Journal of Inherited Metabolic Disease

WILEY

JIMD SSIEM

REVIEW

Are Inherited Metabolic Disorders More Common and Less Predictable Than We Thought?

Nina B. Gold[1,2] | Alanna Strong[3,4] | Harini Somanchi[5] | Jessica Gold[6]

[1]Department of Pediatrics, Mass General Brigham for Children, Boston, Massachusetts, USA | [2]Department of Pediatrics, Harvard Medical School, Boston, Massachusetts, USA | [3]Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA | [4]Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA | [5]Department of Pediatrics, Massachusetts General Hospital, Boston, Massachusetts, USA | [6]Division of Clinical Genetics, Department of Pediatrics, Northwell Health, Great Neck, New York, USA

Correspondence: Nina B. Gold (ngold@mgh.harvard.edu)

Received: 14 July 2025 | **Revised:** 17 September 2025 | **Accepted:** 22 September 2025

Funding: The authors confirm independence from the sponsors; the content of the article has not been influenced by the sponsors. Dr. N.B. Gold's work is supported by K08HG012811 from the National Human Genome Research Institute. Dr. Strong's work is supported by K08DK128606-04 from the National Institute of Diabetes and Digestive and Kidney Diseases.

Keywords: biobanks | genomic screening | inherited metabolic disorders | penetrance | phenotype analysis | prevalence

ABSTRACT

"Genotype-first" approaches, studies that apply genomic sequencing in unselected cohorts of apparently healthy adults or infants, have begun to upend traditional notions about the prevalence and penetrance of inherited metabolic disorders. In this commentary, we discuss how large-scale genomic data from healthy newborns and biobanks of adult research participants, along with clinical testing such as reproductive carrier screening and secondary findings from exome and genome sequencing, have revealed a new category of "genotype positive" cases of IMDs that were previously unrecognized by both clinicians and public health programs. In particular, the prevalence and penetrance of variants linked to IMD have important implications for evaluating the utility of genomic sequencing as a public health screening tool in the newborn period. Although genomic sequencing may allow us to detect treatable disease earlier and identify individuals at risk before irreversible damage occurs, realizing its promise as a screening tool will require an acknowledgment that more genomic data does not always equate to clearer decisions and that disease-associated variants may not universally require intervention.

"Genotype-first" approaches—studies that apply genomic sequencing in unselected cohorts of apparently healthy adults or infants—have begun to upend traditional notions about the prevalence, penetrance, and even pathophysiology of inherited metabolic disorders (IMDs). Historically, IMDs have been diagnosed through two primary pathways: early detection through biomarker-based public health newborn screening (NBS) programs or clinical diagnosis of symptomatic individuals. The global prevalence of IMDs has been based almost exclusively on these clinically ascertained cases and is estimated to be 50.9 per 100 000 live births, or one in approximately 2000 individuals [1]. As we move into an era of prospective, genotype-first approaches, it is becoming increasingly clear that these figures underestimate both the genomic burden and phenotypic spectrum of IMDs in the general population.

Large-scale genomic data from healthy newborns and biobanks of adult research participants, along with clinical testing such as reproductive carrier screening and secondary findings from exome and genome sequencing, have revealed a new category of "genotype positive" cases of IMDs that were previously unrecognized by both clinicians and public health programs. The phenotypes of the individuals identified via these pathways vary widely [2, 3]. Some have classic IMD phenotypes with subtle

biochemical abnormalities, while others are phenotypically normal, raising challenging questions about the definition of disease [4].

A growing number of genotype-first investigations suggest that pathogenic/likely pathogenic (P/LP) variants associated with rare disease are far more common than previously recognized, often outpacing historical estimates of disease prevalence. In a genotype-first approach to data from 72 434 adults in the Mount Sinai BioMe and U.K. Biobanks, Forrest et al. found that thousands of reportedly healthy individuals harbored P/LP variants related to autosomal dominant disorders [5]. Another analysis of the BioMe Biobank suggested that over 2% of adult participants harbored disease-associated variants linked to just nine monogenic disorders, including amyotrophic lateral sclerosis, retinitis pigmentosa, and several cancer predisposition syndromes [6]. Individuals at risk for other rare disorders, including undiagnosed Noonan syndrome, Marfan syndrome [7], and vascular Ehlers-Danlos syndrome [8], have also been identified in both hospital-based and population biobanks. Kingsmore et al. predicted that approximately 2% of adults in the U.K. Biobank harbor P/LP variants associated with 412 severe childhood genetic disorders that are candidates for expanded NBS using genomic sequencing [9].

