RESEARCH ARTICLE

Toward Precision Oculoplastics: Federated Data, Artificial Intelligence, and the Path to Equitable Therapeutic Discovery

Caroline Duncan, MS [1]; Riyona D'Silva, PharmD [1]; Vinod Gauba, MD [1]

[1] GeneVault Lifesciences Inc.

ABSTRACT

Oculoplastic and orbital disorders are rapidly entering a new era of precision medicine, driven by advances in genomics, multi-omic profiling, and artificial intelligence (AI). However, the evidence base supporting these innovations remains overwhelmingly Eurocentric: more than 80% of all genome-wide association study (GWAS) participants globally are of European ancestry, and pivotal trials of teprotumumab, the first targeted biologic for thyroid eye disease (TED), enrolled over 85% European-descended participants. This imbalance constrains understanding of ancestry-specific therapeutic responses, limits the transferability of polygenic and pharmacogenomic models, and risks perpetuating structural bias in emerging AI systems used for diagnostic and surgical decision-making. The high cost of teprotumumab (> USD 200,000 per treatment course) further underscores the need for equitable, predictive stratification. This article outlines how federated learning (FL) and harmonized multi-omic data infrastructures can close these gaps. We highlight the EXAM study as a real-world demonstration of FL's ability to integrate multi-institutional data without transferring sensitive patient information, achieving superior predictive performance while preserving privacy. Building on these principles, we introduce GeneVault Harmony, a federated harmonization and benchmarking framework that integrates genomic, transcriptomic, imaging, and clinical data while maintaining data sovereignty and aligning with GA4GH and WHO standards. By enabling bias quantification, diversity simulation, and cross-institutional interoperability, Harmony provides a scalable foundation for equitable oculoplastic genomics. We conclude by proposing a pathway toward inclusive precision therapeutics in oculoplastics, including multi-regional genomic consortia and federated, multi-omic AI pipelines spanning the Middle East, Africa, South Asia, and Latin America. Integrating globally representative data into future genomic and pharmacogenomic discovery is essential to ensure that precision ophthalmology evolves in a manner that is scientifically robust, clinically equitable, and globally relevant.

Potential Keywords: Oculoplastics • Genomic equity • Federated data • Privacy-preserving AI • Data sovereignty • Teprotumumab • Pharmacogenomics • Precision medicine • Multi - omics • Harmony

1. Introduction

Oculoplastic surgery, the subspecialty of ophthalmology concerned with the reconstruction, oncology, and aesthetic management of the eyelids, orbit, lacrimal system, and periocular region, sitting at the intersection of ophthalmology, oncology, and reconstructive medicine, is undergoing a profound transformation driven by advances in precision medicine and artificial intelligence (AI).[1] The growing integration of genomic and molecular insights into surgical and therapeutic decision-making is redefining how orbital, eyelid, and periocular diseases are diagnosed and treated.

Yet, the promise of these innovations remains unequally distributed. More than 80% of genome-wide association study (GWAS) participants worldwide are of European descent, leaving vast populations across the Middle East, Africa, South Asia, and Latin America underrepresented.[2] This imbalance not only limits the generalizability of genomic discoveries but also risks embedding ancestry bias into emerging AI models that increasingly guide diagnostic algorithms, surgical planning, and therapeutic targeting in oculoplastic care.[3]

These inequities are exemplified by the case of thyroid eye disease (TED), a leading autoimmune orbital disorder. The introduction of teprotumumab, an insulin-like growth factor 1 receptor (IGF-1R) antagonist, represents a landmark advance, offering the first targeted biologic for TED and dramatically altering disease management paradigms. However, most clinical and pharmacogenomic studies of teprotumumab have been conducted in predominantly white European populations (>85% of trial participants), leaving potential ancestry-specific differences in therapeutic response and toxicity largely unexplored.[3] Given the therapeutic cost (>200,000 USD) and variable efficacy, integrating pharmacogenomic and multi-omic data into patient selection could not only enhance clinical precision but also reduce inequities in access and outcomes.

