

Open

A semiquantitative metric for evaluating clinical actionability of incidental or secondary findings from genome-scale sequencing

Jonathan S. Berg, MD, PhD¹, Ann Katherine M. Foreman, MS¹,
Julianne M. O'Daniel, MS¹, Jessica K. Booker, PhD², Lacey Boshe, BA¹,
Timothy Carey, MD³, Kristy R. Crooks, PhD², Brian C. Jensen, MD⁴,
Eric T. Juengst, PhD^{1,3,5}, Kristy Lee, MS¹, Daniel K. Nelson, PhD³,
Bradford C. Powell, MD, PhD¹, Cynthia M. Powell, MD⁶, Myra I. Roche, MS^{1,3,6},
Cecile Skrzynia, MS¹, Natasha T. Strande, PhD¹, Karen E. Weck, MD²,
Kirk C. Wilhelmsen, MD, PhD¹ and James P. Evans, MD, PhD¹

Purpose: As genome-scale sequencing is increasingly applied in clinical scenarios, a wide variety of genomic findings will be discovered as secondary or incidental findings, and there is debate about how they should be handled. The clinical actionability of such findings varies, necessitating standardized frameworks for a priori decision making about their analysis.

Methods: We established a semiquantitative metric to assess five elements of actionability: severity and likelihood of the disease outcome, efficacy and burden of intervention, and knowledge base, with a total score from 0 to 15.

Results: The semiquantitative metric was applied to a list of putative actionable conditions, the list of genes recommended by the American College of Medical Genetics and Genomics (ACMG) for return when deleterious variants are discovered as secondary/incidental

findings, and a random sample of 1,000 genes. Scores from the list of putative actionable conditions (median = 12) and the ACMG list (median = 11) were both statistically different than the randomly selected genes (median = 7) ($P < 0.0001$, two-tailed Mann-Whitney test).

Conclusion: Gene–disease pairs having a score of 11 or higher represent the top quintile of actionability. The semiquantitative metric effectively assesses clinical actionability, promotes transparency, and may facilitate assessments of clinical actionability by various groups and in diverse contexts.

Genet Med advance online publication 13 August 2015

Key Words: actionability; genome-scale sequencing; genomic medicine; incidental findings; secondary findings

INTRODUCTION

Genome-scale sequencing inevitably leads to the identification of many genomic variants with vastly differing clinical relevance, which requires development of innovative categorical approaches for informed consent, analysis, and return of results. Clinical genomic sequencing for suspected monogenic disorders may identify millions of genetic variants in a single patient, with only one or two “diagnostic” variants likely to explain the molecular etiology (“primary” results). Virtually all of the remaining variants are “incidental” to the original indication for analysis, although the term “secondary findings” is now the preferred term for such results when sought in a systematic fashion.¹

We previously proposed a framework for organizing potential incidental/secondary findings into “bins” categorized by clinical validity and clinical utility² and developed provisional lists of binned genes.³ Our goal is to categorize potential findings

before their discovery in a patient to guide informed decision making and return of results. As part of a National Human Genome Research Institute–funded Clinical Sequencing Exploratory Research project called “North Carolina Clinical Evaluation by Next-gen Exome Sequencing (NCGENES),” we assembled a Locus-Variant Binning Committee (LVBC) to refine a category of genomic findings that we call “bin 1”—the list of clinically actionable genes to be analyzed for pathogenic variants and returned as part of the routine results.⁴ Similar efforts are underway at other institutions and organizations.^{5–7}

Recognizing that an expert consensus-based approach without a clear definition and framework for adjudicating actionability could lead to inconsistent and arbitrary results, the LVBC developed a semiquantitative metric for determining the clinical actionability of gene–disease pairs. This metric explicitly recognizes that actionability is a continuum, not a binary

¹Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; ²Department of Pathology and Laboratory Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; ³Department of Social Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; ⁴Division of Cardiology, Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; ⁵Center for Bioethics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; ⁶Division of Genetics and Metabolism, Department of Pediatrics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. Correspondence: Jonathan S. Berg (jsberg@med.unc.edu)

state.^{8,9} That being said, we think that it is vital to define a core set of gene–disease pairs that reach a sufficient threshold of clinical actionability to be considered as part of the routine results of a genome-scale diagnostic test.

In parallel to the efforts of NCGENES, the Evidence-based Genomic Applications in Practice and Prevention Working Group established an evidence-based review procedure consisting of a rapid, sensitive screen for genes with possible actionability; structured data gathering organized around the elements of actionability articulated by the LVBC and detailed herein; and provided assessment by an expert deliberative committee.¹⁰ Such a framework will be most useful, not for definitively determining actionability, but rather for identifying the minority of genes in the human genome that should undergo further scrutiny as possibly actionable in a given specific context.

The 2013 recommendations for analysis and return of certain highly actionable incidental/secondary findings by the American College of Medical Genetics and Genomics (ACMG) used a deliberative consensus method to identify

gene–disease pairs within which clearly pathogenic variants should be returned as part of clinical genome-scale sequencing.⁵ These recommendations were met with criticisms,^{11–13} among which were concerns about the process by which the recommended gene list was developed. Also noted were concerns that some genes on the recommended list may not reach an evidentiary threshold sufficient to justify being returned as incidental/secondary findings. The development of a clear framework for the assignment of clinical actionability is therefore a necessary step toward formalizing such judgments and making assessments reproducible and updatable.

