



# HHS Public Access

Author manuscript

*Genet Med.* Author manuscript; available in PMC 2017 August 07.

Published in final edited form as:

*Genet Med.* 2017 August ; 19(8): 918–925. doi:10.1038/gim.2016.212.

## A Taxonomy of Medical Uncertainties in Clinical Genome Sequencing

**Paul K. J. Han, MD, MPH<sup>1</sup>, Kendall L. Umstead, MS<sup>2</sup>, Barbara A. Bernhardt, MS<sup>3</sup>, Robert C. Green, MD, MPH<sup>4,5,6,7</sup>, Steven Joffe, MD, MPH<sup>8</sup>, Barbara Koenig, PhD<sup>9</sup>, Ian Krantz, MD<sup>10</sup>, Leo B. Waterston, MA<sup>1</sup>, Leslie G. Biesecker, MD<sup>11</sup>, and Barbara B. Biesecker, PhD<sup>2</sup>**

<sup>1</sup>Center for Outcomes Research and Evaluation, Maine Medical Center Research Institute, Portland, ME, USA

<sup>2</sup>Social and Behavioral Research Branch, National Human Genome Research Institute, NIH, Bethesda, MD, USA

<sup>3</sup>Division of Translational Medicine and Human Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

<sup>4</sup>Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA

<sup>5</sup>Personalized Medicine, Partners Healthcare, Cambridge, MA, USA

<sup>6</sup>Department of Medicine, Harvard Medical School, Boston, MA, USA

<sup>7</sup>Broad Institute of MIT and Harvard, Boston, MA, USA

<sup>8</sup>Department of Medical Ethics and Health Policy, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

<sup>9</sup>Institute for Health & Aging, University of California, San Francisco, CA, USA

<sup>10</sup>Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

<sup>11</sup>Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

### Abstract

**Purpose**—Clinical next generation sequencing (CNBS) is introducing new opportunities and challenges into the practice of medicine. Simultaneously, these technologies are generating uncertainties of unprecedented scale that laboratories, clinicians, and patients are required to

---

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

To whom correspondence should be addressed: Barbara B. Biesecker, Ph.D., [barbara.biesecker@nih.gov](mailto:barbara.biesecker@nih.gov), 31 Center Drive, Building 31, Room B1B36, Bethesda, MD 20892-2073.

#### CONFLICTS OF INTEREST

P.K.J.H., K.L.U., B.A.B., S.J., B.K., I.K., L.B.W., and B.B.B. have no conflicts to disclose. 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 for speaking or for advisory services from Illumina, Prudential, Helix, AIA, Invitae, and Roche.

address and manage. We describe in this report the conceptual design of a new taxonomy of uncertainties around the use of CNGS in health care.

**Methods**—Interviews to delineate the dimensions of uncertainty in CNGS were conducted with genomics experts, and themes were extracted in order to expand upon a previously published three-dimensional taxonomy of medical uncertainty. In parallel we developed an interactive website to disseminate the CNGS 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 CNGS 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 CNGS.

### Keywords

uncertainty; genome sequencing; exome sequencing; next-generation sequencing; taxonomy

---

## INTRODUCTION

The arrival of clinical next generation sequencing (CNGS) introduces 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.<sup>1,2</sup> Genomic testing has successfully advanced the diagnosis of previously undiagnosed rare disorders, identification of high risk of disease among healthy individuals, and development of targeted cancer treatments.

Although medicine is replete with uncertainties, genomic technologies introduce uncertainties of unique scale and type.<sup>1,3</sup> 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 CNGS is being used for diagnosis, predispositional testing, preconception screening, prenatal screening, pharmacogenomic profiling, or some other purpose, the introduction of CNGS into the clinic requires laboratories, clinicians, patients, and families to address and manage the many uncertainties it raises. Yet, the first step in this endeavor is

to distinguish and understand these specific uncertainties and 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.<sup>4</sup> It will also guide future research by helping identify and close key knowledge gaps. The current study used as its starting point a conceptual taxonomy of uncertainty in health care.<sup>1</sup> 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.<sup>5</sup> 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.<sup>1</sup> The taxonomy further divides each of these dimensions into more specific sub-dimensions by which various uncertainties can be categorized. The long-term goal of this multi-dimensional 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.<sup>1,6</sup>

