

RESEARCH

Open Access



A new approach to identifying patients with elevated risk for Fabry disease using a machine learning algorithm

John L. Jefferies¹, Alison K. Spencer², Heather A. Lau³, Matthew W. Nelson⁴, Joseph D. Giuliano⁴, Joseph W. Zabinski^{2*} , Costas Boussios², Gary Curhan², Richard E. Gliklich² and David G. Warnock⁵

Abstract

Background: Fabry disease (FD) is a rare genetic disorder characterized by glycosphingolipid accumulation and progressive damage across multiple organ systems. Due to its heterogeneous presentation, the condition is likely significantly underdiagnosed. Several approaches, including provider education efforts and newborn screening, have attempted to address underdiagnosis of FD across the age spectrum, with limited success. Artificial intelligence (AI) methods present another option for improving diagnosis. These methods isolate common health history patterns among patients using longitudinal real-world data, and can be particularly useful when patients experience non-specific, heterogeneous symptoms over time. In this study, the performance of an AI tool in identifying patients with FD was analyzed. The tool was calibrated using de-identified health record data from a large cohort of nearly 5000 FD patients, and extracted phenotypic patterns from these records. The tool then used this FD pattern information to make individual-level estimates of FD in a testing dataset. Patterns were reviewed and confirmed with medical experts.

Results: The AI tool demonstrated strong analytic performance in identifying FD patients. In out-of-sample testing, it achieved an area under the receiver operating characteristic curve (AUROC) of 0.82. Strong performance was maintained when testing on male-only and female-only cohorts, with AUROCs of 0.83 and 0.82 respectively. The tool identified small segments of the population with greatly increased prevalence of FD: in the 1% of the population identified by the tool as at highest risk, FD was 23.9 times more prevalent than in the population overall. The AI algorithm used hundreds of phenotypic signals to make predictions and included both familiar symptoms associated with FD (e.g. renal manifestations) as well as less well-studied characteristics.

Conclusions: The AI tool analyzed in this study performed very well in identifying Fabry disease patients using structured medical history data. Performance was maintained in all-male and all-female cohorts, and the phenotypic manifestations of FD highlighted by the tool were reviewed and confirmed by clinical experts in the condition. The platform's analytic performance, transparency, and ability to generate predictions based on existing real-world health data may allow it to contribute to reducing persistent underdiagnosis of Fabry disease.

Keywords: Fabry disease, AI, Patient identification, Phenotypic biomarker

Background

Fabry disease is an inherited X-linked disorder caused by mutations in the GLA gene that result in deficient or absent lysosomal α -Gal A activity, and intracellular accumulation of globotriaosylceramide (Gb3) and

*Correspondence: jzabinski@om1.com

² OM1, Boston, MA, USA

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

related glycosphingolipids [1–3]. The condition is progressive, due to the cumulative damage done to multiple organ systems, especially the heart, kidney and central and peripheral nervous systems [4]. Fabry disease varies substantially in age of onset and clinical presentation. ‘Classic’ Fabry disease is associated with onset in younger males who first experience neuropathic pain and gastrointestinal symptoms, including abdominal pain and diarrhea, and renal function deterioration. Patients also experience cardiac signs and symptoms (including arrhythmias, myocardial fibrosis, and left ventricular hypertrophy), and frequently suffer strokes and transient ischemic attacks. Later onset Fabry disease affects male and female patients. The presentation can be heterogeneous including a range of symptoms including cardiac, central nervous system (CNS) and renal involvement, mood disorders, hearing loss, neuropathic pain and gastrointestinal (GI) symptoms [5]. Later onset symptoms vary in severity, and progress at different rates. Across the spectrum of Fabry disease, cardiac disease accounts for the majority of deaths in Fabry disease patients [6]. Although Fabry disease was once thought to exclusively affect males, both male and female patients may experience severe clinical manifestations [7].

Estimates of Fabry disease prevalence vary substantially, ranging in occurrence from 1 in 40,000 to 1 in 117,000 live births worldwide [8]. However, due to the variations in multisystemic clinical manifestations, Fabry disease remains substantially underdiagnosed [9]. The challenge of reaching a diagnosis for the full range of patients is made more difficult by a general lack of awareness among clinicians of the true variability in Fabry disease presentation [10]. Definitive diagnosis is achieved only through measured deficiency in α -galactosidase A in males or by detection of a pathogenic *GLA* mutation [10]. The multisystemic nature of the disease’s manifestations compounded by lack of awareness has led to an unmet need as patients often experience significant delays between symptom onset and diagnosis, with symptoms frequently first occurring in childhood or early adolescence but formal diagnosis frequently reached only in patients’ 20’s or 30’s [11].

