

# Deploying whole genome sequencing in clinical practice and public health: Meeting the challenge one bin at a time

*Jonathan S. Berg, MD, PhD<sup>1</sup>, Muin J. Khoury, MD, PhD<sup>2</sup>, and James P. Evans, MD, PhD<sup>1</sup>*

Technological advances often outpace our ability to effectively use them, a situation that certainly could pertain to modern genomics. Breathtaking advances in genetic sequencing technology have the potential to make whole genome sequencing (WGS) available for healthcare and disease prevention. However, current practices in medical genetics are not directly applicable to robust genomic analysis, and new approaches are needed which are “scalable” to this new reality. If the field merely attempts to overlay traditional medical genetic approaches to patient consent and analysis, based on a soon to be obsolete model of testing and analyzing “one-gene-at-a-time,” it threatens to stall our ability to realize the promise of genomic medicine. The informed consent process, data analysis and clinical interpretation, and return of results must be achievable within a reasonable time frame, provide results in a manner consistent with responsible clinical genetics practice, and yet still comport with the realities of modern medicine. This challenge is illustrated by recent reports suggesting that the informed consent process could require 6 hours of face-to-face discussion over the course of several sessions<sup>1</sup> or that delivery of results would require as much as 5 hours of direct patient contact.<sup>2</sup> Thus, new approaches, thoughtfully developed around the unique features of whole genome analysis, are required. The challenges facing the deployment of WGS in clinical practice and public health, while substantial, are not insurmountable if we learn from the way in which other complex technologies are handled in medicine and move forward in an evidence-based manner.

## THE PROMISE OF WGS FOR IMPROVING HEALTH: A UNIVERSAL DIAGNOSTIC AND PUBLIC HEALTH TOOL

### Universal diagnostic testing

In the near future, WGS will transform diagnostic testing in the subset of patients with disorders resulting from disruption of a single gene or chromosomal region. Burgeoning application of WGS in a variety of clinical settings will allow assessment of the diagnostic yield in various subsets of symptomatic patients, guiding its widespread use in this setting. However, although WGS will almost certainly be a powerful diagnostic tool for patients with such disorders, whether such analysis will be a valuable clinical tool for those with common diseases is doubtful, for the simple reason that such disorders have many contributing nongenetic etiologies and because our ability to interpret the combinatorial effects of common genetic variants

remains limited.<sup>3,4</sup> Thus, it is likely that in the clinical setting, the initial use of WGS will have the greatest yield in those with evidence to suggest a highly penetrant, discrete genetic lesion.

### Screening of asymptomatic individuals

Currently, the first sign that an individual may harbor a rare, highly penetrant mutation strongly predictive of disease is when they or a family member manifests disease. In the subset of genetic diseases for which preventive measures are available, detection of such mutations before the onset of disease in a family could be greatly beneficial. The careful application of WGS could be a promising strategy for the identification of such individuals within populations and thus represents a potentially powerful public health application of this technology. If focused on discovering those variants that are medically actionable, the public health impact of WGS could be considerable in the near term as there currently are a number of loci that meet the necessary criteria for utility in the public health context. An illustrative example is Lynch syndrome, which renders approximately 0.1% of the US population at high lifetime risk for colorectal cancer and a variety of other malignancies.<sup>5,6</sup> Effective preventive protocols have been established, which result in decreased morbidity, mortality, and cost.<sup>7-9</sup> Similarly, other highly penetrant conditions exist for which we have effective preventive strategies. In total, approximately 1% of the US population likely harbors such deleterious mutations; detection of these individuals by WGS or multiplex gene panels and the initiation of established preventive strategies would be a promising early application of robust genomic analysis that could immediately impact millions of US citizens.

Other anticipated utilities of WGS in asymptomatic individuals include the preemptive identification of relevant pharmacogenomic (PGx) alleles<sup>10,11</sup> and the identification of carrier status for essentially the entire catalog of autosomal recessive conditions,<sup>12</sup> which would be of considerable potential benefit to couples for reproductive planning.