MATERIALS AND METHODS

Semiquantitative metric categories and scoring rules

The LVBC established five core characteristics of clinical actionability, with particular emphasis on the ramifications of finding a clearly pathogenic variant in a person without signs or symptoms of the disease (Table 1). The five characteristics are reflected by the following questions:

Table 1 Semiquantitative metric framework, questions, scores, and examples

Category	Level	Score	Notes/examples
Severity of disease: “What is the nature of the threat to health for an individual carrying a deleterious allele in this gene?”	Sudden death or inevitable death	3	Cardiac arrhythmia, vascular dissection, fatal infantile neurodegenerative conditions
	Possible death due to illness or comorbidity	2	Cancer, organ failure, moderate to severe intellectual disability
	Modest morbidity	1	Mild to moderate intellectual disability, physical limitations, early-onset neurosensory deficits
	Minimal health impact	0	Benign enzyme defects, nonmedical traits, later-onset neurosensory deficits
Likelihood of disease: “What is the chance that a serious threat will materialize?” (somewhat akin to penetrance)	>50%	3	Most individuals develop the severe outcome
	6–49%	2	Some individuals develop the severe outcome
	1–5%	1	Few individuals develop the severe outcome
	<1%	0	Outcome is very rare or cannot be reasonably estimated
Efficacy of intervention: “How effective are the interventions for preventing harm in a presymptomatic individual?”	Highly effective	3	Nearly all individuals have substantial benefit
	Moderately effective	2	The majority of individuals have some benefit
	Minimally effective	1	The majority of individuals have marginal benefit, or the minority of individuals have substantial benefit
	Ineffective/no interventions available	0	No individuals benefit; only watchful waiting recommended, or symptomatic treatments when disease manifests
Burden of intervention: “What are the burdens or potential harms of initiating interventions in a presymptomatic individual?”	Very low burden	3	Yearly screenings, routine medications, minor dietary/lifestyle modification
	Somewhat burdensome	2	Invasive screening, significant lifestyle alteration, medications with a substantial chance of side effects or more intensive delivery regimens, transplantation with rare complications
	Moderately burdensome	1	Removal of a nonvital organ, transplantation with frequent complications
	Highly burdensome	0	Removal of a vital organ
Knowledge base: “What is the evidence base for decisions about the natural history of the disease and interventions used for preventing serious outcomes?”	Substantial evidence	3	All categories scored confidently, high-quality review or practice guideline
	Moderate evidence	2	Strong primary literature, some details scored by analogy to another well-known disorder
	Minimal evidence	1	Unable to confidently score one or more categories, sparse primary literature or few reported patients
	Controversial or poor evidence	0	Uncertain natural history of disease, primary literature lacking or controversial

1. Severity: “What is the nature of the potential adverse health outcome in an individual carrying a deleterious allele in this gene?” Severity is scored from minimal health impact to modest morbidity to sudden/inevitable death.
2. Likelihood: “What is the chance that this adverse outcome will manifest?” Scoring for this category uses brackets of likelihood and is similar to penetrance.
3. Efficacy of intervention: “How effective are the established interventions for preventing the harmful outcome?” Efficacy of the intervention is scored from lack of demonstrable efficacy to highly effective intervention.
4. Burden of intervention: “How acceptable are the interventions in terms of the burdens or risks placed on the individual?” The burden or acceptability of the intervention is scored from highly consequential to minimally burdensome intervention.
5. Knowledge base: “How much is known about the gene, condition, and intervention to allow scoring in each category?” Knowledge is scored from controversial or poor evidence to substantial evidence.

All five criteria are scored on a scale of 0–3. The “outcome” and “intervention” are defined in advance and the other components of the metric are scored with respect to these parameters. It is critical to consider outcomes together with corresponding interventions to balance the clinical effects expected by natural history against the benefits and harms of these interventions in individuals who have not manifested symptoms of disease.

Gene sets scored

To judge the ability of the metric to distinguish between conditions that vary widely in terms of clinical actionability, three lists of genes were scored: (i) a list of 161 provisionally actionable genes³ (hereafter referred to as “bin 1 genes”); (ii) a list of 57 genes originally recommended by the ACMG⁴ (hereafter referred to as “ACMG genes”); and (iii) a list of 1,000 genes randomly selected from the National Center for Biotechnology Information RefSeq database (hereafter referred to as “random genes”). The random genes were selected from a 7 October 2013 RefSeq download using an in-house python script utilizing the “random” module. They were cross-referenced against Online Mendelian Inheritance in Man (<http://www.omim.org>) and Orphanet (http://www.orphanet.org/cgi-bin/inc/ordo_orphanet.inc.php) and manually curated to identify those with disease associations as of that time. The majority of the random genes (889/1,000) had no documented disease association, were associated only with somatic mutations, or had a modest influence on disease risk based on association study data. These conditions scored 0 by default and were excluded from further analysis.

Although the ACMG’s recommended list has subsequently been reduced to 56, all 57 original genes were analyzed with the expectation that the removed gene (*NTRK1*) would prove to be an outlier with regard to clinical actionability.

In addition, a list of “other” gene–disease pairs were scored, including conditions with phenotypes overlapping those of genes considered to be potentially actionable, disorders that are allelic to others that were scored, or conditions that were selected to evaluate the range of scores obtained for conditions considered not to be actionable. Scores for these “other genes” are included in the overall analysis but were not subject to statistical comparisons between lists because of their heterogeneity.

