatively low. Despite these caveats, TriNetX provides a useful complementary approach. It enables rapid, exploratory analyses on very large patient cohorts, which can generate hypotheses or confirm broad patterns that can later be studied in more detail using specialized ophthalmic databases.

BIG DATA AND GLAUCOMA

Big data has opened new avenues in glaucoma research, one of which is examining the relationship between systemic medications and glaucoma [38,51–55]. For example, studies have investigated the association between calcium channel blockers (CCBs) use and the risk of primary open-angle glaucoma (POAG) [38,52]. Such studies may especially be relevant in African American populations, a group already at an increased risk for POAG and in whom CCBs are often a first-line or second-line treatment option for hypertension [38,56]. A separate set of studies has explored the association between glucagon-like peptide-1 receptor agonists (GLP-1RAs) and glaucoma, given their growing clinical use in promoting weight loss [53,54]. These findings support a potential protective role for GLP-1RAs in glaucoma prevention. Studies like these are generated from real-world data encompassing millions of patients, opening a new line of inquiry into glaucoma prevention. They likely would not have been detected in smaller clinical trials, as glaucoma develops slowly and would have required enormous sample sizes to study prospectively.

Apart from pharmacological associations in glaucoma, one of the most transformative uses of big data in glaucoma has been the development of artificial intelligence tools using multimodal imaging [57▪]. These tools include artificial intelligence-driven visual field analysis, OCT-based diagnosis and disease monitoring, and glaucoma detection and risk of progression [58▪,59,60]. Automating glaucoma detection through the integration of imaging and clinical data offers a promising, cost-effective approach for population-level screening.

Another key advantage of big data in glaucoma research is its ability to reconcile conflicting findings in the literature by enabling large-scale, granular analyses that account for population heterogeneity and clinical complexity. The relationship between statin use and glaucoma provides a representative example. Stein *et al.* initially reported that long-term statin use was associated with a reduced risk of developing open-angle glaucoma [61]. In contrast, Kang *et al.* [62] found no significant association, and Lee *et al.* observed a higher prevalence of glaucoma among statin users, especially in adults aged 60–69 with hyperlipidemia [63]. These conflicting results have contributed to clinical uncertainty, likely driven by differences in study populations, duration of statin exposure, and unmeasured confounding variables.

Larger and more diverse datasets can help address these challenges. In our recent analysis using the TriNetX database, we examined the statin–glaucoma association across racial and ethnic subgroups. We found that statin use was associated with a lower risk of ocular hypertension and primary open-angle glaucoma in non-Hispanic White and Black patients, but not in Asian or Hispanic individuals [55]. This finding emphasizes the importance of considering population-specific effects when interpreting research outcomes and harmonizes the results of the previous statin use papers on glaucoma patients and risk.

By leveraging large datasets with robust matching techniques and careful adjustment for confounders, big data enables more precise phenotyping, improved bias control, and detection of subtle or context-dependent associations. Ultimately, this approach yields more reliable and generalizable evidence to guide clinical practice.

Overall, these studies illustrate how big data can uncover unexpected links between systemic health and eye disease. By analyzing large-scale longitudinal datasets, researchers have identified new associations, potential risk factors, and opportunities for therapeutic advancements in glaucoma. However, while these associations are promising, randomized clinical trials are needed to establish causality, as observational data alone may not be sufficient to confirm definitive relationships.

LIMITATIONS IN BIG DATA

Big data research in ophthalmology faces several core limitations. First, there is a lack of standardized data formats across different databases and clinical systems. This heterogeneity makes it difficult to combine or compare data from multiple sources. Another challenge in big data is ethical and privacy issues. Most big data used in ophthalmology research comes from clinical care (EHRs, insurance claims), where patients did not explicitly consent to have their data included in research databases. While the use of de-identified data is generally permitted under ethical guidelines, it remains important to maintain public trust. There is also the risk of bias and inequity. Big datasets may inadvertently reflect systemic biases in healthcare access or treatment. For instance, if a registry underrepresents certain minority populations or those without insurance (who may not access specialty care as often), studies using that data produce biased conclusions that may not be generalized to certain groups. This can further increase health disparities. Lastly, most real-world databases rely on billing codes such as ICD-9/ICD-10 and CPT codes to identify diagnoses and procedures, so the accuracy of study findings depends heavily on the accuracy of billing code documentation. Tools such as the SOURCE model, which integrates structured billing codes with unstructured clinical text, have shown improved diagnostic accuracy and may help overcome this limitation in future research applications [17].