## DEALING WITH LARGE AMOUNTS OF INFORMATION: SAVED BY OUR IGNORANCE

A significant obstacle to implementing WGS is the almost unimaginable amount of information that will be generated. However, the task is made more manageable when we realize that the majority of data generated from WGS will be useless (at least initially) simply because we have no idea of how to accurately interpret it; thus, it must be disregarded in the clinical context. As illustrated by the above examples, the utility of WGS lies in both targeted application in affected individuals (its use as a diagnostic tool) and in the identification of asymptomatic individuals within populations who have a high risk for preventable disease (its use as a public health tool). In both settings, the use of WGS will lead to the inevitable discovery of incidental findings that have no direct clinical actionability and some that may be frankly unwelcome to many individuals. We

From the <sup>1</sup>Department of Genetics, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; and <sup>2</sup>Office of Public Health Genomics, Centers for Disease Control, Atlanta, Georgia.

Jonathan S. Berg, MD, PhD, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7264. E-mail: JSBerg@med.unc.edu.

Disclosure: The authors declare no conflict of interest.

Published online ahead of print May 9, 2011.

DOI: 10.1097/GIM.0b013e318220aaba

argue that consent, analysis, reporting of results, and policies for dealing with incidental findings must be formulated with respect to the specific context in which WGS is applied. Differing contexts demand distinct approaches.

### Context matters

A tension always exists between the competing benefits and risks of maximizing either sensitivity or specificity. Although both are important in clinical testing, the goal of a diagnostic test is to identify an etiology for a patient's presenting complaints, and thus, it is reasonable to maximize sensitivity to avoid false negatives. However, when used in a public health context, the low a priori chance that any given variant is clinically relevant, the sheer number of variants identified by WGS, and, most critically, the lack of any definitive measures (i.e., "gold standard") by which to evaluate variants of uncertain significance for clinical relevance, all mandate that in the public health context, we maximize specificity and minimize false positives, even when by doing so we give up some sensitivity.

In the individual diagnostic setting, the analysis of WGS data must be undertaken in a way that not only maximizes sensitivity but also avoids overwhelming clinicians and patients alike with uninterpretable information. The interpretation of WGS in the diagnostic setting will provide information that is qualitatively similar to current genetic test results: a "definitive etiology" for the patient's clinical presentation (a positive result), a "possible etiology" (an uncertain result), or "no etiology identified" (a negative or uninformative result). Just as is the case now, the analysis of variants identified by WGS will need to be informed by the patient's presenting symptoms or clinical diagnosis, so that variants can be passed through a computational filtering process that selects only those of possible diagnostic significance for inspection by the molecular diagnostic team, based on whether the affected loci have been demonstrated to have clinically relevant phenotypic implications. In the absence of a definitive etiology for a patient's symptoms, it may be tempting to seek an explanation among novel variants in other candidate genes. However, such variants should not be represented as an etiology for a patient's phenotype until substantial evidence is available to support such a conclusion, and thus, they would not be reported in the diagnostic WGS analysis. Rather, such variants could be funneled (with patient consent) to research studies that seek to illuminate new genotype/phenotype linkages. Such an approach allows individual curation of diagnostic results by those caring for a patient and consideration of a wide range of findings with possible relevance to the patient's presenting complaint, without the distraction of reviewing incidental findings or variants in genes unrelated to known medical conditions.

In the public health setting, the use of WGS will generate massive numbers of variants that can be considered likely clinical "false positives" with respect to the chance that they have health implications for the individual (this usage of the term assumes that the variants are able to be confirmed by alternative methods and are not merely technical false-positive results due to sequencing errors, which is yet another valid concern regarding WGS). To maximize the use of WGS in the public health context, a very high bar must be set for reporting results. Only clearly deleterious mutations in genes known to cause a high risk for preventable disease should be routinely reported. This differs from the public health pursuit of newborn screening, for example, in which sensitivity is maximized at the expense of specificity. The application of WGS demands a different approach for two reasons. (1) The sheer number of variants generated when subjecting individuals to WGS is immense. Reporting each variant would overwhelm any attempt to

harness WGS in this context and unnecessarily dilute the ability to identify those individuals with clear actionable findings. (2) Critically, unlike the newborn screening setting where abnormal screening results can be followed up with highly specific diagnostic tests, there are no definitive confirmatory tests to determine whether novel genomic variants are deleterious. This inability to sort meaningful from irrelevant findings with subsequent testing is an important factor that compels setting a high bar for reporting of variants and limiting such reporting to only clearly deleterious variants in genes which, when mutated, lead to medically actionable recommendations.