Prior efforts to improve rates of diagnosis of Fabry disease and to shorten time-to-diagnosis have been made. Education among healthcare providers to improve familiarity with and awareness of symptoms and disease trajectory has been expanded, with some success in increasing screening and eventual diagnosis rates [12]. Newborn screening for Fabry disease has been implemented in Taiwan, and some US states. Screening programs have detected *GLA* mutations at much higher rates than current estimates of Fabry disease prevalence in the general population, though not all mutations necessarily result

in development of clinically significant Fabry disease [13, 14]. Finally, patient identification through screening of patients with certain conditions associated with Fabry disease (e.g., hypertrophic cardiomyopathy, renal failure) has been attempted [15, 16]. While frequently effective in identifying patients, this method relies on symptoms that only become evident once the disease has progressed substantially. In all efforts to diagnose Fabry disease patients, identifying the first case of the condition within a family is crucial. Other close relatives of the index case patient may be at much greater likelihood of having the disease, and can be evaluated accordingly.

Artificial intelligence (AI) provides a different approach for patient identification. AI methods isolate statistical patterns in large datasets and have been successfully used to predict patient outcomes in clinical settings [17]. These methods can function as ‘general [electronic health record] pattern recognition experts’ [18] and are especially useful in analyzing highly heterogeneous diseases where patients exhibit wide ranges of symptoms and clinical findings over long periods of time. Using AI methods for disease identification relies on assessing the presence of phenotypic patterns in the medical history of an undiagnosed patient to estimate likelihood of disease. These patterns can be learned from known Fabry disease patients’ medical histories. Though clinical applications of this technology are still nascent, AI methods have demonstrated strong performance in patient identification across a range of fields, including ophthalmology, neurology, cardiology, gastroenterology, and hepatology, among others [19–22].

Here, we describe an AI tool used to identify patients with Fabry disease. The tool (OM1 Patient Finder™, OM1 Inc., Boston, MA) was examined to determine its ability to identify Fabry disease cases using health history data. In addition to assessing statistical performance, including in men and women separately, the tool’s use of different features in healthcare records was studied to evaluate correspondence with known clinical signs and symptoms of the disease.

Methods

The data used in this study were drawn from a large cloud-based curated dataset (the OM1 Real World Data Cloud, OM1, Inc, Boston, MA). This dataset is derived from deterministically linked, de-identified, patient-level health care claims, EMR, and other data, and includes medication history and prescription information, laboratory results, symptoms and signs, procedures, and diagnoses. Additional medical and pharmacy claims data are linked to these clinical data to provide further information regarding patients’ clinical care. These data cover January 1, 2013 to the present day, and represent patients

with a wide age and geographic distribution (including patients from all 50 U.S. states). Use of these de-identified data to study patient characteristics and outcomes in retrospective, non-interventional, secondary analyses has been determined to be exempt from institutional review board (IRB) oversight by an independent IRB.

The dataset used in this study contained 4978 patients with confirmed Fabry disease, and 1,000,000 patients without any diagnostic or medication codes that would indicate a Fabry disease diagnosis. The patients with confirmed Fabry disease were identified by the presence of at least one Fabry disease ICD-10 code (E75.21, Fabry (-Anderson) disease), or by evidence of a medication approved to treat Fabry disease (agalsidase beta / Fabrazyme, or migalastat / Galafold) in their medical record. Patients without any of these indicators of a Fabry disease diagnosis were randomly selected from a population of several million patients with evidence of a minimum amount of activity in their medical records. Records were restricted to the period from January 1, 2013 to July 1, 2020. All patients were at least 18 years old for the entirety of the study period.

Following predictive AI modeling methodology [23], the study dataset (including patients with and without confirmed Fabry disease) was divided into two cohorts: a ‘training’ cohort comprising 75% of all patients selected at random, and a ‘testing’ cohort comprising the remaining 25%. The AI algorithm was calibrated to estimate the presence of Fabry disease using records from the training cohort.

