

A taxonomy of medical uncertainties in clinical genome sequencing

Paul K.J. Han, MD, MPH¹, Kendall L. Umstead, MS², Barbara A. Bernhardt, MS³, Robert C. Green, MD, MPH^{4,5,6,7}, Steven Joffe, MD, MPH⁸, Barbara Koenig, PhD⁹, Ian Krantz, MD¹⁰, Leo B. Waterston, MA¹, Leslie G. Biesecker, MD¹¹ and Barbara B. Biesecker, PhD²

Purpose: Clinical next-generation sequencing (CNBS) is introducing new opportunities and challenges into the practice of medicine. Simultaneously, these technologies are generating uncertainties of an unprecedented scale that laboratories, clinicians, and patients are required to address and manage. We describe in this report the conceptual design of a new taxonomy of uncertainties around the use of CNBS in health care.

Methods: Interviews to delineate the dimensions of uncertainty in CNBS were conducted with genomics experts and themes were extracted in order to expand on a previously published three-dimensional taxonomy of medical uncertainty. In parallel, we developed an interactive website to disseminate the CNBS taxonomy to researchers and engage them in its continued refinement.

Results: The proposed taxonomy divides uncertainty along three axes—source, issue, and locus—and further discriminates the uncertainties into five layers with multiple domains. Using a hypothetical clinical example, we illustrate how the taxonomy can be applied to findings from CNBS and used to guide stakeholders through interpretation and implementation of variant results.

Conclusion: The utility of the proposed taxonomy lies in promoting consistency in describing dimensions of uncertainty in publications and presentations, to facilitate research design and management of the uncertainties inherent in the implementation of CNBS.

Genet Med advance online publication 19 January 2017

Key Words: exome sequencing; genome sequencing; next-generation sequencing; taxonomy; uncertainty

INTRODUCTION

The arrival of clinical next-generation sequencing (CNBS) has introduced new opportunities and challenges. It offers increased diagnostic precision and improved understanding of the causes and outcomes of an array of diseases at the individual level and promises to advance predictive and preventive medicine on an unprecedented scale. Genomic technologies have already begun to deliver on this promise. Although these technologies have been available for only a few years, their integration into clinical practice for the evaluation of patients with a variety of disorders has been swift and their demonstrated utility impressive.^{1,2} Genomic testing has successfully advanced the diagnosis of previously undiagnosed rare disorders, the identification of high risk for disease among healthy individuals, and the development of targeted cancer treatments.

Although medicine is replete with uncertainties, genomic technologies introduce uncertainties of unique scales and

types.^{1,3} Genomic testing is technically complex, involving multiple steps—from sample acquisition to clinical report generation to communication of results to the patient—each of which introduces uncertainty regarding, for example, the accuracy and reliability of test results. The clinical uses of genomic testing introduce other uncertainties regarding the benefits and harms of genomic test information, the optimal strategies for communicating this information to patients, and the consequences of genomic testing for patients, family members, the health-care system, and society.

Whether CNBS is being used for diagnosis, predispositional testing, preconception screening, prenatal screening, pharmacogenomic profiling, or some other purpose, the introduction of CNBS into the clinic requires laboratories, clinicians, patients, and families to address and manage the many uncertainties it raises. However, the first step in this endeavor is to distinguish and understand these specific uncertainties and

¹Center for Outcomes Research and Evaluation, Maine Medical Center Research Institute, Portland, Maine, USA; ²Social and Behavioral Research Branch, National Human Genome Research Institute, NIH, Bethesda, Maryland, USA; ³Division of Translational Medicine and Human Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA; ⁴Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA; ⁵Personalized Medicine, Partners Healthcare, Cambridge, Massachusetts, USA; ⁶Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA; ⁷Broad Institute of MIT and Harvard, Boston, Massachusetts, USA; ⁸Department of Medical Ethics and Health Policy, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA; ⁹Institute for Health & Aging, University of California, San Francisco, California, USA; ¹⁰Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; ¹¹Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA.
Correspondence: Barbara B. Biesecker (barbara.biesecker@nih.gov)

Submitted 11 August 2016; accepted 17 November 2016; advance online publication 19 January 2017. doi:10.1038/gim.2016.212

the extent to which they are reducible by the acquisition of further knowledge. This, in turn, enables thoughtful consideration of their practical and ethical implications. Clinicians can then help patients appreciate the uncertainties pertaining to their testing experience and draw on personal resources to cope with them.