### The imperative to ignore variants of unknown significance

Given our limited understanding of genetic variants at present, >99.9% of any individual's estimated 3–4 million variants must be ignored from any reasonable clinical or public health perspective.<sup>13</sup> Diverting attention of the clinician and patient by exhaustive analysis and reporting of information for which we have no understanding and which contains no known medical relevance would represent a disservice to both the clinical and public health endeavors. This is not to say that such information will not eventually be useful or that it should not be used in a research setting with proper consent. However, it does mean that such information should not be part of the primary clinical record and that providers should not waste time discussing it with patients. An analogous situation exists in the realm of imaging. When a magnetic resonance imaging is performed, we do not waste time documenting each variable structure, pixel by pixel. Instead, we appropriately concentrate on those aspects of the image that we currently understand to have clinical meaning. Similarly, with WGS we must concentrate only on those variants that have been demonstrated to have meaningful implications for patients and the public.

### AVOIDING INFORMATION OVERLOAD THROUGH A STANDARDIZED, CLINICALLY ORIENTED STRUCTURED ANALYSIS

The vast amount of incidentally generated WGS information must be organized in a clinically oriented manner to facilitate shared decision making by patients and clinicians. This can be accomplished by assigning variants identified in the course of WGS into predetermined clinically relevant "bins," defined by utility (or lack thereof), and making this structured analysis, described later, an explicit aspect of clinical WGS. Such a categorical approach moves us away from an untenable model based on the analysis of one gene at a time and facilitates its use at every level, including the consent process, interpretation, and patient decision making regarding return of results. By constructing an a priori categorical framework, the massive amounts of information generated by WGS can be dealt with in a way that is scalable to the realities of whole genome analysis. Such a framework can serve as an initial basis for the interpretation of incidental results and will set appropriately stringent criteria for the reporting of variants in the public health setting to avoid overwhelming patients and physicians and distracting them from the small amount of truly meaningful information that will be generated.

### To report or not to report

There are two relevant parameters to consider when determining whether a variant should be reported and acted on: first, the gene or locus in which the variant resides, and second, the nature of the variant itself. A first pass at the analysis of incidental WGS findings should identify variants that exist in

known disease-relevant loci. A second-order analysis can then determine whether the variants existing in such clinically relevant loci represent deleterious mutations or innocuous (or indefinable and thus nonreportable) variants. As discussed earlier, a key to surmounting the challenge of the vast amount of WGS data is to define the appropriate thresholds for the designation and reporting of a variant as deleterious in the different contexts in which WGS may be applied. In the public health context, any variant found in an asymptomatic individual inherently has a low a priori chance of being deleterious. Similarly, in the clinical diagnostic setting, when an incidental variant is discovered that is unrelated to the referring diagnosis, the chance of it being significant is low from an a priori perspective. Therefore, such variants must be triaged, such that only those found in clearly medically relevant loci and known to cause disease or strongly predicted to disrupt function are designated potentially causative and acted on, thus maximizing specificity and avoiding “false positive” results.

**A “binning” system for incidental findings**

Figure 1 depicts one possible scheme for the categorization of incidental findings, here defined as variants unrelated to a patient’s clinical presentation (i.e., those variants which remain after possible diagnostic variants are extracted when WGS has been applied in a diagnostic context) or any variants detected when WGS is applied to an asymptomatic individual (i.e., in the public health context).

The binning system proposed here allows for a scalable approach to the analysis and return of incidental results identified

during diagnostic testing, as well as WGS performed in a public health setting among asymptomatic individuals, permitting a category-based approach to consent and patient education. Deleterious variants in bin 1, by definition, have immediate clinical utility and would be reported, regardless of the context of the WGS evaluation (clinical/diagnostic or presymptomatic) just as medically actionable incidental findings are now dealt with in medicine as a whole.<sup>14</sup> Indeed, it is the identification of rare individuals with clearly deleterious bin 1 variants that would be the goal of WGS when applied to asymptomatic individuals in a public health context. Known or presumed-deleterious variants in bin 2, despite being reliably associated with a disease or relevant trait, are not medically actionable. As their perceived utility will differ among individuals, their potential return when derived in the clinical setting may be dealt with in a risk-stratified manner by shared decision making between patients and their providers. Incidental variants of uncertain significance or presumed benign variants in bin 1 or bin 2 would not be reported, as this information would, by definition, have unclear implications and would essentially represent “false positive” results. Finally, bin 3 variants (the majority of findings in the context of WGS) have—by definition—no known medical relevance and clinical reporting is not warranted.