Assessment and consensus scoring

The multidisciplinary LVBC included clinical geneticists, genetic counselors, physicians from other specialties such as cardiology and neurology, a primary care physician, clinical laboratorians, and ethicists. Information about gene–disease relationships was obtained from OMIM, GeneReviews,¹⁴

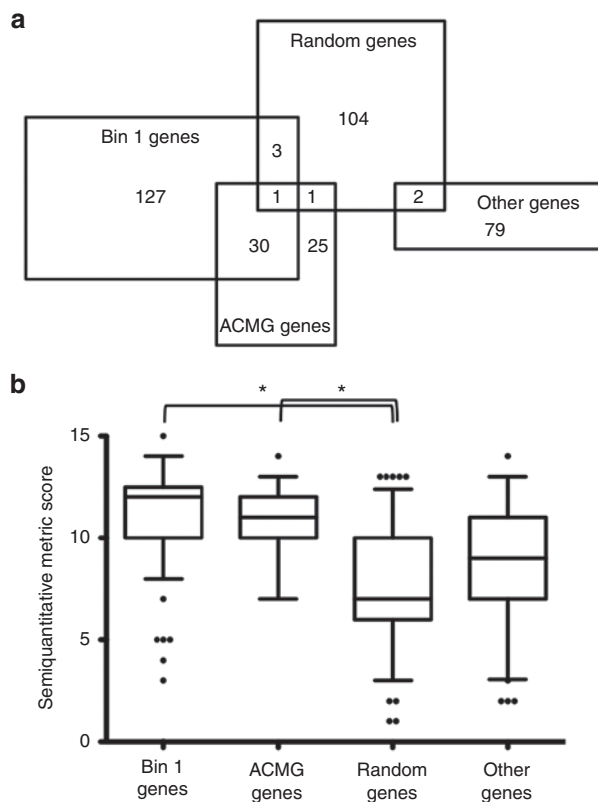


Figure 1 Semiquantitative metric scores. (a) Summary of overlap between the gene lists analyzed. The American College of Medical Genetics and Genomics (ACMG) genes included 26 gene–disease pairs (25 not including *NTRK*) that were not among the bin 1 genes. Conversely, 130 bin 1 genes were not among the ACMG genes. Of the 111 random genes with a defined disease association, 4 overlapped with the bin 1 genes (*ANK2*, *BRIP1*, *COL1A2*, *PROC*), 1 overlapped with the ACMG genes (*NTRK1*), and 1 gene overlapped both lists (*PTEN*). The Locus-Variant Binning Committee also evaluated 80 other gene–disease pairs, including alternative phenotypes for some genes, or different genes with similar disease phenotypes. One of these genes was on the ACMG list (*NF2*) and two were on the random list (*CASQ2* and *MAX*). (b) Distribution of semiquantitative metric scores. Box-whisker plots showing the median, 25th–75th percentiles (box), and 5th–95th percentiles (whiskers) of the scores for the bin 1 gene list, ACMG gene list, random gene list, and other gene list. Asterisks indicate statistically significant differences ($P < 0.0001$, two-tailed Mann-Whitney test).

Table 2 Examples of semiquantitative metric scores for selected genes

Gene	OMIM phenotype	List	Severity	Likelihood	Efficacy	Burden	Knowledge base	Total
<i>ACADM</i>	201450	Bin 1	3	3	3	3	3	15
Notes: Medium-chain acyl-CoA dehydrogenase deficiency is associated with a high lifetime risk of hypoglycemic episodes provoked by fasting or illness. These episodes can be fatal but are highly preventable through avoidance of fasting or provision of intravenous fluids with dextrose during illnesses. ¹⁸								
<i>MLH1</i>	609310	Bin 1, ACMG	2	3	3	2	3	13
Notes: Lynch syndrome is a highly penetrant cancer predisposition syndrome in which individuals are at risk for colorectal cancer as well as other cancers. Increased colonoscopic screening is highly effective at preventing deaths due to colon cancer. ¹⁹								
<i>FBN1</i>	154700	Bin 1, ACMG	3	2	2	3	3	13
Notes: Marfan syndrome is characterized by skeletal, eye, and cardiovascular involvement; the most severe aspect is increased risk for aortic aneurysm and dissection. Screening of the aortic root and arch for evidence of dilation can effectively identify individuals at high risk for dissection and allow initiation of preemptive definitive surgical management. ²⁰								
<i>LDLR</i> (Heterozygous)	143890	Bin 1, ACMG	2	3	2	3	3	13
Notes: Familial hypercholesterolemia results in elevated risk for death due to coronary artery disease. Treatment with lipid-lowering medications can somewhat mitigate the risk in affected individuals. ²¹								
<i>F8</i>	306700	Bin 1	2	3	3	2	3	13
Notes: Hemophilia presents a risk for severe bleeding that can result in severe chronic morbidity or fatality. Recombinant clotting factor is an effective preventive measure. ²²								
<i>HFE</i> (C282Y homozygous)	235200	Bin 1	2	1	3	3	3	12
Notes: Hemochromatosis leads to iron overload that can cause cirrhosis, cardiomyopathy, and endocrine dysfunction. Although biochemical evidence of abnormal iron homeostasis is seen in the majority of patients, less than 10% develop severe end-organ manifestations. Biochemical screening followed by therapeutic phlebotomy is highly effective for reducing morbidity. ²³								
<i>PTEN</i>	153480	Bin 1, ACMG, random	2	3	2	3	2	12
Notes: PTEN hamartoma syndrome is associated with an increased risk for several malignancies, including breast cancer. Penetrance is high, and it is expected that increased screening will benefit at-risk individuals, by analogy to other cancer predisposition syndromes. However, specific data regarding screening protocols are lacking. ²⁴								
<i>CASQ2</i>	611938	Random, other	3	3	2	3	1	12
Notes: Catecholaminergic polymorphic ventricular tachycardia is characterized by episodes of arrhythmia induced by adrenergic stress. Patients can present with syncope or sudden cardiac death. Limited data suggest that intervention with β -blockers and flecainide can be effective, and in some cases implantable cardioverter-defibrillator placement is required. ^{25,26}								
<i>MYH7</i>	115196	ACMG	3	1	3	3	2	12
Notes: Hypertrophic cardiomyopathy can lead to sudden cardiac death in a small proportion of affected individuals. Management includes cardiology surveillance and risk stratification, with more aggressive interventions in those who develop symptoms. ^{27,28}								
<i>SLC2A1</i>	606777	Random	2	3	2	2	2	11
Notes: Glucose transporter type 1 deficiency syndrome has a broad phenotypic spectrum including individuals with seizures and/or complex movement disorder. Symptoms show substantial improvement with a ketogenic diet. ^{29,30}								
<i>F5</i> Leiden (homozygous)	188055	Other	2	3	2	2	2	11
Notes: This specific mutation in the <i>F5</i> gene causes resistance to cleavage and inactivation by protein C. In the homozygous state, the result is a substantially elevated risk for venous thrombosis, which in some cases can lead to mortality due to pulmonary embolism. Awareness of this tendency allows measures to be taken to prevent immobility, reducing the chance of a clot developing. ^{31,32}								
<i>SDHB</i>	115310	Bin 1, ACMG	2	2	1	3	2	10
Notes: Heterozygous pathogenic variants in <i>SDHB</i> cause a syndrome of predisposition to paragangliomas, which can become malignant. Biochemical screening and imaging are recommended in asymptomatic individuals, with the expectation that this protocol would be somewhat effective for detecting tumors at smaller size and earlier stage. ^{33,34}								
<i>ALB</i>	615999	Random	1	1	3	3	2	10
Notes: Individuals with familial dysalbuminemic hyperthyroxinemia are clinically euthyroid, although a variant in the albumin gene leads to preferential affinity for thyroxine (T4). Some patients have been mistakenly treated for hyperthyroidism, leading to modest morbidity, which could be avoided simply by foreknowledge about this relatively benign phenotype. ³⁵								