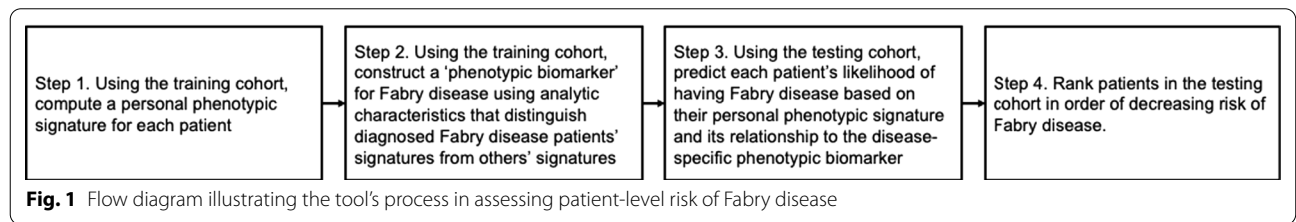
The AI tool first assessed patients by computing a personal phenotypic signature for each patient, using longitudinal health history data. This signature comprised a collection of related phenotypic characteristics (e.g. reports of symptoms, use of medications, records of procedures), grouped together, and reviewed and labeled with clinically descriptive signifiers by authors with medical expertise. For example, a phenotypic signal labeled ‘neuropathy’ contained diagnosis codes for polyneuropathy and skin paresthesia, and procedure codes indicating nerve conduction studies and needle electromyography, as well as many others. Most relevant signals calculated by the AI tool corresponded to organ systems and categories of pathologies. The statistical strength and relevance

of a signal for each individual patient’s predicted likelihood of undiagnosed Fabry disease is dependent on that signal’s manifestation in that patient’s history.

The tool then learned analytic relationships between patients’ phenotypic signatures and the outcome of interest—here, diagnosed Fabry disease—using the confirmed Fabry disease and non-confirmed patient data cohorts. Differences between the cohorts were used to construct a statistical ‘phenotypic biomarker’ for Fabry disease. As a final step, the platform predicted a likelihood of having Fabry disease for each patient in the testing cohort based on that patient’s phenotypic signature profile and its relationship to the phenotypic biomarker. Analytic performance in this classification task was quantified using receiver operating characteristic (ROC) curve analytic process represented in Fig. 1.

In addition, predictive performance was studied by examining effective Fabry disease ‘prevalence’ in groups of patients identified by the algorithm as at highest risk of having Fabry disease following stratification by risk. The tool rank-ordered patients in the testing cohort from greatest to least predicted Fabry disease likelihood and counted the number of confirmed Fabry disease patients in different risk strata (e.g., within the 1% of patients at greatest risk by predicted likelihood). Dividing these counts by the total number of patients in a risk group yielded an effective ‘prevalence’ for that risk group.

The study dataset was highly enriched with confirmed Fabry disease patients, containing approximately 1 confirmed patient for every 50 patients without confirmed Fabry disease (roughly three orders of magnitude greater than expected background prevalence). As such, calculated ‘prevalence’ values in risk groups were much higher than expected background prevalence of Fabry disease in the general population. To normalize, projected prevalence within these higher-risk strata was extrapolated using a conservative baseline population-wide assumption of 1 in 50,000 based on existing prevalence estimates [8]. For example, if calculated prevalence in a higher-risk stratum was 1 in 10 (that is, five times greater than the study population overall), projected ‘real-world’ prevalence after correcting for study population enrichment would be five times greater than the 1 in 50,000 background assumption, or 1 in 10,000.



The algorithm’s performance was also assessed by examining relative presence of phenotypic signals in different higher-risk strata. Within each, the fraction of patients with a particular phenotypic signal in their health history was calculated, and these frequencies of occurrence were compared across risk strata to identify patterns in phenotypic signal presentation in higher- and lower-risk groups.

Finally, these analyses were repeated after dividing the testing cohort into male-only and female-only subcohorts to examine robustness of predictive power while stratifying by sex.

Following generation and assessment of the tool’s analytic performance, a group of Fabry disease experts was assembled to review outputs, including components from patients’ individual phenotypic signatures and the platform’s phenotypic biomarker for Fabry disease. This review was intended to establish concordance between medical understanding of the condition and the algorithm’s outputs across results, especially considering the phenotypic variety in disease presentation in patients’ medical records.