Cataloguing the dimensions of uncertainty will contribute to the field of clinical genomics by facilitating recognition of the uncertainties inherent in each step of genomic testing and helping researchers, clinicians, patients, and relatives establish realistic expectations of its processes and outcomes.⁴ It will also guide future research by helping to identify and close key knowledge gaps. The current study used as its starting point a conceptual taxonomy of uncertainty in health care.¹ This taxonomy begins with a working definition of “uncertainty” as the conscious awareness of ignorance—a self-awareness of incomplete knowledge of some aspect of the world.⁵ It then distinguishes three major dimensions of uncertainty: (i) source, (ii) issue, and (iii) locus. *Source* refers to the cause of a given uncertainty or the fundamental reason for a specific knowledge gap. *Issue* refers to the substantive situation, outcome, or alternative to which a given uncertainty applies. *Locus* is the particular party or stakeholder in whose mind(s) a given uncertainty resides.¹ The taxonomy further divides each of these dimensions into more specific subdimensions by which various uncertainties can be categorized. The long-term goal of this multidimensional conceptual taxonomy is to facilitate a clearer understanding of the uncertainties inherent in the diagnosis and prognosis of various disorders and a rational approach to their management by researchers, clinicians, patients, and other stakeholders.^{1,6}

MATERIALS AND METHODS

The current project was conceptual in nature and its aim was to develop a systematic and logically coherent working taxonomy of uncertainty in CNGS that can be refined and modified over time. It was empirically informed by a synthesis of insights from open-ended qualitative interviews undertaken with leaders in the clinical genomics field. These interviews were conducted between June and August 2014 by one of the investigators (P.K.J.H.) with six genomics, genetic counseling, and bioethics experts from among the National Human Genome Research Institute–funded Clinical Sequencing Exploratory Research consortium investigators who volunteered to participate. The interviews explored experts’ views on the major sources and issues of uncertainty in clinical genomics, the reducibility or irreducibility of these uncertainties, and their implications for clinical care and the lives of patients and families. Interviews were carried out until saturation of novel topics was achieved. The interviews were transcribed verbatim and the transcripts coded for these topics. Themes were then synthesized from the coded transcripts by two of the study investigators (B.B.B. and P.K.J.H.) and used to expand a draft taxonomy of uncertainties specific to genome sequencing and clinical genomic testing that was modeled on Han and colleagues’ three-dimensional

conceptual taxonomy. Four of the expert interviewees also reviewed and helped to iteratively refine the taxonomy.

In parallel with our initial effort to represent the key uncertainties of clinical genomics within a conceptual taxonomy, we developed an interactive taxonomy platform website (<http://research.nhgri.nih.gov/taxonomy/>). The purpose of this website is to disseminate the new taxonomy of uncertainty and collaboratively engage researchers in its further refinement. The interactive Web page is curated by three of the authors (B.B.B., L.G.B., P.K.J.H.) and enables individuals to suggest revisions. Using a platform that allows ongoing input enables the taxonomy to reflect the current experiences and knowledge of a broad variety of individuals and to be responsive to technological and conceptual advances. In a field as dynamic as clinical genomics, this capacity is essential and enhances the representativeness and utility of the taxonomy.

RESULTS

Dimension 1: sources of uncertainty

The first of the three major dimensions of uncertainty is *source*, which is further subdivided into three main types: probability, ambiguity, and complexity (**Figure 1**). *Probability* (otherwise known as risk) refers to the fundamental indeterminacy or stochastic nature of future outcomes and has also been termed first-order or aleatory uncertainty; an example is the point estimate of risk (e.g., “20% probability of benefit from treatment”). *Ambiguity* refers to “the lack of reliability, credibility, or adequacy of information” (ref. 1) about probability and is also known as second-order or epistemic uncertainty. “Ambiguity arises in situations in which risk information is unavailable, inadequate, or imprecise” (ref. 6); a predominant representation is the confidence interval around a point estimate (e.g., “10% to 30% probability of benefit from treatment”). *Complexity* refers to features of risk information that make it difficult to understand; examples include conditional probabilities or multiplicity in risk factors, outcomes, or choice options, which diminish their comprehensibility or produce information overload.