Achieving this requires harmonized, federated approaches to data sharing, such as emerging frameworks like GeneVault Harmony, that enable secure, privacy-preserving integration of genomic, transcriptomic, and imaging data across diverse populations while maintaining data sovereignty and ethical governance.

2. Genomic Insights in Ophthalmology and Oculoplastics: Current Landscape

2.1. REVIEW OF MAJOR OPHTHALMIC GWAS:

Over the past decade, ophthalmic genomics has transformed the understanding of complex eye disease. Across ophthalmic phenotypes, heritability estimates span a wide range, indicating strong but variable genetic influence. In a comprehensive meta-analysis, the lowest heritability was observed for diabetic retinopathy ($h^2 \approx$ 0.18–0.27), while the highest heritability reached 0.99 for eyes with $\geq$ 20 small hard drusen in age-related macular degeneration, reflecting near-complete genetic determination of that phenotype.[4] Large-scale GWAS have uncovered hundreds of loci for age-related macular degeneration (AMD), glaucoma, keratoconus, myopia, and diabetic retinopathy, revealing critical pathways in complement activation, extracellular-matrix remodelling, angiogenesis, and neurodegeneration.[5,6] These discoveries have accelerated functional annotation, polygenic risk prediction, and the development of targeted interventions such as complement inhibitors for AMD and gene therapy for inherited retinal disorders. Despite racial and ethnic disparities in oculoplastic disorders, progress remains overwhelmingly Eurocentric, where the majority of ophthalmic studies are performed in European-based cohorts such as the UK Biobank and FinnGen.[7–9] When studies are replicated in diverse populations, discovery is boosted, and novel loci are then discoverable.[10]

This representation gap creates scientific and clinical blind spots. AI models trained on homogeneous, European-only genomic and imaging data risk encoding ancestry bias, while pharmacogenomic algorithms for biologics such as teprotumumab cannot accurately predict efficacy or toxicity in diverse patients.[11] Inclusive, federated data infrastructures that enable trans-ancestry fine mapping and secure integration of genomic, transcriptomic, and imaging data across regions are therefore essential to achieve equitable precision therapeutics.

Oculoplastic-related disorders such as Graves' orbitopathy, congenital craniofacial syndromes, and lacrimal gland tumours remain among the most genetically underexplored conditions in ophthalmology. While large-scale genomic studies have advanced understanding of AMD, glaucoma, and myopia, orbital and periocular diseases are largely absent from global GWAS catalogues. Most published studies involve small, European-dominant cohorts, with minimal inclusion from Africa, South Asia, or the Middle East. This limits identification of ancestry specific loci, modifier variants, and gene–environment interactions that could inform reconstructive, oncologic, and autoimmune oculoplastic care.[10,12,13]

2.2. MIDDLE EASTERN, AFRICAN, AND ASIAN GENOMIC DIVERSITY ARE UNDEREXPLORED IN ORBITAL DISEASE GENETICS.

Genomic diversity across the Middle East, Africa, and Asia remains vastly underexplored in the context of orbital and oculoplastic disease genetics. While regional initiatives such as the Qatar Genome Program, H3Africa, and the Japan Biobank have revealed millions of population-specific variants, a large gap in intelligence remains.[14–16] This underrepresentation limits understanding of ancestry-linked risk loci, modifier variants, and gene-environment interactions that may influence disease susceptibility or surgical outcomes. Integrating genomic data from these regions is therefore essential to capture global diversity, improve transferability of genetic discoveries, and ensure equitable translation of precision-medicine advances in oculoplastic and orbital disease care.

