

## **Genomic Medicine:**

### **A Future Flooded with Risk Information\***

**Edward R.B. McCabe\*\* and Linda L. McCabe\*\*\***

#### INTRODUCTION

Risk prediction in the practice of medical genetics is probabilistic and not absolute. The complexity of risk

---

© 2007 Edward R.B. McCabe and Linda L. McCabe.

\* This article results from a recent symposium entitled "The Risks Posed by New Biomedical Technologies: How Do We Analyze, Communicate & Regulate Risk?" presented by the University of Minnesota's Consortium on Law and Values in Health, Environment & the Life Sciences, co-sponsored by the Joint Degree Program in Law, Health & the Life Sciences; the Center for Bioethics; and the Center for Environment and Health Policy in May 2006.

\*\* Edward R.B. McCabe, M.D., Ph.D. serves as Professor and Mattel Executive Endowed Chair of the Department of Pediatrics at the David Geffen School of Medicine at UCLA. He holds a joint appointment in the Department of Human Genetics. He is responsible for establishing the Mattel Children's Hospital at UCLA, where he serves as Physician-in-Chief. He is a Co-Director of the UCLA Center for Society and Genetics. He was the President of the American Board of Medical Genetics, the President of the American College of Medical Genetics and the President of the Western Society for Pediatric Research. He is the current President of the American Pediatric Society. In 2009, he will be the President of the International Congress of Inborn Errors of Metabolism. Dr. McCabe received his BS in biology with Honors from Johns Hopkins University and his M.D. and Ph.D. from the University of Southern California. He completed his Pediatric Residency at the University of Minnesota and his Pediatric Metabolism Fellowship at the University of Colorado Health Sciences. He is a Member of the Institute of Medicine and a Fellow of the American Association for the Advancement of Science.

\*\*\* Linda L. McCabe, Ph.D. is an Adjunct Associate Professor in the Departments of Human Genetics and Pediatrics at the David Geffen School of Medicine at UCLA. She teaches courses on ethical issues in human genetics and on the impact of genetics on our daily lives. She serves as the Managing Editor for the journal *Molecular Genetics and Metabolism*. Dr. McCabe received her BA from Towson State College and her Ph.D. from the University of Southern California.

predictions is increased by the complexity of systems biology and an incomplete understanding of genetic and environmental modifiers. Genotype does not predict phenotype even for rare “single” gene or “simple” Mendelian disorders. The challenges facing risk assessment in the medical context include lack of understanding of the complexity of common, complex disorders, the influence of genetic and environmental effects in pharmacogenetics, and the difficulty of whole genome approaches. The inability to give absolutes to our patients means that often we give misinformation to our patients. This article discusses how risk predictions in the genetics context are compounded by complexity and incompleteness in our understanding of systems biology.

### GENETICS AND RISK PREDICTION

Genotype<sup>1</sup> cannot necessarily predict phenotype<sup>2</sup> for rare “single” gene disorders<sup>3</sup> due to the complexity of the genome,<sup>4</sup> the proteome<sup>5</sup> interacting with the transcriptome,<sup>6</sup> and the dynamically coupled systems that are involved. Beyond the rare single gene disorders are common “complex” disorders, which occur when the primary mutated gene effect diminishes and two or more mutated genes have co-equal effects. Familiar examples of complex disorders include obesity, heart disease, diabetes, and cancer. Examining and predicting complex traits is even more complicated than single gene traits, and poses the question of how to deal with them in providing medical care.<sup>7</sup>

---

1. Genotype is the genetic makeup of an individual. SCIENCE AND TECHNOLOGY ENCYCLOPEDIA 231 (2000).

2. Phenotype is the physical characteristics of an individual resulting from heredity. SCIENCE AND TECHNOLOGY ENCYCLOPEDIA, *supra* note 1, at 400.

3. A single gene disorder is a genetic disorder caused by one gene. The Human Genome, <http://genome.wellcome.ac.uk> (last visited Mar. 28, 2007).

4. The genome is the entire complement of genetic material of an individual coded in sequence by the DNA that makes up the chromosomes. SCIENCE AND TECHNOLOGY ENCYCLOPEDIA, *supra* note 1, at 231.