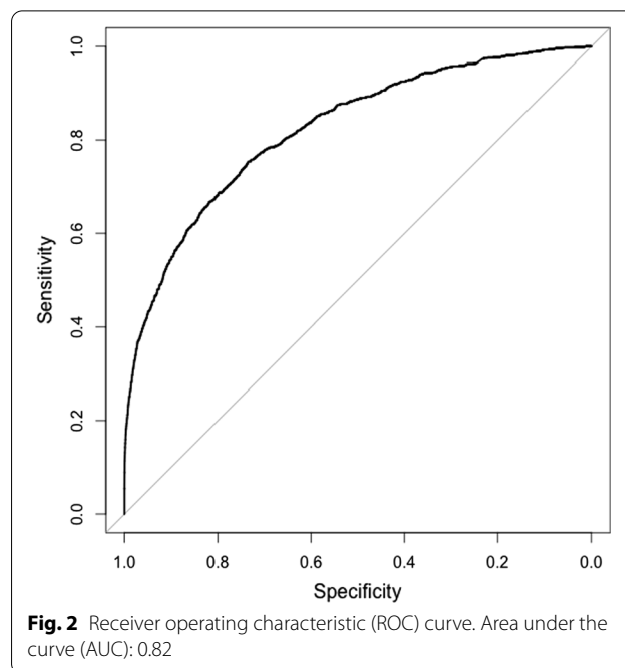
Results

The study population comprised a total of 1,004,978 patients. Relevant summary characteristics of the population are presented in Table 1.

The tool demonstrated very strong analytic performance in identifying Fabry disease patients in the test cohort, with an overall area under the receiver operating curve (AUROC) of 0.82 (Fig. 2).

Following rank-ordering of test set patients by predicted risk of Fabry disease, the tool performed very well in concentrating Fabry disease patients at the riskier end of this ranking. Confirmed Fabry disease patient presence in the riskiest 1% of patients as identified by the algorithm was nearly 24-fold greater than the baseline prevalence level. Using a true population assumption of Fabry disease prevalence of 1 in 50,000, simulated prevalence in the 1% identified by the platform would be roughly 24 times greater, or 1 in 2100. Additional amplifications and projected prevalence calculations are provided in Table 2.

Many phenotypic signals contributed to the tool’s overall performance. These signals all displayed differences in frequency of occurrence between patients across risk



strata; in general, patients in higher-risk strata had higher rates of occurrence of these phenotypic signals. These differences were quantified by calculating signal prevalence in specific higher-risk strata. Figure 3 displays these calculations for several representative phenotypic signals corresponding to clinically meaningful aspects of Fabry disease presentation. These provide a sample of the broader group of signals contributing to the algorithm’s predictive performance.

Sensitivity analysis examining the tool’s predictive performance in the test set after stratification by sex demonstrated strong analytic results for both men and women. Performance was slightly stronger in the male-only cohort relative to the female-only cohort, but amplification remained substantial within higher-risk cohorts for both groups relative to the overall study population. Amplification within the riskiest 1% subgroup is shown in Table 3.

Performance in the test set also remained very strong within male and female subsets, as illustrated by the respective ROC curves. The AUC for the male-only

Table 1 Age and sex distribution of study population

	Patients with confirmed Fabry disease (n = 4978)	Patients with no confirmed Fabry disease diagnosis (n = 1,000,000)
Sex (% female)	54.5%	55.1%
Mean age (SD)	47.0 (21.8)	44.8 (24.6)

Table 2 Amplification in riskier strata of the testing set, following rank-ordering by predicted likelihood of Fabry disease

Risk group	Amplification relative to total population	Projected prevalence
Riskiest 10%	5.4x	1 in 9261
Riskiest 1%	23.9x	1 in 2090
Riskiest 0.1%	109.8x	1 in 455

