The Incidentalome
A Threat to Genomic Medicine

Isaac S. Kohane, MD, PhD

Daniel R. Masys, MD

Russ B. Altman, MD, PhD

GENOMIC MEDICINE IS POISED TO OFFER A BROAD ARray of new genome-scale screening tests. However, these tests may lead to a phenomenon in which multiple abnormal genomic findings are discovered, analogous to the "incidentalomas" that are often discovered in radiological studies. If practitioners pursue these unexpected genomic findings without thought, there may be disastrous consequences. First, physicians will be overwhelmed by the complexity of pursuing unexpected genomic measurements. Second, patients will be subjected to unnecessary follow-up tests, causing additional morbidity. Third, the cost of genomic medicine will increase substantially with little benefit to patients or physicians (but with great financial benefits to the genomic testing industry), thus throwing the overall societal benefit of genome-based medicine into question. In this article, we discuss the basis for these concerns and suggest several steps that can be taken to help avoid these substantive risks to the practice of genomically personalized medicine.

Diagnostic Testing and Incidental Findings

Physicians are generally trained to order tests carefully and only if such tests will result in a change in management. For this reason, much time is spent deciding if a renal panel with 7 blood measurements should be expanded to a comprehensive panel with 20 or more measurements. Physicians know that as the number of tests increases, the chance that a spurious abnormal test result will arise also increases. They also know that it is difficult to ignore abnormal findings, and they often must embark on a sequence of more expensive tests to investigate the findings. Furthermore, the significance of an abnormal finding is related to the prevalence of disease in the population from which the tested patient is drawn. Therefore, if the risk associated with the finding was established in a population with a high prevalence of disease, the rate of false-positive results when testing in a population with a lower rate of disease will be much higher.

CME available online at www.jama.com

There is a rich literature in radiology on the "incidentaloma," which is a finding (most commonly a mass) found on computed tomography or magnetic resonance imaging studies ordered for symptoms or concerns totally unrelated to the gland in which the mass is found. The workup of an incidentaloma is complicated by concerns that it may be associated with malignant disease and, at least initially, the lack of good data on the prevalence of malignant disease in the general population. Incidentalomas occur because imaging modes do not only report on the areas of direct clinical concern but, incidentally, report on all organs in the field of view.[1]

This phenomenon of possible incidental genomic findings—the *incidentalome*—threatens to undermine the promise of molecular medicine. In particular, the application of comprehensive genotype and functional genomic measurements across the general population is likely to yield unexpected incidental findings for nearly everyone. Of course, there are important differences in the interpretation of genomic data and radiological data (eg, discovering incidentalomas may be lifesaving), but the potential similarity is that the clinician and patient are confronted with results that they did not anticipate when the test was ordered.

The sequencing of the human genome has brought increasing interest in the use of genome-scale technologies to measure individual variation in the human genome. A variety of technologies have emerged that make it economically attractive to assess the structure and function of hundreds of thousands of genes simultaneously. Although all humans share more than 99.8% of their genome DNA sequence, the remaining 0.2% (along with environmental exposures) is responsible for much of the variation in risk of disease and response to therapies. Recent reports indicate that more than 300 000 single-nucleotide polymorphisms can be measured on an individual genome for a few hundred dollars.[2] Clinical genomics studies have shown that the expression pattern of thousands of genes can differentiate cancer cells from normal cells and can distinguish sub-

Author Affiliations: Center of Biomedical Informatics and Division of Health Sciences Technology, Harvard Medical School, Boston, Mass (Dr Kohane); Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, Tenn (Dr Masys); and Department of Genetics, Stanford University, Stanford, Calif (Dr Altman).
Corresponding Author: Isaac S. Kohane, MD, PhD, Center of Biomedical Informatics and Division of Health Sciences Technology, Harvard Medical School, 10 Shattuck St, Boston, MA 02115 (isaac_kohane@harvard.edu).

types of cancer,[3-5] inflammatory bowel disease,[6,7] neurodegenerative disease,[8,9] and many others.[10]

Many companies are competing to create high-reliability, high-throughput assays for measuring thousands of genotypes and cellular phenotypes in order to create the infrastructure for molecular medicine. In many cases, comprehensive panels are being created that provide measurements of large numbers of genomic variables. These panels raise the possibility that clinical use of genomics may not be entirely focused on specific gene variants but may also include a number of related variants. For example, one panel accurately measures more than 30 polymorphisms in just 2 genes. These genes come from a group of more than 200 that may be important for drug dosing, so future generations of this panel may test for $200 \times 15 = 3000$ polymorphisms. In the best case, this infrastructure will provide the basis for genome-informed medical decision making that will lead to diagnoses and therapies that are more targeted, have reduced variability in outcome, maximize efficacy, and minimize adverse effects.