## MATERIALS AND METHODS

The current project was conceptual in nature, and aimed 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 by one of the investigators (PKJH) between June and August 2014 with six genomics, genetic counseling, and bioethics experts from among the NHGRI-funded Clinical Sequencing Exploratory Research (CSER) 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 (BBB, PKJH), 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 to 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/](http://research.nhgri.nih.gov/taxonomy/)). The purpose of this website is to disseminate the new uncertainty taxonomy and collaboratively engage researchers in its further refinement. The interactive web page is curated by three of the authors (BBB, LGB, PKJH), 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 range 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 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; 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 sub-types: 1) conceptual, 2) methodological, and 3) clinical. *Conceptual ambiguity* arises from two main sources. The first is *model inadequacy*: limitations in either the theoretical or 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.<sup>4</sup>

*Methodological ambiguity*, the second principal sub-type of ambiguity, arises from five main sub-sources: 1) sample or data integrity problems, 2) test limitations, 3) unmeasured factors, 4) procedural variability or error, and 5) test misinterpretation. *Sample or data integrity problems* represent deficiencies in laboratory samples or processing, which result in diagnostic error. *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 sub-type 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 sub-types reflecting distinct origins: 1) multiplicity of causes, 2) multiplicity of effects, and 3) effect modification. *Multiplicity of causes*, in turn, arises from three specific sub-sources. 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 *non-genetic causation*: the presence of non-genetic (i.e., environmental) determinants of disorders or phenotypic characteristics, which may interact with genetic determinants and have poorly quantified effects.

*Multiplicity of effects*, the second principal sub-type of complexity, originates from *pleiotropy*: a phenomenon in which a single gene mutation causes multiple apparently unrelated disorders or phenotypic characteristics (e.g., Neurofibromatosis).

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

Importantly, any or all of these major sources of uncertainty—i.e., probability, ambiguity, complexity—and the more specific sub-types 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 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.”<sup>1</sup>

In CNGS *scientific uncertainty* pertains to many different issues—diagnostic, prognostic, causal, or therapeutic. *Diagnostic issues* include: 1) *Gene-phenotype association*—identification of variants in a gene that cause disease and detection of a phenotype; 2) *Pathogenicity of variants*—whether or not a variant causes disease; and 3) *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 on 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 over-simplified 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 ought to 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 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 a number of different stakeholders: patients, clinicians, researchers, or other individuals including family members, regulators, payers, or health policymakers (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, 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 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**

To illustrate the potential value of our provisional conceptual taxonomy of uncertainty, we present a case (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—resulting in varying degrees of uncertainty—by individual stakeholders, each of whom represents a distinct locus 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 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 fall in several places in our conceptual taxonomy, and influence the content of the counseling she provides to the 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—notably his personal uncertainties—initially reside in only his mind and not the internist’s, because it is Mark whose own life 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, encompassing more specific questions and precisely defined sources of ignorance. Yet it is also more generic—i.e., focused on generalizable knowledge

affecting populations more than specific knowledge affecting individuals. Whereas the uncertainties experienced by Mark and his internist focus on the implications of CNGS for Mark himself, the uncertainties of laboratory professionals, researchers, and policymakers 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 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 on 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 a reminder of important and salient issues that the discussion should address to maximize the 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—e.g., doubts about whether having his variant will affect his lifestyle, or about 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 that reduces 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 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. Researchers, on the other hand, may focus on ambiguity arising from conceptual problems (e.g., the distinctions between normal and abnormal phenotypes, non-pathogenic and pathogenic genotypes) or methodological problems (e.g., how sequencing is conducted, how sequencing data are analyzed and interpreted). Payers may be most concerned about whether recommended testing meets their internal criteria for reimbursement. Health policymakers 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.<sup>1,6</sup> The appropriate breadth and depth of uncertainty to convey to different stakeholders remains to be determined, and 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 upon an existing taxonomy of medical uncertainty,<sup>1</sup> 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 exist but may escape their awareness, and to thereby determine the best course of action—e.g., information seeking and further scientific research for uncertainties that are reducible, and strategies to promote coping and adaptation for uncertainties that are irreducible.

Indeed, we believe that a taxonomic approach to uncertainty may benefit other domains of medicine, and 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.<sup>7</sup> 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' understanding of the uncertainties in genome sequencing—e.g., contrasting reducible and non-reducible uncertainties of different types and significance. Furthermore, in the disclosure of CNGS 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 non-genomic, can help various stakeholders have a clearer and more realistic understanding of the value and limitations of both CNGS and other medical interventions, and 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 continuously 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 makes 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 in 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. Dr. Green is supported by NIH grants: U01HD006500, U19HD077671, R01HG005092, R01HG02213, U01HG008685 and R01-HG006615, R01CA154517 and U41HG006834.