ACMG, American College of Medical Genetics and Genomics; PTEN, phosphatase and tensin homolog.

^a*NTRK1* was included in an original version of the ACMG publication but removed from the final published list.

Table 2 Continued on next page

Table 2 Continued

Gene	OMIM phenotype	List	Severity	Likelihood	Efficacy	Burden	Knowledge base	Total
<i>FLCN</i>	135150	Bin 1	2	1	2	3	1	9
Notes: Birt-Hogg-Dubé syndrome is a hereditary cancer predisposition syndrome associated with benign hamartomatous skin lesions, benign and malignant kidney neoplasms, and lung cysts leading to spontaneous pneumothorax. Screening protocols are recommended and assumed to be somewhat effective at detecting kidney tumors, but the evidence base for screening is limited. ³⁶								
<i>TNNI3</i>	613690	ACMG	3	1	1	3	1	9
Notes: As with <i>MYH7</i> , the <i>TNNI3</i> gene is associated with cardiomyopathy and risk for sudden death. ²⁷ However, the degree to which screening can be an effective strategy for preventing this outcome in individuals with <i>TNNI3</i> pathogenic variants is less well known.								
<i>MYLK</i>	603776	ACMG	3	0	1	3	1	8
Notes: As with <i>FBN1</i> , pathogenic variants in <i>MYLK</i> are reported to be associated with increased risk of aortic dissection. However, the penetrance is essentially unknown and the pathophysiology seems to involve dissections and not aneurysms, limiting the efficacy of screening. ³⁷ The overall knowledge base about <i>MYLK</i> -associated disease is somewhat limited.								
<i>GCK</i>	125851	Bin 1	1	1	1	3	2	8
Notes: Maturity-onset diabetes of the young results in a rare form of insulin-dependent diabetes. However, these individuals rarely exhibit diabetic ketoacidosis, and overall complications of diabetes are low in this disorder. ³⁸								
<i>NAGLU</i>	252920	Random	3	3	0	0	2	8
Notes: Pathogenic variants in <i>NAGLU</i> cause mucopolysaccharidosis type IIIB (Sanfilippo B), a lysosomal storage disease that leads to significant morbidity and mortality but has no effective preventive measures at this time. Supportive treatments for symptomatic manifestations are the mainstay of care but would not be expected to substantially alter outcomes. ³⁹								
<i>NTRK1</i>	155240	ACMG ^a	2	0	3	2	0	7
Notes: The <i>NTRK1</i> gene has been implicated in predisposition to medullary thyroid carcinoma, which in principle could be effectively prevented through prophylactic thyroidectomy. However, the evidence supporting causality of germ-line variants in cancer predisposition is weak, and there is insufficient data with which to estimate penetrance. ⁴⁰								

ACMG, American College of Medical Genetics and Genomics; PTEN, phosphatase and tensin homolog.

^a*NTRK1* was included in an original version of the ACMG publication but removed from the final published list.

PubMed searches, and clinical guidelines, when available. Members of the LVBC prepared the evidence review, typically with a single member assigned primary responsibility for each gene–disease phenotype pair. The committee met regularly to review the evidence and to agree on a score for each element of the semiquantitative metric or direct additional review.

