



## **Editorial**

---

Scand J Work Environ Health [1999;25\(1\):1-3](#)

doi:10.5271/sjweh.376

**Biomarkers in the identification of risks, especially with regard to susceptible persons and subgroups**

by [Vainio H](#)

This article in PubMed: [www.ncbi.nlm.nih.gov/pubmed/10204664](http://www.ncbi.nlm.nih.gov/pubmed/10204664)

---



This work is licensed under a [Creative Commons Attribution 4.0 International License](#).

## *Biomarkers in the identification of risks, especially with regard to susceptible persons and subgroups*

Epidemiologic associations between exposure and disease outcomes have usually been based on the assumption that all people are equally susceptible to the effects of the exposure, although these effects may be modified by factors such as age, gender, ethnicity, and hormonal status. Pharmacogenetic studies have shown, however, that people vary in their ability to metabolize drugs (1). There is also increasing evidence that variations in DNA (deoxyribonucleic acid) repair capacity, cell cycle control, and immune response affect the risk for disease. In view of the genetic differences in a number of factors that predict the probability that disease will result from an exposure, risk factors may be identifiable only if the association between an exposure and a disease is strong, as in the case of tobacco smoking and lung cancer.

Assessments of the risks of populations that have heterogeneous responses may thus be biased and result in risk estimates that are diluted or masked. In studies of the etiology of multicausal diseases, such as cancer, a simplistic approach in which only single factors are evaluated is insufficient, and a multifactorial model is required to evaluate the environmental exposures and genetic and hormonal factors that affect susceptibility. Identifying susceptible subsets of the population on the basis of polymorphic genes involved in the line of defense between exposure and the initiation of disease processes in cells may more clearly delineate the factors that increase health risks among some, but not all, people. New strides may be made in understanding disease etiology and the role of particular factors in etiopathogenesis if molecular epidemiologic studies were conducted; however, incorporating molecular markers of susceptibility into epidemiologic studies may pose methodological problems that must be addressed by the research community. In addition, the double-edged sword of polymorphisms should be kept in mind. For instance, an increased risk for cancer might be associated with one genotype and an increased risk for toxicity with another. Or the same genotype may be associated with an increase in the occurrence of one, or a subtype of one, disease and a decreased risk for another.

Effect modification, also described as interaction, occurs when the association between an exposure and disease varies with different levels of a third variable. Therefore, data may be stratified according to the variable that is thought to modify the effect. Molecular epidemiology has extended the need for stratification to polymorphisms of putative risk-modifying genes. Within specific genetic categories, associations may be evaluated between groups who are putatively "at risk" and those who are not. This method of studying gene-environment interactions may more clearly elucidate cause and reveal previously unidentified risk factors, by the effects to be detected in subgroups when no main effect is observed overall.

A recent paper by Ambrosone et al (2) illustrates this concept. Although several studies have shown that tobacco smoking does not increase the risk for breast cancer overall (3), postmenopausal women with the slow *N*-acetyltransferase 2 (NAT2) genotype and who smoked were at increased risk for breast cancer in comparison with postmenopausal women without that genotype (2). Aromatic amines present in cigarette smoke, such as 4-aminobiphenyl, can first be *N*-oxidized and then be *O*-acetylated by NAT2 to activated forms that can bind to DNA. *N*-Acetylation, also catalyzed by the NAT2 isoenzyme, may also occur before *N*-oxidation, and result in the less active *N*-hydroxyarylamide. This conversion is viewed as a detoxification reaction.

The findings of Ambrosone and her co-workers (2) were not, however, corroborated in a study on the health of nurses (4). In that study, cigarette smoking was not appreciably associated with the risk for breast cancer among either slow or fast acetylators. Furthermore, a recent study by Millikan et al (5) gave little evidence that the effects of smoking are modified by NAT genotype, except among postmenopausal women. The relative risk for breast cancer associated with smoking within the previous 3 years was greater for women with the NAT2 rapid genotype than for those with the NAT2 slow genotype; these results are in direct contrast to those of Ambrosone et al (2). Thus the results of studies of modification of the effects of smoking by the NAT2 genotype are largely negative. Since the NAT isozymes catalyze both the activation and inactivation of tobacco-derived aromatic amines, however, any differences in risk associated with genotype may be difficult to detect.

A similar concept was addressed in 3 studies of breast cancer, menopausal status, and the activity of catechol-*O*-methyltransferase (COMT), which catalyzes methylation of the catechol estrogens (6—8). The association tended to be null in heterozygous populations, but clear relationships between dependent and independent variables were found when the data were stratified. While Thompson et al (7) found no association between the COMT genotype in a case-referent study of pre- and postmenopausal women together, clear inverse associations were found when the women were stratified by menopausal status. A similar phenomenon was noted by Lavigne et al (6), except that the associations by menopausal status were inverted. Millikan and his co-workers (8) did not find any association between COMT genotype and breast cancer risk, nor did the menopausal status modify the risk estimates. Thus the results of the 3 studies are inconclusive.

A further illustration of the complexities of such studies is provided by recent studies of colorectal cancer and the NAT2 genotype. Some of the studies suggest that the fast phenotype of NAT2 confers susceptibility to colorectal cancer because of greater activation of dietary heterocyclic amines, particularly in people who consume well-done red meat. In the study of Welfare et al (9), no association was observed between the NAT2 genotype and cancer risk, but the risk was increased among recent smokers with the slow NAT2 genotype (odds ratio 2.3). In contrast, the fast NAT2 genotype was associated with an increased risk among frequent consumers of red meat, the odds ratio being 6.0 (95% confidence interval, 1.6—26). This study suggests therefore that there are different risk factors for colorectal cancer in slow acetylators (ie, smoking) and fast acetylators (consumption of red meat).