Potential Implications of Genomic Testing

The following 3-part hypothetical scenario illustrates some of the implications of application of genomic testing:

There exists a single genomic test that has 99.9% sensitivity (true-positive rate) and a false-positive rate of 0.1% (ie, specificity or a true-negative rate of 99.9%) for a rare treatable disease, "X." For comparison, cystic fibrosis tests have been reported with 99% sensitivity[11] and *BRCA1* testing shows a sensitivity of 81% when the false-positive rate is 42%.[12] The hypothetical test in this case was developed in a study of families with disease X, and it works well on this population, which has a disease prevalence of 1 in 1000, much greater than in the general population. Specifically, if 1000 individuals from this population are tested, then there will be 1 true-positive, 1 false-positive, and 998 true-negative results. Two individuals (those with the true-positive and false-positive results) are tested further at some expense; the one with the true-positive result is given a diagnosis and is treated. The person with the false-positive result is tested and disease is ruled out. The conclusion in this case is that genome measurements are useful for diagnosing disease X. This occurs because the pretest probability of disease of 0.001 yields a posttest probability of 0.5 if the test result is positive, the sensitivity of the test is 0.999, and the false-positive rate of the test is 0.001, using the Bayes theorem.[13]

However, if this same genomic test is applied to the general population (with no such occurrence of the disease in their kinship), the overall disease prevalence is 1 in 100 000, or a pretest disease probability of 0.00001. If a general population of 10 000 000 individuals is tested, 100 will have the disease, 10 000 people will test positive with no disease, 100 people with disease will be missed, and 9 989 900 people will have a negative test and no disease. Thus, 10 100 people test positive and require follow-up. One hundred are accu-

Figure. Percentage of Total Population With a False-Positive Test Result
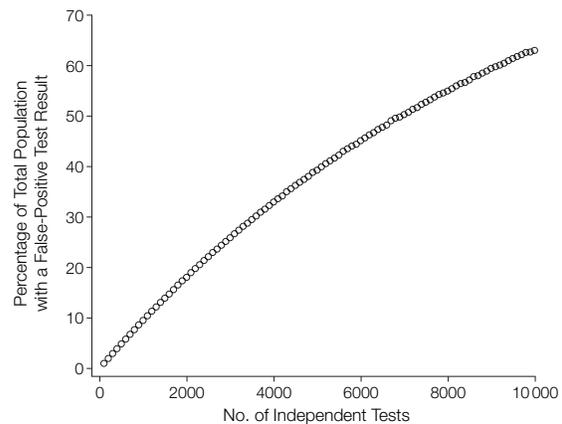
As the number of tests increases to 10 000, the fraction of the population that has a false-positive test result increases to more than 60%. Any large-scale genomic panel is therefore likely to routinely report false-positive results. The data for this figure were generated by running a simulation in which a population of 100 000 was tested with 1 through 10 000 tests, each with a sensitivity of 100% and a false-positive rate of 0.01%. That is, 10 individuals with false-positive tests were randomly selected from the population for each test. Because some individuals could be selected more than once with a larger panel of tests, the increase in the number of individuals with false-positive test results is less than linear.

rately identified as having the disease, but the cost of doing so is very high because 10 000 have a workup and are found to not have the disorder. Thus, in this population a positive test result raises the posttest probability that an individual has disease only from 0.00001 to 0.0099, or less than 1 in 1000. Now, the conclusion is that this is a poor test for screening and leads to too many false-positive results. This problem will be replayed with individuals coming from different ethnic and geographic backgrounds. As demonstrated by the HapMap project,[14] these populations can differ in the frequencies of several genomic markers.

The first example illustrates the use of a single genomic test. What if the general population is screened for several genetic variants at once? Suppose there is a panel of genomic tests, each with superb testing performance: a sensitivity of 100% and a false-positive rate of 0.01%. That is, of 100 000 individuals, each test will only produce 10 false-positive results. Assuming a disease prevalence of 1 in 100 000, in a population of 100 000 the number of true-positive results will increase by 1 with each additional test. The increase in the number of false-positive results will be 10 with each independent test, but some individuals will be subject to multiple false-positive results; therefore, the increase of the number of individuals with a false-positive result will be slightly less than 10 per test. The **FIGURE** shows the increase in the proportion of individuals with a false-positive test result under these assumptions. As illustrated, with 10 000 independent tests, more than 60% of the entire population tested would have false test results.

There now exist genomic panels with hundreds of thousands of tests (eg, whole-genome single-nucleotide polymorphism microarrays currently used in experimental settings). Not all of these tests are independent, but they also do not have the stringent sensitivity and false-positive rate of this example. Even an order of magnitude fewer falsely "diagnosed" cases would extract enormous financial and health costs from our society.