## References

1. Han PKJ, Klein WMP, Arora NK. Varieties of uncertainty in health care: a conceptual taxonomy. *Med Decis Making*. 2011; 31:828–838. [PubMed: 22067431]

2. de Ligt J, Willemsen MH, van Bon BWM, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med.* 2012; 367:1921–1929. [PubMed: 23033978]
3. Biesecker LG, Green RC. Diagnostic clinical genome and exome sequencing. *N Engl J Med.* 2014; 370:2418–2425. [PubMed: 24941179]
4. Bowker, GC., Star, SL. *Sorting things out: classification and its consequences.* Cambridge: MIT Press; 1999.
5. Smithson, M. *Ignorance and uncertainty: emerging paradigms.* Berlin: Springer Verlag; 1989.
6. Han PKJ. Conceptual, methodological, and ethical problems in communicating uncertainty in clinical evidence. *Med Care Res Rev.* 2013; 70:14S–36S. [PubMed: 23132891]
7. Vassy JL, Korf BR, Green RC. How to know when physicians are ready for genomic medicine. *Sci Transl Med.* 2015; 7:287fs19.

**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 Mark 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.

**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 by 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 and what dose.*

**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, multi-faceted, 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 and 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.*



**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 by environment interactions: *Environmental and behavioral factors may moderate the risk of drug-induced myopathy.*

**II. Issue****A Scientific****1. Diagnostic**

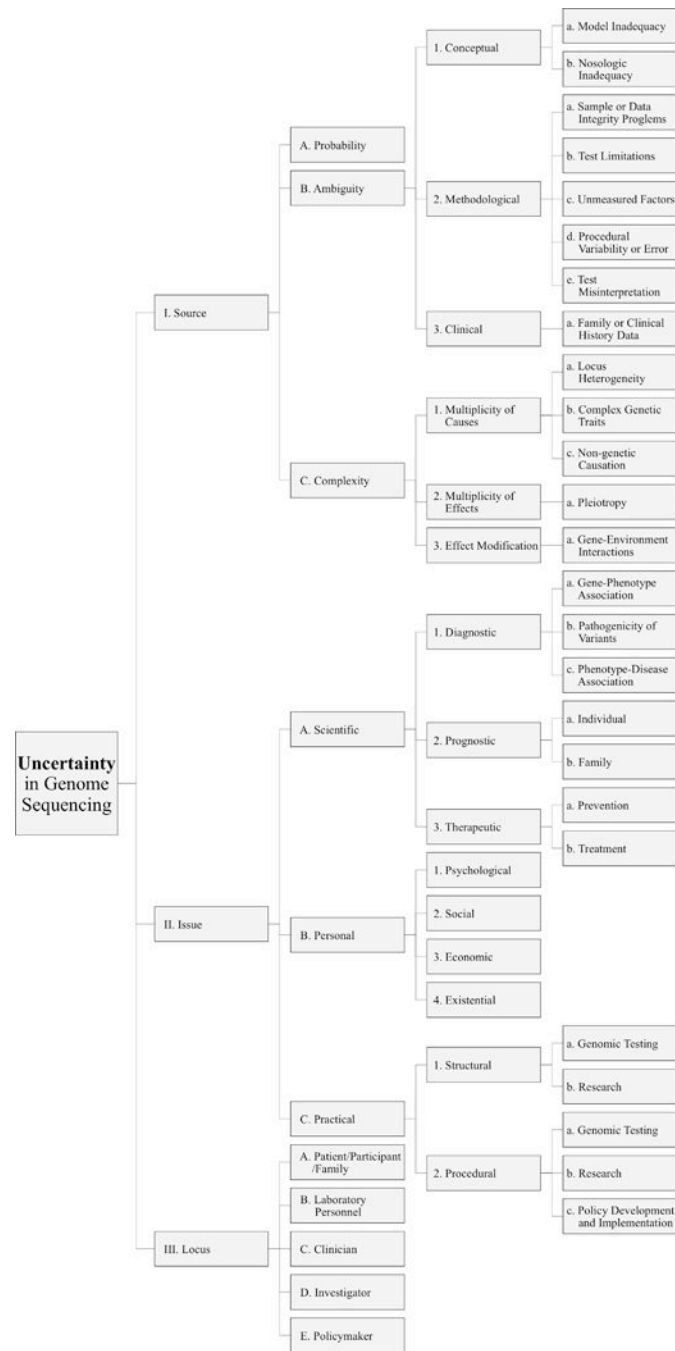
- b** Pathogenicity of variants: *Whether a given genomic variant will result in drug-induced myopathy.*

**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.*



**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.