One of the major limitations in many genotype-first studies has been the paucity of phenotypic data, which is essential to assessing disease penetrance and ultimately the positive predictive value of genomic screening in the general population [10]. While genomic sequencing is high-throughput and has become relatively inexpensive, accurate phenotyping remains labor-intensive and variable in quality. Phenotyping methods that rely on ICD-10 codes or abstracting notes from electronic medical records (EMR) likely underestimate true disease penetrance, as subtle disease manifestations may go unrecognized by clinicians, be undocumented, or lack appropriate or accurately applied codes [7, 8, 11, 12]. This disconnect limits our ability to define variant pathogenicity, natural history, and the true burden of disease in the general population. Detailed clinical follow-up, including targeted histories and physical examinations, diagnostic biochemical testing, and imaging is essential if we are to effectively utilize population-scale data and genotype-first ascertainment. Indeed, our own approach of recontacting and phenotyping participants in dedicated outpatient visits has yielded meaningful clinical insights and revealed cases of missed diagnosis and true non-penetrance [2, 3, 13].

We used a hospital-based biobank containing whole exome sequencing data from 53 345 participants [14] to identify adults at risk for childhood-onset diseases that are high-priority targets of genomic NBS [3, 15]. In a preliminary study across 54 genes, we identified 82 individuals (0.15%, or 1 in 650 participants) with P/LP variants in a dosage sufficient to cause disease (e.g., one variant in disorders associated with autosomal dominant inheritance or two variants in those associated with autosomal recessive inheritance) in 10 genes. The vast majority of participants (58/82, 70.7%) had not been previously diagnosed with the associated genetic condition. Through EMR review, patient recontacting, and informed clinical phenotyping, we showed that 46 of 82 (57.9%) individuals had suggestive signs of the associated disorder, much higher than estimates of penetrance for other monogenic disorders that are based solely on EMR data or diagnostic codes [5, 16]—though perhaps still quite a bit lower than might be expected based on historical "phenotype-first" methods of patient ascertainment.

Carrier screening, a clinically available test which has been used for decades to assess the risk of parents having offspring affected by monogenic disease, has also revealed unexpected diagnoses of IMD in apparently healthy adults. One in 43 women was found to harbor disease-causing variants in 12 genes associated with dominant or X-linked phenotypes, including 1 in 2545 at risk for Fabry disease and 1 in 8331 at risk for ornithine transcarbamylase (OTC) deficiency [17]. In another study, 70 of 91 637 (0.76% or 1 in 1309) women were found via carrier screening to be at personal risk of biotinidase deficiency, including one woman with two alleles associated with profound disease [18]. More information about the phase of these variants, laboratory test results, and clinical symptoms of individuals at risk for IMD identified incidentally through carrier screening is needed, but these cases may challenge the dichotomy between which health conditions should or should not be screened and force a reconsideration of the variants that should be reported in apparently healthy individuals, particularly newborns. Because these research programs include relatively small cohorts and currently lack longitudinal clinical follow-up data through adolescence and adulthood [19], their estimates of penetrance remain limited. While some individuals with canonical disease genotypes may develop classic biochemical and clinical features, others may remain asymptomatic well into adulthood, or indefinitely.

Understanding the prevalence and penetrance of P/LP variants linked to IMD has important implications for evaluating the utility of genomic sequencing as a public health screening tool in the newborn period. Prevalence can be used to estimate the expected number of positive results, while predicted penetrance informs the positive predictive value of screening—that is, the proportion of individuals who will actually develop clinical symptoms in their lifetime. Studies such as BabySeq, GUARDIAN, and nearly 30 others worldwide have shown that sequencing can identify newborns with P/LP variants in genes linked to monogenic disorders that are not included in current NBS panels [20–22]. Genomic sequencing may be the only way to ascertain infants at risk for specific IMDs that do not have pathognomonic biochemical markers, such as glycogen storage disorders types Ia and Ib, or OTC deficiency, for which biochemical NBS is unreliable.