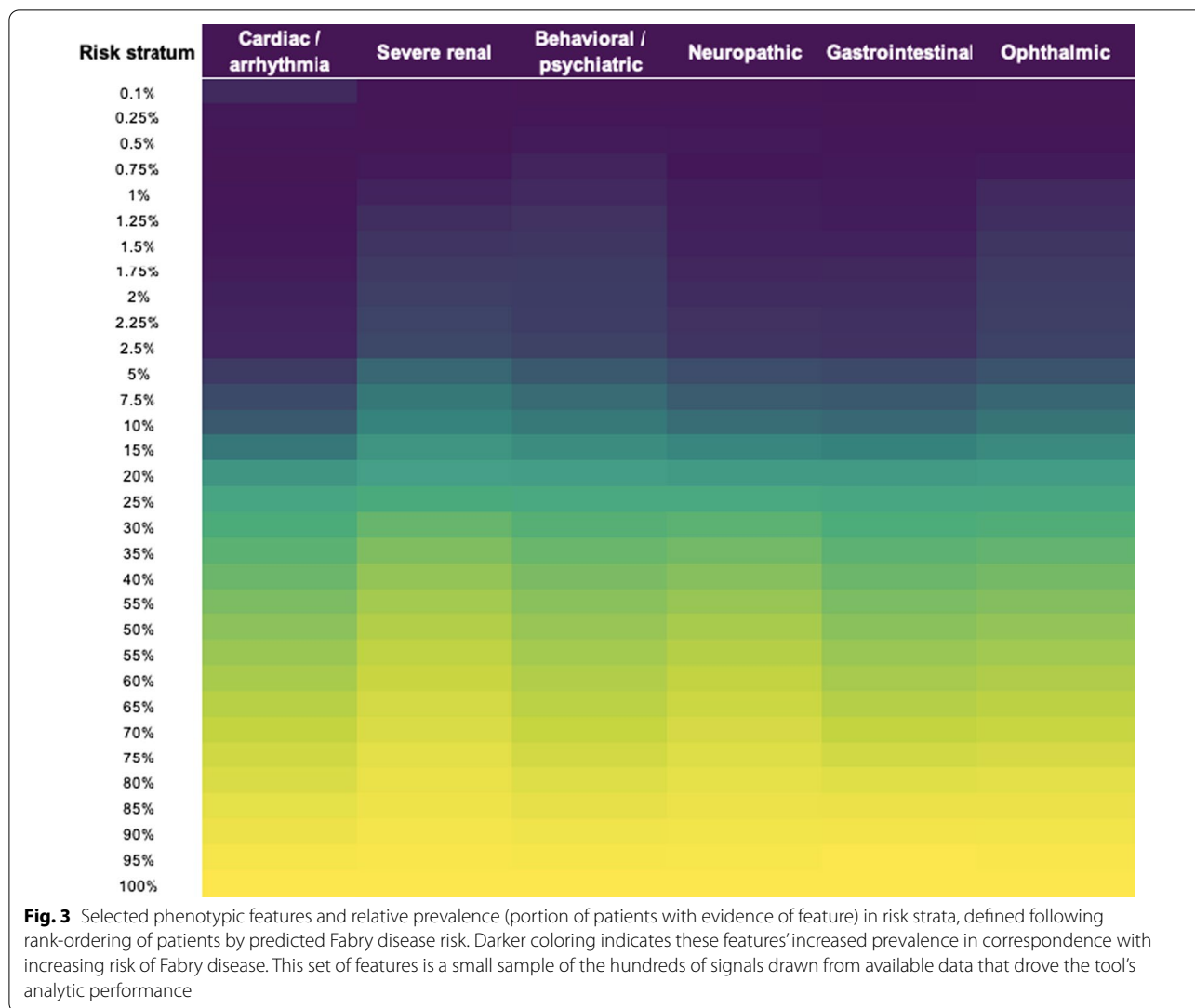


Table 3 Amplification of Fabry disease occurrence in riskiest 1% stratum of the testing set, grouped by sex

Patient subcohort	Amplification relative to assumed 1 in 50,000 prevalence	Projected prevalence
All patients	23.9x	1 in 2090
Male	26.8x	1 in 1867
Female	21.8x	1 in 2291

subcohort was 0.83; for the female-only subcohort, it was 0.82, reflecting balanced performance.

Discussion

Using a predictive model-based AI approach to pre-screen potential undiagnosed patients may present a more efficient way to focus expensive diagnostic

screening efforts on patients at greatest risk of Fabry disease. The AI tool analyzed in this study demonstrated strong overall performance in identifying Fabry disease patients at the individual level using structured medical record data for calibration and prediction. Fabry disease prevalence in higher-risk strata identified by the algorithm was substantially greater than in the background population. This performance was achieved using existing structured real-world health data; no additional Fabry disease-specific data were gathered, nor was information from unstructured data (i.e., clinical notes) used. Even though the data analyzed were de-identified, the AI tool's ability to recognize relevant phenotypic patterns in individual patients' histories preserves its potential for patient-specific real-world application. Future applications to identified patient populations will face challenges around consent and data privacy not addressed in this current study, and overcoming these challenges will be

crucial to identifying individual patients so proper treatment can be pursued.

The phenotypic signals driving the platform's predictions of Fabry disease risk correspond very well to clinical knowledge about Fabry disease and its presentation, including both 'classic' and late-onset phenotypes. The tool isolated differences in occurrence of phenotypic signals associated with the clearest clinical presentation of advanced disease—specifically, severe renal damage, cardiac arrhythmias, and neuropathies—with more frequent occurrence among those patients determined to be at greatest risk of Fabry disease by the tool. At the same time, the algorithm captured and utilized many more subtle statistical signals as well, including phenotypic signals associated with behavioral health, cerebrovascular damage, and auditory and balance-related symptoms. This abundance of phenotypic signals provided the tool with a rich set of contributing factors to predict Fabry disease risk, reflecting the diversity of patient experience without relying on a limited set of disease characteristics. The authors intend to explore these phenotypic signals at greater length in a subsequent publication.

