means that 55% of all subjects who will not develop the disease will be classified falsely. In a population in which 95% of the individuals will not develop the disease, as in the study of Yang et al., this means that 52% will undergo unnecessary preventive treatment. When a sensitivity of 0.90 is chosen, the percentage of all subjects who are unnecessarily selected is 73%. In comparison, the sensitivity and specificity of mammography in a large population–based breast cancer screening program were 0.75 and 0.92, respectively (Carney et al. 2003). Thus, the multiplex genetic tests of Yang et al. are by no means efficient screening strategies.

In conclusion, the clinical usefulness of genetic testing should be evaluated by ROC analysis. Using this approach for the data of Yang et al., we found that the discriminative ability of the multiplex genetic test increased by the addition of more genes but that its performance for use as a screening instrument was rather inefficient. It remains to be investigated whether these results are representative of the prediction of common disease by multiplex genetic tests that include genetic factors with low mutation prevalence and low relative risks. In that case, alternative statistical strategies are needed to increase the potential clinical application of selective genetic testing.

Acknowledgments

The study was financially supported by the Netherlands Organization for Scientific Research (NWO Pioneer and ZonMW; grant number 945-10-039) and the Center for Medical Systems Biology (CMSB).

A. Cecile J. W. Janssens,[1] M. Carolina Pardo,[2] Ewout W. Steyerberg,[1] and Cornelia M. van Duijn[2]

[1]*Department of Public Health and* [2]*Department of Epidemiology and Biostatistics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands*

References

Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, Geller BM, Abraham LA, Taplin SH, Dignan M, Cutter G, Ballard-Barbash R (2003) Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med 138:168–175

Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29–36

Holtzman NA, Marteau TM (2000) Will genetics revolutionize medicine? N Engl J Med 343:141–144

Thompson ML, Zucchini W (1989) On the statistical analysis of ROC curves. Stat Med 8:1277–1290

Vineis P, Schulte P, McMichael AJ (2001) Misconceptions about the use of genetic tests in populations. Lancet 357: 709–712

Yang Q, Khoury MJ, Botto L, Friedman JM, Flanders WD (2003) Improving the prediction of complex diseases by testing for multiple disease-susceptibility genes. Am J Hum Genet 72:636–649

Address for correspondence and reprints: Dr. Cecile Janssens, Center for Clinical Decision Sciences, Department of Public Health, Erasmus MC, P.O. Box 1738, 3000 DR The Netherlands. E-mail: a.janssens@erasmusmc.nl

0002-9297/2004/7403-0029$15.00

Am. J. Hum. Genet. 74:588–589, 2004

Revisiting the Clinical Validity of Multiplex Genetic Testing in Complex Diseases: Reply to Janssens et al.

To the Editor:

We appreciate the comments by Janssens and her associates (2004 [in this issue]) regarding our study on the use of likelihood ratios to improve the prediction of complex diseases by testing for multiple-susceptibility genes (Yang et al. 2003). As Janssens et al. correctly point out, our study considers only the predicted probability of disease for subjects who have all positive testing results, and this is likely to be an infrequent occurrence. We think that the suggestion made by Janssens et al. to use receiver-operating–characteristic (ROC) curves to assess multiple genetic testing is very useful. The ROC curves provide a valuable way of evaluating the accuracy and discriminatory ability of diagnostic tests (Hanley 1989). Janssens et al. use the ROC curves to assess the classification of patients into a disease group, but multiplex genetic testing is likely also to be of value in identifying people who are at lower-than-average risk for developing a particular disease. This might allow them to put off receiving a more expensive intervention for some time—for example, to defer mammography for breast cancer detection for 10 years (Fletcher 1997) or to avoid screening for prostate cancer until $\geqslant$60 years of age (Harris and Lohr 2002).

The predictive value of combining tests obviously does depend on the relative risk associated with each component test, with a bigger effect resulting from tests that make larger independent contributions. Janssens et al. suggest that an odds ratio of 1.5–1.7 for each test is more likely than an odds ratio of 3.5. This might be true, but we do not yet know what the relative frequency of genes of larger or smaller effect will turn out to be for any common multifactorial disease. We used five genetic tests and an environmental factor as a simplified illustration in our analysis, but, in the near future, 50

or 100 genetic tests might be available for many common diseases. If there are numerous predisposing alleles and each has an independent odds ratio of only 1.5–1.7, the overall effect would still be substantial. We simulated models of 10, 15, and 20 genes with a risk of 1.5–1.7 each and found the areas under the ROC curves (AUCs) to be 0.70, 0.74, and 0.77, respectively. The discriminatory ability of 20 gene tests, each with an odds ratio of 1.5–1.7, is comparable with the test of total cholesterol level for prediction of coronary heart disease (Wilson et al. 1998). The effect would be even greater if only 5% or 10% of all alleles tested had odds ratios in the range of 2.5–3.5 or if we could identify combinations of a few genes and/or gene-environment interactions that are strong predictors of the disease.

