Inherited Genetic Markers and Cancer Outcomes: Personalized Medicine in the Postgenome Era

Timothy R. Rebbeck, Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, and Abramson Cancer Center, University of Pennsylvania School of Medicine, Philadelphia, PA

The report of Tsuchiya et al\(^1\) represents an area of research that has the potential to impact personalized genomic medicine; this area is the use of inherited genetic markers to predict clinical outcomes, including prognosis and recurrence, in cancer patients. Tsuchiya et al\(^1\) report that longer repeat polymorphisms in the *IGF1* and/or *CYP19* genes are significantly associated with poorer prostate cancer-specific survival among men with metastatic prostate cancer. This research complements other reports that inherited genetic markers predict prostate cancer outcomes, including associations with *CYP3A4*,\(^2\) *SRD5A2*,\(^3\) the androgen receptor gene (*AR*),\(^4\)\(^-\)\(^7\) and the vitamin D receptor gene (*VDR*).\(^8\) Numerous reports have also been made that relate inherited genetic markers and prognosis or recurrence with tumors at other sites. Prediction of clinical outcome by inherited genetic markers is likely to reflect the natural history of disease progression and, therefore, may involve genes and biochemical pathways that act along the continuum from disease causation, recurrence, and mortality (Fig 1). Indeed, *IGF1*, *CYP19*, *CYP3A4*, *SRD5A2*, *AR*, and *VDR* have all been implicated in prostate cancer etiology and outcomes. Inherited genetic markers may also provide predictive or prognostic information in addition to somatic tumor markers or inherited genetic markers of treatment response and toxicity (eg, pharmacogenetic markers; Fig 1).

Despite reports associating inherited genetic markers with clinical outcomes, the existing literature is probably insufficient to warrant clinical translation of these markers at this time. What evidence must be gathered before inherited genetic markers could be applied in a clinical setting? Establishing a causal relationship between risk factor and outcome is a minimal requirement, and a variety of criteria for establishing causal inferences has been proposed.\(^9\)\(^-\)\(^1\)\(^1\) As outlined in the following sections, additional factors must be considered before an inherited genetic marker can be applied to predict cancer prognosis, recurrence, or other clinical outcomes.

**What Should We Know Before We Can Apply an Inherited Marker Clinically?**

Is There A Priori Support for the Association?

If an association is reported, external data should be available to further support the observed association. For example, information about the functional significance of a marker should be available that supports its association with clinical outcome. These data include information about the putative function and hypothesized direction of an association from in vitro or in vivo experimental data or from the use of computational tools that evaluate evolutionary conservation or predicted structural effects of genetic variants.\(^1\)\(^2\) An important consideration regarding the putative function of an inherited marker is whether its function has been adequately validated in humans. For example, some in vitro or animal model data may not accurately reflect the biologic environment in which a marker may be acting in humans. Finally, associations should be based on well-constructed a priori hypotheses to avoid opportunistic or spurious relationships found only in data subgroups that were not part of the original study aims.

In What Context Can Markers Be Applied?

Interindividual and tumor heterogeneity is likely to exist that may affect the utility of an inherited marker to predict clinical outcomes. Similarly, inherited markers are unlikely to act alone in prediction of disease outcome, but instead, they may interact with other markers (Fig 1), exposures, behavior, demographic characteristics, or other factors. Thus, prediction of outcomes using single markers may be inadequate to fully characterize the predictive outcomes, and more complex predictive models may be required. For example, Powell et al\(^2\) reported that *CYP3A4* genotypes predicted clinical outcomes in European American prostate cancer patients but not in African American patients. Shimbo et al\(^7\) suggested *AR*-CAG genotypes predict prognosis but optimally in combination with knowledge of serum testosterone levels. Thus, population and disease heterogeneity may need to be considered before clinical application of inherited markers.
**What Is the Relationship Between Marker and Outcome?**

Consideration should be given to how large the effect of an inherited marker should be before it can be used to make clinical decisions. In general, it is likely that a common disease, common variant model is present, in which inherited markers associated with outcome and prognosis are common in the population and predict relatively common events. Markers occurring at a rare frequency that confer large effects are less likely to be observed, although some may exist. For example, in the pharmacogenetics literature, TPMT genotype can be used to predict major toxicities after thiopurine drug treatment, but the homozygous mutation that confers these large risks is rare.\(^{14}\) If the common disease, common variant hypothesis is present, then studies linking inherited markers with outcomes require sufficiently large sample sizes to detect small- to moderate-sized effects of relatively commonly occurring variants. Therefore, it may be desirable to identify relative risk increases conferred by a marker of 50% or less. How low might the relative risk be before it is deemed “uninteresting” for clinical translation? If a particular genetic marker confers a 10% increased risk of poor prognosis, it is unlikely that this information alone would be used to make a clinical decision. Therefore, the most likely variants to consider are those that confer a relative risk sufficiently high to warrant intervention, and these variants may be rare or common. However, it is likely that interactions of multiple genotypes, acting on the background of exposures, treatment, or other factors, may need to be considered to achieve adequate prediction.

**Is the Association Consistent Across Studies?**

A basic requirement of an association that could result in a clinically relevant intervention should be the reproducibility of the association across multiple studies and consistency lines of investigation. Because the literature in this field remains limited, most associations of genotypes with outcomes have not been reproduced. However, a consistent pattern of prediction has been observed with AR-CAG repeats and prostate cancer prognosis.\(^{4-7}\) As in the context of disease etiology studies, most associations may not be reproducible,\(^ {15}\) but only those that can be reliably reproduced will be candidates for clinical translation.