5. The proteome is the entire complement of proteins in a given biological organism or system. TOM STRACHAN & ANDREW P. READ, HUMAN MOLECULAR GENETICS 3 553 (Garland Sci. 2004) (1996).

6. The transcriptome is the set of all messenger ribonucleic acid (mRNA) molecules in a cell or set of cells. STRACHAN & READ, *supra* note 5, at 545.

7. Common complex diseases involve multiple genes and gene and environment interactions, making diagnosis more difficult. Lori B. Andrews & Erin Shaughnessy Zuiker, *Ethical, Legal and Social Issues in Genetic Testing*

Despite their differences, however, common complex disorders are not distinct from their rare single gene counterparts; both are part of a continuum of increasing structural complexity. Therefore, the tools developed for “single” gene disorders will be adaptable to common complex diseases. Yet, some of the challenges with common complex traits stem from the fact that we do not understand the predictive significance of either the genetics or the environment at this point.

Consider risk prediction for cystic fibrosis (CF). CF is an autosomal recessive multisystem disease which can affect the lungs, liver, pancreas, small intestine, reproductive tract, and sweat glands of the skin.<sup>8</sup> CF is considered to be a single gene disorder caused by mutations in the cystic fibrosis conductance regulator (CFTR) chloride channel, which is encoded on chromosome seven.<sup>9</sup> Lung disease is the leading cause of death for patients with CF.<sup>10</sup> Thick mucus clogs the air passages of the lungs and these clogged passages lead to chronic respiratory infections and progressive destruction of the lungs.<sup>11</sup> Mucus can also obstruct the pancreatic ducts, which prevents digestive enzymes from entering the intestines, eventually leading to loss of insulin production and diabetes mellitus.<sup>12</sup> Research suggests a complexity of CFTR regulation based on post-translational modification, protein kinases A and C, and the availability of adenosine triphosphate (ATP).<sup>13</sup>

The most common CFTR mutation in Caucasians is  $\Delta F508$ .<sup>14</sup> Originally, data suggested that  $\Delta F508$  was associated

---

for *Complex Genetic Diseases*, 37 VAL. U. L. REV. 793, 795 (2003).

8. See Francis S. Collins, *Cystic Fibrosis: Molecular Biology and Therapeutic Implications*, 256 SCIENCE 774, 774 (1992).

9. See Collins, *supra* note 8, at 774.

10. See Am. Lung Ass'n, *Cystic Fibrosis Fact Sheet*, Nov. 2006, <http://www.lungusa.org/site/pp.asp?c=dvLUK900E&b=35042> (indicating that respiratory failure is the “primary cause” of death for 90% of adults with CF).

11. See *id.*

12. See Paulus Ripa et al., *The Relationship Between Insulin Secretion, the Insulin-Like Growth Factor Axis and Growth in Children with Cystic Fibrosis*, 56 CLINICAL ENDOCRINOLOGY 383 (2002).

13. For an extensive survey of research on the relationships between CFTR functioning and translational modification, protein kinase, and ATP levels, see generally Nat'l Ctr. for Biotech Info., Online Mendelian Inheritance in Man, *Cystic Fibrosis Transmembrane Conductance Regulator; CFTR*, [www.ncbi.nlm.nih.gov/entrez/dispmim.cgi?id=602421](http://www.ncbi.nlm.nih.gov/entrez/dispmim.cgi?id=602421) (last visited Mar. 28, 2007).

14. Anselm A. Zdebik et al., *Additional Disruption of the C1C-2 C1-Channel Does Not Exacerbate the Cystic Fibrosis Phenotype of Cystic Fibrosis*