Because probability arises from the fundamental indeterminacy of future events, there is no additional source into which it can be further subdivided. However, the other main sources of uncertainty, ambiguity, and complexity can be further categorized according to more specific fundamental types and sources (**Figure 1**). Ambiguity in clinical genome sequencing has three principal subtypes: (i) conceptual, (ii) methodological, and (iii) clinical. *Conceptual ambiguity* arises from two main sources. The first is *model inadequacy*: limitations in either the theoretical or the empirical models (e.g., genes, animal systems) used to represent gene–disease mechanisms. The second is *nosological inadequacy*: limitations in the conceptual systems used to classify diseases and phenotypes.⁴

Methodological ambiguity, the second principal subtype of ambiguity, arises from five main subsources: (i) sample or data integrity problems, (ii) test limitations, (iii) unmeasured factors, (iv) procedural variability or error, and (v) test misinterpretation. *Sample or data integrity problems* represent deficiencies in

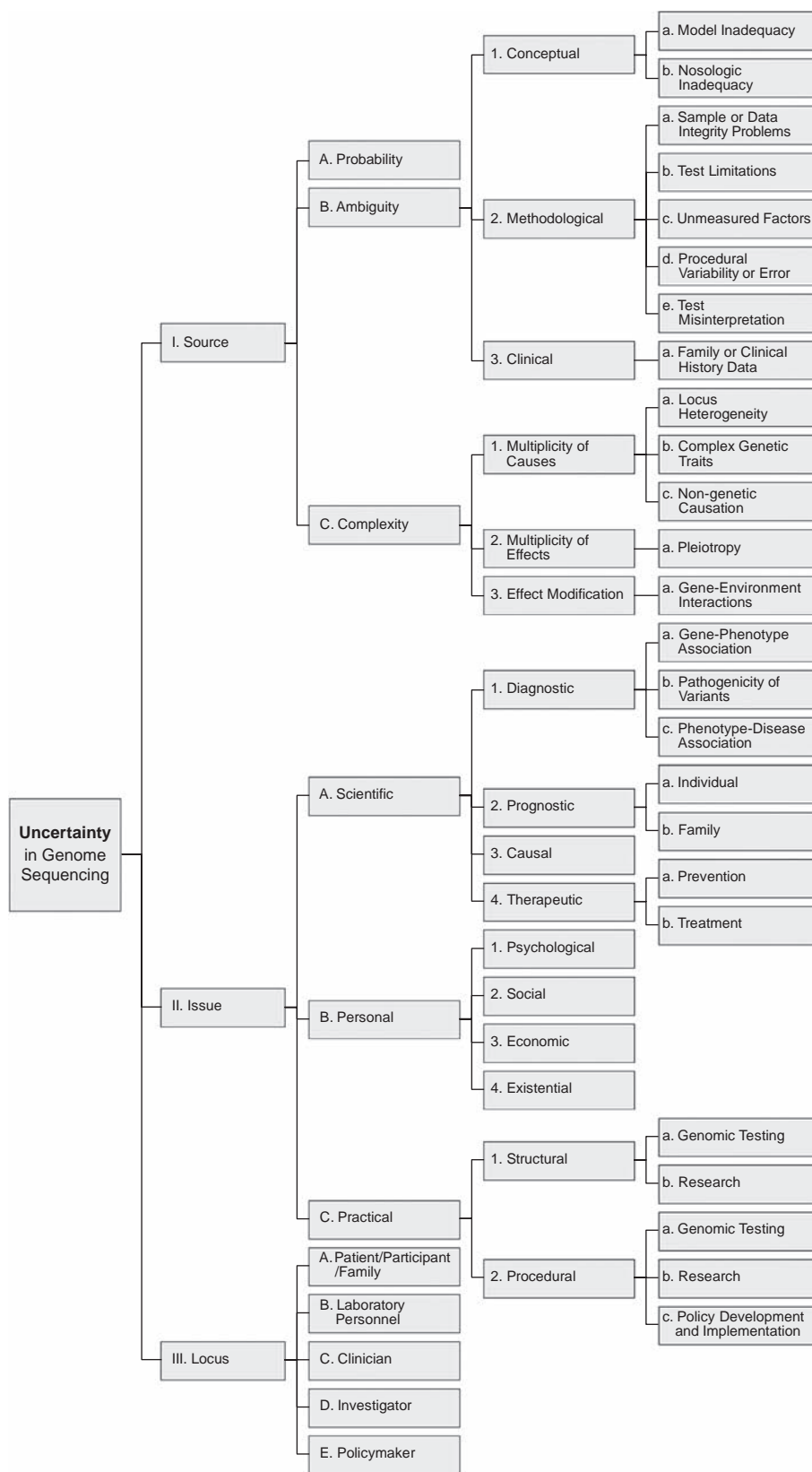


Figure 1 A taxonomy of uncertainty. This is a visual representation of a three-dimensional taxonomy of medical uncertainty in clinical genome sequencing. The three major dimensions are source, issue, and locus.