This identification by the tool of phenotypic signals confirmed by the group of Fabry disease experts to be associated with Fabry disease is strong evidence that the algorithm is operating in concordance with medical, biological, and epidemiological knowledge about the condition. AI models do not always agree with clinical or scientific knowledge in this way and may generate predictions without obvious explanations. This 'black box' problem, where AI technology makes predictions that cannot be linked to expert knowledge, does not characterize the tool examined in this study. On the contrary, the algorithm elicited and relies on characteristics known to be associated with Fabry disease. This quality makes the platform an especially promising candidate for real-world clinical application.

This study has several limitations. First, real-world health history data were used for analysis. These data may suffer from missingness or incomplete capture, in addition to errors resulting from mistakes in data recording or transcription. However, this limitation is partly mitigated by the size of the cohort analyzed. The nearly 5000 patients in the confirmed Fabry disease patient group is substantially larger than Fabry disease cohorts found in much of the literature published in this disease area; to the authors' knowledge, this is one of the largest known Fabry disease patient cohorts studied to date [24, 25].

Second, the training cohort used to calibrate the AI algorithm relied on patients labeled as having Fabry disease—that is, those diagnosed with the disease, or with evidence of a disease-specific treatment in their health history record. Because Fabry disease diagnosis in

general does not necessarily correspond to true prevalence, bias may exist within the confirmed Fabry disease cohort. Somewhat surprisingly, the cohort is relatively balanced between male and female patients. The explanation for this balance is not immediately obvious. Diagnosis bias towards 'classic' Fabry disease could result in overrepresentation of male patients. However, since female patients experience later onset and longer lifespan than male patients overall [4, 26], survivorship bias could have contributed to greater female representation. Further research on the gender balance in Fabry disease epidemiology is necessary to better contextualize and address these questions. It is important to note that the AI tool in this study draws statistical information from all available aspects of Fabry disease patients' health histories, including late-onset patterns from dominantly 'classic' patients, and can use information in male patients' histories to inform predictions for females (and vice-versa). Pediatric patients were not analyzed in this study due to data limitations, but future research in pediatric populations may provide additional clarity around patterns in disease manifestation as detected and utilized for AI-driven patient identification.

Finally, the projected prevalence results presented herein assume a baseline population Fabry disease prevalence of 1 in 50,000. This assumption aligns with existing literature around Fabry disease [5] but is likely an underestimate of true prevalence due to underdiagnosis, and is reflective of classic Fabry disease rather than late onset variants. Consequently, projected prevalence estimates in higher-risk strata identified by the tool are likely conservative. If true prevalence is greater than 1 in 50,000, estimates of prevalence in higher-risk groups identified would increase as well.

This study also has many strengths. We evaluated a novel AI tool for identification of undiagnosed Fabry disease patients. The algorithm demonstrated strong analytic performance in identifying patients with Fabry disease, achieving an out-of-sample AUROC of 0.82. In the 1% of patients labeled by the platform as at greatest risk, Fabry disease prevalence was nearly 24-fold greater than in the population overall. The phenotypic characteristics of tool-identified patients correspond to existing literature, represent the multisystemic nature of the disease, and were clinically validated by a group of Fabry experts. These signals' range covers the full severity spectrum of the disease. This diversity of phenotypic signals provides robustness to the algorithm's predictive power, which it maintained when tested separately in all-male and all-female subcohorts.

Fabry disease patients continue to face long, difficult journeys from initial presentation of symptoms to eventual diagnosis. These challenges burden Fabry disease

patients with clinical manifestations outside the ‘classic’ presentation of the disease, including women, for whom barriers to accurate diagnosis can be especially high. AI technology offers a promising opportunity for earlier diagnosis of Fabry disease by drawing on statistical patterns in from large datasets of patients known to have the condition. Earlier diagnosis, in turn, could result in earlier monitoring, treatment if needed, slowed progression, and better outcomes for patients.

Conclusions

This study demonstrated that use of a novel AI tool may lead to improved identification of patients with undiagnosed Fabry disease. By labeling patients at disproportionate risk of having the condition, using existing medical record data, the AI tool tested may substantially improve the efficiency of more determinative approaches to Fabry disease diagnosis while continuing to generate new insights into patient characteristics. Future research will focus on clinical implementations of this technology to examine its performance in real-world settings.

Acknowledgements

The authors thank Danielle Cooke, Cristi Cavanaugh, and Bagirathy Ravishankar for their contributions to the preparation of this manuscript.

