## Genetics inMedicine COMMENT

## Genome-first findings require precision phenotyping

Monica A. Giovanni, MS<sup>1</sup> and Michael F. Murray, MD<sup>1</sup>

We are in a time of transition. A longstanding phenome-first world for ascertainment of genetic conditions is now sharing the stage with genome-first ascertainment; it is reasonable to believe that in this 21<sup>st</sup> century that genome-first ascertainment could become the predominant method of genetic risk ascertainment. Incidental and secondary findings obtained as part of diagnostic testing or research studies are currently the primary source of "genome-first", however if population screening becomes routine so, too, will genome-first ascertainment. In this moment of transition, we believe it is critically important for the field of clinical genomics to focus on the essential task of creating the phenotyping standards needed in order to rule-in or rule-out a genotype-phenotype correlation for any given case of genome-first ascertainment.

It is worth noting that the advent of next-generation sequencing was not the start of this genome-first transition. An important and relevant precedent was the application of CFTR screening of prospective parents in the 1990s. This "carrier screening" consistently identified cases of prospective parents carrying two pathogenic variants.<sup>1</sup> These were individuals who typically had never been evaluated for cystic fibrosis, but who unexpectedly found themselves undergoing a clinical evaluation to follow up on dual variant identification. In 2018, candidate cases of genome-first cystic fibrosis benefit from having a robust precision phenotyping process that includes: (1) a well-described "expanded phenotypic range" that has come to be appreciated in the greater than 25 year period since the ready availability of CFTR genotype testing; (2) the availability of phenotyping experts; and (3) a set of clear diagnostic criteria for the condition.<sup>2</sup> This robust process stands as a model as genome-first phenotyping is developed for other gene-condition pairs.

A paper by Thompson and colleagues in *Genetics in Medicine*<sup>3</sup> describes the superb use of available sequence data to return genome-first findings to the parents of children with intellectual disability who were sequenced as part of a trio study.<sup>4</sup> This work demonstrates the state of the art for "precision" approaches to sequencing, variant calling, and genetic counseling. The patient engagement included in this report includes the collection of self-reported (family and personal) medical history. Importantly, the methods do not include expert review of medical records or new clinical evaluations. In this instance, without targeted phenotyping,

the authors are not able to claim any definitive genotype-phenotype correlations in the participants.

In a genome-first world, the concept of penetrance is extremely important. Having an accurate determination of penetrance in the reported literature and available databases going forward is essential if we are to avoid the well-described problem of misattribution of DNA variant pathogenicity that frequently found its way into the literature in the late 20<sup>th</sup> century and initial years of the 21<sup>st</sup> century.<sup>5</sup> One key to avoiding a misattribution of genotype–phenotype correlation in the genome-first literature is to reach consensus early and enshrine in the peer-review literature an understanding that "self-reporting" of personal and family history is not sufficient for establishing a genotype–phenotype correlation.<sup>3</sup>

The nature of classic monogenic disease descriptions is that they are usually first described in patients and families with phenotypes on the severe end of the phenotypic spectrum and, conversely, the nature of monogenic disease that are found in previously undiagnosed patients uncovered through genome-first screening is that it includes "subclinical" cases on the mild end of the phenotypic spectrum. An example of genome-first screening uncovering an extended phenotypic range relates to familial hypercholesterolemia (FH), for which we have seen evidence of the condition without dramatic elevations of measured low-density lipoprotein cholesterol, but still associated with increased end-organ disease.<sup>6,7</sup>

The specific training of phenotyping experts for any particular condition can certainly vary, as can the types of practitioners who can eventually apply well-defined diagnostic criteria. For any given condition, there are clinicians who can meet the challenge. The broad genomics community is one of diverse skills: laboratorians have pushed the field forward immeasurably with sequencing techniques and variant classification efforts; and genetic counselors are poised within their scope of practice to assemble family histories and communicate unexpected genetic information to patients engaged in genomic screening.<sup>8</sup> For the clinical evaluation task, it will be others who can bring advanced phenotyping skills to bear in order to establish the presence or absence of disease-associated signs and symptoms. In rare and ultra-rare diseases, this will often be medical geneticists, in collaboration with other specialists, who for decades have carried out detailed genotype-phenotype correlations and syndrome

<sup>&</sup>lt;sup>1</sup>Department of Genetics, Yale School of Medicine, New Haven, Connecticut, USA. Correspondence: Michael F. Murray (michael.murray@yale.edu)

Submitted 6 March 2018; accepted: 20 March 2018; https://doi.org/10.1038/s41436-018-0026-4 Published online: 8 June 2019

delineations.<sup>9,10</sup> In more common diseases, it will be medical specialists, primary care providers, and advanced practitioners.