laboratory samples or processing that result in diagnostic errors. *Test limitations* represent inherent constraints in the accuracy or precision of laboratory instruments or techniques. *Unmeasured factors* are causal biological variables that affect phenotype but are not currently recognized or assayed. *Procedural variability or error* refers to methodological problems that leave existing variants undetected. *Test misinterpretation* refers to failures of laboratory or clinical personnel to correctly annotate or analyze genomic test results or of health-care providers to accurately understand the clinical implications of the results.

Clinical ambiguity, the final principal subtype of ambiguity, arises from incomplete or erroneous family or clinical history data. These data limitations have multiple causes, including missing information, faulty information gathering, or problems of recall.

Complexity, the third main source of uncertainty in clinical genomic testing, has three principal subtypes reflecting distinct origins: (i) multiplicity of causes, (ii) multiplicity of effects, and (iii) effect modification. *Multiplicity of causes*, in turn, arises from three specific subsources. The first is *locus heterogeneity*: the phenomenon in which a single disorder or phenotypic characteristic is caused by mutations in heterogeneous genes in different individuals (e.g., autism). The second is *complex genetic traits*: the determination of single disorders or phenotypic characteristics by variation at multiple genetic loci (e.g., height). The third is *nongenetic causation*: the presence of nongenetic (i.e., environmental) determinants of disorders or phenotypic characteristics that may interact with genetic determinants and have poorly quantified effects.

Multiplicity of effects, the second principal subtype of complexity, originates from *pleiotropy*, a phenomenon whereby a single gene mutation causes multiple apparently unrelated disorders or phenotypic characteristics (e.g., neurofibromatosis).

Effect modification, the final principal subtype of complexity, reflects the presence of moderating or mediating causal pathways such as *gene–environment interactions*, which affect the extent to which genomic factors affect health.

Importantly, any or all of these major sources of uncertainty—probability, ambiguity, and complexity—and the more specific subtypes and problems from which uncertainty ultimately originates may apply, in principle, to any given CNGS test. However, these uncertainty sources do not manifest in the abstract, but rather in particular issues that constitute the second major dimension of uncertainty.

Dimension 2: issues of uncertainty

The second of the three major dimensions of uncertainty is issue—the specific substantive matter about which an individual lacks knowledge. Issues of uncertainty can further be subdivided into three main types: scientific, practical, and personal (Figure 1). “Scientific uncertainty is data-centered, whereas practical and personal uncertainties are system- and patient-centered, respectively.”¹

In CNGS, *scientific uncertainty* pertains to many issues—diagnostic, prognostic, causal, or therapeutic. *Diagnostic issues*

include (i) *gene–phenotype association*, (identification of variants in a gene that cause disease and detection of a phenotype), (ii) *pathogenicity of variants* (whether a variant causes disease), and (iii) *phenotype–disease association* (whether a given phenotypic manifestation is part of a disease or syndrome). *Prognostic issues* include the future outcomes of individuals and their family members. *Causal issues* include the underlying factors and mechanisms that determine or explain a given genomic variant or its ultimate phenotypic manifestations. *Therapeutic issues* include the prevention and treatment of the diseases or health consequences of genomic variants.

Personal uncertainty pertains to psychosocial and existential issues, including the effects and implications of genomic test results for one’s own goals or outlook on life, personal relationships, sense of meaning or one’s future well-being and that of one’s family members or social groups. An oversimplified understanding of results by patients and their relatives would also be reflected in this domain. Personal uncertainties may also encompass moral questions—e.g., whether the results of CNGS testing should be disclosed to particular individuals or shared with family members or other parties.

Practical uncertainty pertains to lack of knowledge about both the structures of health care (i.e., the institutional facilities and resources of the health-care system) and the processes of health care (i.e., the procedures and actions required to deliver, access, or utilize health-care services related to CNGS testing).