Authors' contributions

AS, CB, and JZ developed, refined, and analyzed the artificial intelligence technology described in this work. MN and JG provided substantial contributions to research design input and oversight. GC and RG provided clinical expert review and refinement of the technology's outputs. JJ, DW, and HL provided Fabry disease expert review, oversight, and substantially contributed to linking technology outputs to real-world clinical management of Fabry disease. JZ prepared the manuscript. All coauthors contributed to manuscript revision. All authors read and approved the final manuscript.

Funding

Funding for this work was provided by Amicus Therapeutics.

Availability of data and materials

The data that support the findings of this study are available from OM1, Inc. but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of OM1, Inc.

Declarations

Ethics approval and consent to participate

Use of the de-identified data examined in this initiative to study patient characteristics and outcomes in retrospective, non-interventional, secondary analyses has been determined to be exempt from institutional review board (IRB) oversight by an independent IRB.

Consent for publication

Not applicable.

Competing interests

JJ reports honoraria from Sanofi Genzyme, Chiesi Pharma, Amicus, Abbott, CHFSolutions, Audentes, AstraZeneca, and IPS Heart. AS reports employment and stock options at OM1, Inc. HL reports grant/research support and advisory board membership with grant, honoraria, and reimbursed travel/food for Amicus and Biomarin, advisory board membership with honoraria, reimbursed

travel/food for Chiesi and Actellion, grant/research support, consultant, advisory board membership, fellowship grant with research support, honoraria, and reimbursed travel/food from Sanofi, grant/research support, advisory board membership with honoraria, reimbursed travel/food from Takeda, grant/research support, consultant, advisory board membership, employment with honoraria, reimbursed travel/food, and salary at Ultragenyx, grant/research support from Mallinckrodt, Intrabio, Pfizer, Protalix, Ovid Therapeutics, and Denali, grant research support, consultant at Prevail Therapeutics and Aspa Therapeutics. Advisory board membership with honoraria for Taysa, Advisory board membership and lecturer with honoraria at National Gaucher Foundation and National Fabry Disease Foundation, advisory board member, lecturer, and grant support from National Tay Sachs and Allied Diseases, lecturer with honoraria for FSIG, advisory board member, researcher for Adult Polyglucosan Body Disorders Research Foundation. MN reports salary, stock, and benefits from Amicus Therapeutics, Inc. JG reports salary, stock, and stock options at Amicus Therapeutics, Inc. JZ reports employment and stock options at OM1, Inc. CB reports employment and stock options at OM1, Inc. GC reports stock options from OM1, Inc., employment at Brigham and Women's Hospital, consultant activities at and stock options from Allena Pharmaceuticals, holding a grant from Decibel Therapeutics and GSK, and consultant activities for Dicerna, Shire-Takeda, and AstraZeneca. RG reports employment and stock options at OM1, Inc. DW reports stock ownership in Reata Pharma, advisory-board membership at Chiesi Pharma, and consultant activities for Amicus Biotherapeutics and Protalix Biotherapeutics.

Author details

¹Division of Cardiovascular Diseases, University of Tennessee Health Science Center, Memphis, TN, USA. ²OM1, Boston, MA, USA. ³Division of Neurogenetics, Department of Neurology, New York University, New York, NY, USA. ⁴Amicus Therapeutics, Inc., Cranbury, NJ, USA. ⁵Department of Medicine, Division of Nephrology, University of Alabama at Birmingham, Birmingham, AL, USA.

Received: 21 July 2021 Accepted: 6 December 2021

Published online: 20 December 2021

References

1. Sweeley CC, Klionsky B. Fabry's disease: classification as a sphingolipidosis and partial characterization of a novel glycolipid. *J Biol Chem*. 1963;238:3148–50.
2. Brady RO, Gal AE, et al. Enzymatic defect in Fabry's disease—ceramidetrihexosidase deficiency. *N Engl J Med*. 1967;276:1163–7.
3. Kint JA. The enzyme defect in Fabry's disease. *Nature*. 1970;227(5263):1173.
4. Ortiz A, Germain DP, Desnick RJ, Politei J, Mauer M, Burlina A, et al. Fabry disease revisited: management and treatment recommendations for adult patients. *Mol Genet Metab*. 2018;123(4):416–27.
5. Tuttolomondo A, Pecoraro R, Simonetta I, Miceli S, Pinto A, Licata G. Anderson-Fabry disease: a multiorgan disease. *Curr Pharm Des*. 2013;19(33):5974–96.
6. Schiffmann R, Hughes DA, Linthorst GE, Ortiz A, Svarstad E, Warnock DG, et al. Screening, diagnosis, and management of patients with Fabry disease: conclusions from a “kidney disease: improving global outcomes” (KDIGO) controversies conference. *Kidney Int*. 2017;91(2):284–93.
7. Arends M, Wanner C, Hughes D, Mehta A, Oder D, Watkinson OT, et al. Characterization of classical and nonclassical fabry disease: a multicenter study. *J Am Soc Nephrol*. 2017;28(5):1631–41.
8. Mehta A, Beck M, Eyskens F, Feliciani C, Kantola I, Ramaswami U, et al. Fabry disease: a review of current management strategies. *QJM*. 2010;103(9):641–59.
9. Ranieri M, Bedini G, Parati EA, Bersano A. Fabry disease: recognition, diagnosis, and treatment of neurological features. *Curr Treat Options Neurol*. 2016;18(7):33.
10. Curiati MA, Aranda CS, Kyosen SO, Varela P, Pereira VG, D'Almeida V, et al. The challenge of diagnosis and indication for treatment in fabry disease. *J Inborn Errors Metab Screen*. 2017;5:232640981668573.
11. Eng CM, Fletcher J, Wilcox WR, Waldek S, Scott CR, Sillence DO, et al. Fabry disease: baseline medical characteristics of a cohort of 1765 males and females in the Fabry registry. *J Inheret Metab Dis*. 2007;30(2):184–92.