Unfortunately, even if genomic tests were to achieve the unrealizable—100% sensitivity and a false-positive rate of 0—the risks of the incidentalome remain. A significant pathological disease burden never reaches clinical significance and is unrelated to the ultimate cause of death. For example, a high number of incidental pituitary microadenomas are found in cadavers,[15] and a large number of prostate carcinomas accurately diagnosed after the finding of an elevated prostate-specific antigen level in all likelihood would not contribute to an individual's death.[16] It is highly likely that a substantial proportion of these incidental findings will be partly accounted for by genetic risk factors measured in an accurate comprehensive genomic test panel. One hundred percent accurate identification of such incidental pathologies will lead to iatrogenic pathology. That is, these real findings without realistic clinical importance can lead to aggressive diagnostic and therapeutic investigations in an otherwise healthy individual.

Genomic Medicine in Clinical Practice

For genomic medicine to achieve economies of scale, it may be tempting for vendors to offer technologies for measuring multiple genomic measurements simultaneously, not just the few that are relevant to the clinical question. Physicians ordering these tests will be put in a difficult position: ignore the results of incidental findings (and take on the risk of liability should they be clinically meaningful), refrain from ordering the tests (and bring genome-informed medicine to a standstill), or feel compelled to spend millions of health care dollars with more expensive "gold standard" tests (either ordering individual tests immediately or alerting the patient and watching and testing them over long periods) to follow-up on these test results. Patients would have printouts from the Internet describing all the diseases associated with their own genetic testing results. Physicians would soon realize that these tests are not appropriate for their patients and would probably stop using them altogether. Some authors have suggested that a generic genetic consent process could limit the requirement for following up on genetic information of limited value.[17] Lacking these, insurance companies, employers, and governments could conclude that genetic testing is noisy and imprecise and should not be pursued as a matter of fiscal responsibility and in the best interests of patients.

Can this scenario be avoided and benefit still be derived from the tremendous promise of genomic medicine? The following key actions may help to avoid these genomic doomsday scenarios caused by the incidentalome.

First, the overall disease prevalence for all diseases with a genetic component must be estimated in the general population per ethnic group. Of course, the definition of an ethnic group is difficult but can be informally defined as a collection of individuals with substantially shared ancestry. Operationally, each such human population should be identified and defined at a granularity compatible with available resources. This will allow the sensitivity and false-positive rate of each individual genomic test to be combined with prevalence to estimate the real overall risk of a positive test result based on approximate ancestry.

This is not an easy task, but an initial analysis would at least indicate those disease prevalences that vary widely across populations, and thus deserve more consideration when creating genetic tests. This will enable physicians to understand the incredibly low risks for most incidental findings, and appropriately ignore them. The Online Mendelian Inheritance in Man resource currently has 16 600 entries describing abnormal genotypes, phenotypes,[18] or both, which probably represent only a small fraction of diseases that must be characterized with respect to genetic risk. In addition to cataloging the prevalence of disease, it will be important to document the population-based prevalence of polymorphisms, not just at the 5% or 10% level, but at the level that will be measured in the population. Rare disease-associated genetic variants with 1-in-10 000 frequency (0.01%) will be present in 35 000 persons in the United States (assuming a population of 350 million). Obtaining this level of detail will require mobilizing a large proportion of the population as study participants for estimating risk. The sociological and legislative challenges of such an effort will dwarf the technical challenges.[19]

Second, information systems must be created for use in the clinic and at the bedside for estimating and explaining the risks associated with various incidental genomic findings. The current interest and emphasis on a national health information infrastructure in the United States is reassuring in this regard, but the need for this infrastructure to enable genome-informed medical decision making is not generally cited as a prime driver. Physicians must have ready access to the significance and risks of positive genomic results so they can understand the real risks to their patients, explain them, and make cost-effective decisions about subsequent testing. The alternative of nonphysician or even direct-to-consumer automated decision support[20-22] may well grow if the health care establishment fails to take the lead in adopting such tools. With regard to genetic information, the mode of delivery of genetic testing results to physicians and patients must be carefully considered. It is already possible for patients to order genetic tests directly,[23] and the Internet provides a medium for delivery of whole-genome data. One can even imagine that Internet businesses may offer analyses of genomic data to help patients understand the results.

Third, physicians and other health care professionals must be educated in the importance of rational interpretation of ge-

nomic tests. Attitudes of genetic determinism—the belief that genetics completely determines phenotypic outcome—must be informed by an understanding that most genetic measurements only shift the probability of an outcome, which often depends on other environmental triggers and chance. The importance of prevalence in probability computations must be understood and supported in clinical decision support systems. Educators must make clear the unfavorable implications of ordering tests "just to be on the safe side."

