
Genomic Research and Incidental Findings

Brian Van Ness

The Complexity of the Human Genome

Medical practice is poised to incorporate genome-scale testing into treatment decisions. However, broad genome testing in laboratories may lead to discoveries not anticipated, yet highly significant to the health of the patient. Understanding the complexity of our genome and its relationship to our health is an overwhelming task. Currently, much of the effort to unravel this complexity is in the realm of research. However, researchers are often neither qualified nor prepared to deal with incidental findings of genetic abnormalities that influence health and disease. These incidental observations, referred to recently by Issac Kohane et al. as the “incidentalome,” may lead to complex legal, ethical, and financial problems that may seriously complicate the role of the research community in medical genomics.¹ Currently, most genetics researchers, while aware of the potential for incidental findings, simply do not want to deal with them.

The draft sequence of the human genome published in 2001,² followed by the more refined sequence released by the International Human Genome Consortium,³ have revealed large genetic variations in the human genome (polymorphisms). Although our genomes are 99.9% identical, it is the variation that makes each of us unique. As you look around a crowded room, there is clear, visible evidence of the differences that separate us. Not surprisingly, those differences include all the biologic functions that create variations in the risk for disease and how that disease may progress or respond to therapy. Unlike somatic mutations, polymorphisms are stable and heritable. Polymorphisms include single changes in the sequence of the 4 bases representing the genomic code, called single nucleotide polymorphisms (SNPs), as well as larger chromosomal alterations such as deletions, insertions, and regional copy number variations. The central concept is that as we reproduce, the population is continually mixing the pot of DNA sequences. And the pot is not equally distributed in variation. Consider that a genetic variation that alters a drug response in 20% of the Japanese population is found in less than 1% of the Northern European, Caucasian population.⁴ To a physician prescribing a treatment, this knowledge can be important in understanding potential response.

While our genes define much of who we are, they obviously do not provide this information in any single gene. The estimated 25,000 gene sequences form a complex interactive set of biologic pathways. There

Brian Van Ness, P.h.D., is currently the Head of the Department of Genetics, Cell Biology & Development at the University of Minnesota.

are examples of single-gene variations that cause disease (e.g., sickle cell anemia, cystic fibrosis, and Huntington disease). There are genetic variations that increase the risk for disease (e.g., alterations in the BRCA1 gene that increase the risk for developing breast cancer). But most of what influences our health is a complex interplay of several or even hundreds to thousands of genes (e.g., genes affecting cancer or cardiovascular diseases). So understanding the importance of incidental findings requires understanding the genomic data that are rapidly accumulating. The impact of a genetic variation may be unknown today, but understood in the future. Thus, the *incidentalome* is time-sensitive.

Sources of Incidental Findings in Genomic Research

The working definition of an incidental finding (IF) is “a finding concerning an individual research participant that has potential health or reproductive importance and is discovered in the course of conducting research but is beyond the aims of the study.”⁵

with significant dilemmas if the pedigree shows misattributed paternity. There is a very active and growing body of research identifying new disease genes in linkage studies that depend on accurately defining

Currently, most genetics researchers, while aware of the potential for incidental findings, simply do not want to deal with them.

pedigrees. The incidental finding of misattributed paternity may have health, legal, and financial implications. Researchers may well wish to avoid the social complications of revealing misattributed paternity in a family, but accurate genetic counseling on disease risk may be difficult without addressing the paternity issue.

Technologies developed in just the past 10 years have significantly increased the potential for incidental findings of significance to a research subject’s health. We now have the capability to examine up to a million genetic variations in an individual’s genome.

Small chromosomal abnormalities (e.g., insertions, deletions, and copy number changes), changes in the expression pattern of thousands of genes, and single changes in the 3 billion bases that constitute the genome of an individual are being tested. Researchers are interested in associations of single or multiple genetic variations with disease as well as with drug response. Most phase II and phase III clinical drug trials now include collection of

Understanding the importance of incidental findings requires understanding the genomic data that are rapidly accumulating. The impact of a genetic variation may be unknown today, but understood in the future. Thus, the *incidentalome* is time-sensitive.