While there are clear diagnostic criteria for some genetic conditions, there are many that do not have diagnostic criteria or have criteria written for a phenome-first world with too much weight given to genetic results. The current lack of diagnostic criteria for a condition as common as hereditary breast and ovarian cancer syndrome is not surprising given that since the discovery of BRCA1/2, the focus has been on devising approaches for identifying elevated pretest probability to determine who should be genetically tested. In a genome-first world, the development of hereditary breast and ovarian cancer syndrome diagnostic criteria has become essential. In the case of FH, a reweighting of genome-first results will need to be considered in the light of the frequency of cardiovascular disease in the population and the weight given to genetic results in the established FH diagnostic criteria.<sup>6</sup> An answer to the important question of whether and how a positive family history fits into genome-first diagnostic criteria will need to evolve as the evidence emerges. As diagnostic standards are proposed and then tested, distinctions between syndromic disease and common disease in the presence of monogenic risk will become easier; for example, in determining when prostate cancer is syndromic versus sporadic in the setting of BRCA2 risk.

With advancing applications for next-generation sequencing, the past decade has focused on generating high-quality sequence data and developing algorithms and methods for accurate variant classification. These advances have pushed the field forward; the community is now poised for large-scale exome and genome sequencing in unselected cohorts, even rigorously investigating potential benefits and harms from population sequencing. Accurate sequence data can be rapidly generated and variants can be queried within established, open databases.<sup>11</sup> What lies ahead is a broad effort to develop an understanding of the predictive value of genome-first approaches. The success of this effort will be contingent on clear standards for descriptions of penetrance.

While we have focused on getting the sequence and variant interpretation right, in many cases, sufficient phenotypic data have not yet been gathered for it to be possible to comment on the true presence of disease penetrance. The clinical genomics community must refocus its efforts on standardizing the gathering of appropriate phenotypic data in the setting of incidental and secondary findings.<sup>12</sup>

## DISCLOSURE

Dr Murray reports personal fees from Invitae and grants from Regeneron outside the submitted work. The remaining author declares no conflict of interest.

## REFERENCES

- 1. Strom CM, Crossley B, Redman JB, et al. Cystic fibrosis screening: lessons learned from the first 320,000 patients. *Genet Med.* 2004;6:136–40.
- Paranjape SM, Zeitlin PL. Atypical cystic fibrosis and CFTR-related diseases. Clin Rev Allergy Immunol. 2008;35:116–23.
- Thompson ML, Finnila CR, Bowling KM, et al. Genomic sequencing identifies secondary findings in a cohort of parent study participants. *Genet Med.* 2018. https://doi.org/10.1038/gim.2018.53
- Bowling KM, Thompson ML, Amaral MD, et al. Genomic diagnosis for children with intellectual disability and/or developmental delay. *Genome Med.* 2017;9:43.
- Xue Y, Chen Y, Ayub Q, et al. 1000 Genomes Project Consortium Deleterious- and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. *Am J Hum Genet*. 2012;91:1022–32.
- Abul-Husn NS, Manickam K, Jones LK, et al. Genetic identification of familial hypercholesterolemia within a single U.S. health care system. *Science*. 2016 Dec 23;354. pii: aaf7000.
- Khera AV, Won HH, Peloso GM, et al. Diagnostic yield and clinical utility of sequencing familial hypercholesterolemia genes in patients with severe hypercholesterolemia. J Am Coll Cardiol. 2016;67:2578–89.
- Practice-Based Competencies for Genetic Counselors 2015. Accreditation Council For Genetic Counseling. http://www.gceducation.org/ Documents/ACGC%20Core%20Competencies%20Brochure\_15\_Web. pdf. Accessed 28 Feb 2018
- Ramoni RB, Mulvihill JJ, Adams DR, et al. Undiagnosed Diseases Network, Wise AL The Undiagnosed Diseases Network: accelerating discovery about health and disease. *Am J Hum Genet*. 2017;100:185–92.
- Johnston JJ, Lewis KL, Ng D, et al. Individualized iterative phenotyping for genome-wide analysis of loss-of-function mutations. *Am J Hum Genet*. 2015;96:913–25.
- 11. Rehm HL, Berg JS, Brooks LD, et al. ClinGen ClinGen—the Clinical Genome Resource. N Engl J Med. 2015;372:2235–42.
- Murray MF Your DNA is not your diagnosis: getting diagnoses right following secondary genomic findings. *Genet Med.* 2016;18:765–7.