Importantly, for all these specific *issues* of uncertainty—scientific, practical, and personal—the underlying cause may be any of the *sources* comprising the first dimension of uncertainty (probability, ambiguity, complexity). Any or all of these sources may engender uncertainty about not only diagnosis, prognosis, causal explanations, and treatment recommendations (scientific uncertainty) but also the expected quality of care and the procedures required to access care (practical uncertainty), as well as the effects of illness or treatment on one’s personal relationships and goals in life (personal uncertainty). In theory, probabilities exist for all of these outcomes, although these probabilities are unknown to varying degrees—and thus ambiguous—and are further compounded by varying degrees of complexity.

Dimension 3: loci of uncertainty

The third dimension of the taxonomy, the *locus* of uncertainty (i.e., the party in whose mind uncertainty resides) can be any of several stakeholders: patients, clinicians, researchers, or other individuals, including family members, regulators, payers, or health policy makers (Figure 1). Uncertainty may be shared, as when a physician effectively educates a patient about limitations in scientific evidence regarding the benefits, harms, and clinical utility of CNGS testing. At other times, however, uncertainty may not be shared between parties, such as when physicians (but not their patients) are aware of scientific ignorance or when patients (but not their physicians) are aware of ignorance about their own values and preferences.

The deeper challenge, however, is that the differential salience of specific uncertainties to individual stakeholders can

Box 1 Pharmacogenetics case of statin-induced myopathy risk

A 37-year-old man, Mark, has moderate, untreated hypercholesterolemia that is being managed by diet with some degree of success, but he has not been able to achieve his target cholesterol levels. Mark's internist ordered CNGS because Mark's son has autism. Mark's sequence was accessed primarily for filtering variants to exclude candidates for the cause of his son's autism. However, Mark's internist requests an extended genome analysis that identifies him as having a chr12:21,331549T>C, *SLCO1B1* p.(Val174Ala) heterozygous variant, which predicts an increased risk for myopathy from statin (especially simvastatin). After learning about the possible side effect of myopathy, Mark is concerned that this treatment approach could compromise his running ability and interfere with his regular social contact with running companions.

Note: The authors devised this hypothetical case to illustrate application of the taxonomy. It was not based on an actual person, and any potential resemblance to an actual case is coincidental.

perpetuate ignorance. When individuals ignore their uncertainty about a given issue, they diminish their awareness of what they do not know. This can be consequential, particularly in clinical and policy endeavors, where different domains of knowledge (and ignorance) intersect. In the domain of clinical care, for example, patients who do not grapple with scientific uncertainties surrounding the benefits and harms of medical interventions cannot benefit fully from shared decision making.

A comprehensive taxonomy of uncertainty in clinical genomic testing can mitigate this problem by making different stakeholders more aware of—and better able to understand—areas of uncertainty that lie beyond their usual focus of attention. The sharing of uncertainties that such a taxonomy can facilitate may, in turn, help people make better informed decisions about how to manage these uncertainties.

A case example: varieties of uncertainty in clinical genomic testing

We present a case to illustrate the potential value of our provisional conceptual taxonomy of uncertainty (Box 1). It demonstrates the variety of uncertainties that arise in CNGS testing and how the taxonomy may provide a useful organizational framework that can make existing uncertainties more explicit to different stakeholders. This case involves numerous areas of uncertainty that are differentially perceived by individual stakeholders, each of whom represents a distinct locus of uncertainty, thereby resulting in varying degrees of uncertainty. We now use our taxonomy to compare and contrast the uncertainties experienced by three of the different stakeholders: clinicians, patients, and genetics laboratory professionals.

Clinician perspective. Mark's internist's primary concern is the care of her patient. She wonders about what clinical actions

Box 2 Clinician perspective

I. Source

- A. Probability: Mark's genomic variant confers some risk of drug-induced myopathy
- B. Ambiguity: the strength of the scientific evidence linking Mark's genomic variant with drug-induced myopathy is not clear

2. Methodological

- c. Unmeasured factors: associations with additional biological variables that may play a causal role have not been established

C. Complexity

3. Effect modification

- a. Gene–environment interactions: Mark's exercise program and other environmental and behavioral factors may moderate the risk of drug-induced myopathy

II. Issue

A. Scientific

1. Diagnostic

- a. Phenotype–disease association: whether Mark's hypercholesterolemia will result in cardiovascular disease; whether Mark's genomic variant will result in drug-induced myopathy

3. Prognostic

- b. Individual: whether Mark will develop myopathy and, if so, how severe it will be