This discussion focuses on the IFs discovered in the course of research, not clinical care. The distinction is important, for most research studies are not done under the Clinical Laboratory Improvement Amendments (CLIA)⁶-certified laboratory testing guidelines of quality assurance, and often researchers are not clinically certified to evaluate the health significance of genetic or genomic incidental findings.

One of the earliest examples of incidental findings in genetic research emerged in the course of establishing family pedigrees — the incidental discovery of misattributed paternity. Rough estimates suggested as high as 10% of research subjects included in a pedigree analysis manifested misattributed paternity.⁷ Family histories of disease have become routine in the initial assessment of risk, and genetic counselors are faced

blood DNA in order to identify SNPs that may predict response, non-response, and adverse effects in trial participants. Often the approach is global (genome-wide scans) and unbiased by pre-selection of candidate genes. Computer programs then sift through the complex array of data sets, looking for genomic patterns associated with a specific outcome. However, additional genetic variations associated with conditions such as blood disorders, cancer, immune disorders, and neurologic disease may be included in the “scan” of the genome because researchers know that identical genetic abnormalities often can lead to multiple outcomes by deregulating common pathways in different tissues. For example, variations in the BRCA1 gene were included in a scan for genes asso-

ciated with a certain type of bone marrow cancer in our own laboratory, because in mammary cells certain BRCA1 variations can lead to signal proliferation (leading to breast cancer risk); thus, the rationale was to see if BRCA1 had a role in other disease tissues. Recent analysis has shown that BRCA1 variations do not appear to be associated with bone marrow cancer,⁸ but what if the data in such a study revealed subjects with increased risk of breast or ovarian cancer due to BRCA1? This is a clear example of a potential incidental finding, as the study had no intent to determine risk for breast cancer; yet, the information is part of the subjects' database.

Most researchers understand that results that may show genetic associations with an outcome are not precise, but rather shift the probability of an outcome. Is there an obligation to inform a research participant of findings despite their lack of precision?

This leads to some compelling questions: (1) What is the responsibility of such a researcher to even look at the results of the BRCA1 (or other) genotypes? (2) What is the responsibility of the researcher to know what genotypes are high-risk in large genome screens? (3) If the research testing is not done under CLIA-certified laboratory conditions, how reliable is the test? (4) What if genetic variation in the database has no known impact on health today, but is discovered three years later to predict a life-threatening disease? (5) What if a genetic variation has significant implications for health, but only when it occurs with 2, 4, or 24 other genetic variants? and (6) To whom should the researcher report a genetic "concern" — the subject, a physician, or the Institutional Review Board (IRB) that approved the study?

One of the common approaches in consenting participants in a research study is to state "no individual information will be given." Yet, even if the researcher is protected legally by this approach, what is the ethical obligation? And participants themselves may seek information, especially if they suspect it has health importance. Indeed, David Shalowitz and Frank Miller in a 2005 *JAMA* article say that "participants have presumptive entitlement to information about themselves."⁹ This presents a significant challenge to a researcher who may have thousands of genotypes in hand, if participants expect an interpretation of each. For genome-wide screening, understanding the impli-

cations of genotypes for health risks or outcomes is an enormous task.

The breadth of data collection in genomic studies is becoming daunting, even to the most sophisticated supercomputers. Consider a study of 1 million genetic variations in 5,000 individuals (not an uncommon research project today). This results in a database of 5 billion entries. As described by Issac Kohane et al.,¹⁰ consider the reliability of such a huge data set. If a screen has 99.99% accuracy, we consider that a pretty reliable screen. But, the converse is that for a data set of 5 billion, a 0.01% inaccuracy is 500,000 data points. For most studies examining group/data associations, the inaccuracies may be distributed among all data points, and present very little background noise. But for the one individual for whom an error may have occurred in a critical health-related gene, it is a very large concern. Thus, the health concern in large research studies may be the result of small but significant inaccuracies. In contrast, targeted diagnostic testing in certified laboratories has specific quality controls built in to the targeted assay. The answer may be simply to suggest that the subject be informed of the possible concern and urged to obtain a CLIA-certified, targeted laboratory test. However, the subject may experience burdensome concern and expense for follow-up testing.

