Louisiana Law Review

Volume 66
Number 5 *Special Issue*
Symposium: Proceedings of "The Genomics Revolution? Science, Law and Policy"

Article 5

12-1-2005

Risk Prediction for Common Diseases

Paula W. Yoon

Risk Prediction for Common Diseases

Paula W. Yoon[*]

My topic for discussion today is whether the promising field of genomics is going to allow us to more accurately predict who in the population is at future risk for common chronic diseases. Before attempting to answer this question, I will provide some general background and will then give some specific examples in the area of heart disease, diabetes, and obesity.

I work in the Office of Genomics and Disease Prevention at the Centers for Disease Control ("CDC") and, like most of CDC, our focus is on disease prevention. Our office has a crosscutting role, working with all the programs at CDC to help them integrate genomics into what they do, whether it is research, policy, or practice.

The year 2003 was certainly the year of the human genome. We celebrated the completion of the human genome project and the fiftieth anniversary of Watson and Crick's discovery of the double helix structure of DNA.[1] One headline claimed, "DNA has changed the world: But now what?" That is the big question.

As scientists, we are pretty good at collecting DNA. We can get DNA from all kinds of body fluids and parts; it is amazing where we can find DNA.[2] Our technology for analyzing DNA, which now includes whole genome scans and chips, is also improving immensely, rapidly becoming faster and cheaper.[3] But what does this mean for the health of individuals, families, and communities?

The work that the scientific community is doing can be thought of on a continuum, a translation continuum from gene discovery to tools and processes that can be used to prevent and treat disease. Along that continuum, there are various research methods that are used to study genomics. First, there are family studies or linkage

---

[*] ScD, MPH, Epidemiologist, Office of Genomics and Disease Prevention, Centers for Disease Control and Prevention (CDC). This work is based upon a live presentation made on February 5, 2004 and does not necessarily reflect events and changes thereafter.

1. Horace F. Judson, *The Greatest Surprise for Everyone—Notes on the 50th Anniversary of the Double Helix*, 348 New Eng. J. Med. 1712, 1712–14 (2003).

2. Lin Zhang et al., *Whole Genome Amplification From a Single Cell: Implications for Genetic Analysis*, 89 Proc. Natl. Acad. Sci. USA 5847, 5847–51 (1992).

3. Alan E. Guttmacher & Francis S. Collins, *Genomic Medicine—A Primer*, 347 New Eng. J. Med. 1512, 1512–20 (2002).

studies, which look for genetic differences within small groups. Next, there are population studies of different kinds—case-control studies, cohorts, and so forth. When established associations lead to new genetic tests or tools, there are clinical trials. Finally, after studying a test or tool under controlled circumstances, we need to evaluate its validity and utility in real circumstances. This last step—which should occur before it is introduced for widespread clinical or public health application—is often skipped or glossed over. I am going to come back to this validation gap at the end of my talk. Certainly, the use of emerging technology like whole genome scans and haplotype mapping [4] make the field of genomic research very exciting. The promise is there but we have a long way to go from gene discovery to treatment and prevention of disease.

When genomic applications, like genetic tests, are carried successfully through the research continuum what are their potential uses for treating and preventing disease? There are basically four areas of potential application. First is the possibility of guiding drug therapy. Pharmacogenomics, or individualized medication based on genetically determined variation in effects, is probably one of the most promising areas in the field of genomics today. There is individual variation in the enzymes that metabolize, absorb, and transport pharmaceuticals. There has also been some success in identifying genes that cause some individuals to have extremely positive or extremely negative effects from drugs. Another related example is the use of DNA probes to identify pathogens. For example, in the diagnosis of meningitis,[5] DNA probes can be used to determine fairly rapidly whether the illness is due to bacteria or a virus and, specifically, which pathogen. This allows for quicker application of the most effective therapy.

A second way in which genomics can be used to prevent disease is by modifying the environment. If we can identify individuals who are more susceptible to disease because of their genetic make-up, it may be possible to reduce their risk through individual behavior changes, such as diet, exercise, or smoking cessation. Environmental modification can also occur through community interventions, such as spraying for mosquitoes or

---

4. Francis S. Collins et al., *A Vision for the Future of Genomics Research: A Blueprint for the Genomic Era*, 22 Nature 835, 839 (2003).