To mitigate the subjective nature of assessing certain categories and to enhance consistency between scores, the LVBC arrived at a series of scoring conventions (examples in Table 1). Scores for categories 1 (severity) and 2 (likelihood) are linked to the same specific outcome, either the most severe potential outcome or what is generally considered the primary outcome for the disease. However, scores for a given gene–disease pair can be calculated for more than one outcome of interest to account for disease pleiotropy. For example, different scores can be calculated for *BRCA1* depending on whether the outcome of interest is breast cancer or ovarian cancer. In effect, categories 1 and 2 reflect the medical implications of disease faced by an individual with a pathogenic finding. Scores for categories 3 (efficacy) and 4 (burden) reflect specific presymptomatic interventions targeted to the outcome described in categories 1 and 2 (e.g., bilateral risk-reducing mastectomy or bilateral salpingo-oophorectomy per the example of *BRCA1*). The semiquantitative metric thus approximates the concept of clinical utility by balancing the potential benefits and harms of intervention when an incidental/secondary finding is discovered in a presymptomatic individual.

RESULTS

A total of 1,213 unique genes were evaluated using the semi-quantitative metric (Supplementary Table S1 online). After removing genes not implicated in a single-gene disorder, 324 unique genes representing 372 gene–disease pairs were scored. These genes included 161 bin 1 genes, 57 ACMG genes, and 111 random genes associated with defined monogenic disorders. There was some degree of overlap between these lists, as depicted in Figure 1a. In cases where the random genes were associated with more than one condition, the highest of the scores was chosen to represent the random gene–disease pair; scores for additional gene–disease pairs were tallied in the “other” category.

The median score of the bin 1 genes was 12 (range 0–15); 84/161 gene–disease pairs scored ≥ 12 , while 29/161 pairs scored < 10 . The median score for the ACMG genes was 11 (range 7–14); 25/57 gene–disease pairs scored ≥ 12 , while 11/57 pairs scored < 10 . The *NTRK1* gene, originally included on ACMG’s preliminary recommended list and subsequently dropped, scored 7. In comparison, the median score of the 111 random genes was 7 (range 1–13); only 14/111 gene–disease pairs scored ≥ 12 , while 81/111 pairs scored < 10 . Figure 1b shows the distribution of scores for all of the pairs. The distributions of scores for the bin 1 genes and the ACMG genes are not significantly different from each other, but both lists are significantly different than the random genes ($P < 0.0001$, two-tailed Mann-Whitney test), indicating that the semiquantitative

metric effectively distinguishes between gene–disease pairs that were qualitatively deemed to be actionable in earlier efforts from those that would not be enriched for actionability. **Table 2** presents several scoring examples, and all scores are included in **Supplementary Table S1** online.

Among the gene–disease pairs scoring highest using this metric were *MLH1* (associated with Lynch syndrome) at 13 and *RYR1* (malignant hyperthermia) at 12. Despite its low penetrance, the *HFE* gene (implicated in hereditary hemochromatosis) scored 11 because of the availability of highly effective and noninvasive preventive measures. By contrast, some genes that were considered actionable by the ACMG, such as *SDHB*, *SDHC*, and *SDHD* (associated with hereditary pheochromocytoma/paraganglioma susceptibility) received scores below 11 because of limited evidence that biochemical screening in an otherwise asymptomatic individual would produce better long-term outcomes than treatment upon onset of symptoms. Other genes, such as *MYH11* and *MYLK* (which have recently been implicated in familial thoracic aortic aneurysm and dissection), could be considered to have effective interventions by analogy to other well-known conditions, but they scored lower because of a limited knowledge base, which precludes accurate assessment of penetrance.

As demonstrated by the range of scores observed for the selected gene–disease pairs, actionability is a continuum rather than a binary state. Using the random genes as a benchmark, 21/111 (19%) scored ≥ 11 , while 30/111 (27%) scored ≥ 10 . The LVBC chose to consider genes with a score ≥ 11 , essentially the top quintile, as meeting the threshold of actionability for inclusion in the revised bin 1 list. This yields a list of 168 genes representing 176 gene–disease pairs from the total 372 pairs scored. The fact that 19% of random genes associated with single-gene disorders scored ≥ 11 suggests that as many as 500 genes of the $\geq 3,000$ single-gene disorders might rise to this threshold of actionability. Thus, we have not yet identified all of the “actionable” gene–disease pairs, and a systematic screen of single-gene disorders is needed. Furthermore, scores are subject to change depending on advances in medical genetics, which will likely increase some scores over time.

DISCUSSION

Management of the vast range of heterogeneous information generated when genomic analysis is undertaken remains one of the most challenging aspects of applying genomics in the clinical realm. Patient preferences must be taken into account, especially with regard to genomic findings that have limited clinical actionability. Individuals may make greatly varying choices regarding whether they want to learn about different types of genomic findings; we are studying these preferences and the parameters that influence them as part of the NCGENES study. However, just as there are incidentally discovered laboratory values that are flagged as “critical” levels, or radiographic findings that require clinical action, it follows that when certain genomic findings exceed a threshold of actionability, the default procedure should be to provide those results as part of

the routine protocol when performing clinical genome-scale sequencing tests.

It is thus critical to define a subset of clinically actionable genomic findings that are likely to be accepted by most individuals and allow a standardized and streamlined process for informed decision making in clinical genome-scale diagnostic testing. Otherwise, the decision about returning genomic findings could (at the reductio ad absurdum extremes) be relegated to an all-or-none choice irrespective of the actionability of the information, or a nearly infinite menu of potential findings organized at the level of certain genes or even specific variants. Neither of these options seems tenable in the current clinical setting.

It should be stressed that a policy of routine return of a small subset of genomic findings does not preclude patient choice by means of an informed “opt-out” at the initiation of testing, as now endorsed by the ACMG. In addition, this policy does not prohibit laboratories from offering additional categories of non–medically actionable genomic information (what we refer to as “bin 2”) as an “opt-in” to those who desire such information, with appropriate education and decision making. Nevertheless, any such menu of options needs to be articulated before consent and analysis, which calls for an a priori process to define which gene–disease pairs fall into any given category.

