Contents

From the editor

E•nig•ma *n*
1: an obscure speech or writing 2: something hard to understand or explain 3: an inscrutable or mysterious person Etymology: *Latin aenigma, from Greek* ainigmat-, ainigma, *from* ainissesthai *to speak in riddles, from* ainos *fable, Date: 1539*
Adapted from the Merriam-Webster Online Dictionary[1]

For many oncologists, a biostatistician's work can be easily characterized as an enigma. We know that all clinical trial or laboratory data must be framed in statistical language and have deep respect for someone who knows which test to apply when. This issue of *CCR Focus* is aimed at bridging the gap between biostatistician and oncologist in our world of targeted therapy and genomic analyses.

Over the past decade, the biostatistician's work in oncology has steadily expanded as we ask not only which clinical trial results are valid, but also what sets of genes among thousands evaluated are relevant to the clinical question. As one example, among the increasing numbers of gene signatures evaluating breast cancer prognosis derived from cDNA microarray work, there appears to be some prognostic concordance but with little overlap in gene identity (Haibe-Kains B, et al. BMC Genomics 9:394, 2008). With an Affymetrix chip containing over 20,000 genes, at $P = 0.05$, 1000 genes will reach statistical significance by random chance for high expression in breast cancer samples if no additional filters are applied. This simple concept may explain the lack of overlap. How, then, can investigators determine which signature, or which biomarker has been appropriately validated, and how can that signature or biomarker be used in translational clinical trials?

The articles in this issue of *CCR Focus*, contributed by experts in the field, address several facets of this problem, beginning with an overview by Stephen L. George as Guest Editor that highlights the contributions that follow. Chau, MacLeod, and Figg discuss validation of analytic methods for biomarkers used in drug development. Owzar et al. evaluate the necessary preprocessing of microarray data - those early steps in microarray analysis that have major impact on final outcome. Taylor, Ankerst, and Andridge consider validation of prediction models using genomic analyses. Finally, both George in the *Focus* overview and Simon explore methods of clinical trial design that incorporate biomarker detection. As with every issue of *CCR Focus*, our goal is that these articles will inform the clinical cancer researcher who is interested but not expert in the field, but also stimulate the thinking of the expert. In the case of this *CCR Focus* on Biostatistics and Biomarkers, it is hoped that these articles for translational cancer researchers can begin to unravel the enigma.

Susan E. Bates, M.D.
Senior Editor
CCR Focus

[1] http://www.merriam-webster.com/dictionary/enigma