*Bin 1: “Clinically actionable”*

Bin 1 holds variants within genes/loci that have direct clinical utility based on the current medical literature (e.g., in terms of disease prevention or established treatment guidelines) and

Criteria:		<i>Clinical Utility</i>	<i>Clinical Validity</i>			<i>Unknown Clinical Implications</i>
<b>Genes</b>	Bins:	<b>Bin 1 Medically actionable incidental information</b>	<b>Bin 2A Low risk incidental information</b>	<b>Bin 2B Medium risk incidental information</b>	<b>Bin 2C High risk incidental information</b>	<b>Bin 3  All other loci</b>
	Examples:	<i>BRCA1/2 MLH1, MSH2 FBN1 NF1</i>	PGx variants and common risk SNPs	<i>APOE</i> Carrier status for recessive Mendelian disorders	Huntington Prion diseases ALS (SOD1)	
	Estimated number of genes/loci:	10s	10s (eventually 100s–1000s)	1000s	10s	
<b>Alleles that would be reportable (YES) or not reportable (NO) in a clinical context</b>						
<b>Variants</b>	Known deleterious	YES	YES/NO <sup>1</sup>	YES/NO <sup>1</sup>	YES/NO <sup>1</sup>	N/A <sup>2</sup>
	Presumed deleterious	YES	N/A <sup>3</sup>	YES/NO <sup>1</sup>	YES/NO <sup>1</sup>	NO <sup>4</sup>
	VUS	NO	N/A <sup>3</sup>	NO	NO	NO <sup>4</sup>
	Presumed benign	NO	N/A <sup>3</sup>	NO	NO	NO
	Known benign	NO	NO	NO	NO	NO

N/A: not applicable; VUS: Variant of uncertain significance

<sup>1</sup> Reporting through decision making with an appropriate provider if elected by the patient.

<sup>2</sup> By definition, variants in genes with unknown implications could not be considered deleterious.

<sup>3</sup> By definition, SNPs or PGx variants will either be present or absent.

<sup>4</sup> Variants in genes with unknown clinical implications would not be reported; however, they may serve as an important substrate for research, potentially uncovering new disease genes.

**Fig. 1.** Proposed system for “binning” of incidental WGS results

must therefore be acted on. The types of variants falling within this bin include highly penetrant rare variants in genes associated with Mendelian disorders for which there are established clinical management recommendations (e.g., neurofibromatosis type 1 or Marfan disease) or those that confer a high risk of a preventable disease (e.g., Lynch syndrome or *BRC1/2*). By setting appropriately stringent requirements for inclusion based on clinical utility, this category will (at least initially) be small. Thus, we expect that only infrequently will an individual have a deleterious variant assigned to bin 1. Given their clinical actionability, such variants would be flagged for confirmation by Sanger sequencing and officially reported to patients (or asymptomatic individuals if undergoing screening for the purpose of detecting such actionable lesions).

### Bin 2: "Clinically valid but not directly actionable"

Bin 2 contains variants within genes/loci demonstrated to have clinical validity but which have no strongly actionable implications (i.e., a lack of demonstrated clinical utility). The lack of any clinical utility argues that in the public health setting, these results would not be reported. In the individual clinical setting, however, it might be that some patients are interested in receiving such information. Again, the categorical approach that we envision would facilitate potential return of such results to patients if desired. As such results vary widely regarding their potential to cause patient distress and anxiety, bin 2 is subdivided into three subcategories that are calibrated to the potential for distress on return of results in the clinical setting.

- Bin 2A: "Low risk, clinically valid results"—This category contains common single-nucleotide polymorphisms (SNPs) with well-documented associations with disease risk by genome-wide association studies, which have clinical validity but lack proven clinical utility. Many PGx variants are also included in this subcategory.

At present, considerable evidence supports the association between a variety of common SNPs and risk for certain common medical conditions or health-related quantitative traits.<sup>15,16</sup> Such variants are unlikely to cause significant distress or harm on their return,<sup>17</sup> might confer some amount of "personal utility,"<sup>18</sup> and are likely therefore to be of interest to some patients, even though evidence is lacking that these results have a significant impact on health behavior outcomes.<sup>18,19</sup> Nevertheless, despite recent attempts to model disease risks based on common SNPs as part of the clinical assessment of a genome sequence,<sup>20</sup> at present such information has dubious clinical utility<sup>1,19,21</sup> and the clinical validity of aggregate risk scores is thus far lacking.<sup>22</sup> Thus, we argue that it would be premature to incorporate this information into the medical record. Over time, the development of robust models and prospective clinical studies may demonstrate clinical validity and even utility of such variants; if so, we expect this category to expand.