with the most serious pulmonary symptoms, although this putative association has not been supported by further research based on information obtained from an increased numbers of patients.<sup>15</sup> There is a lack of genotype-phenotype correlation, and modifier genes determine the severity of  $\Delta F508$  homozygosity.<sup>16</sup> Drumm et al. studied the sequence variants in ten genes previously reported to be modifiers of CF in patients separated into groups with severe or mild lung disease.<sup>17</sup> Among the ten genes investigated they found significant associations only for transforming growth factor  $\beta 1$  (*TGF $\beta$ 1*).<sup>18</sup> *TGF $\beta$ 1* is involved in cell signaling and gene expression.<sup>19</sup> The functions of *TGF $\beta$ 1* include immune responses, proinflammatory and anti-inflammatory effects, growth and differentiation, and extracellular matrix production.<sup>20</sup> *TGF $\beta$ 1* is also associated with other forms of lung disease, including the progression of idiopathic pulmonary fibrosis, chronic obstructive pulmonary disease, and asthma.<sup>21</sup> Drumm et al. identified thirty-one additional markers, referred to as single nucleotide polymorphisms (SNPs), around *TGF $\beta$ 1*, and the association between disease severity and *TGF $\beta$ 1* SNP variants was confirmed.<sup>22</sup> Increased levels of the TGF $\beta$ 1 protein were associated with more severe lung disease.<sup>23</sup>

Research has also found that *TGF $\beta$ 1* reduces CFTR

---

*Transmembrane Conductance Regulator Mouse Models*, 279 J. BIOLOGICAL CHEMISTRY 22276 (2004).

15. See, e.g., Joseph Zabner et al., *CFTR  $\Delta F508$  Mutation Has Minimal Effect on the Gene Expression Profile of Differentiated Human Airway Epithelia*, 289 AM. J. PHYSIOLOGY - LUNG CELLULAR & MOLECULAR PHYSIOLOGY L545 (2005), available at <http://ajplung.physiology.org/cgi/reprint/289/4/L545>.

16. See, e.g., Mitchell L. Drumm et al., *Genetic Modifiers of Lung Disease in Cystic Fibrosis*, 353 NEW ENG. J. MED. 1443, 1449-53 (2005).

17. Drumm et al. split a population of over eight hundred Delta F508 positive patients into two groups based upon the severity of their lung disease symptoms. *Id.* at 1444. The first group belonged in the top quartile of severity, while the second group represented the bottom quartile of severity. *Id.* Drumm et al. then looked at variance of ten genes that allegedly influenced the severity of CF and compared those results between these two groups with the severe and mild lung disease. *Id.* at 1444-45.

18. *Id.* at 1446-48.

19. See *id.* at 1451.

20. *Id.* at 1451.

21. *Id.*

22. *Id.* at 1448-49.

23. See *id.* at 1449-51.

expression in nasal polyps in patients who do not have cystic fibrosis.<sup>24</sup> In 2005, a group from Spain headed by Virginie Prulière-Escabasse furthered earlier work with *TGFβ1*, looking at the connection between CF and nasal polyposis. Nasal polyposis is a chronic inflammatory disease of the nasal mucosa causing polyps in the nose and sinuses.<sup>25</sup> It can be primary, meaning we do not know what it is associated with, but it also has 50% prevalence with CF.<sup>26</sup> The hypothesis of Prulière-Escabasse et al. was that altered CFTR expression and function was involved in the pathogenesis of nasal polyposis.<sup>27</sup> They looked only at CFTR expression in non-CF patients and compared the nasal polyp to the control mucosal epithelial cells, finding that CFTR was expressed in the controlled nasal epithelial cells but not in the nasal polyps.<sup>28</sup> Thus, *TGFβ1* down-regulated CFTR expression and function measured by chloride currents in the control but not the nasal polyps.<sup>29</sup> From this finding, they concluded that this was consistent with studies in CF patients that showed high *TGFβ1* secretors had more severe lung disease.<sup>30</sup> These results are consistent with another study that found that patients with CF who secreted high levels of *TGFβ1* had more severe lung disease.<sup>31</sup> Based on these similar findings, there may be a story developing for *TGFβ1* as a genetic modifier. The research has found that *TGFβ1* functions in immune responses,<sup>32</sup> it has both pro-inflammatory and anti-inflammatory effects,<sup>33</sup> it is involved

---

24. Virginie Prulière-Escabasse et al., *TGF-β1 Downregulates CFTR Expression and Function in Nasal Polyps of Non-CF Patients*, 288 AM. J. OF PHYSIOLOGY - LUNG CELLULAR & MOLECULAR PHYSIOLOGY L77 (2005), available at <http://ajplung.physiology.org/cgi/reprint/288/1/L77> (comparing CFTR expression in nasal polyps to normal mucosa in the same individual and finding that CFTR is expressed in normal mucosa, but not in the nasal polyps).