Despite rapid advances in ophthalmic and oculoplastic genomics, a substantial translational gap persists between discovery and clinical implementation in underrepresented populations. The overwhelming

dominance of European-ancestry cohorts in GWAS restricts the portability of findings, with polygenic risk scores (PRS) derived from European data performing up to 40% less accurately in African, South Asian, and Middle Eastern individuals.[17,18]

3. Precision Therapeutics in Oculoplastics: The Case of Teprotumumab

Teprotumumab (Tepezza), an insulin-like growth factor 1 receptor (IGF-1R) inhibitor, is the first FDA-approved biologic therapy for thyroid eye disease (TED) and represents a major milestone in precision oculoplastic medicine.[19] By blocking IGF-1R signalling and cross-talk with the thyroid-stimulating hormone receptor, teprotumumab reduces orbital fibroblast activation and inflammatory tissue expansion, resulting in marked improvements in proptosis and diplopia. The pivotal OPTIC trials demonstrated significant clinical benefit, with 83% of patients achieving a $\geq$ 2 mm reduction in proptosis and 78% showing overall clinical response.[20] Despite its transformative potential, teprotumumab's clinical evidence base remains demographically narrow, with over 85% of trial participants being of European ancestry, limiting insight into ancestry-specific pharmacogenomic or immunologic determinants of treatment response. Potential variability in IGF-1R expression, downstream signalling, and immune profiles across ancestries has not been systematically characterized, raising concern that efficacy and adverse-event rates may differ in underrepresented populations.[21] Moreover, the high treatment cost (> USD 200,000 per course) underscores the importance of pharmacogenomic and multi-omic stratification to identify likely responders and minimize non-beneficial exposure.[22] Integrating federated AI and multi-omic

frameworks by linking genomic, transcriptomic, proteomic, and digital-pathology data across international sites could enable discovery of predictive biomarkers of response and adverse-event risk while maintaining patient privacy and data sovereignty. [23,24]

4. The Role of Federated Data Infrastructures in Equitable Oculoplastic Research

Federated learning (FL) enables multi-institutional collaboration across the globe without transferring sensitive data, creating a foundation for equitable discovery in ophthalmology and oculoplastics. In contrast to centralized repositories, each institution trains machine learning models locally on its genomic, imaging, or clinical data and shares only the learned parameters, as demonstrated in Figure 1. This privacy-preserving architecture maintains data sovereignty and regulatory compliance with HIPAA, GDPR, and emerging AI governance frameworks, allowing partners in regions such as the Middle East, Africa, and South Asia to directly participate in discovery while retaining control over local data. A compelling technical validation of FL's power is the EXAM Study, a collaboration among 20 hospitals across five continents, which trained a federated deep learning model to predict COVID-19 outcomes from chest CTs and clinical variables. The model achieved an AUC of 0.94, outperforming any single-site model while maintaining full privacy compliance.[25] Similar federated frameworks are now being adapted for multi-omic integration, where distributed learning mitigates ancestry bias, increases statistical power in rare-variant discovery, and enhances AI generalizability across diverse populations.[26]
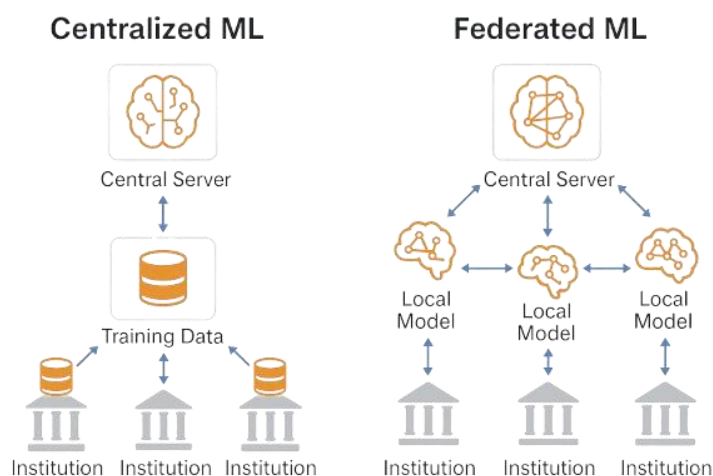