For some PGx variants, there may be a high level of evidence that supports a role in influencing the efficacy of certain drugs, and in a few cases, there may be evidence-based recommendations for altering medical therapy. Given that some patients and their providers may be interested in this information, that there may be some presumption of clinical (or at least "personal") utility in certain circumstances, and that the provision of such data are unlikely to be harmful, such results will initially be included in bin 2A. In a practical sense, these bin 2A variants ultimately might need no specific genetic counseling but rather might be incorporated into an electronic medical record (EMR) in a form that enables "just in time" prompting of physicians

when the information is relevant (e.g., when prescribing a medication or when completing health screening activities).

- Bin 2B: "Medium risk results"—This category encompasses a broad range of genomic results that would be generally considered neither completely innocuous nor truly shocking. Some such results have robust clinical validity (e.g., significant association with a genetic disorder) but are nondeterministic due to incomplete penetrance. Reporting of this class of variants, by definition, implies no specific medical recommendations or imperative medical management changes as do deleterious variants falling within "bin 1" loci.

An example of such information is the presence of an *APOE4* allele, which despite its clear association with Alzheimer disease risk (i.e., clinical validity) is not routinely used clinically given the lack of interventions to reduce risk (i.e., lack of clinical utility). Another example is carrier status for an autosomal recessive condition in which the carrier state is phenotypically inconsequential but may be important for family planning to some individuals. The incidental results that fall into this category have greater potential for causing distress or having significant repercussions; thus, the manner of their return (when elected by patients) and patient education and counseling about the results can be calibrated to this increased level of risk.

- Bin 2C: "High risk results"—This category includes the small number of variants within genes/loci for which the return of incidental positive results could be harmful to patients. Although we generally eschew genetic exceptionalism, and the likelihood of identifying a variant in this category is extremely small in the absence of a suggestive family history, it would be unacceptable to cause harm to patients by casually providing such devastating information in the absence of adequate preparation. This select group of loci would include, e.g., Huntington disease and Cruetzfeld-Jakob disease, conditions with high penetrance and no available treatment for which individuals known to be at risk might decline predictive genetic testing. Reporting of this category of information in the clinical setting would be contingent on demonstrated and sustained interest by an individual and adequate counseling. Thus, by taking care in how this category of information is delivered, we can protect patients from casual return of potentially disruptive information while avoiding excessive paternalism and preserving patient autonomy.

### Bin 3: "Unknown or no clinical significance"

Bin 3 contains all other variants within genes/loci that have not been strongly linked to a phenotype, clinical outcome, or intervention. The majority of variants identified in each individual (including risk SNPs with low odds ratios) will fall into this category, and thoughtful clinical judgment mandates that patients and clinicians not waste time and effort clinically analyzing this large category of results with no known clinical relevance. Bin 3 variants will, however, form an important substrate for future research.

### Summary of bins

In summary, by taking a categorical approach to classifying such information, the wide range of inherently heterogeneous incidental findings discovered in patients can be classified in a way that enables patients and providers to discuss each category of possible results and, with sufficient interest and risk-calibrated counseling, learn of their own individual results if they desire.

## PRACTICAL CONSIDERATIONS

### Sensitive information

One of the more troublesome aspects of WGS is that inevitable—and potentially unwelcome—surprises will occur, such as learning about risk for adult-onset disorders or behavioral predilections. Addressing individual preferences regarding the storage and return of sensitive incidental information is best dealt with in the initial consent process for WGS. Individuals would be apprised of the possibility that sensitive information could be forthcoming, educated about its potential implications, and allowed an ongoing opportunity for an appointment to receive such results, separate from the discussion of diagnostic results. The binning process will greatly facilitate such discussions as counseling can be streamlined by being oriented around categories of possible results as opposed to an impractical prospective discussion of each possible locus in which one might be found to have a variant. Moreover, education and counseling can be patient driven and calibrated to possible risk by using such a categorical approach, while allowing scaling of genetic counseling.