CONCLUSION

Big data has immense potential to transform ophthalmology research and ultimately improve patient outcomes. Over the past decade, the ophthalmic community has assembled several major datasets, including TriNetX, Cosmos, All of Us, IRIS Registry, SOURCE, and UK Biobank. These datasets collectively cover hundreds of millions of patients and a vast array of clinical information. They have already enabled important discoveries, such as identifying systemic risk factors for glaucoma and developing artificial intelligence algorithms for early disease detection. Comparative analysis of the databases reveals that each has its niche: some offer depth in clinical ophthalmic data, while others provide ophthalmic imaging. The most impactful research often arises when multiple data sources are used in a complementary fashion.

Acknowledgements

None.

Financial support and sponsorship

None.

Conflicts of interest

There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

■ of special interest

■■ of outstanding interest

1. Murdoch TB, Detsky AS. The inevitable application of big data to healthcare. JAMA 2013; 309:1351–1352.
2. Knapp AN, Leng T, Rahimy E. Ophthalmology at the forefront of big data integration in medicine: insights from the IRIS Registry Database. Yale J Biol Med 2023; 96:421–426.
3. Soh ZD, Cheng CY. Application of big data in ophthalmology. Taiwan J Ophthalmol 2023; 13:123–132.
4. Parke DW 2nd, Rich WL 3rd, Sommer A, Lum F. The American Academy of Ophthalmology's IRIS((R)) Registry (Intelligent Research in Sight Clinical Data): a look back and a look to the future. Ophthalmology 2017; 124: 1572–1574.
5. Chiang MF, Sommer A, Rich WL, *et al.* The 2016 American Academy of Ophthalmology IRIS((R)) Registry (Intelligent Research in Sight) Database: characteristics and methods. Ophthalmology 2018; 125: 1143–1148.
6. Lidder AK, Vanner EA, Chang TC, *et al.* Intraocular pressure spike following stand-alone phacoemulsification in the IRIS(R) Registry (Intelligent Research in Sight). Ophthalmology 2024; 131:780–789.
7. Zarbin MA, MacCumber MW, Karcher H, *et al.* Real-world safety outcomes with brolucizumab in neovascular age-related macular degeneration: findings from the IRIS(R) Registry. Ophthalmol Ther 2024; 13:1357–1368.
8. Ciociola EC, Sekimitsu S, Smith S, *et al.* Racial disparities in glaucoma vision outcomes and eye care utilization: an IRIS Registry Analysis. Am J Ophthalmol 2024; 264:194–204.
9. Maturi J, Maturi V, Scott AW, *et al.* Effect of race and insurance status on treatment and outcomes in diabetic retinopathy: analysis of 43 274 eyes using the IRIS Registry. J Vitreoretin Dis 2024; 8:270–279.
10. Morano MJ, Khan MA, Zhang Q, *et al.* Incidence and risk factors for retinal detachment and retinal tear after cataract surgery: IRIS(R) Registry (Intelligent Research in Sight) Analysis. Ophthalmol Sci 2023; 3:100314.
11. Zhang LJ, Dana R, Lorch AC, *et al.* Complications after pediatric penetrating keratoplasty: an IRIS Registry Study. Cornea 2024; 43:1555–1559.
12. Oke I, Elze T, Miller JW, *et al.* The incidence of strabismus after upper and lower blepharoplasty in the United States. Ophthalmic Plast Reconstr Surg 2024; 40:449–452.
13. Pershing S, Lum F. The American Academy of Ophthalmology IRIS Registry (Intelligent Research In Sight): current and future state of big data analytics. Curr Opin Ophthalmol 2022; 33:394–398.
14. Bernstein IA, Fernandez KS, Stein JD, *et al.* Big data and electronic health ■ records for glaucoma research. Taiwan J Ophthalmol 2024; 14:352–359.

Highlights the potential of big data to advance glaucoma research, while addressing key challenges in data quality, standardization, and access.

15. (SOURCE) SORC. SOURCE Collaborative 2025. Available at: https://www.sourcecollaborative.org/. [Accessed 1 April 2025].
16. Juniat V, Athwal S, Khandwala M. Clinical coding and data quality in oculoplastic procedures. Eye (Lond) 2019; 33:1733–1740.
17. Stein JD, Rahman M, Andrews C, *et al.* Evaluation of an algorithm for identifying ocular conditions in electronic health record data. JAMA Ophthalmol 2019; 137:491–497.
18. Chaudhury AS, Ige M, Marwah S, *et al.* Sight Outcomes Research Collaborative (SOURCE) Consortium. Race, social determinants of health, and the quality of diabetic eye care. JAMA Ophthalmol 2024; 142:961–970.
19. Tarabichi Y, Frees A, Honeywell S, *et al.* The Cosmos Collaborative: a vendor-facilitated electronic health record data aggregation platform. ACI open 2021; 5:e36–e46.
20. Deputy NP, Deckert J, Chard AN, *et al.* Vaccine Effectiveness of JYNNEOS against Mpox Disease in the United States. N Engl J Med 2023; 388: 2434–2443.
21. Gottlieb M, Moyer E, Bernard K. Epidemiology of pulmonary embolism diagnosis and management among United States emergency departments over an eight-year period. Am J Emerg Med 2024; 85:158–162.
22. Gibson KS, Olson D, Lindberg W, *et al.* Postpartum blood pressure control and the rate of readmission. Am J Obstet Gynecol MFM 2024; 6:101384.
23. Varghese JS, Ravi Kumar V, Bartelt J, *et al.* The role of urban residence, race and ethnicity, and glycemic control in receiving standards of care and progression to vision-threatening diabetic retinopathy. Diabetes Care 2025; 48:29–37.
24. Swaminathan SS, Medeiros FA. Socioeconomic disparities in glaucoma severity at initial diagnosis: a nationwide electronic health record cohort analysis. Am J Ophthalmol 2024; 263:50–60.
25. Denny JC, Rutter JL, *et al.* All of Us Research Program. The "All of Us" Research Program. N Engl J Med 2019; 381:668–676.
26. All of Us Research Program Genomics I. Genomic data in the All of Us ■ Research Program. Nature 2024; 627:340–346.