Finally, physicians and medical specialty groups must decide if a genome-wide panel (that is, a panel of 500 000 genetic polymorphisms all ordered and measured together), however cost-effective to measure, has any role in clinical medicine, or if a series of more focused genomic-based panels, with clear indications for use and proper protocols for workup of unexpected findings, are more attractive. In our opinion, it is imprudent to use testing panels comprising a sizable fraction of the genome for clinical care or screening. There are considerable logistical and financial implications for companies involved in creating diagnostic tests in how these tests are delivered in practice. It will be incumbent on practitioners to ensure that there is appropriate clinical justification and convincing market pressure to perform these tests in a manner that ushers in the era of genome-informed medical decision making and does not allow the incidentalome to block its arrival.

REFERENCES

1. Illes J, Kirschen M, Edwards E, et al. Incidental findings in brain imaging research. *Science*. 2006;311:783-784.

2. Affymetrix. http://www.affymetrix.com/products/pricing/index.affx. Accessed June 20, 2006.

3. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*. 1999;286: 531-537.

4. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Expression profiling predicts outcome in breast cancer. *Breast Cancer Res*. 2003;5:57-58.

5. Mullighan CG, Flotho C, Downing JR. Genomic assessment of pediatric acute leukemia. *Cancer J*. 2005;11:268-282.

6. Heller RA, Schena M, Chai A, et al. Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proc Natl Acad Sci U S A*. 1997;94:2150-2155.

7. Csillag C, Nielsen OH, Borup R, Nielsen FC. Microarrays and Crohn's disease: collecting reliable information. *Scand J Gastroenterol*. 2005;40:369-377.

8. Ermak G, Davies KJ. Gene expression in Alzheimer's disease. *Drugs Today (Barc)*. 2002;38:509-516.

9. Miller RM, Federoff HJ. Altered gene expression profiles reveal similarities and differences between Parkinson disease and model systems. *Neuroscientist*. 2005; 11:539-549.

10. Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet*. 2006;7:55-65.

11. Rock MJ, Hoffman G, Laessig RH, Kopish GJ, Litsheim TJ, Farrell PM. Newborn screening for cystic fibrosis in Wisconsin: nine-year experience with routine trypsinogen/DNA testing. *J Pediatr*. 2005;147(3 suppl):S73-S77.

12. James PA, Doherty R, Harris M, et al. Optimal selection of individuals for *BRCA* mutation testing: a comparison of available methods. *J Clin Oncol*. 2006;24:707-715.

13. Ogino S, Wilson RB. Bayesian analysis and risk assessment in genetic counseling and testing. *J Mol Diagn*. 2004;6:1-9.

14. Kruglyak L. Power tools for human genetics. *Nat Genet*. 2005;37:1299-1300 Med.

15. Teramoto A, Hirakawa K, Sanno N, Osamura Y. Incidental pituitary lesions in 1000 unselected autopsy specimens. *Radiology*. 1994;193:161-164.

16. Hugosson J, Aus G, Becker C, et al. Would prostate cancer detected by screening with prostate-specific antigen develop into clinical cancer if left undiagnosed? a comparison of two population-based studies in Sweden. *BJU Int*. 2000;85:1078-1084.

17. Elias S, Annas GJ. Generic consent for genetic screening. *N Engl J Med*. 1994; 330:1611-1613.

18. Online Mendelian Inheritance in Man. http://www.ncbi.nlm.nih.gov/Omim /mimstats/html. Accessed June 20, 2006.

19. Kohane IS, Altman RB. Health-information altruists—a potentially critical resource. *N Engl J Med*. 2005;353:2074-2077.

20. Miller SM, Fleisher L, Roussi P, et al. Facilitating informed decision making about breast cancer risk and genetic counseling among women calling the NCI's Cancer Information Service. *J Health Commun*. 2005;10(suppl 1):119-136.

21. Mouchawar J, Hensley-Alford S, Laurion S, et al. Impact of direct-to-consumer advertising for hereditary breast cancer testing on genetic services at a managed care organization: a naturally-occurring experiment. *Genet Med*. 2005; 7:191-197.

22. Mouchawar J, Laurion S, Ritzwoller DP, Ellis J, Kulchak-Rahm A, Hensley-Alford S. Assessing controversial direct-to-consumer advertising for hereditary breast cancer testing: reactions from women and their physicians in a managed care organization. *Am J Manag Care*. 2005;11:601-608.

23. Javitt GH. Policy implications of genetic testing: not just for geneticists anymore. *Adv Chronic Kidney Dis*. 2006;13:178-182.