It should be emphasized that although it would be perfectly reasonable to divulge none of the variants in bin 2 if a patient so chooses, medically actionable (bin 1 results) would be reported just as we routinely report medically important but incidental findings that occur, for example, in the course of medical imaging or laboratory assessment. With respect to handling potentially sensitive information in the medical record, there already exists abundant precedent for the special treatment of psychiatric information; there is no reason why sensitive genetic data could not be treated in a similar manner in the medical record.

### How will binning of genes/loci be determined?

Given our current limited understanding of the genome, the designation of loci to specific bins must be an iterative, centralized, evidence-based, and consensus-driven process.<sup>23–25</sup> There will be a need for broad pooling of epidemiological and clinical data on phenotypes and genotypes as it is unlikely that a single provider will be able to make sense of WGS data. We propose that a broad coalition of experts and stakeholders engage in an ongoing, collaborative, and open forum similar to the process used to develop newborn screening guidelines<sup>26</sup> and for the evaluation of genomic applications in practice and prevention (EGAPP; <http://www.egappreviews.org/>).<sup>27</sup> Such an independent multidisciplinary panel would conduct periodic evidence-based reviews and designate the loci (and nature of alleles) to be included in bin 1 (possessing clinical utility), bin 2 (possessing only clinical validity), or bin 3 (all other variants). Requests to move a given locus from one bin to another could be made by anyone, with decisions based on emerging data.

The process of WGS analysis will ultimately depend on a well-curated database of genes, phenotypes associated with mutations in those genes, their inheritance patterns, and the previously described pathogenic and benign variants within them. Other critical requirements will be the population frequencies of variants identified within disease genes and the ability to accurately predict the effects of protein-coding variants. Some of these elements already exist (OMIM, HGMD, and dbSNP) but not in a form that is readily amenable to clinical WGS analysis. Efforts are underway to accomplish clinically meaningful genomic annotation and to establish criteria for determining the clinical utility of genomic information at particular loci.<sup>23</sup> Although the challenge is great, it can only be met by actually collecting and analyzing WGS data in representative

populations and clinical groups. Consortia such as the “muta-DATABASE” ([www.mutadatabase.org](http://www.mutadatabase.org))<sup>28</sup> may ultimately meet this need.

### Disseminating information to providers

The medical workforce is unprepared to deal with genomic information, and only a multifaceted approach can meet this challenge. We estimate that bin 1 variants currently rise to a cumulative population prevalence of 1–5% (a number that will increase as more loci achieve clinical utility), underlining the public health potential of identifying such variants. As genetics permeates general medicine and its specialties, it will be critical that practitioners have access to information resources such as GeneReviews (<http://www.ncbi.nlm.nih.gov/sites/GeneTests/>) and EGAPP reviews. Moreover, “just in time” technology must be used to facilitate appropriate decisions, for example, by prompting providers to review an individual’s PGx profile when prescribing certain medications. Realizing this vision will require an EMR capable of incorporating genomic information. However, the paucity of actionable information (at least initially) from WGS will facilitate EMR development as most information (bin 3) need not be accessed outside of a research setting and, therefore, need not be included in a patient’s formal medical record.

Knowledgeable personnel will be critical to help integrate WGS into patient care and there is tremendous potential for clinical geneticists, genetic counselors, and adequately trained nurses to fill this niche. However, they will need to be capable of earning their keep. Thus, securing licensing and adequate reimbursement for counselors (and reimbursement for physicians’ cognitive services) will be critical. It would be ironic if the clinical fruits of our scientific advances in genomics were thwarted due to soluble reimbursement issues.

### A work in progress

Clearly, our understanding of the clinical relevance of most genomic information is woefully inadequate. As we accrue more experience, our ability to define the meaning of variants will gradually improve, and we expect that assignment of loci to any given bin will be subject to ongoing revision. Implicit in the strategy described earlier is the expectation that the organization of loci and of specific variants within this framework will change as new scientific data are generated. For example, one’s *APOE* status is currently considered both nonactionable and “sensitive,” placing it firmly in bin 2. However, if medications or other strategies are identified that mitigate the risk of Alzheimer disease in those with an *APOE4* allele, this locus would be reassigned to bin 1 as medically actionable. Similarly, as treatments emerge for specific Mendelian disorders, those loci would move from bin 2 to bin 1. Finally, as bin 3 variants are shown to be medically relevant, they will be shifted to bin 2 (or bin 1 depending on whether actionability exits).