4. Therapeutic

- a. Prevention: what interventions can be undertaken to avert drug-induced myopathy or cardiovascular disease
- b. Treatment: whether statin therapy should be prescribed and, if so, which statin at what dose

are appropriate and how to counsel Mark at this point. She recognizes that multiple uncertainties apply to Mark's case, but her own greatest uncertainties revolve around the strength of the evidence about the risk of myopathy associated with Mark's particular genomic variant and thus the appropriateness of recommending a lifestyle change and/or medication for Mark (Box 2). The most salient sources and issues of uncertainty for this internist apply to several places in our conceptual taxonomy and influence the content of the counseling she provides to her patient. She focuses her discussion on the evidence for the risk of drug-induced myopathy and the known approaches to mitigating this risk.

Patient perspective. Mark's primary concern focuses on the implications of his genomic variant and the actions he needs to take to preserve his health. Note that many of the uncertainties that Mark experiences (Box 3) are identical to those that his internist faces. However, because of Mark's relative scientific naïveté, the sources of these uncertainties are less differentiated or precise. At the same time, some of Mark's uncertainties—

Box 3 Patient perspective

I. Source

- A. Probability: Mark's genomic variant confers some risk of drug-induced myopathy
- C. Complexity: information about Mark's genomic variant is unfamiliar and multifaceted and thus difficult for him to understand

II. Issue

A. Scientific

- 1. Diagnostic: the meaning of Mark's genomic variant is unclear to him
- 3. Prognostic
 - b. Individual: whether Mark will develop myopathy and, if so, how severe it will be
- 4. Therapeutic
 - a. Prevention: what interventions can be undertaken to avert drug-induced myopathy or cardiovascular disease
 - b. Treatment: whether statin therapy should be prescribed and, if so, which statin at what dose

B. Personal

- 1. Psychological: whether and to what extent Mark will experience adverse effects of knowledge of the genomic variant on his psychological well-being
- 2. Social: whether discontinuing running with Mark's social group will have adverse effects on his social well-being

- C. Practical: what actions Mark needs to undertake, both now and in the future, to access the care he needs is unclear

notably his personal uncertainties—initially reside only in his mind and not the internist's because it is Mark's own life that is affected and he is the expert on his own values and perspectives.

Genetics laboratory perspective. In contrast to clinicians and patients, laboratory professionals are primarily concerned with technical and scientific aspects of CNGS and its implications that extend beyond the care of any individual patient. The laboratory professional's perspective on uncertainty is focused on the bioinformatics, the testing pipeline, and the classification of variants (Box 4). Laboratory geneticists have access to greater knowledge about all of the methodological issues that produce uncertainty in CNGS, and their broader concerns also make other uncertainties more salient. For these reasons, it is also deeper scientifically and encompasses more specific questions and precisely defined sources of ignorance. However, it is also more generic—i.e., focused on generalizable knowledge affecting populations more than on specific knowledge affecting individuals. Whereas the uncertainties experienced by Mark and his internist focus on the implications of CNGS for Mark himself, the uncertainties

Box 4 Laboratory scientist perspective

I. Source

B. Ambiguity

- 1. Conceptual
 - a. Model inadequacy: the myopathy risk model does not account for all causal pathways and mechanisms
- 2. Methodological
 - c. Unmeasured factors: other pharmacogenomic variants that may increase or decrease the likelihood of drug-induced myopathy (e.g., ABCB1 efflux transporter) are not detected by the assay

C. Complexity

- 1. Multiplicity of causes
 - b. Complex genetic trait: unrecognized variants may affect the risk of drug-induced myopathy
- 3. Effect modification
 - a. Gene–environment interactions: environmental and behavioral factors may moderate the risk of drug-induced myopathy

II. Issue

A. Scientific

- 1. Diagnostic
 - b. Pathogenicity of variants: whether a given genomic variant will result in drug-induced myopathy
- 2. Causal: how and why statins cause myopathy
- 3. Prognostic
 - b. Individual: whether any specific individual characteristics are significant predictors of myopathy onset and severity

C. Practical

- 1. Procedural
 - a. Genomic testing: the extent to which testing facilities adhere to standards for the conduct and interpretation of CNGS

of laboratory professionals, researchers, and policy makers extend beyond Mark or any given individual.

