Commentary

The Value of Small Observations in the Era of Big Science

Stephen B. Gruber

Division of Molecular Medicine and Genetics, Departments of Internal Medicine, Epidemiology, and Human Genetics, University of Michigan, Ann Arbor, Michigan

We are now unequivocally in the era of big science. Genotyping costs have become so trivial and genetic variation is so important that it is almost heresy to consider launching an epidemiologic investigation without collecting germ line DNA to evaluate the contributions of genetic diversity to the pathogenesis of disease. As genetics steps into the spotlight of epidemiology, studies are bursting in size (and cost) to evaluate interactions of genes and the environment and to precisely measure the magnitude of small associations. Technology has given us epidemiology on steroids. This big science approach has evolved as a consistent research strategy even for cancers where the environmental contributions are unequivocal and dominant. Smoking causes lung cancer, human papillomavirus causes cervical cancer, and excessive UV exposure causes melanoma, but, of course, the story is a bit more complicated. There is no question that we need to learn more about the causal pathways exploited by these carcinogens and how they can be modified even when epidemiology has successfully elucidated the main causes of disease. The real question is, is there still room for small science?

An example of the kind of insight that can be gained from small science is published in a recent issue of Cancer Epidemiology Biomarkers and Prevention. The study of MC1R variants and the risk of melanoma by Goldstein et al. (1) examined 16 families with melanoma that carry mutations of CDKN2A. In this study, we learn that the presence of multiple MC1R variants is associated with the risk of melanoma and that these variants contribute to the clinical presentation of patients with multiple primary melanoma. Some might argue that a study of 395 individuals is not small, but perhaps what best captures the idea of small science is the reliance on an extremely well-characterized phenotype and thoughtful comparisons that yield insight. In the Goldstein study, clinical and epidemiologic data collected over 26 years provide a unique resource to measure the genetic and environmental contributions to multiple primary melanoma, specifically in melanoma-prone families who carry a mutation in CDKN2A. This study is a wafer-thin slice of a very big pie.

Yet, it is precisely the care in choosing the thin slice that can provide generalizable insight. Numerous studies have shown a relationship between MC1R and the risk of melanoma (2-4) and two studies have shown that the presence of an MC1R variant increases the penetrance of CDKN2A mutation carriers (5, 6). Other studies have shown that carrying multiple variants is associated with more risk than a single variant (7). Here, as in prior studies, the relative risk for MC1R was stronger for those variants designated as red hair color compared with non–red hair color variants, and the risk seemed cumulative with additional variants. In addition, by focusing on a comparison of multiple primary versus single primary melanoma, the authors showed that MC1R is a risk factor for multiple primary melanoma and that cumulative mutations are associated with decreasing age at diagnosis, beyond the effects of a major gene like CDKN2A. This approach is another good example of how studies that concentrate on comparisons of multiple primary and single primary cancers can provide an informed perspective on risk factors from a different vantage point (8).

Phenotype Is the Key

Phenotype is the key to small science, and perhaps the best example is Knudson’s classic study of retinoblastoma (9). By recognizing the unique phenotype of children with young onset, bilateral, familial retinoblastoma, Knudson’s work led to the “two-hit” hypothesis and the eventual discovery of the first tumor suppressor gene. It is worthwhile to remind ourselves that the sample size of Knudson’s study was only 48 children (9). Similarly, the success in identifying DES as the cause of clear cell adenocarcinoma of the vagina relied on focusing on an extremely rare phenotype. The original case-control study by Herbst et al. (10) included a grand total of 8 cases and 32 controls.

Epidemiology will continue to rely on carefully defined phenotypes and this will require more than our reflexive reliance on “pathologically confirmed diagnoses” as the hallmark of well-designed cancer studies to minimize the possibility of misclassification of disease. Technology is once again providing us with the muscle to define phenotypes more clearly, and it is likely that molecular classification of tumors will play an even more prominent role in the future of cancer epidemiology.

Interactions

Goldstein et al.’s article also serves as an example of how one can intelligently examine phenotype to gain perspective on the contributions of multiple genes. Indeed, the term “gene-gene interaction” is not used consistently in genetic research. Does this only refer to biochemical evidence of protein binding or direct modulation of a downstream effect by multiple DNA variants? Or could it refer to multiple genes that contribute to the risk of a disease with evidence of departure from pure additive or multiplicative effects? The evidence from studies of melanoma, including this one, suggests that classic epidemiologic definitions for “interactions” are still appropriate even when it seems clear that the regulatory pathways of MC1R and CDKN2A do not directly overlap. We are likely to gain additional insight with further inquiries into higher-level interactions, especially for systems like melanoma, where sun exposure is clearly an independent risk factor but still needs to be assessed within the context of individuals who carry mutations of CDKN2A and variants of MC1R.

Experimental data need to complement evidence of epidemiologic interactions to elaborate potential mechanisms of
disease and we might be surprised by the results. For example, the development of a genetically engineered human tissue assay that recapitulates melanoma allows individual mutations in candidate genes to be tested alone or in combination for their oncogenic potential (11). One surprising result from this approach is that the most common somatic mutation of
\textit{BRAF} identified in melanoma turns out to have little oncogenic potential on its own, despite its critical position in the RAS-RAF-MEK-ERK-MAP kinase pathway. Coexpression of mutant forms of genes now permits epidemiologic hypotheses to be tested \textit{in vivo} and has yet to be fully harvested as a strategy in melanoma or other tumor systems. Applying technologies like these to integrate the evidence from experimental and observational paradigms will be particularly helpful in elucidating the meaning of interactions especially when biochemistry fails to enlighten.

Genetically engineered human tissue model systems and massive epidemiologic investigations dedicated to whole-genome association studies certainly qualify as big science, but small studies that ask smart questions about a carefully defined phenotype can still make an impact. Thankfully, there is still room for small science.

References