Reports the release of over 245 000 whole-genome sequences from the All of Us program, highlighting its diverse cohort and how researchers can access linked genomic and health data through a user-friendly, cloud-based platform.

27. Delavar A, Radha Saseendrakumar B, Lee TC, *et al.* Associations between thyroid eye disease and glaucoma among those enrolled in the National Institutes of Health All of Us Research Program. Ophthalmic Plast Reconstr Surg 2023; 39:336–340.
28. Yoo K, Lee C, Baxter SL, Xu BY. Relationship between glaucoma and chronic stress quantified by Allostatic Load Score in the All of Us Research Program. Am J Ophthalmol 2025; 269:419–428.
29. McDermott JJt, Lee TC, Chan AX, *et al.* Novel association between opioid use and increased risk of retinal vein occlusion using the National Institutes of Health All of Us Research Program. Ophthalmol Sci 2022; 2:100099.
30. Zhao AT, He J, Lei Y, *et al.* Associations between dry eye disease and mental health conditions in the All of Us Research Program. Am J Ophthalmol 2025; 270:61–66.
31. Chen JS, Esko JD, Walker E, *et al.* High-density lipoproteins associate with age-related macular degeneration in the All of Us Research Program. Ophthalmology 2025; 132:684–691.
32. Delavar A, Radha Saseendrakumar B, Weinreb RN, Baxter SL. Racial and ethnic disparities in cost-related barriers to medication adherence among patients with glaucoma enrolled in the National Institutes of Health All of Us Research Program. JAMA Ophthalmol 2022; 140:354–361.
33. Sudlow C, Gallacher J, Allen N, *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015; 12:e1001779.
34. Chua SYL, Thomas D, Allen N, *et al.* UK Biobank Eye & Vision Consortium. Cohort profile: design and methods in the eye and vision consortium of UK Biobank. BMJ Open 2019; 9:e025077.
35. Warwick AN, Curran K, Hamill B, *et al.* UKBB Eye and Vision Consortium. UK Biobank retinal imaging grading: methodology, baseline characteristics and findings for common ocular diseases. Eye (Lond) 2023; 37:2109–2116.
36. Kim J, Aschard H, Kang JH, *et al.* Modifiable Risk Factors for Glaucoma Collaboration. Intraocular pressure, glaucoma, and dietary caffeine consumption: a gene-diet interaction Study from the UK Biobank. Ophthalmology 2021; 128:866–876.
37. Madjedi KM, Stuart KV, Chua SYL, *et al.* Modifiable Risk Factors for Glaucoma Collaboration and the UK Biobank Eye and Vision Consortium. The Association of Physical Activity with Glaucoma and Related Traits in the UK Biobank. Ophthalmology 2023; 130:1024–1036.
38. Kastner A, Stuart KV, Montesano G, *et al.* Calcium channel blocker use and associated glaucoma and related traits among UK Biobank Participants. JAMA Ophthalmol 2023; 141:956–964.
39. Khawaja AP, Cooke Bailey JN, Wareham NJ, *et al.* UK Biobank Eye and Vision Consortium, NEIGHBORHOOD Consortium. Genome-wide analyses identify 68 new loci associated with intraocular pressure and improve risk prediction for primary open-angle glaucoma. Nat Genet 2018; 50:778–782.
40. Poplin R, Varadarajan AV, Blumer K, *et al.* Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nat Biomed Eng 2018; 2:158–164.
41. Palchuk MB, London JW, Perez-Rey D, *et al.* A global federated real-world data and analytics platform for research. JAMIA Open 2023; 6:ooad035.
42. Topaloglu U, Palchuk MB. Using a federated network of real-world data to optimize clinical trials operations. JCO Clin Cancer Inform 2018; 2:1–10.
43. Hackl CM, Lee VA, Dunnigan JK, *et al.* A narrative review of TriNetX in ■ ophthalmology: applications, advantages, and pitfalls. Ann Eye Sci 2024; 9.