25. *Id.* at L77.

26. *Id.*

27. *Id.*

28. *Id.* at L78-L79, L81.

29. *Id.*

30. *Id.* at L82.

31. Peter D. Arkwright et al., *TGF-β1 Genotype and Accelerated Decline in Lung Function of Patients with Cystic Fibrosis*, 55 THORAX 459 (2000), available at <http://thorax.bmj.com/cgi/reprint/55/6/459>.

32. See, e.g., S. Kakumu et al., *Effect of Recombinant Human Transforming Growth Factor β1 on Immune Responses in Patients with Chronic Hepatitis B*, 13 LIVER 62 (1993).

33. See Sharon M. Wahl, *Transforming Growth Factor Beta (TGF-β) in Inflammation: A Cause and a Cure*, 12 J. CLINICAL IMMUNOLOGY 61 (1992).

in growth and differentiation,<sup>34</sup> and it is involved in extra-cellular matrix production.<sup>35</sup> In lung disease it has been shown to be associated with progression of idiopathic pulmonary fibrosis,<sup>36</sup> increased activity, chronic obstructive pulmonary disease,<sup>37</sup> and asthma.<sup>38</sup> We have more work to do, however, to confirm that there is a real biological relationship.

#### THE COMPLEXITY OF SYSTEMS BIOLOGY

Risk predictions are compounded by the complexity of systems biology.<sup>39</sup> Systems biologists would argue that these dynamic systems can be modeled and the models are amenable to mathematical analyses, moving biology from a descriptive science to a quantitative discipline. In the twentieth century, physics developed the fundamental scientific basis for rigorous applications in engineering by using mathematics. In the twenty-first century, biology will provide the foundation for applications in medicine, and these applications will increasingly rely on the mathematical algorithms of systems biology.

Medicine has begun to rely on mathematical algorithms in ways that are transparent to both physician and patient. For example, in the field of imaging, X-rays are no longer analogue, which involves direct film exposure, but are now digital, and the images, including sharpness and contrast, are based on calculations from the raw digital data.<sup>40</sup> Mathematical

---

34. See Mary E. Dickinson et al., *Chromosomal Localization of Seven Members of the Murine TGF- $\beta$  Superfamily Suggests Close Linkage to Several Morphogenetic Mutant Loci*, 6 GENOMICS 505 (1990).

35. See Chie-Pein Chen et al., *Hypoxia and Transforming Growth Factor- $\beta$ 1 Act Independently to Increase Extracellular Matrix Production by Placental Fibroblasts*, 90 J. CLINICAL ENDOCRINOLOGY & METABOLISM 1083 (2005).

36. See Moira K.B. Whyte, *Genetic Factors in Idiopathic Pulmonary Fibrosis: Transforming Growth Factor- $\beta$  Implicated at Last*, 168 AM. J. OF RESPIRATORY & CRITICAL CARE MED. 410 (2003).

37. See L. Wu et al., *Transforming Growth Factor- $\beta$ 1 Genotype and Susceptibility to Chronic Obstructive Pulmonary Disease*, 59 THORAX 126 (2004).

38. See Judith C.W. Mak et al., *Analysis of TGF- $\beta$ 1 Gene Polymorphisms in Hong Kong Chinese Patients with Asthma*, 117 J. ALLERGY CLINICAL IMMUNOLOGY 92 (2006).

39. Systems biology is the study of the relationships and interactions between and among parts of biological systems. Ekat Kritikou, Forward, *All Systems Go!*, 7 NATURES REVS. MOLECULAR CELL BIOLOGY 801, 801 (2006).

40. H.K. Huang et al., *Digital Radiology at the University of California*,

algorithms are also crucial for ultrasonography, computerized tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET). These algorithms are invisible to the physician, and a physician need not be a mathematician to examine a digital image.