The comments of Janssens et al. also raise several interesting points regarding different perspectives on multiple genetic testing. Epidemiologic studies, including those on the utility of ROC curves for screening, provide a useful population perspective. In contrast, clinicians usually focus on individual patients rather than on the population as a whole, and this focus will be enhanced by the development of personalized genomic medicine (Roses 2000; Jain 2002). It is true that no more than a few people per million might turn out to have a very high risk defined by positive results for multiple genetic tests for a particular disease. However, it might be very important to these few people to know that they are at high risk if an intervention is available to prevent the disease. Our likelihood-ratio–based method provides an approach that is useful for individual patients and their physicians in predicting the probability of developing disease.

QUANHE YANG,[1] MUIN J. KHOURY,[2] LORENZO BOTTO,[1] J. M. FRIEDMAN,[4] AND W. DANA FLANDERS[3]

[1]*National Center on Birth Defects and Developmental Disabilities and* [2]*Office of Genomics and Disease Prevention, Centers for Disease Control and Prevention (CDC), and* [3]*Department of Epidemiology, School of Public Health, Emory University, Atlanta; and* [4]*Department of Medical Genetics, University of British Columbia, Vancouver*

References

Fletcher SW (1997) Whither scientific deliberation in health policy recommendations? Alice in the Wonderland of breast-cancer screening. N Engl J Med 336:1180–1183

Hanley JA (1989) Receiver operating characteristic (ROC) methodology: the state of the art. Crit Rev Diagn Imaging 29:307–335

Harris R, Lohr KN (2002) Screening for prostate cancer: an update of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 137:917–929

Jain KK (2002) Personalized medicine. Curr Opin Mol Ther 4:548–558

Janssens ACJW, Pardo MC, Steyerberg EW, van Duijn CM (2004) Revisiting the clinical validity of multiplex genetic testing in complex diseases. Am J Hum Genet 74:585–588 (in this issue)

Roses AD (2000) Pharmacogenetics and the practice of medicine. Nature 405:857–865

Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB (1998) Prediction of coronary heart disease using risk factor categories. Circulation 97:1837–4187

Yang Q, Khoury MJ, Botto L, Friedman JM, Flanders WD (2003) Improving the prediction of complex diseases by testing for multiple disease-susceptibility genes. Am J Hum Genet 72:636–649

Address for correspondence and reprints: Dr. Quanhe Yang, National Center on Birth Defects and Developmental Disabilities (CDC), 1600 Clifton Road, MS E-86, Atlanta, GA 30333. E-mail: qay0@cdc.gov

0002-9297/2004/7403-0030$15.00

Am. J. Hum. Genet. 74:589–591, 2004

Impact of Genotyping Errors on Type I Error Rate of the Haplotype-Sharing Transmission/Disequilibrium Test (HS-TDT)

To the Editor:

In a recent issue of the *Journal,* Zhang et al. (2003) proposed a haplotype-sharing transmission/disequilibrium test (HS-TDT) for the null hypothesis of no linkage or no association between a disease and a chromosomal region in which several tightly linked markers have been typed. Their method is applicable to data of nuclear families without phase information. The general idea of their approach is to compare the similarity of the transmitted haplotypes with the similarity of the nontransmitted haplotypes. If the chromosomal region contains a susceptibility locus, it is expected that the haplotypes being transmitted to affected children are more similar than parental haplotypes that have not been transmitted. This reasoning seems intuitively appealing. However, it may be supposed that a larger observed similarity for transmitted than for nontransmitted haplotypes is not necessarily due to the presence of a disease-susceptibility locus but can be a consequence of undetected genotyping errors. The proportion of genotyping errors that result in a Mendelian inconsistency (MI) is relatively small for family trios (Gordon et al. 1999). More important, in the context of HS-TDT, is the fact that the chance to detect a genotyping error differs for transmitted and nontransmitted haplotypes. Obviously, mistyping of an allele on a nontransmitted parental haplotype can never