Bridging perspectives on uncertainty in CNGS

The foregoing case illustrates how the various stakeholders who represent the loci of uncertainty vary in the breadth, depth, and precision of their awareness of areas for which there is a lack of knowledge. This variation in uncertainty reflects differences in the prior knowledge that each stakeholder brings to CNGS and in their primary concerns, both professional and personal. These differences can leave stakeholders unaware of important areas of ignorance and thus to have false certainty. For example, Mark may be uninformed about scientific ambiguity regarding the strength of the association between his variant

and his risk for adverse health outcomes. He may assume that if information about his variant was returned to him with the recommendation that he avoid certain statin therapy, then there is relative certainty about the risk for myopathy. For him, the consequences of medication therapy may be most salient. At the same time, Mark's internist may be uninformed about outcomes important to Mark, such as the personal, social, and economic consequences of interventions undertaken to reduce Mark's disease risks. Finally, genetics laboratory professionals may be uninformed about the concerns, values, and needs of individual patients.

Our conceptual taxonomy can help mitigate these problems by bridging the perspectives of different stakeholders, thereby promoting shared awareness of otherwise unconsidered sources and issues of uncertainty and enabling stakeholders to approach these uncertainties in an organized manner. For example, the taxonomy may facilitate clinicians' awareness of the various uncertainties identified by genetics laboratories, in order to better differentiate areas of greater or lesser certainty and to weigh the evidence more explicitly and precisely in making clinical recommendations for patients. This exercise may also prompt the internist to return to the primary literature to update her knowledge of the areas of uncertainty about penetrance. It can also help genetics professionals identify important knowledge gaps to frame their next study related to use of CNGS or efforts to improve procedures for implementing CNGS.

The taxonomy can also help clinicians organize their approach to counseling patients and promoting shared decision making. It supplies a framework for the discussion and serves as a reminder of important and salient issues that the discussion should address to maximize a patient's ability to share in the making of an informed decision about further action. At the same time, the taxonomy can draw clinicians' attention to potential sources or issues of uncertainty that may not affect clinicians themselves but may have profound effects on patients. For example, Mark's internist may not feel at all uncertain about her recommendations to mitigate risks of the identified variant; however, questions may remain in Mark's mind. Similarly, the internist may not experience uncertainty about personal issues that are important to Mark, such as doubts about whether having his variant will affect his lifestyle or whether he should share this information with his wife, who tends to worry about his health. The internist can use the taxonomy as a mental checklist to reduce the potential for her to overlook uncertainties that exist in her patient's mind and that should be explored further to promote shared decision making. The taxonomy can help clinicians and patients reach a greater mutual understanding of not only which uncertainties exist in each other's minds but also which uncertainties each party judges to be important.

At the same time, the taxonomy can also facilitate a broader understanding of CNGS-related uncertainties among other stakeholders, whose unique concerns and perspectives might limit their awareness of existing sources or issues of uncertainty. For example, laboratory professionals may naturally focus more on ambiguity arising from technical limitations in

genomic sequencing tests and error rates in the procurement and processing of biological specimens. By contrast, researchers may focus on ambiguity arising from conceptual problems (e.g., the distinctions between normal and abnormal phenotypes or nonpathogenic and pathogenic genotypes) or methodological problems (e.g., how sequencing is conducted or how sequencing data are analyzed and interpreted). Payers may be most concerned about whether recommended testing meets their internal criteria for reimbursement. Health policy makers may focus primarily on ethical uncertainties (e.g., whether and when sequencing data should be disclosed to patients and family members and used as a basis for medical intervention) as well as the societal consequences of implementing genomic sequencing. The primary value of the new taxonomy is to enable all of these different stakeholders to better acknowledge—and thereby understand—the uncertainties that apply to not only their own domains of interest but also those of other stakeholders.

A critical question, however, is how broad and deep the awareness of ignorance ought to be for any given stakeholder. Human attention spans are limited, and uncertainty can have adverse psychological effects on individuals, causing fear, diminished sense of well-being, and avoidance of decision making.^{1,6} The appropriate breadth and depth of uncertainty to convey to different stakeholders remain to be determined and most likely depends on their unique goals and concerns.

DISCUSSION

We have proposed a new taxonomy of medical uncertainty in genome sequencing as a tool to help standardize how uncertainty is conceptualized, defined, and described by various stakeholders motivated by differing goals, perspectives, and concerns. Building on an existing taxonomy of medical uncertainty,¹ the current effort aims to help diverse stakeholders undertake a rational, logically coherent approach to understanding and managing uncertainty in CNGS. The new taxonomy can enable stakeholders to identify and assess uncertainties that might otherwise escape their awareness and to determine the best course of action, such as information seeking and further scientific research for uncertainties that are reducible and strategies to promote coping and adaptation for uncertainties that are irreducible.