5. A. Backman et al., *Evaluation of an Extended Diagnostic PCR Assay for Detection and Verification of the Common Causes of Bacterial Meningitis in CSF and Other Biological Samples*, 13 Molecular and Cellular Probes 49, 49–60 (1999).

having more sidewalks in communities so people can walk to school and work. Workplace interventions are another possibility. For example, if we know that within a particular workplace certain people are more susceptible to substances like chemicals and pesticides, exposures levels can be lowered for the safety of everyone.

Gene therapy is another potential option for treating disease, but this is still experimental. There has been some success with using gene therapy to treat single gene disorders, such as primary immune deficiency, but results are mixed.[6] The National Institutes of Health ("NIH") is funding a number of studies in gene therapy research to develop treatments for cancer, heart disease, and AIDS. This is an area of research that looks promising, but we have a long way to go.

A fourth avenue for disease prevention is offering targeted screening and interventions based on increased susceptibility.[7] For example, a person at risk for hereditary colorectal cancer may be encouraged to be screened earlier or more frequently than is normally recommended and may benefit from more intensive screening methods, such as colonoscopy. Presymptomatic medical therapies may also be helpful, whether they are as simple as taking an aspirin a day or a prescribed medication.

These four main areas—drug therapy, environmental modification, gene therapy, and targeted interventions—are the key strategies we think of when we talk about using genetics and genomics to help us treat and prevent disease.

Where are we today on the research continuum for preventing common chronic diseases like heart disease, diabetes, and obesity? The last estimate from the Human Genome Project ("HGP") was that humans have between twenty and twenty-five thousand genes,[8] as well as millions of variants of these genes. Common chronic diseases result from the interactions of multiple genes with multiple environmental and behavioral factors. Our epidemiological and statistical methods are currently limited in their ability to find meaningful associations among all of these interacting factors.

Occasionally, chronic diseases may be due to the influence of single genes. More often, smaller effects are reported in a way that

---

6. Rebecca H. Buckley, *Primary Immunodeficiency Diseases: Dissectors of the Immune System*, 185 Immunol. Rev. 206, 206–19 (2002).

7. Muin J. Khoury et al., *Do We Need Genomic Research for the Prevention of Common Diseases with Environmental Causes?*, 161 Am. J. Epidemiol. 799, 799–805 (2005).

8. Human Genome Project, http://ww.ornl.gov/sci/techresources/ Human_ Genome/home.shtml.

can be misleading. For example, the popular press reported in late 2003 that doctors had found a link between MEF2A and heart disease.[9] Some of the statements that appeared in print were along the lines of: researchers identify defects in a gene that are directly linked to heart attacks. Everyone who has this gene mutation is destined to have the disease. If you do not have this gene, you appear to be free from developing the disease. When findings are presented this way, they can be misunderstood by the public.

If you look at the article itself, you will find that mutation of the MEF2A is an inherited disorder with features of coronary artery disease.[10] This MEF2A mutation is very rare, occurring in very few families. The study was done in Iowa, where researchers found a fairly large extended family that carried this lethal mutation and, yes, almost all family members who had this mutation ended up with coronary artery disease.[11]

Findings from this study may contribute to our understanding of factors that cause heart disease in the wider population but the public needs to understand that it is not as simple as finding the "heart disease gene," the "cancer gene," or the "diabetes gene." However, after this news appeared in the popular press, there were probably people going into their physicians' offices asking to be tested for the "coronary artery disease gene."

Another very interesting area in genomic research is the search for genetic factors that play a role in the effectiveness of diet and physical activity. Why, given the same body mass index, are some people who engage in an exercise program successful while others are not? Certainly it is more than just determination and metabolism!

I wanted to show the Obesity Gene Map Database ("OGMD")[12] because this comes from the Pennington Biomedical Research Center and Dr. Bouchard is the senior author on these papers. The OGMD is a wonderful example of an informatics tool for genomics population health research, the challenge of which extends well beyond genotyping and collecting the data.[13] We are deluged with information, and we need better tools to organize this information and determine what is meaningful. The OGMD presents information on the different obesity syndromes and phenotypes identified thus far, along with lots of information about

9. Lejin Wang et al., *Mutation of MEF2A in an Inherited Disorder with Features of Coronary Artery Disease*, 302 Science 1578, 1578–81 (2003).

10. *Id.*

11. *Id.*

12. Obesity Gene Map Database, http://obesitygene.pbrc.edu/.

13. Stephen Chanock & Sholom Wacholder, *One Gene and One Outcome? No Way*, 8 Trends in Molecular Med. 266, 266–69 (2002).

linkage studies, candidate genes, and mouse models. To date, it looks like there are over 300 genes and markers that have been identified and associated with obesity.