In the public health setting, as the goal is simply to identify individuals with deleterious bin 1 variants, these would be the only variants reported. An argument can be made that a multiplex sequencing panel consisting of only bin 1 genes would be the most straightforward approach to applying next generation sequencing in the public health setting. However, it may well be that WGS will soon be less expensive than specific gene capture and subsequent sequencing, an important consideration in the public health context. Moreover, with the rapid pace of progress in genomics, by using a WGS approach, one gains the considerable advantage that the test need only be done once and simply subjected to reanalysis of data as more genes are designated to have “bin 1” status.

Despite the lack of current medical management implications for common risk SNPs, a major focus of research in the post-genome-wide association study era will be developing an understanding of the interplay between genetic variants and the interactions between genes and the environment.<sup>29–31</sup> The binning structure proposed earlier may be most appropriate for the analysis of individual variants typical of Mendelian medical genetics. In contrast, the underlying biology of complex diseases will require more complex models, which could nevertheless be incorporated into the structured analysis of WGS over time as genomic medicine matures and combinations of risk SNPs are demonstrated to have clinical utility for stratifying the population with respect to health screening or preventive interventions.

Thus, an individual's WGS data will need to be periodically reanalyzed, necessitating mechanisms to communicate new information to the patient and medical providers, and a key aspect of informed consent for WGS should be ensuring that patients understand that the clinical significance of their genetic results will almost certainly change over time. Other challenges exist before we can expect to fully realize the medical promise of WGS, including streamlining informatics and improving the accuracy of next-generation sequencing<sup>32</sup>; addressing third-party coverage of WGS testing; and addressing ELSI-related issues.<sup>33,34</sup>

## CONCLUSION

There currently exists a profound mismatch between our ability to interrogate the human genome and our ability to use that information to improve health. However, this is to be expected given the astounding complexities (and high stakes) inherent in clinical medicine. The implementation of genomic medicine requires few qualitatively novel approaches; medicine has long dealt with large amounts of unambiguous data. However, we must develop new models by which clinical implementation of genomics can be scaled to the new reality of whole-genome analysis. We feel that some version of the binning process described in this study, by formulating clinically relevant categories into which loci are assigned on an a priori basis, will allow for a categorical, patient-driven, and streamlined approach to implementing genomic medicine and will facilitate realization of its public health potential. Such an approach will facilitate both the clinical evaluation of novel variants and patient involvement in determining return of results. As we construct such a system, we must keep a strong focus on evidence, expand health-oriented phenotypic annotation of genomic variants, create a centralized, evidence-based, iterative process to define clinically significant genomic findings, and ensure that the cognitive expenditures by providers who grapple with genomic issues are adequately reimbursed. We are up to meeting these challenges and indeed must do so if we are to realize the promise of genomic medicine.