This article describes the delineation of a novel semiquantitative metric providing a transparent definition of clinical actionability and a framework for evaluating criteria in a streamlined fashion. We outline criteria for actionability generally similar to other expert deliberative processes.^{5,15} However, this framework is unique in that the dimensions of actionability can be assessed consistently across different types of disorders. The results indicate that, as expected, the ACMG list is enriched for genes that achieve high scores for actionability, both supporting their inclusion in a recommended list and generally reinforcing the parameters used in the current assessment. Future versions of the ACMG list could be informed by this scoring metric, or one similar to it, in order to remove genes that fall below a stringent threshold and to include additional genes with scores equivalent to those on the current recommended list. The percentage of individuals who will have such findings is predictable and depends on the list of gene–disease pairs being evaluated and the stringency with which variants are selected for return.^{3,15–17}

Nuances in application of the semiquantitative metric

The subcategories of the metric reflect the clinical impact of a condition (severity and likelihood of a given outcome) while balancing the potential benefits (effectiveness of interventions) and harms (burden of intervention), thus approximating the clinical utility of revealing incidental/secondary findings in a presymptomatic individual. Each of these facets is necessary to include, despite the subjectivity inherent in scoring some of them. For example, both periodic phlebotomy (as in the case of hemochromatosis) and surgical removal of the stomach to prevent diffuse gastric cancer (in the case of *CDH1* mutations) are highly effective measures to prevent morbidity and mortality,

yet the burdens of these interventions are dramatically different and therefore greatly affect the concept of actionability when considering the return of genetic information in a setting in which the individual is not likely to have overt symptoms.

Scoring each category on a 0–3 scale allows for a limited degree of granularity that captures the qualitative nature of certain categories (e.g., effectiveness and burdens of intervention). It would be difficult, and potentially problematic, to spread the range of scores into finer subdivisions or to develop a more complex nonlinear scoring system. That said, different scoring systems could be explored should there be a compelling rationale to do so. In addition, customized gene lists could be generated for other contexts by applying weighting schemes or selecting a different threshold. For example, one might envision that selection of genes for primary screening of the general population should demand extremely high knowledge and efficacy scores, and those components of the metric could be weighted accordingly. In general, differential weighting of criteria will change the rank order of scores that are close to one another, and will primarily affect gene–disease pairs near the threshold used to define actionability. However, those with scores farther away from the chosen threshold would be much less likely to have their position above or below the threshold affected by changes in weighting. Finally, the evidence used in establishing the scores can be explicitly defined, and the scores can be updated to incorporate new evidence. Thus, the semiquantitative metric provides a structured framework and a more nuanced and transparent method for defining a list of clinically actionable genes than would be possible with expert consensus approaches.

In practice, the LVBC established a set of conventions for scoring different categories (see [Table 1](#) for examples). It is challenging to directly compare the severity of conditions that lead to bodily harm, such as death or organ failure, with conditions that lead to physical or cognitive impairment. The severity score is thus intended to judge the relative severity between disparate conditions. The likelihood of a given outcome is the most quantitatively definable component of the metric, although data are lacking for many conditions, requiring either an estimate with some uncertainty (reflected in a lower score for knowledge base) or a score of 0 when the available data are simply too limited to make a reasonable assessment, as in the case of autosomal-dominant conditions with only a few patients reported in the medical literature.

In the absence of definitive end points, the effectiveness of interventions for different clinical outcomes often relies on expert opinion. The LVBC generally considered the screening or preventive measures that all individuals with a positive finding would be expected to undergo, rather than more definitive treatments that would be required only in those who manifest symptoms. By any measure, the burden of intervention is the most subjective and personally nuanced aspect of the metric. It is likely that different individuals hold different views on what is acceptable and what constitutes an unreasonable burden in the

context of their own life experiences. Thus, while we fully recognize that it can be difficult to assess the burden of a particular intervention for an individual, we attempted to define a scoring rubric that could roughly define the relative burdens of interventions across the population. This score could be replaced in the future by a more quantitative measure derived from discrete choice experiments or other means of assessing relative values, such as the methods used to measure quality-adjusted life-years.

Finally, the knowledge base score was applied as a single measure reflecting the degree to which each component of the score could be confidently defined. Alternatively, a knowledge score could be assigned to each component separately based on the knowledge base for that element. While more complicated, such an approach would provide greater granularity. The strength of the gene–disease association itself is embedded in the knowledge base score, since less well-described conditions rarely have sufficient knowledge to accurately score certain elements of the metric. However, we do not intend for this metric to be used as a stand-alone measure of the clinical validity of a gene–disease association.

Disorders that predispose to thoracic aortic aneurysm and dissection illustrate certain nuances of scoring. These conditions convey an increased risk of sudden death due to dissection, and the typical intervention is to implement a vascular imaging screening program. This intervention is a highly effective and noninvasive means to detect and monitor the size of an aneurysm before it poses significant danger of acute dissection. More invasive intervention (i.e., vascular surgery) is required only if the individual develops a clinically significant aneurysm. Overall, this combination of interventions would be a highly effective and generally acceptable way to manage the risk of sudden death due to an aortic dissection, although in some conditions the risk of dissection is not directly related to aneurysm size, in which case screening would be less effective. In addition, the effectiveness of an intervention for more rare conditions, such as *MYLK*-associated thoracic aortic aneurysm and dissection, must be extrapolated based on analogy to related conditions because of the lack of information available regarding the effectiveness of interventions specific to that condition.