Systems dynamics are based on network architecture.<sup>41</sup> We used to think that biology was composed of random networks, in which every node was connected to a similar number of nodes.<sup>42</sup> However, biology is made up of scale-free networks in which some nodes are more highly connected, while others may have only one connection, a structure described as “hub and spoke.” If a node that is a hub does not function because of a mutation, then this takes out a significant portion of the network. If a node with only one connection does not function, however, the network is not affected as much, or perhaps not at all. The internet is an example of a scale-free network comprised of interconnected nodes. A number of these nodes are highly connected; these are the critical servers. Other nodes have just one connection, such as a desktop computer or laptop device. If one computer stops working, it will not affect the network. If a server node at a major university goes down, all of the associated computers can no longer access the other nodes, and are rendered useless for networking purposes at that time.

Our knowledge of modifiers is incomplete. If we look at cystic fibrosis we know about genetic modifiers, such as *TGFβ1*, and environmental modifiers, such as colonization with *Pseudomonas* bacteria. These individual modifiers will influence risk in a probabilistic manner. However, risk influences will vary depending upon the genetic make-up of the individual. Therefore, to better understand the probabilistic risks associated with a particular individual, we will need whole genome information on large populations.

#### PHARMACOGENETICS: AT THE INTERFACE OF GENETICS AND ENVIRONMENT

Pharmacogenomics, the interface of genetics and

---

*Los Angeles: A Feasibility Study*, 16 J. DIGITAL IMAGING (2003).

41. See Albert-László Barabási & Zoltán N. Oltvai, *Network Biology: Understanding the Cell's Functional Organization*, 5 NATURE REVIEWS GENETICS 101 (2004) (discussing a parallel between cellular function and architectural features of other complex systems).

42. *Id.* at 104.

environment, is one of the first places where genomic medicine is beginning to impact routine medical care.<sup>43</sup> This is the case for molecular microbiology in particular, dealing with viruses and bacteria for rapid diagnosis. The incidence of adverse drug reactions (ADRs) in U.S. hospitals is 6.7%, and the incidence of fatal ADRs is 0.3%.<sup>44</sup> Many drug side effects are not stochastic events, because individuals are genetically predisposed to experience an ADR.<sup>45</sup> The physician's goal is to use pharmacogenetics to predict and prevent ADRs and to optimize an individual's therapy.

Genomic medicine will be predictive, preventive, and personalized. Testing devices will be required to acquire genomic information in order to utilize pharmacogenomics. The first microarray-based<sup>46</sup> testing device that was approved by the Food and Drug Administration (FDA) on January 12, 2005 was the Roche AmpliChip CYP450.<sup>47</sup> The AmpliChip CYP450 tests for genetic variation in the cytochrome P450 enzymes, CYP2D6 and CYP2C19, which metabolize drugs.<sup>48</sup> Drugs which impact these enzymes include the following classes: anti-depressants, anti-psychotics, anticonvulsants, cardiovascular medications, cancer chemotherapeutics, and anti-malarial drugs.<sup>49</sup> Physicians typically give the same dose to all adults, assuming identical drug metabolism in all adults.

The AmpliChip CYP450 identifies different rates of metabolism and classifies people as poor, intermediate, normal,

---

43. Pharmacogenomics is the study of how genes modify drug metabolism. STRACHAN & READ, *supra* note 6, at 610.

44. Jason Lazarou, et al., *Incidence of Adverse Drug Reactions in Hospitalized Patients*, 279 J. AM. MED. ASS'N 1200 (1998).

45. Taisei Mushiroda et al., *A Model of Prediction System for Adverse Cardiovascular Reactions by Calcineurin Inhibitors Among Patients with Renal Transplants Using Gene-Based Single-Nucleotide Polymorphisms*, 50 J. HUM. GENETICS 442 (2005).

46. A DNA microarray is a two-dimensional array of genes or gene fragments set on a solid surface, such as glass, plastic or silicon chip for use in monitoring gene expression levels. STRACHAN & READ, *supra* note 6, at 175.

47. Press Release, F. Hoffman-La Roche Ltd., Roche's AmpliChip CYP450 Test Receives FDA Clearance-Microarray-Based Diagnostic Test. (Jan. 12, 2005), *in Medical News Today*, available at [medicalnewstoday.com/medicalnews.php?newsid=18822](http://medicalnewstoday.com/medicalnews.php?newsid=18822).