With the exception of a few hundred genetically determined diseases, most disease predictions based on genetics are probability estimates. The beauty of human biology is that evolution has resulted in redundancy and genetic modifiers that shift the gamble on good or bad outcomes; outcomes often depend on environmental exposures. Most researchers understand that results that may show genetic associations with an outcome are not precise, but rather shift the probability of an outcome. Is there an obligation to inform a research participant of findings despite their lack of precision? Certainly, as research continues, we hope to hone the ability to predict outcomes from modest probability shifts to more precise predictions. Moreover, outcomes may rely on gene-gene, gene-environment, or complex interactions. Thus, understanding the true health impact of incidental findings for many research results awaits much more precision in data analysis.

The Process of Making Recommendations to Researchers

The past decade leading to the genomic data explosion has resulted in a variety of national and international groups attempting to establish ethical guidelines about informing individual participants participating in genetic research of individual research results,¹¹ including reporting incidental findings. One extreme view may be that all research results should be disclosed to the individual. This is problematic, because early research results are often relevant to group predictions, but not individual predictions. The same may be true for an incidental finding in the context of a result that has statistical significance in group comparisons, but is not highly predictive on an individual basis. For example, in a large study, a genetic variant may be highly associated with disease progression. Data points are evaluated as a group and the mean (or median) value for progressive disease may be found to be statistically significant in the group carrying one allele of a gene, but not in the group carrying the variant of the same gene. Yet, the value of any one individual analysis is not very predictive because there is lot of variation around that highly significant group result. In addition, early genetic research results often are imprecise in predicting the development and severity of a condition, and may not be validated in subsequent studies. Early reporting of potentially important IFs to individuals in this context could be very misleading. Indeed, we are seeing exhaustive requirements for multiple validation studies before the Food and Drug Administration will approve a diagnostic test that results in an accurate assessment of a disease or clinical response to a drug. Incidental findings will need to meet the same rigorous burden of proof — that an incidental finding has significant health impact before it is considered important enough to inform the subject.

In 2004, the National Heart, Lung, and Blood Institute (NHLBI) of the National Institutes of Health developed a working group set of recommendations on reporting individual genetic results in research studies.¹² The working group unanimously agreed that there are conditions in which genetic results should be reported to the research participants. To summarize: (1) the risk for disease should be significant; (2) the disease should have important health implications (i.e., fatal or substantial morbidity); and (3) there should be a proven therapeutic or preventative intervention available. For all studies, the NHLBI working group recommended that this reporting plan should be explicit in the study design, with IRB approval after consideration of validity and risks and benefits. Nota-

bly, the working group recommended that only CLIA-certified tests should be reported as clinically valid; if not CLIA-certified, the results should only be reported as clinically valid by confirming them in a CLIA-certified laboratory.¹³ If research results are reported without confirmation in a CLIA-certified laboratory, they should be clearly labeled as “research only.” Some of these recommendations may be problematic. If a genetic variation increases the odds ratio to 2, then the population risk for a disease that may be 1 in 1,000, now doubles to 2 in 1,000. This may be statistically valid, but is it a precise predictor worth reporting? Further, the NHLBI recommendations were driven by study designs targeting specific outcomes and genetic associations. How do these recommendations fit in the context of incidental findings unrelated to the study design? The incidentalome looms in the background of such studies. The researcher may be unaware of the incidental impact of every genetic variation included in large genome-wide studies, the predictive value may be uncertain, and the process of informing the subject becomes problematic.

International bodies also recognize the ethical duty to disclose individual genetic research results. In 1991, and again in 2002, the Council for International Organizations of Medical Sciences' (CIOMS) *International Guidelines for Ethical Review of Epidemiological Studies* recommended that “individual subjects... be informed of any finding that relates to their health status.”¹⁴ The definition of “health status,” however, was nebulous, and the potential for incidental findings within a research study were not addressed. A similar position was expressed by the World Health Organization (WHO).¹⁵ Notably, the CIOMS guidelines also recognized the subject's “right not to know.”¹⁶ In 2002, the Consortium on Pharmacogenomics argued that “researchers are obligated to offer the research participant the option of disclosure of research information when its reliability has been established and when the disclosure is of potential benefit.”¹⁷ Again, definitions of “reliable” and “potential benefit” were left to professional judgment, and incidental observations within a research study were not addressed. The Consortium also recognized the importance of quality assurance, measures to maintain confidentiality, and options for appropriate counseling — all within a legal and ethical framework.¹⁸ Further, the Consortium recommended that the consent form identify who has the responsibility for informing participants in a genetic study of possible health-related findings.¹⁹ Recognizing that researchers have limited expertise in handling medically relevant information, the American Society of Human Genetics,²⁰ as well as the Canadian College