If universal genomic NBS becomes a reality, confirmatory biochemical testing will continue to be essential to guide diagnosis. The sensitivity of genomic sequencing for IMDs has not yet been firmly established, but several studies have demonstrated that it is less than that of biochemical testing for certain IMDs that are currently included on NBS [23–25]. Furthermore, even with biochemical validation, when to treat is not always straightforward. For some IMDs, treatment is invasive, high-burden, or carries iatrogenic risk. This includes conditions such as lysosomal storage disorders, which may prompt the recommendation of an early hematopoietic stem cell transplant (HSCT), or even dietary therapies that may affect quality of life, highlighting the continued need for biochemical evaluation to guide management. In the gnomAD database, we previously identified apparently

healthy adults at risk for a range of severe, early-onset monogenic conditions treatable with HSCT [26]. Treatment decisions will require a judicious balance between early action and clinical restraint, guided by shared decision-making and longitudinal outcome data, which will continue to grow over time.

The path forward will involve an assessment of prevalence and penetrance of IMD-associated variants in population- and hospital-based biobanks, prospective studies that capture the natural history and treatment outcomes identified by sequencing, and strong clinical and policy frameworks that support shared decision-making between clinicians and families of children at risk for rare disease. We may face difficult decisions about whether to screen and report all risk-associated variants, especially those with low penetrance, in settings with limited health resources. In the long term, population-scale genomic screening has the potential to redefine our understanding of the variable expressivity of IMD. While it may allow us to detect treatable disease earlier and identify individuals at risk before irreversible damage occurs, realizing this promise will require an acknowledgment that more genomic data does not always equate to clearer decisions, and that disease-associated variants may not universally require intervention.

Author Contributions

Nina B. Gold drafted the original manuscript. Alanna Strong and Harini Somanchi contributed to manuscript editing and revision. Jessica Gold provided critical oversight, contributed references, and assisted with manuscript refinement. All authors reviewed and approved the final version of the manuscript.

Ethics Statement

The authors have nothing to report.

Conflicts of Interest

Dr. Nina Gold has received consulting fees from Guidepoint Global LLC, MyOme Inc., and RCG Consulting, has received speaking honoraria from Ambry Genetics, and has been paid for expert testimony at the Risk Management Foundation of the Harvard Medical Institutions Inc. Dr. Strong, Ms. Somanchi, and Dr. Jessica Gold declares no conflicts of interest.

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

1. D. Waters, D. Adeloye, D. Woolham, E. Wastnedge, S. Patel, and I. Rudan, "Global Birth Prevalence and Mortality From Inborn Errors of Metabolism: A Systematic Analysis of the Evidence," *Journal of Global Health* 8 (2018): 021102.

2. J. I. Gold, S. Madhavan, J. Park, et al., "Phenotypes of Undiagnosed Adults With Actionable OTC and GLA Variants," *Human Genetics and Genomics Advances* 4 (2023): 100226.

3. N. Gold, J. Yeo, E. Perez, et al., "O35: A Genotype-First Approach Among Adults at Risk for Treatable Monogenic Disorders Reveals Underdiagnosis and Incomplete Penetrance," *Genetics in Medicine Open* 3 (2025): 102111.

4. I. S. Forrest, G. N. Nadkarni, and R. Do, "Penetrance of Deleterious Clinical Variants-Reply," *JAMA* 327 (2022): 1927.

5. I. S. Forrest, K. Chaudhary, H. M. T. Vy, et al., "Population-Based Penetrance of Deleterious Clinical Variants," *JAMA* 327 (2022): 350–359.

6. I. S. Forrest, Á. Duffy, J. K. Park, et al., "Genome-First Evaluation With Exome Sequence and Clinical Data Uncovers Underdiagnosed Genetic Disorders in a Large Healthcare System," *Cell Reports Medicine* 5 (2024): 101518.

7. B. M. Wenger, N. Patel, M. Lui, et al., "A Genotype-First Approach to Exploring Mendelian Cardiovascular Traits With Clear External Manifestations," *Genetics in Medicine* 23 (2021): 94–102.

8. M. M. Lui, M. Shadrina, B. D. Gelb, and A. R. Kontorovich, "Features of Vascular Ehlers-Danlos Syndrome Among Biobank Participants Harboring Predicted High-Risk COL3A1 Genotypes," *Circulation: Genomic and Precision Medicine* 16 (2023): e003864.

9. S. F. Kingsmore, M. Wright, L. D. Smith, et al., "Prequalification of Genome-Based Newborn Screening for Severe Childhood Genetic Diseases Through Federated Training Based on Purifying Hyperselection," *American Journal of Human Genetics* 111 (2024): 2618–2642.