Potential limitations

The current metric does not account for certain contextual factors, such as the age of the individual, the typical age at onset of disease or the age at which clinical actions would be implemented in a presymptomatic individual, the sex of the individual, the general availability and cost of recommended preventive strategies, or the ability of relevant genetic lesions to be detected. Given these limitations and the necessarily subjective nature of any assessment of actionability, the scores and evidence base generated by the application of the semiquantitative metric are best considered an initial starting point for more nuanced discussions about individual conditions or particular clinical applications.

Certain features of the metric could lead to minor irregularities in scoring. For example, if the LVBC decided that a proposed intervention for a given condition was considered ineffective (score = 0), then no score could be assigned to reflect the “burden” of that intervention because the additional points would inflate the final score. However, if the proposed intervention was considered even “minimally” effective, the total score could swing by as much as four points (effectiveness = 1 and burden = 3 points for a minimally invasive intervention). This impact could be partially mitigated by weighting schemes that emphasize the contribution of the effectiveness of the intervention over the burden score to the overall total.

The final scores for each gene–disease pair were determined through consensus of the LVBC. This process was not amenable to evaluation of interrater variability in scoring, but in practice we found that the semiquantitative metric facilitated more efficient discussions and greater consistency than earlier attempts to arrive at consensus without a structured framework. It could be argued that this process simply replaces a single idiosyncratic expert consensus about “actionability” (the current state of other deliberative processes) with several different potentially idiosyncratic decisions. However, assessment of scores for each criterion permits more systematic evidence curation and updating, as well as a more flexible approach to weighting the importance of different criteria, than would be possible otherwise.

Finally, in some cases review of the scores by domain experts may prompt revisions based on deeper understanding of the clinical scenario or greater awareness of literature that was not captured in our review process. In other cases medical advances may increase the overall scores by improving the knowledge base and perhaps the efficacy of interventions for many conditions. It is expected that there will be a need for ongoing assessment of clinical actionability and updating of results, which again is streamlined by the existence of a structured framework.

Conclusions and future questions

We have presented a framework that defines five aspects of clinical actionability, evaluates them qualitatively, and effectively distinguishes between lists of genes deemed to be potentially actionable by expert opinion versus randomly selected genes. This framework is flexible and can be adapted to different contexts. It is too early to know whether it is more efficient or more reliable than other expert consensus approaches, or whether it ultimately leads to better clinical outcomes. In addition, it remains to be determined whether other groups would arrive at the same scores. However, the inherent transparency of the framework facilitates comparison between different efforts, critical evaluation, and updating of the scores as new knowledge accrues as a result of the constantly evolving medical literature. We anticipate using this or a similar metric to evaluate all human disease genes to guide the application of genomic medicine.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

ACKNOWLEDGMENTS

This work was funded by the National Institutes of Health grant U01 HG006487.

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

1. Presidential Commission for the Study of Bioethical Issues. Anticipate and Communicate: Ethical Management of Incidental and Secondary Findings in the Clinical, Research, and Direct-to-Consumer Contexts. 2013. http://bioethics.gov/sites/default/files/FINALAnticipateCommunicate_PCSBI_0.pdf. Accessed 2 December 2014.
2. Berg JS, Khoury MJ, Evans JP. Deploying whole genome sequencing in clinical practice and public health: meeting the challenge one bin at a time. *Genet Med* 2011;13:499–504.
3. Berg JS, Adams M, Nassar N, et al. An informatics approach to analyzing the incidentalome. *Genet Med* 2013;15:36–44.
4. Berg JS, Amendola LM, Eng C, et al.; Members of the CSER Actionability and Return of Results Working Group. Processes and preliminary outputs for identification of actionable genes as incidental findings in genomic sequence data in the Clinical Sequencing Exploratory Research Consortium. *Genet Med* 2013;15:860–867.
5. Green RC, Berg JS, Grody WW, et al.; American College of Medical Genetics and Genomics. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* 2013;15:565–574.
6. Holm IA, Taylor PL. The Informed Cohort Oversight Board: From Values to Architecture. *Minn J Law Sci Technol* 2012;13:669–690.
7. van El CG, Cornel MC, Borry P, et al.; ESHG Public and Professional Policy Committee. Whole-genome sequencing in health care: recommendations of the European Society of Human Genetics. *Eur J Hum Genet* 2013;21:580–584.
8. Lindor NM, Johnson KJ, McCormick JB, Klee EW, Ferber MJ, Farrugia G. Preserving personal autonomy in a genomic testing era. *Genet Med* 2013;15:408–409.
9. Berg JS. Response to Lindor et al. *Genet Med* 2013;15:409–410.
10. Goddard KA, Whitlock EP, Berg JS, et al. Description and pilot results from a novel method for evaluating return of incidental findings from next-generation sequencing technologies. *Genet Med* 2013;15:721–728.
11. Burke W, Antommaria AH, Bennett R, et al. Recommendations for returning genomic incidental findings? We need to talk! *Genet Med* 2013;15:854–859.
12. Ross LF, Rothstein MA, Clayton EW. Premature guidance about whole-genome sequencing. *Per Med* 2013;10.
13. Wolf SM, Annas GJ, Elias S. Point-counterpoint. Patient autonomy and incidental findings in clinical genomics. *Science* 2013;340:1049–1050.
14. Pagon RA, Adam MP, Ardinger HH et al., eds. *GeneReviews*. University of Washington: Seattle, WA, 1993–2014.
15. Dorschner MO, Amendola LM, Turner EH, et al.; National Heart, Lung, and Blood Institute Grand Opportunity Exome Sequencing Project. Actionable, pathogenic incidental findings in 1,000 participants' exomes. *Am J Hum Genet* 2013;93:631–640.
16. Johnston JJ, Rubinstein WS, Facio FM, et al. Secondary variants in individuals undergoing exome sequencing: screening of 572 individuals identifies high-penetrance mutations in cancer-susceptibility genes. *Am J Hum Genet* 2012;91:97–108.
17. Lawrence L, Sincan M, Markello T, et al. The implications of familial incidental findings from exome sequencing: the NIH Undiagnosed Diseases Program experience. *Genet Med* 2014;16:741–750.
18. Wilcken B. Fatty acid oxidation disorders: outcome and long-term prognosis. *J Inher Metab Dis* 2010;33:501–506.
19. Järvinen HJ, Aarnio M, Mustonen H, et al. Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. *Gastroenterology* 2000;118:829–834.