**What Factors Affect Reproducibility?**

Laboratory and genotype quality could affect reproducibility of studies. Nondifferential genotype misclassification among study participants can result in bias toward the null hypothesis, but differential genotype misclassification can confer unpredictable biases. Appropriate laboratory methods with adequate quality control, including open availability of assay conditions and freely shared control samples, should be used to maximize the study comparability and reproducibility.

Appropriate study design and data analysis are critical to achieving reproducibility among studies. The epidemiologic design of these studies must be of the highest integrity, both to ensure minimal bias and to maximize the potential that results will be generalizable to the appropriate patient population. Adequate statistical power to avoid false-negative reports is critical. Large sample sizes may be required to detect associations with small magnitude effects, to detect these effects in patient subsets (eg, by tumor stage, race, sex, or other relevant clinical characteristics), and to identify relevant interactions among genes and between genes and exposures that have been generated as the result of biologically plausible a priori hypotheses. Similarly, appropriate statistical analysis will maximize the chances that relevant and sufficiently complex relationships of inherited markers, tumor markers, treatment, and other exposures are revealed. If many hypotheses are tested, corrections for multiple comparisons may be required, with appropriately stringent and justified levels of statistical significance.

Population structure (ie, population stratification)\(^{16}\) has been suggested as a potential source of bias in association studies of disease etiology, and similar arguments may be raised for studies of disease outcome. Population stratification is a form of bias that may result in incorrect estimates of association when both of the following conditions are fulfilled: the genotype of interest varies by race and the outcome of interest varies by race. Because the frequency of many genotypes varies across race and disease outcomes also vary by race, population stratification may lead to either false-positive or false-negative associations. To date, there is little evidence in the disease etiology or outcome literature that substantial biases are conferred by population stratification within populations such as European Americans\(^ {16}\) or African Americans.\(^ {17}\) Despite the lack of theoretical or empirical observations of serious biases caused by population stratification, race and ethnicity cannot be ignored. Aside from the potential for bias, race and ethnicity may be important considerations in the appropriate application of genetic data in clinical practice.

The rapidly expanding knowledge base that has accompanied the characterization of the human genome has facilitated our understanding of genes that predict cancer outcomes. It is likely that the majority of these inherited markers will be predictive in concert with environmental exposures, behavioral factors, social environments, and treatment. In addition, it is likely that these genetic variants will be predictive of outcome in concert with somatic tumor markers and pharmacogenetic markers (Fig 1).

Once sufficient data are accumulated that indicate an inherited marker predicts outcome, when is it reasonable to consider translation of that marker into a clinical setting? For some markers, it may be sufficient for epidemiologic data alone to direct clinical translation. For example, the strong epidemiologic relationship of cigarette smoking and lung cancer allows us to advise people not to smoke, even though we do not fully understand the mechanism of cigarette-induced lung carcinogenesis. The reason smoking cessation can be clinically applied is because of the large and consistent magnitude of the epidemiologic relationship between smoking and lung cancer. An understanding of the disease mechanism remains important, and continued fundamental research to obtain better insights into the disease and its treatment or prevention is still warranted. However, it may be possible to apply interventions even with limited knowledge of the biologic mechanism if the epidemiologic relationship is sufficiently robust.

However, this paradigm must be approached with great caution when considering markers that may confer lower levels of risk or that are acting in a complex multifactorial environment. The lessons of the Alpha-Tocopherol Beta-Carotene Cancer Prevention Study\(^ {18}\) and Beta-Carotene and Retinol Efficacy Trial\(^ {19}\) provide cautionary examples. Beta-carotene was widely believed to act as a chemopreventive agent, but both the Alpha-Tocopherol Beta-Carotene Cancer Prevention Study and the Beta-Carotene and...
Retinol Efficacy Trial reported increased lung cancer incidence in the treated groups in their studies.\textsuperscript{18,19} Considering the large amount of research leading to these trials, this unexpected finding suggests that detailed information about the complex biologic interactions of genetics, exposures, treatments, disease subtypes, and other factors may also be required before inherited genetic markers may be appropriately applied to predict clinical outcome.

The advent of personalized medicine using inherited markers could drastically change the practice of medicine and improve clinical outcomes. As this research matures, consideration must be given to the appropriate application of this information in the design of treatment trials, the economic impact and benefit of marker utilization, and the potential ethical, legal, and social consequences of the clinical application of inherited genetic markers. Initially, a major need is to promote appropriate molecular epidemiologic research that identifies biologically plausible, replicable marker-disease associations that have clinical utility in predicting treatment type, prognosis, and outcomes. These markers will only be identified through the appropriate multidisciplinary integration of knowledge from population, clinical, and fundamental sciences.

© 2006 by American Society of Clinical Oncology

REFERENCES

Acknowledgment
This work was supported by Grant No. P50-CA105641 from the Public Health Service and the University of Pennsylvania Abramson Cancer Center. We thank Dr Susan Domchek for her comments on the concepts presented in this editorial.

Author’s Disclosures of Potential Conflicts of Interest
The author indicated no potential conflicts of interest.

Author Contributions
Conception and design: Timothy R. Rebbeck
Financial support: Timothy R. Rebbeck
Administrative support: Timothy R. Rebbeck
Manuscript writing: Timothy R. Rebbeck
Final approval of manuscript: Timothy R. Rebbeck