Data-mining tools[14] are another invaluable resource. There are so many publications that it is becoming virtually impossible for researchers working in most areas of genomics to sort through the literature and figure out what is important. Data-mining techniques can be used to identify pertinent literature as well as find gene-gene, gene-environment, and gene-disease associations that may not have been explicitly noted before. When they work well, these data-mining techniques pull the most relevant materials to the forefront and flag what is most interesting.

So, apart from a few obesity syndromes, are there any clinical applications for obesity based on what we have learned from genomic research so far? We have learned that "fat stores are regulated over long periods of time by complex systems that involve input and feedback from fatty tissues, the brain and endocrine glands, like the pancreas and thyroid."[15] We also know that these systems cannot be studied in a static fashion. We cannot just measure a person's BMI at one point in time, and then look at their susceptibility genes. The research is much more complex. The most promising area currently appears to be pharmacogenomics, developing new drug strategies to target satiety and numerous metabolic pathways. In the meantime, it is important to realize that genetic susceptibilities are not destiny. We can make an impact on obesity by modifying behaviors and environment. Of course, for the most part, this approach at the population level has not been easy or particularly effective.

What about Type 2 diabetes? Again, there are single gene defects, such as those resulting in MODY, or Maturity Onset Diabetes of Youth, as well as syndromes like insulin resistance, where we have learned quite a bit. Like obesity, diabetes is due to complex interactions of multiple genes and environmental factors. The candidate gene approach has been used for the most part to find genes that regulate insulin signaling and secretion. Genome-wide scans in high-risk families have searched for major susceptibility loci and studies of specific ethnic groups with high rates of diabetes, such as the Pima Indians, have generated more insight than we have into many other chronic diseases. However,

---

14. Ying Liu et al., *Text Mining Functional Keywords Associated with Genes*, 11 Medinfo 292, 292–96 (2004).

15. CDC Office of Genomics and Disease Prevention, *Obesity and Genetics: What We Know, What We Don't Know and What It Means*, Feb. 2002, http://www.cdc.gov/genomics/info/perspectives/files/obesknow.htm.

it is still not clear why diabetes clusters in some specific racial and ethnic groups—for example, Native American, Hispanic, and some Scandinavian populations.[16] Other factors that help to predict the onset of diabetes include obesity, insulin resistance, hyperglycemia, dyslipidemia, and family history.[17]

Family history as a risk factor for common diseases is receiving renewed interest. A review of about fifteen different studies looking at family history as a risk factor for diabetes shows a consistent, positive association, with relative risk ranging between 1.5 and six.[18] The variation is due to the types of relatives that were included in the family history assessment, the number of relatives affected, age at disease onset, and so forth.[19] Although the studies were not entirely comparable, the evidence indicates that family history is a very strong risk factor for diabetes. Family history is a genomic tool that reflects shared genetic susceptibilities, environment, and behaviors.

One of the groundbreaking studies in family history was done by Roger Williams and his group in Utah, where they conducted a family history assessment in the school system.[20] They worked with the Education Department to include a family history module in their health education programs in the schools, and they covered nearly the entire state.[21] Students took home a family history questionnaire and called their relatives and solicited all kinds of health information.[22] The researchers calculated a family history risk score for cardiovascular disease based on this information and identified families at high risk.[23] Seventy-two percent of early coronary heart disease ("CHD") occurred in just fourteen percent

16. Jared Diamond, *The Double Puzzle of Diabetes*, 423 Nature 599, 599–602 (2003).

17. The Expert Comm. On the Diagnosis and Classification of Diabetes Mellitus, *Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus*, 26 Diabetes Care S5, S5-20 (2003).

18. Tabitha A. Harrison et al., *Family History of Diabetes as a Potential Public Health Tool*, 24 Am. J. Prev. Med. 152, 153–55 (2003).

19. Maren T. Scheuner et al., *Family History: A Comprehensive Genetic Risk Assessment Method for the Chronic Conditions of Adulthood*, 71 Am. J. Med. Genetics 315, 315–24 (1997).

20. Roger R. Williams et al., *Usefulness of Cardiovascular Family History Data for Population-Based Preventive Medicine and Medical Research (The Health Family Tree Study and the NHLBI Family Heart Study)*, 87 Am. J. Cardiol. 129, 129–35 (2001).

21. *Id.* at 130.