Stratification by a putative effect modifier in epidemiologic studies should provide clearer answers than an overall analysis; however, in many studies of polymorphisms, in, for example, xenobiotic-metabolizing enzymes, and cancer risk, conflicting results have been obtained. The literature of molecular epidemiology is thus filled with inconclusive data, indicating that it is time the molecular epidemiologic community explored the areas of bias and flaws in study design and analysis that result in such inconsistent results.

Inadequate power to detect a true effect is one reason for inconsistent results. Statistical power depends on sample size, the size of the effect to be detected, and the variation within the study population. Small sample sizes are common in molecular epidemiologic studies, not only because molecular assays are expensive and the number of subjects recruited is thus restricted, but also because the method of stratified analysis automatically reduces the population by half or more. For studies of gene-environment interactions, cases and referents are stratified by genotype, and associations between the risk factor and disease status are evaluated separately within each stratum. Even in large studies, therefore, the numbers of subjects in each cell are drastically reduced, increasing the odds for type 1 and type 2 errors. In studies that are necessarily small because of stratification, movement of even a few persons from one category to another can greatly skew the results.

Medicine is concerned a priori with the health of people (10). When this direction is applied to the prevention of disease, people are identified who are liable to develop illness; then, they can be given help

in obtaining protection or taking avoiding action. The neat classification into 2 groups — susceptible deviants and the healthy (normal) remainder — is, of course, a considerable oversimplification. Susceptibility is rarely confined to a distinct high-risk minority, and the ability to predict individual outcomes is weak. As Walker (11) has pointed out, however, the more information we can gather on the sources of individual susceptibility, the better we can bridge the risk inferences we make for populations and obtain the kind of information that will enable people to make their own judgments about risk.

The potential for the unethical use of susceptibility markers is great. If these markers correlate with particular racial or cultural characteristics, individual people and groups who are already burdened by discrimination may face further problems. In addition, knowledge about individual variability may lead to the false assurance that, for example, workers who “pass” a susceptibility “screen” constitute a hardier group, and the temptation to place them in more hazardous jobs or to relax controls might increase. Inadvertent or intentional disclosure of findings on the susceptibility of a person could have a chilling effect on his or her ability to keep or get jobs or to obtain health insurance.

Before long, the whole human genome will have been mapped. DNA chips will be used to identify polymorphisms in a multitude of genes, so that it will become possible to evaluate the complex interactions of numerous genetic polymorphisms and environmental exposures. New oligonucleotide array-based powerful tools for genome analysis are already available to yield information on the inherited determinants of susceptibility to common multifactorial diseases, such as myocardial infarction (12). With the advances in the development of tools to identify susceptible subgroups of people, however, comes the responsibility for devising strategies to ensure valid results, to anticipate the ethical and legal implications of these tools, and to ensure that technology lives up to the promise. Environmental and occupational health can reap the benefits of advances in molecular biology and genetics, but it will take a concerted and cautious effort.

#### References

1. Nebert DW. Role of genetics and drug metabolism in human cancer risk. *Mutat Res* 1991;247:267—81.
2. Ambrosone CB, Freudenheim JL, Graham S, Marshall JR, Vena JE, Brasure JR, et al. Cigarette smoking, *N*-acetyltransferase 2 genetic polymorphisms, and breast cancer risk. *JAMA* 1996;276:1494—501.
3. Baron JA, Newcomb PA, Longnecker MP, Mittendorf R, Storer BE, Clapp RW, et al. Cigarette smoking and breast cancer. *Cancer Epidemiol Biomarkers Prev* 1996;5:399—403.
4. Hunter DJ, Hankinson SE, Hough H, Gertig DM, Garcia-Closas M, Spiegelman D, et al. A prospective study of *NAT2* acetylation genotype, cigarette smoking, and risk of breast cancer. *Carcinogenesis* 1997;18:2127—32.
5. Millikan RC, Pittman GS, Newman B, Tse J-KJ, Selmin O, Rockhill B, Savitz D, et al. Cigarette smoking, *N*-acetyltransferases 1 and 2, and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 1998;7:371—8.
6. Lavigne JA, Helzlsouer KJ, Huang H, Strickland PT, Bell DA, Selmin O, et al. An association between the allele coding for a low activity variant of catechol-*O*-methyltransferase and the risk of breast cancer. *Cancer Res* 1997;57:5493—7.
7. Thompson PA, Shields PG, Freudenheim JL, Stone A, Vena JE, Marshall JR, et al. Genetic polymorphisms in catechol-*O*-methyltransferase, menopausal status, and breast cancer risk. *Cancer Res* 1998;15:2107—10.
8. Millikan RC, Pittman GS, Tse J-KJ, Duell E, Newman B, Savitz D, et al. Catechol-*O*-methyltransferase and breast cancer risk. *Carcinogenesis* 1998;19:1943—7.
9. Welfare MR, Cooper J, Bassendine MF, Daly AK. Relationship between acetylator status, smoking, and diet and colorectal cancer risk in the north-east of England. *Carcinogenesis* 1997;18:1351—4.
10. Rose G. Sick individuals and sick populations. *Int J Epidemiol* 1985;14:32—8.
11. Walker VR. Direct inference, probability, and a conceptual gulf in risk communication. *Risk Anal* 1995;15:603—9.
12. Pastinen T, Perola M, Niini P, Terwilliger J, Salomaa V, Vartiainen E, et al. Array-based multiplex analysis of candidate genes reveals 2 independent and additive genetic risk factors for myocardial infarction in the Finnish population. *Hum Mol Genet* 1998;7:1453—62.

Harri Vainio, MD  
 Division of Health Risk Assessment  
 National Institute of Environmental Medicine  
 Karolinska Institutet  
 Box 210  
 S-171 77 Stockholm Sweden