20. Dean JC. Marfan syndrome: clinical diagnosis and management. *Eur J Hum Genet* 2007;15:724–733.
21. Versmissen J, Oosterveer DM, Yazdanpanah M, et al. Efficacy of statins in familial hypercholesterolaemia: a long term cohort study. *BMJ* 2008;337:a2423.
22. National Hemophilia Foundation. MASAC Recommendation Concerning Prophylaxis (Regular Administration of Clotting Factor Concentrate to Prevent Bleeding). <http://www.hemophilia.org/sites/default/files/document/files/masac179.pdf>. Accessed 2 December 2014.
23. Bacon BR, Adams PC, Kowdley KV, Powell LW, Tavill AS; American Association for the Study of Liver Diseases. Diagnosis and management of hemochromatosis: 2011 practice guideline by the American Association for the Study of Liver Diseases. *Hepatology* 2011;54:328–343.
24. Eng C. PTEN hamartoma tumor syndrome (PHTS). In: *GeneReviews® [Internet]*, Pagon RA, Adam MP, Ardinger HH et al., eds. University of Washington: Seattle, WA, 2014.
25. Pflaumer A, Davis AM. Guidelines for the diagnosis and management of catecholaminergic polymorphic ventricular tachycardia. *Heart Lung Circ* 2012;21:96–100.
26. Khoury A, Marai I, Suleiman M, et al. Flecainide therapy suppresses exercise-induced ventricular arrhythmias in patients with CASQ2-associated catecholaminergic polymorphic ventricular tachycardia. *Heart Rhythm* 2013;10:1671–1675.
27. Cirino AL, Ho C. Hypertrophic cardiomyopathy overview. In: *GeneReviews*, Pagon RA, Adam MP, Ardinger HH et al., eds. University of Washington: Seattle, WA, 2014.
28. Gersh BJ, Maron BJ, Bonow RO, et al.; American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines; American Association for Thoracic Surgery; American Society of Echocardiography; American Society of Nuclear Cardiology; Heart Failure Society of America; Heart Rhythm Society; Society for Cardiovascular Angiography and Interventions; Society of Thoracic Surgeons. 2011 ACCF/AHA guideline for the diagnosis and treatment of hypertrophic cardiomyopathy: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation* 2011;124:2761–2796.
29. Klepper J, Diefenbach S, Kohlschütter A, Voit T. Effects of the ketogenic diet in the glucose transporter 1 deficiency syndrome. *Prostaglandins Leukot Essent Fatty Acids* 2004;70:321–327.
30. Leen WG, Klepper J, Verbeek MM, et al. Glucose transporter-1 deficiency syndrome: the expanding clinical and genetic spectrum of a treatable disorder. *Brain* 2010;133(Pt 3):655–670.
31. Simioni P, Tormene D, Prandoni P, et al. Incidence of venous thromboembolism in asymptomatic family members who are carriers of factor V Leiden: a prospective cohort study. *Blood* 2002;99:1938–1942.
32. Varga E. Inherited thrombophilia: key points for genetic counseling. *J Genet Couns* 2007;16:261–277.
33. Andersen KF, Altaf R, Krarup-Hansen A, et al. Malignant pheochromocytomas and paragangliomas - the importance of a multidisciplinary approach. *Cancer Treat Rev* 2011;37:111–119.
34. Young AL, Baysal BE, Deb A, Young WF Jr. Familial malignant catecholamine-secreting paraganglioma with prolonged survival associated with mutation in the succinate dehydrogenase B gene. *J Clin Endocrinol Metab* 2002;87:4101–4105.
35. Heufelder AE, Klee GG, Wynne AG, Gharib H. Familial dysalbuminemic hyperthyroxinemia: cumulative experience in 29 consecutive patients. *Endocr Pract* 1995;1:4–8.
36. Pavlovich CP, Grubb RL 3rd, Hurley K, et al. Evaluation and management of renal tumors in the Birt-Hogg-Dubé syndrome. *J Urol* 2005;173:1482–1486.
37. Pomianowski P, Elefteriades JA. The genetics and genomics of thoracic aortic disease. *Ann Cardiothorac Surg* 2013;2:271–279.
38. Aijian RA, Owen KR. Glucokinase MODY and implications for treatment goals of common forms of diabetes. *Curr Diab Rep* 2014;14:559.
39. Valstar MJ, Ruijter GJ, van Diggelen OP, Poorthuis BJ, Wijburg FA. Sanfilippo syndrome: a mini-review. *J Inherit Metab Dis* 2008;31:240–252.
40. Gimm O, Greco A, Hoang-Vu C, Dralle H, Pierotti MA, Eng C. Mutation analysis reveals novel sequence variants in NTRK1 in sporadic human medullary thyroid carcinoma. *J Clin Endocrinol Metab* 1999;84:2784–2787.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>