## REFERENCES

- Mayer AN, Dimmock DP, Arca MJ, et al. A timely arrival for genomic medicine. *Genet Med* 2011;13:195–196.
- Ormond KE, Wheeler MT, Hudgins L, et al. Challenges in the clinical application of whole-genome sequencing. *Lancet* 2010;375:1749–1751.
- Janssens AC, Gwinn M, Bradley LA, Oostra BA, van Duijn CM, Khoury MJ. A critical appraisal of the scientific basis of commercial genomic profiles used to assess health risks and personalize health interventions. *Am J Hum Genet* 2008;82:593–599.
- Janssens AC, van Duijn CM. Genome-based prediction of common diseases: advances and prospects. *Hum Mol Genet* 2008;17:R166–R173.
- de la Chapelle A. The incidence of Lynch syndrome. *Fam Cancer* 2005;4:233–237.
- Hampel H, Frankel WL, Martin E, et al. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). *N Engl J Med* 2005;352:1851–1860.
- Guillem JG, Wood WC, Moley JF, et al. ASCO/SSO review of current role of risk-reducing surgery in common hereditary cancer syndromes. *J Clin Oncol* 2006;24:4642–4660.
- Lindor NM, Petersen GM, Hadley DW, et al. Recommendations for the care of individuals with an inherited predisposition to Lynch syndrome: a systematic review. *JAMA* 2006;296:1507–1517.
- Mvundura M, Grosse SD, Hampel H, Palomaki GE. The cost-effectiveness of genetic testing strategies for Lynch syndrome among newly diagnosed patients with colorectal cancer. *Genet Med* 2010;12:93–104.
- Weinshilboum RM, Wang L. Pharmacogenetics and pharmacogenomics: development, science, and translation. *Annu Rev Genomics Hum Genet* 2006;7:223–245.
- Wang L, McLeod HL, Weinshilboum RM. Genomics and drug response. *N Engl J Med* 2011;364:1144–1153.
- Bell CJ, Dinwiddie DL, Miller NA, et al. Carrier testing for severe childhood recessive diseases by next-generation sequencing. *Sci Transl Med* 2011;3:65ra4.
- Kohane IS, Masys DR, Altman RB. The incidentalome: a threat to genomic medicine. *JAMA* 2006;296:212–215.
- Lumbreras B, Donat L, Hernandez-Aguado I. Incidental findings in imaging diagnostic tests: a systematic review. *Br J Radiol* 2010;83:276–289.
- Manolio TA. Genomewide association studies and assessment of the risk of disease. *N Engl J Med* 2010;363:166–176.
- Hindorf LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 2009;106:9362–9367.
- Bloss CS, Schork NJ, Topol EJ. Effect of direct-to-consumer genomewide profiling to assess disease risk. *N Engl J Med* 2011;364:524–534.
- Foster MW, Mulvihill JJ, Sharp RR. Evaluating the utility of personal genomic information. *Genet Med* 2009;11:570–574.
- Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: genomic profiling to assess cardiovascular risk to improve cardiovascular health. *Genet Med* 2010;12:839–843.
- Ashley EA, Butte AJ, Wheeler MT, et al. Clinical assessment incorporating a personal genome. *Lancet* 2010;375:1525–1535.
- Li-Wan-Po A, Farndon P, Cooley C, Lithgow J. When is a genetic test suitable for prime time? Predicting the risk of prostate cancer as a case-example. *Public Health Genomics* 2010;13:55–62.
- Ng PC, Murray SS, Levy S, Venter JC. An agenda for personalized medicine. *Nature* 2009;461:724–726.
- Khoury MJ, Rich EC, Randhawa G, Teutsch SM, Niederhuber J. Comparative effectiveness research and genomic medicine: an evolving partnership for 21st century medicine. *Genet Med* 2009;11:707–711.
- Veenstra DL, Roth JA, Garrison LP Jr, Ramsey SD, Burke W. A formal risk-benefit framework for genomic tests: facilitating the appropriate translation of genomics into clinical practice. *Genet Med* 2010;12:686–693.
- Khoury MJ, Coates RJ, Evans JP. Evidence-based classification of recommendations on use of genomic tests in clinical practice: dealing with insufficient evidence. *Genet Med* 2010;12:680–683.
- Green NS, Rinaldo P, Brower A, et al. Advisory Committee on Heritable Disorders and Genetic Diseases in Newborns and Children. Committee report: advancing the current recommended panel of conditions for newborn screening. *Genet Med* 2007;9:792–796.
- Teutsch SM, Bradley LA, Palomaki GE, et al. EGAPP Working Group. The evaluation of genomic applications in practice and prevention (EGAPP) initiative: methods of the EGAPP Working Group. *Genet Med* 2009;11:3–14.
- Bale S, Devisscher M, Van Criekinge W, et al. MutaDATABASE: a centralized and standardized DNA variation database. *Nat Biotechnol* 2011;29:117–118.
- Khoury MJ, Wacholder S. Invited commentary: from genome-wide association studies to gene-environment-wide interaction studies—challenges and opportunities. *Am J Epidemiol* 2009;169:227–230; discussion 234–235.
- Chatterjee N, Wacholder S. Invited commentary: efficient testing of gene-environment interaction. *Am J Epidemiol* 2009;169:231–233; discussion 234–235.
- Bookman EB, McAllister K, Gillanders E, et al. Gene-environment interplay in common complex diseases: forging an integrative model—recommendations from an NIH workshop. *Genet Epidemiol* 2011;35:217–225.
- Koboldt DC, Ding L, Mardis ER, Wilson RK. Challenges of sequencing human genomes. *Brief Bioinform* 2010;11:484–498.
- Guttmacher AE, McGuire AL, Ponder B, Stefánsson K. Personalized genomic information: preparing for the future of genetic medicine. *Nat Rev Genet* 2010;11:161–165.
- Ali-Khan SE, Daar AS, Shuman C, Ray PN, Scherer SW. Whole genome scanning: resolving clinical diagnosis and management amidst complex data. *Pediatr Res* 2009;66:357–363.