12. Savary A-L, Morello R, Brasse-Lagnel C, Milliez P, Bekri S, Labombarda F. Enhancing the diagnosis of Fabry disease in cardiology with a targeted information: a before–after control–impact study. *Open Heart*. 2017;4(1):e000567.
13. Laney DA, Bennett RL, Clarke V, Fox A, Hopkin RJ, Johnson J, et al. Fabry disease practice guidelines: recommendations of the national society of genetic counselors. *J Genet Couns*. 2013;22(5):555–64.
14. Spada M, Pagliardini S, Yasuda M, Tukul T, Thiagarajan G, Sakuraba H, et al. High incidence of later-onset Fabry disease revealed by newborn screening. *Am J Hum Genet*. 2006;79(1):31–40.
15. Monserrat L, Gimeno-Blanes JR, Marín F, Hermida-Prieto M, García-Honrubia A, Pérez I, et al. Prevalence of Fabry disease in a cohort of 508 unrelated patients with hypertrophic cardiomyopathy. *J Am Coll Cardiol*. 2007;50(25):2399–403.
16. Maruyama H, Takata T, Tsubata Y, Tazawa R, Goto K, Tohyama J, et al. Screening of male dialysis patients for Fabry disease by plasma globotriaosylsphingosine. *Clin J Am Soc Nephrol*. 2013;8(4):629–36.
17. Lindberg DS, Prosperi M, Bjarnadottir RI, Thomas J, Crane M, Chen Z, et al. Identification of important factors in an inpatient fall risk prediction model to improve the quality of care using EHR and electronic administrative data: a machine-learning approach. *Int J Med Inf*. 2020;143:104272.
18. Dias R, Torkamani A. Artificial intelligence in clinical and genomic diagnostics. *Genome Med*. 2019;11(1):70.
19. Lin W-C, Chen JS, Chiang MF, Hribar MR. Applications of artificial intelligence to electronic health record data in ophthalmology. *Transl Vis Sci Technol*. 2020;9(2):13.
20. Myszczyńska MA, Ojames PN, Lacoste AMB, Neil D, Saffari A, Mead R, et al. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nat Rev Neurol*. 2020;16(8):440–56.
21. Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial intelligence in precision cardiovascular medicine. *J Am Coll Cardiol*. 2017;69(21):2657–64.
22. Le Berre C, Sandborn WJ, Aridhi S, Devignes M-D, Fournier L, Smail-Tabbone M, et al. Application of artificial intelligence to gastroenterology and Hepatology. *Gastroenterology*. 2020;158(1):76–94.e2.
23. Scheinost D, Noble S, Horien C, Greene AS, Lake EMR, Salehi M, et al. Ten simple rules for predictive modeling of individual differences in neuroimaging. *Neuroimage*. 2019;193:35–45.
24. Arends M, Hollak CEM, Biegstraaten M. Quality of life in patients with Fabry disease: a systematic review of the literature. *Orphanet J Rare Dis*. 2015;10(1):77.
25. Elliott PM, Germain DP, Hilz MJ, Spada M, Wanner C, Falissard B. Why systematic literature reviews in Fabry disease should include all published evidence. *Eur J Med Genet*. 2019;62(10):103702.
26. Waldek S, Patel MR, Banikazemi M, Lemay R, Lee P. Life expectancy and cause of death in males and females with Fabry disease: findings from the Fabry registry. *Genet Med*. 2009;11(11):790–6.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