Figure 1: Centralized versus Federated Machine Learning

GeneVault Harmony builds on these principles, serving as a federated harmonization and benchmarking framework that operationalizes data equity within global precision medicine ecosystems. Developed by GeneVault Lifesciences Inc., Harmony integrates multi-omic, transcriptomic, and clinical imaging data through a privacy-preserving metadata layer, leveraging standards such as GA4GH Beacon v2, OMOP CDM, and HL7 FHIR. The platform's core modules — Harmonization Readiness Scoring, Diversity Gap Quantification, and a

Global Diversity Simulation Engine — quantify how inclusion of underrepresented ancestries improves variant discovery, pharmacogenomic prediction, and AI model robustness. Harmony's architecture employs secure connectors, ontology-matching AI, and generative modelling to map and correct representational bias, while maintaining compliance with ISO 27001 and cross-jurisdictional privacy frameworks. By benchmarking diversity and interoperability across federated networks, Harmony provides a quantifiable, auditable pathway to

measure and improve equity in data assets, empowering ophthalmic and oculoplastic consortia to implement trustworthy, globally inclusive precision medicine.

5. Integrating Multi-Omic and AI-Assisted Discovery in Oculoplastic Oncology

Precision oncology principles are increasingly applicable to orbital and adnexal tumours: in ocular adnexal sebaceous carcinoma, next-generation sequencing consistently implicates TP53 and RB1 as common drivers, highlighting biologic heterogeneity relevant to surveillance and therapy.[27] Integrating somatic and germline variation with imaging/radiomics and digital pathology improves prognostic modelling for head and neck and ocular analogues, enabling more accurate prediction of recurrence and treatment outcomes than single-modality approaches; multiple studies and reviews document gains when radiomic features are combined with clinical or molecular data.[28]

Because most centres hold small, regional cohorts, federated learning can pool signals across sites without moving data, mitigating site bias and revealing subtle patterns, such as variant clusters or immune signatures, that are invisible locally; large multi-institution studies and methods reviews show federated models improve generalization while preserving privacy. We suggest pilot federated consortia linking academic hubs across the Gulf Cooperation Council (GCC), Sub-Saharan Africa, and Europe, using GA4GH aligned discovery and authorization layers to maintain data sovereignty while enabling multi-omic, imaging-aware models for sebaceous gland carcinoma and lacrimal-gland ACC; this design follows best practice guidance from FL deployments and GA4GH governance for cross border genomics.[29]

6. Ethical, Economic, and Societal Implications

Advancing global genomics requires data equity, ensuring that underrepresented populations share fairly in the benefits of research through access to results, return of value, and governance roles. The GA4GH Framework for Responsible Sharing of Genomic and Health-Related Data likewise articulates benefit-sharing, proportionality, and accountability as global norms for cross-border data use. Inclusive AI trained and validated on diverse, ancestry rich datasets can reduce health expenditure waste by identifying likely non-responders and optimizing use of high cost therapies; health economic evaluations show that predictive biomarkers improve cost-effectiveness when they avert ineffective treatment courses. Federated and privacy preserving approaches should be aligned with GA4GH standards (e.g., DUO, DRS, AAI, Beacon/Data Connect) to ensure interoperable consent, authorization, and data exchange across jurisdictions, and GA4GH explicitly encourages federated models when data cannot legally move.[23] In parallel, the WHO has issued global principles for ethical human genomic data collection, access, use, and sharing, emphasizing equity, transparency, and public trust as prerequisites for responsible international genomics.[30]