Provides a narrative review of TriNetX use in ophthalmology, detailing its strengths in large-scale, real-world data analysis and limitations due to lack of ophthalmic exam findings and imaging.

44. Chauhan MZ, Elhusseiny AM, Kishor KS, *et al.* Association of primary open-angle glaucoma with diabetic retinopathy among patients with type 1 and type

2 diabetes: a large global database study. Ophthalmology 2024; 131: 827–835.
45. Mammo DA, Wai K, Rahimy E, *et al.* Association of cutaneous keloids, hypertrophic scarring, and fibrosis with risk of postoperative proliferative vitreoretinopathy. Ophthalmology 2024; 131:961–966.
46. Wai KM, Mishra K, Koo E, *et al.* Impact of GLP-1 agonists and SGLT-2 inhibitors on diabetic retinopathy progression: an aggregated electronic health record data study. Am J Ophthalmol 2024; 265:39–47.
47. Kuo HT, Chen CY, Hsu AY, *et al.* Association between immune checkpoint inhibitor medication and uveitis: a population-based cohort study utilizing TriNetX database. Front Immunol 2023; 14:1302293.
48. Wai KM, Ludwig CA, Koo E, *et al.* Risk of stroke, myocardial infarction, deep vein thrombosis, pulmonary embolism, and death after retinal vein occlusion. Am J Ophthalmol 2024; 257:129–136.
49. Chauhan MZ, Eleiwa TK, Abdelnaem S, *et al.* Risk of abducens nerve palsy following COVID-19 vaccination. J AAPOS 2024; 28:103867.
50. Elhusseiny AM, Chauhan MZ, Jabbehdari S, *et al.* Using real-world data to assess the association of retinal detachment with topical pilocarpine use. Am J Ophthalmol 2025; 271:1–6.
51. Chu YY, Lee CY, Huang WY, *et al.* Association of atopic dermatitis and risk of glaucoma surgery: a global population-based study. J Glaucoma 2024; 33: 735–741.
52. Lehrer S, Rheinstein PH. Amlodipine increases risk of primary open-angle glaucoma. Clin Hypertens 2024; 30:33.
53. Vasu P, Dorairaj EA, Weinreb RN, *et al.* Risk of glaucoma in patients without diabetes using a glucagon-like peptide 1 receptor agonist. Ophthalmology 2025; 132:859–868.
54. Muayad J, Loya A, Hussain ZS, *et al.* Comparative effects of glucagon-like peptide 1 receptor agonists and metformin on glaucoma risk in patients with type 2 diabetes. Ophthalmology 2025; 132:271–279.
55. Elhusseiny AM, Eleiwa TK, Dihan QA, *et al.* Racial and ethnic differences in the association between statin use and the risk of ocular hypertension and open-angle glaucoma. Am J Ophthalmol 2025; 276:117–125.
56. Siegfried CJ, Shui YB. Racial disparities in glaucoma: from epidemiology to pathophysiology. Mo Med 2022; 119:49–54.
57. ■ Wu JH, Lin S, Moghimi S. Big data to guide glaucoma treatment. Taiwan J Ophthalmol 2024; 14:333–339.
Summarizes key studies using big data and machine-learning approaches to improve risk assessment, predict outcomes, and personalize glaucoma treatment strategies.
58. ■ Pham AT, Pan AA, Yohannan J. Big data in visual field testing for glaucoma. Taiwan J Ophthalmol 2024; 14:289–298.
Provides a comprehensive review of how large-scale visual field datasets are leveraged to improve glaucoma diagnosis, monitor progression, optimize clinical decision-making, machine learning, real-world analytics, and multimodal data integration.
59. Mehta P, Petersen CA, Wen JC, *et al.* Automated detection of glaucoma with interpretable machine learning using clinical data and multimodal retinal images. Am J Ophthalmol 2021; 231:154–169.
60. Chaurasia AK, Greatbatch CJ, Han X, *et al.* Highly accurate and precise automated cup-to-disc ratio quantification for glaucoma screening. Ophthalmol Sci 2024; 4:100540.
61. Stein JD, Newman-Casey PA, Talwar N, *et al.* The relationship between statin use and open-angle glaucoma. Ophthalmology 2012; 119:2074–2081.
62. Kang JH, Boumenna T, Stein JD, *et al.* Association of statin use and high serum cholesterol levels with risk of primary open-angle glaucoma. JAMA Ophthalmol 2019; 137:756–765.
63. Lee SY, Paul ME, Coleman AL, *et al.* Associations between statin use and glaucoma in the All of Us Research Program. Ophthalmol Glaucoma 2024; 7: 563–571.