10. B. D. Gelb, "2024 ASHG Presidential Address: Incomplete Penetrance and Variable Expressivity: Old Concepts, New Urgency," *American Journal of Human Genetics* 112 (2025): 461–466.

11. L. Bastarache and J. F. Peterson, "Penetrance of Deleterious Clinical Variants," *JAMA* 327 (2022): 1926–1927.

12. T. Nakanishi, V. Forgetta, T. Handa, et al., "The Undiagnosed Disease Burden Associated With Alpha-1 Antitrypsin Deficiency Genotypes," *European Respiratory Journal* 56 (2020): 2001441.

13. A. Dickey, L. Rebeiz, R. K. Leaf, et al., "The Prevalence and Penetrance of Acute Intermittent Porphyria in the Mass General Brigham Biobank. In: Session 9—Rational Approach(Es) to Diagnosis and Care in AHPS (September 25th, 2024)," *BMJ Publishing Group Ltd* 11 (2024): A10.

14. C. L. Blout Zawatsky, N. Shah, K. Machini, et al., "Returning Actionable Genomic Results in a Research Biobank: Analytic Validity, Clinical Implementation, and Resource Utilization," *American Journal of Human Genetics* 108 (2021): 2224–2237, https://doi.org/10.1016/j.ajhg.2021.10.005.

15. N. B. Gold, S. M. Adelson, N. Shah, et al., "Perspectives of Rare Disease Experts on Newborn Genome Sequencing," *JAMA Network Open* 6 (2023): e2312231.

16. C. F. Wright, B. West, M. Tuke, et al., "Assessing the Pathogenicity, Penetrance, and Expressivity of Putative Disease-Causing Variants in a Population Setting," *American Journal of Human Genetics* 104 (2019): 275–286.

17. V. Souter, B. Prigmore, E. Becraft, et al., "Reproductive Carrier Screening Results With Maternal Health Implications During Pregnancy," *Obstetrical & Gynecological Survey* 79 (2024): 247–249.

18. P. Benn, Y. Wang, J. Gray, et al., "Evaluating Reproductive Carrier Screening Using Biotinidase Deficiency as a Model: Variants Identified, Variant Rates, and Management," *Genetics in Medicine* 27 (2025): 101345.

19. R. C. Green, N. Shah, C. A. Genetti, et al., "Actionability of Unanticipated Monogenic Disease Risks in Newborn Genomic Screening: Findings From the BabySeq Project," *American Journal of Human Genetics* 110 (2023): 1034–1045.

20. O. Ceyhan-Birsoy, J. B. Murry, K. Machini, et al., "Interpretation of Genomic Sequencing Results in Healthy and Ill Newborns: Results From the BabySeq Project," *American Journal of Human Genetics* 104 (2019): 76–93.

21. T. Minten, S. Bick, S. Adelson, et al., "Data-Driven Consideration of Genetic Disorders for Global Genomic Newborn Screening Programs," *Genetics in Medicine* 27 (2025): 101443.

22. A. Ziegler, C. Koval-Burt, D. M. Kay, et al., "Expanded Newborn Screening Using Genome Sequencing for Early Actionable Conditions," *JAMA* 333 (2025): 232–240.

23. S. L. Bick, A. Nathan, H. Park, R. C. Green, M. H. Wojcik, and N. B. Gold, "Estimating the Sensitivity of Genomic Newborn Screening for Treatable Inherited Metabolic Disorders," *Genetics in Medicine* 27 (2025): 101284.

24. S. Cook, E. Dunn, J. Kornish, et al., "Molecular Testing in Newborn Screening: VUS Burden Among True Positives and Secondary Reproductive Limitations via Expanded Carrier Screening Panels," *Genetics in Medicine* 26 (2024): 101055.

25. M. H. Wojcik, T. Zhang, O. Ceyhan-Birsoy, et al., "Discordant Results Between Conventional Newborn Screening and Genomic Sequencing in the BabySeq Project," *Genetics in Medicine* 23 (2021): 1372–1375.

26. N. B. Gold, S. M. Harrison, J. H. Rowe, et al., "Low Frequency of Treatable Pediatric Disease Alleles in gnomAD: An Opportunity for Future Genomic Screening of Newborns," *HGG Adv* 3 (2022): 100059.