7. Conclusion

The future of oculoplastic surgery and orbital disease management lies in precision medicine guided by diverse genomic representation. By moving beyond Eurocentric discovery toward inclusion of Middle Eastern, African, Latin American, and Asian ancestries, the field can uncover ancestry-specific genetic mechanisms that inform risk prediction, pharmacogenomics, and therapeutic innovation.[31] Federated harmonization frameworks, such as Harmony, exemplify how innovation, equity, and sovereignty can coexist allowing multi-institutional analyses that preserve local data control while enabling global discovery through privacy preserving AI.[23,32] These approaches are essential to ethically and effectively integrate multi-omic and imaging data across populations. To close the diversity gap, we call for multi-regional genomic consortia in ophthalmology and oculoplastics, linking clinical imaging, histopathology, and omics datasets under shared governance. By embracing federated collaboration and inclusive data representation, oculoplastic research can transition toward truly personalized, globally equitable therapeutics.

References

1. Mehta S, Belliveau MJ, Oestreicher JH. Oculoplastic surgery. *Clin Plast Surg*. 2013;40(4):631-651. doi:10.1016/j.cps.2013.08.005

2. Fatumo S, Chikowore T, Choudhury A, Ayub M, Martin AR, Kuchenbaecker K. A roadmap to increase diversity in genomic studies. *Nat Med*. 2022;28(2):243-250. doi:10.1038/s41591-021-01672-4

3. Smith TJ, Kahaly GJ, Ezra DG, et al. Teprotumumab for Thyroid-Associated Ophthalmopathy. *N Engl J Med*. 2017;376(18):1748-1761. doi:10.1056/NEJMoa1614949

4. Sanfilippo PG, Hewitt AW, Hammond CJ, Mackey DA. The Heritability of Ocular Traits. *Surv Ophthalmol*. 2010;55(6):561-583. doi:10.1016/j.survophthal.2010.07.003

5. Han X, Gharahkhani P, Hamel AR, et al. Large-scale multitrait genome-wide association analyses identify hundreds of glaucoma risk loci. *Nat Genet*. 2023;55(7):1116-1125. doi:10.1038/s41588-023-01428-5

6. Xue Z, Yuan J, Chen F, et al. Genome-wide association meta-analysis of 88,250 individuals highlights pleiotropic mechanisms of five ocular diseases in UK Biobank. *eBioMedicine*. 2022;82:104161. doi:10.1016/j.ebiom.2022.104161

7. 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(10):1438-1447. doi:10.1016/j.ophtha.2021.03.010

8. Emanuele N, Sacks J, Klein R, et al. Ethnicity, Race, and Baseline Retinopathy Correlates in the Veterans Affairs Diabetes Trial. *Diabetes Care*. 2005;28(8):1954-1958. doi:10.2337/diacare.28.8.1954

9. Stein JD, Kim DS, Niziol LM, et al. Differences in Rates of Glaucoma among Asian Americans and Other Racial Groups, and among Various Asian Ethnic Groups. *Ophthalmology*. 2011;118(6):1031-1037. doi:10.1016/j.ophtha.2010.10.024

10. Han X, Gharahkhani P, Hamel AR, et al. Large-scale multitrait genome-wide association analyses identify hundreds of glaucoma risk loci. *Nat Genet*. 2023;55(7):1116-1125. doi:10.1038/s41588-023-01428-5

11. Demographic distributions of teprotumumab prescription for thyroid eye disease | IOVS | ARVO Journals. Accessed October 21, 2025. https://iovs.arvojournals.org/article.aspx?articleid=2804452

12. A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants | Nature Genetics. Accessed October 21, 2025. https://www.nature.com/articles/ng.3448

13. New insights into craniofacial malformations - PMC. Accessed October 21, 2025. https://pmc.ncbi.nlm.nih.gov/articles/PMC4571997/

14. Elfatih A, Saad C, Mifsud B, Mbarek H. Analysis of 14,392 whole genomes reveals 3.5% of Qataris carry medically actionable variants. *Eur J Hum Genet*. 2024;32(11):1465-1473. doi:10.1038/s41431-024-01656-1