of Medical Geneticists,²¹ recommend that results of DNA analyses should be reported first to an appropriate health care professional, who in turn can decide whether and how to inform the research subject. How medical professionals should be trained to evaluate research data and incidental findings, and trained to report all of this to research participants, is not addressed. This places the medical professional in a position of potential liability.

screens at the time of study submission to the IRB. It should be agreed by all researchers involved in the study that a directed effort to find genetic variations unrelated to the study goals is neither necessary, nor expected. However, when an incidental finding has clear health implications that may be life-threatening or have a strong impact on quality-of-life, and there is an identifiable treatment option, then researchers should have a means to communicate these findings to

The potential for IFs needs to be addressed to provide useful benefits to the research subjects. However, plans for notifying participants of genetic concerns should be evaluated by IRBs in advance and addressed in the consent form. Realizing the predictive power of genomics requires research that, in turn, relies on public trust. That trust depends on appropriate and respectful communication between the researcher and the participant.

Establishing Guidelines for Incidental Findings in Medical Genomic Research

In discussing the *incidentalome* as a potential threat to medical genomic research, Issac Kohane et al. made some compelling recommendations²² that are worth considering and expanding. As the number of genetic tests increase, the likelihood of unanticipated abnormal results will also increase. The simplest standard would be for consent forms to explicitly state that no individual results will be given. For early genetic screening studies, in which there is imprecision as well as lack of clear validation in the results, this may be appropriate.

First, Kohane et al. contend that standards for genetic associations with a disease must be estimated in the general population per ethnic group.²³ With regard to IFs, this may mean that the health impact of a variant may be more or less significant in the context of the genetic background of different populations. That is, variations in genetic modifiers (e.g., other interacting or functionally redundant genes) may be unequally distributed in different ethnic groups, and thus, may alter the health impact of the IF in one group more than another.

Second, Kohane et al. contend that reliable information and data management systems must be established that clearly identify the risks of each genetic variant included in research studies.²⁴ This would establish important guidelines for agreement on the impact of reporting an incidental finding. One might expand this to recommend that a medical professional review potential risk and benefits of the genetic

research subjects. Thus, researchers should establish a pathway for handling incidental findings in proposals to the IRB.

One way to avoid the necessity for reporting incidental findings in research is to maintain complete anonymity of the subject in the data set. This is problematic, because anonymized data do not allow follow-up associations. Often, the most valuable associations are derived from longitudinal studies that follow the course of a disease or treatment. Forcing anonymity into a genetic study design may limit responsibility for reporting IFs, but it can compromise the goals of the study.

Researchers should state how they will deal with IFs in large genomic screens in IRB submissions, and should also define when and how they will disclose IFs to study participants. Any IF disclosed should come with a strong disclaimer that the research was not done under the quality assurance of CLIA certification. While recommendations for further testing may be made, the researcher should not bear the financial responsibility. Moreover, the limited medical expertise of the researcher should remove liability for failure to report all possible relevant genetic associations. A subject should not consider it an “all clear” if no concerns are reported.

Conclusion

Genomic technologies are increasing the range of genome scrutiny. Millions of data points can be collected on every individual. Whole genome sequencing (6 billion bases) is not far behind. Once the entire

sequence of an individual is known, every genetic variation will be revealed. Finding the useful information will involve identifying variations highly associated with disease or significant health issues. Research tools are being developed, but the complexity, the redundancy, and the interactions will not provide simple predictions. These predictions will come in increments, and one researcher's predictive genetic variation will become another's incidental finding. There are currently inadequate federal guidelines addressing IFs and thoughtful consideration is needed so that benefits are realized, but without onerous requirements and threats to the genomic research community. The potential for IFs needs to be addressed to provide useful benefits to the research subjects. However, plans for notifying participants of genetic concerns should be evaluated by IRBs in advance and addressed in the consent form. Realizing the predictive power of genomics requires research that, in turn, relies on public trust. That trust depends on appropriate and respectful communication between the researcher and the participant.