22. *Id.*

23. *Id.* at 130–31.

of the families, and eighty-six percent of early stroke occurred in eleven percent of families.[24]

These and similar studies have been the impetus for a number of initiatives around the country to promote the use of family history as a genomic tool for disease prevention. You may ask, "Why focus on family history? Doctors have been collecting family histories forever." They have, but it has not been systematic and often the information is not carefully assessed or used for prevention. Usually a form is filled out in the waiting room and then filed in a medical record. How can we better utilize this family history? Perhaps we can use it to assess risks for common chronic diseases and to influence early screening, or to educate people about prevention measures, such as behavior changes.

At the CDC, we began an initiative in 2002 in collaboration with several NIH institutes, a number of professional organizations, and academic institutions to evaluate the use of family history for assessing risk of common diseases and influencing early detection and prevention strategies. We began with a review of the literature to identify gaps in our knowledge of family history as a risk factor for chronic diseases. We brought together a group of experts to present the evidence at a workshop in February 2003 and published a series of papers based on the presentations in the *American Journal of Preventive Medicine*.[25] We then formed a work group and started examining the family history tools and strategies that were currently in use. We came up with some criteria for what we thought would make a good family history tool for public health. Now we are in the process of developing a new tool with a contractor and will conduct extensive pilot testing and evaluation studies. If we find that this tool is valid and useful, we will follow with public health campaigns about the importance of knowing your family history and provide education programs to improve the use of family history in clinical settings.

The tool will include a data collection component where individuals can answer questions about the health history of their close relatives, risk assessment algorithms that will provide a qualitative assessment of risk for each disease included in the tool, and a report that will include suggested prevention strategies. For

---

24. Steven C. Hunt et al., *Family History Assessment: Strategies for Prevention of Cardiovascular Disease*, 24 Am. J. Prev. Med. 136, 137–38 (2003).

25. Paula W. Yoon et al., *Research Priorities for Evaluating Family History in the Prevention of Common Chronic Diseases*, 24 Am. J. Prev. Med. 128, 128–35 (2003).

example, someone at moderate familial risk for heart disease who smokes, is overweight, does not exercise, and eats high levels of saturated fat might be given a priority list of behaviors to change based on their family history. Individuals with a high familial risk for colorectal cancer may need a genetic work-up, and earlier, more frequent, and intensive screening.

The prototype of the tool focuses on six diseases: heart disease, stroke, diabetes, colorectal cancer, breast cancer, and ovarian cancer. Three research groups—the University of Michigan School of Medicine, Evanston Northwest Health Care Research Institute, and Case Western Reserve University—will conduct an evaluation of the tool when it is completed. The study will determine if familial risk assessment and personalized prevention recommendations are effective in motivating people to change their lifestyle and screening behaviors.

In summary, specific measures must be taken to translate this new science of genomics into tools and processes that are going to help us treat and prevent disease. We need large-scale, population-based collaborative research because, when you start looking at multiple genes and multiple environmental factors to stratify risks, you need big numbers to find meaningful associations. The new biobanks now being created in some countries may help provide the large population samples that will be needed. We also need to develop complementary statistical methods because those we have now are not sufficient to grasp the complexity of the data. We need public health and clinical assessments of emerging genetic technologies. We need to look seriously at the validity and utility of what we are doing and the tools we are developing—including genetic tests. Just because we can do something does not mean that we should do it, especially on a population basis. And, we need to thoroughly evaluate new genomic tools and tests in order to develop evidenced-based policies and practices.

Returning to the translation continuum, there is a particular need to focus on the validation gap. Studies must be done to validate what is being undertaken, in terms of analytic validity, clinical validity, clinical utility, and the ethical, legal, and social issues that are associated with the technology. There are some areas where this has been done fairly well—for example, in the area of breast cancer testing. I think most people would agree that *BRCA1* and *BRCA2* testing as presently done—in appropriate groups of women, who are at very high risk and meet the criteria for testing based on family history—has been shown to be valid and useful.

Other tests have totally skipped the validation process and have gone directly to the consumer, such as a cardiogenomic and

nutrigenomic profiles offered by several companies. You send your buccal sample by mail; they do a series of analyses on certain genetic susceptibility genes and then give you a profile of your risk of disease with some suggested preventive measures. The problem with these tests is that the data showing validity and utility, if they even exist, are not in the public domain. How can we develop effective and safe policies and practice guidelines with no information about their validity and utility? These are some of the challenges that will face public health in the new genomic era.

The research is very exciting and there is much promise for finding new strategies to prevent and treat disease, but we must continue to emphasize the importance of validation and the adoption of evidenced-based practices.