48. Li Li et al., *New Cytochrome P4250 2D6\*56 Allele Identified by Genotype/Phenotype Analysis of Cryopreserved Human Hepatocytes*, 34 DRUG METABOLISM & DISPOSITION 1411 (2006).

49. Press Release, *supra* note 47.

and ultra-rapid metabolizers.<sup>50</sup> Poor metabolizers suffer the effects of toxic drug levels from prolonged and/or excessive drug levels. Ultra-rapid metabolizers may not achieve any benefit from a drug because they break down the drug before it has time to work. The opposite phenomenon occurs for pro-drugs that are metabolized to the therapeutically active compound.

#### WHOLE GENOME APPROACHES

The same kind of technology that is used for the Cytochrome P450, the AmpliChip, can also be used for whole-genome approaches. Whole genome applications currently use SNP microarrays or SNP chips.<sup>51</sup> SNPs (for example, where a base A is changed to a C), usually occur in a non-coding region that does not make a difference.<sup>52</sup> It may be in a site that does not change to an amino acid. An SNP is not a mutation, but a genetic marker. SNPs are distributed throughout the genome, and SNP chips were originally designed to identify SNP variation associated with different phenotypes. SNP chips can also be used to look at the presence, absence or amplification of a genomic region. SNP chips combine molecular cytogenetics with the identification of modifiers so that SNPs influencing the location of a mutation in the DNA of an individual or a deletion in the DNA of an individual can be examined.

We have been using SNP chips to screen for individuals with the chromosome 22q11.2 deletion syndrome.<sup>53</sup> This syndrome may include congenital cardiovascular defects, craniofacial abnormalities, and hypoplasia of the thymus and parathyroid glands.<sup>54</sup> This is the most frequent interstitial deletion in humans – one in 4,000 live births – and 90% of affected individuals have a three megabase region deleted.<sup>55</sup> This disorder shows a lack of genotype/phenotype correlation, meaning that patients with identical deletions have different

---

50. *Id.*

51. Itsik Pe'er et al., *Evaluating and Improving Power in Whole-Genome Association Studies Using Fixed Marker Sets*, 38 *NATURE GENETICS* 663 (2006).

52. *Id.*

53. C.M. Stanczak et al., *Single Chromosomal Copy Deletion in DiGeorge Syndrome Using a SNP Mapping Array*, *HUMAN MUTATION* (forthcoming 2007).

54. *Id.*

55. *Id.*

phenotypes, and patients with different deletions have identical phenotypes. We have successfully used SNP chips to identify patient deletions. Combining information about the deletion with whole-genome SNP variation will enable us to improve prediction of phenotype. However, risk predictions will remain probabilistic and not absolute.

With a whole-genome approach, as an example, we can identify patients' deletions. This can be combined with the knowledge of these polymorphisms, the SNPs in the deleted region and genome-wide. We can utilize whole-genome information to improve prediction of the phenotype, that is, all other SNPs. Despite all of this, risk predictions are still going to be probabilistic and not absolute. It is just the way biology is.

#### UNDERSTANDING AND UTILIZING COMPLEXITY

In order to understand and utilize complexity we need systems biology. Systems biology is an attempt to quantify the complexity of dynamically coupled robust networks. At this time, however, we have limited information and methods; much remains unknown. Systems biology will involve an incredible volume of data and corresponding mathematical algorithms for analysis. Despite the mathematical reduction of data, the output will be voluminous. We need to find a way to represent the output and the risks by genomic position or network diagrams that physicians can easily interpret and patients can easily understand.

#### CONCLUSION

Genomic medicine represents a future flooded with risk information. We already recognize that we do not know all of the information needed to counsel our patients in the genetics clinic because this risk information is inherently probabilistic and not absolute. We must develop comprehensive representations of this risk information in order to better transfer that information to the general pediatrician and the general practitioner for both the care of their patients and for informing their patients. In pediatrics, parents tell us they do not want a physician to dictate the care of their child, but want to be partners in the care of their child. Parents want to understand what this risk information means. Developing comprehensive representations of risk information will require

2007] *FUTURE FLOODED WITH RISK INFORMATION* 439

multi-disciplinary teams of geneticists, mathematicians, engineers, and others to work together.