We believe that a taxonomic approach to uncertainty may benefit other domains of medicine and that the novel nature of CNGS provides an opportunity to model such an approach and demonstrate its broader utility. Applying a conceptual taxonomy of uncertainty to CNGS may highlight questions and concerns related to the novel nature of this technology, and applying the taxonomy to other novel technologies might have similar effects. The taxonomy will also require continual revision as the state of the science in CNGS advances, diminishing some uncertainties.⁷ Meanwhile, for patients and research participants deciding whether to undergo CNGS, uncertainties about its benefits and risks need to be explicitly acknowledged and communicated to ensure informed decision making—a

requirement both in CNGS and in medicine more broadly. Applying our conceptual taxonomy in consent discussions between genetic counselors or research nurses and patients may enhance informed decision making by increasing the precision of patients' understandings of the uncertainties in genome sequencing, such as contrasting reducible and nonreducible uncertainties of different types and significance. Furthermore, in the disclosure of CGNS results, the taxonomy can help clinicians and patients understand what uncertainties need to be accepted and managed clinically and psychologically, and facilitate actions to accomplish these goals.

Identifying the dimensions of uncertainty in a consistent manner applicable to all kinds of medical interventions, genomic and nongenomic, can help various stakeholders to have a clearer and more realistic understanding of the value and limitations of both CNGS and other medical interventions as well as to make better informed decisions about their use. Our taxonomy of uncertainty in CNGS is an initial step toward this broader goal.

Furthermore, because medical knowledge is continually evolving and limited by the perspectives of different stakeholders, any taxonomy of uncertainty is at best provisional, incomplete, and subject to revision. Accordingly, we have launched a website that provides stakeholders with the opportunity to edit and refine the taxonomy to enhance its usefulness to the broader community. Comments will be stored, curated, and used to periodically update the taxonomy. It will thus be a dynamic and consensual effort undertaken by many different stakeholders.

The novelty of CNGS and the vast amount of genetic code yet to be deciphered make the uncertainty surrounding this technology inevitable. Our current conceptual taxonomy draws attention to this uncertainty; however, it is offered not to discourage use of CNGS but to provide a framework that can inform efforts to reduce uncertainties and ultimately promote

its confident and appropriate use. It remains for future work to refine and apply the taxonomy to efforts to study and utilize CNGS to improve health.

ACKNOWLEDGMENTS

The authors thank Dave Kanney and Tyra Wolfsberg for designing the Web-based platform for the taxonomy, Sharon Plon for her interview and suggestions for the manuscript, and Barbara Athens for early diagrams of the taxonomy. The Intramural Research Program at the National Human Genome Research Institute, National Institutes of Health, funded this study. R.C.G. is supported by NIH grants U01HD006500, U19HD077671, R01HG005092, R01HG02213, U01HG008685, R01-HG006615, R01CA154517, and U41HG006834.

DISCLOSURE

L.G.B. receives royalties from Genentech Corp., is an unpaid advisor to Illumina Corp., and receives honoraria from Wiley-Blackwell Inc. R.C.G. has received compensation from Illumina, Prudential, Helix, AIA, Invitae, and Roche for speaking or for advisory services. The other authors declare no conflict of interest.

REFERENCES

1. Han PK, Klein WM, Arora NK. Varieties of uncertainty in health care: a conceptual taxonomy. *Med Decis Making* 2011;31:828–838.
2. de Ligt J, Willemsen MH, van Bon BW, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 2012;367:1921–1929.
3. Biesecker LG, Green RC. Diagnostic clinical genome and exome sequencing. *N Engl J Med* 2014;370:2418–2425.
4. Bowker GC, Star SL. *Sorting things out: classification and its consequences*. MIT Press: Cambridge, MA, 1999.
5. Smithson M. *Ignorance and uncertainty: emerging paradigms*. Springer: Berlin, Germany, 1989.
6. Han PKJ. Conceptual, methodological, and ethical problems in communicating uncertainty in clinical evidence. *Med Care Res Rev* 2013;70:145–365.
7. Vassy JL, Korf BR, Green RC. How to know when physicians are ready for genomic medicine. *Sci Transl Med* 2015;7:287fs19.