Acknowledgement

This publication was made possible by National Human Genome Research Institute (NHGRI) grant # R01 HG003178 (PI, S. M. Wolf; Co-Is, J. P. Kahn, F. Lawrenz, C. A. Nelson). Its contents are solely the responsibility of the author and do not necessarily represent the official views of NHGRI.

References

1. I. S. Kohane et al., "The Incidentalome: A Threat to Genomic Medicine," *JAMA* 296, no. 2 (2006): 212-215; E. S. Lander et al., "Initial Sequencing and Analysis of the Human Genome," *Nature* 409, no. 6822 (2001): 860-921.
2. J. C. Venter et al., "The Sequence of the Human Genome," *Science* 291, no. 5507 (2001): 1304-1351.
3. International Human Genome Consortium, "Finishing the Euchromatic Sequence of the Human Genome," *Nature* 431, no. 7011 (2004): 931-945.
4. E. Sugiyama et al., "Pharmacokinetics of Gemcitabine in Japanese Cancer Patients: The Impact of a Cytidine Deaminase Polymorphism," *Journal of Clinical Oncology* 25, no. 1 (2007): 32-42.
5. S. M. Wolf et al., "Managing Incidental Findings in Human Subjects Research: Analysis and Recommendations," *Journal of Law, Medicine & Ethics* 36, no. 2 (2008): 219-248.
6. Centers for Medicare and Medicaid Services, "Clinical Laboratory Improvement Amendments," available at <<http://www.cms.hhs.gov/CLIA/>> (last visited December 30, 2007).
7. L. Turney, "The Incidental Discovery of Nonpaternity Through Genetic Carrier Screening: An Exploration of Lay Attitudes," *Qualitative Health Research* 15, no. 5 (2005): 620-634.
8. Unpublished observation of author (November 11, 2007).
9. D. I. Shalowitz and F. G. Miller, "Disclosing Individual Results of Clinical Research: Implications of Respect for Participants," *JAMA* 294, no. 6 (2005): 737-740.
10. See Kohane et al., *supra* note 1.
11. B. M. Knoppers et al., "The Emergence of an Ethical Duty to Disclose Genetic Research Results: International Perspectives," *European Journal of Human Genetics* 14, no. 10 (2006): 1170-1178.
12. National Heart, Lung & Blood Institute (NHLBI), *NHLBI Working Group on Reporting Genetic Results in Research Studies*, Bethesda, MD, 2004, available at <<http://www.nhlbi.nih.gov/meetings/workshops/gene-results.htm>> (last visited November 11, 2007).
13. *Id.*
14. Council for International Organization of Medical Sciences (CIOMS), *International Guidelines for Biomedical Research Involving Human Subjects*, Geneva, 2002, available at <http://www.cioms.ch/frame_guidelines-nov_2002.htm> (last visited November 11, 2007).
15. World Health Organization (WHO), *Genetic Databases: Assessing the Benefits and the Impact on Human and Patient Rights*, Geneva, 2003, available at <<http://www.law.ed.uk/ahrb/publication/online/whofinalreport.doc>> (last visited November 11, 2007).
16. *Id.*
17. Consortium on Pharmacogenetics, *Ethical and Regulatory Issues in Research and Clinical Practice*, Minneapolis, 2002, available at <http://www.bioethics.umn.edu/News/pharm_report.pdf> (last visited November 11, 2007).
18. *Id.*
19. *Id.*
20. American Society of Human Genetics (ASHG), *DNA Banking and DNA Analysis: Points to Consider*, Bethesda, MD, 1998, available at <<http://genetics.faseb.org/genetics/ashg/policy/pol-02.htm>> (last visited November 11, 2007).
21. Canadian College of Medical Geneticists, "Policy Statement Concerning DNA Banking and Molecular Diagnosis," *Clinical Investigation of Medicine* 14 (1991): 363-365.
22. See Kohane et al., *supra* note 1.
23. *Id.*
24. *Id.*