15. Nagai A, Hirata M, Kamatani Y, et al. Overview of the BioBank Japan Project: Study design and profile. *J Epidemiol*. 2017;27(3, Supplement):S2-S8. doi:10.1016/j.je.2016.12.005

16. The H3Africa Consortium, Matovu E, Bucheton B, et al. Enabling the genomic revolution in Africa. *Science*. 2014;344(6190):1346-1348. doi:10.1126/science.1251546

17. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 2019;51(4):584-591. doi:10.1038/s41588-019-0379-x

18. Duncan L, Shen H, Gelaye B, et al. Analysis of polygenic risk score usage and performance in diverse human populations. *Nat Commun*. 2019;10(1):3328. doi:10.1038/s41467-019-11112-0

19. Men CJ, Amarikwa L, Pham B, et al. Teprotumumab for the Treatment of Recalcitrant Thyroid Eye Disease. *Ophthal Plast Reconstr Surg*. 2024;40(3):276-285. doi:10.1097/IOP.0000000000002564

20. Teprotumumab for Thyroid-Associated Ophthalmopathy | New England Journal of Medicine. Accessed October 21, 2025. https://www.nejm.org/doi/10.1056/NEJMoa1614949

21. Zong Y, Qiu S, Yang M, et al. Teprotumumab for Thyroid Eye Disease: Mechanism, Clinical Efficacy, and Current Challenges. *Antibodies Basel Switz*. 2025;14(3):55. doi:10.3390/antib14030055

22. Tamhankar MA, Raza S, Brutsaert E, et al. The burden of illness in thyroid eye disease: current state of the evidence. *Front Ophthalmol*. 2025;5:1565762. doi:10.3389/fopht.2025.1565762

23. Rehm HL, Page AJH, Smith L, et al. GA4GH: International policies and standards for data sharing across genomic research and healthcare. *Cell Genomics*. 2021;1(2):100029. doi:10.1016/j.xgen.2021.100029

24. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring Fairness in Machine Learning to Advance Health Equity. *Ann Intern Med*. 2018;169(12):866-872. doi:10.7326/M18-1990

25. Dayan I, Roth HR, Zhong A, et al. Federated learning for predicting clinical outcomes in patients with COVID-19. *Nat Med*. 2021;27(10):1735-1743. doi:10.1038/s41591-021-01506-3

26. Drainakis G, Katsaros KV, Pantazopoulos P, Sourlas V, Amditis A. Federated vs. Centralized Machine Learning under Privacy-elastic Users: A Comparative Analysis. In: *2020 IEEE 19th International Symposium on Network Computing and Applications (NCA)*. 2020:1-8. doi:10.1109/NCA51143.2020.9306745

27. Peterson C, Moore R, Hicks JL, et al. NGS Analysis Confirms Common TP53 and RB1 Mutations, and Suggests MYC Amplification in Ocular Adnexal Sebaceous Carcinomas. *Int J Mol Sci*. 2021;22(16):8454. doi:10.3390/ijms22168454

28. Peng Z, Wang Y, Wang Y, et al. Application of radiomics and machine learning in head and neck cancers. *Int J Biol Sci.* 2021;17(2):475-486. doi:10.7150/ijbs.55716

29. Sheller MJ, Edwards B, Reina GA, et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Sci Rep.* 2020;10(1):12598. doi:10.1038/s41598-020-69250-1

30. Guidance for human genome data collection, access, use and sharing. Accessed October 21, 2025. https://www.who.int/publications/i/item/9789240102149?utm_source=chatgpt.com

31. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet.* 2019;51(4):584-591. doi:10.1038/s41588-019-0379-x

32. Kaissis GA, Makowski MR, Rückert D, Braren RF. Secure, privacy-preserving and federated machine learning in medical imaging. *Nat Mach Intell.* 2020;2(6):305-311. doi:10.1038/s42256-020-0186-1