Projection of Molecular Epidemiology in Medicine

P.A. Schulte*

Resumen

La epidemiología Molecular es un término que describe la incorporación de biomarcadores moleculares y de otro tzpo en la epidemiología. La Epidemiología Molecular tiene el mismo paradigma de la Epidemiología tradicional, pero representa la oportunzdad de emplear el alto poder de la Biología Moleculary de la Bioquímica contemporánea, para identificar las relaciones con las enfermedades por exposición ambiental Hay tres tipos de biomarcadores que pueden ser usados en este sentido marcadores de exposición, de efecto, y de susceptibilidad. Estos marcadores pueden usarse como variables dependientes o como independientes en la mayoría de los estudios epidemiológicos. Para su empleo es necesario que sean validados y cuenten con pruebas de campo. Esto requiere estrecha colaboración entre los científicos de los laboratorios con los de la Salud Pública. Especial atención merecen la interpretación y la difusión de los datos de los biomarcadores, así como la notable repercusión ética que tienen estos aspectos.

Palabras clave: *Epidemiologia Molecular, Biología Molecular, Salud Pública, biomarcadores, enfermedades por exposición.*

Summary

Molecular epidemiology is a term to describe the incorporation of molecular and other types of biomarkers into epidemiology Molecular epidemiology uses the same paradigm as traditional epidemiology but the former represents the opportunity to use the enhances resolving power of molecular biology and contemporaiy biochemical science to assess exposure disease relationships. There are three types of biomarkers that can be used in this regard: They include markers of exposure, effect and susceptibility. These markers can be used as dependent and independent variables in most epidemiologic study designs. Critica1 in their use is that they are validated and field tested. This requires extensive collaboration between laboratoiy and public health scientist. Special attention also needs or be paid to the interpretarion and communication of biomarker data and the ethical issues attendant to their use.

Keywords: *Molecular epidemiology, molecular biology, public health, biomarkers, exposure disease.*

In 1987 the US National Academy of Sciences published a model for a continuum of biological markers between external exposure to a xenobiotic and resultant disease.[1] Biomarkers can be defined as biochemical, molecular, genetic, immunologic or physiologic signals of events in biological systems. The incorporation of biomarkers in epidemiologic research has been referred to as "molecular epidemiology". Molecular epidemiology and traditional epidemiology use the same paradigms but the former represents the opportunit~ to use ihe enhanced resolving power of molecular biology and contemporary biochemical science in the assessment of exposure-disease realtionship.[2] The impetus for the National Academy effort to evaluate biological markers (biomarkers) in part comes from the needs of the government environmental agencies.[3-5] They have a concern

* National Institute for Occupational Safety and Health Cincinnati, Ohio.

as to whether biomarkers can be used to assess some of the environmental problems. In this paper a framework will be presented for considering biomarkers for use in epidemiologic studies. Additionally, some of the issues involves in transferring or graduating markers from the laboratory to the field will be discussed.

Epidemiology has historically involved assessment of the associations between exposures and diseases and epidemiologists have made great strides and contributions to environmental health even without knowing what the mechanism that links an exposure and disease. For example, John Snow, without knowing the mechanism of cholera development, could take the handle off the Broad Street pump, thereby stopping consumption of cholera-contaminated water.[6] Now with new tools from the basic sciences we can now look into this "black box" between exogenous exposures and diseases. A couple of other developments from basic sciences are also available to enhance epidemiologic research. One of these is the ability to identify smaller and amounts of xenobiotics. For example, one study of DNA adducts of polycyclic aromatic hydrocarbons in foundry workers demonstrated levels detected in fentamoles of adducts.[7] Similarly, in a study of hospital workers exposed to ethylene oxide the dose was measured at the picomole.[8] Although very low levels of xenobiotics can be measured with current technologies the ability to detect is far ahead of our ability to interpret. The health risk of such low levels of exposure is not known.

The ability to measure low levels of a xenobiotic also requires attention to the origins of a measured dose in population studies. It is necessary to understand that a biological marker of interest represents the accumulation of exposure from various sources and by various routes. Consequently it is necessary to account for the background of the marker that might occur in the nominally "unexposed" group.

Another contribution that basic sciences are providing to epidemiology and environmental health is the ability to look earlier in the natural history of the disease, so instead of using morbidity or mortality as an outcome, it is possible to identify intermediate outcomes (preclinical effect markers earlier in the natural history). This is the area where interventions in the disease process might be efficacious in reducing disease.

Due to these scientific developments, it is now possible to describe a new kind of framework in which to consider epidemiologic research.[9] Instead of associating exogenous exposure with clinical disease, it is now possible to be more specific about exposure. It is possible to identify the internal dose - the amount that gets into the organism, and the biologically effect dose. the amount that actually interacts with critical macromolecules. These are essentially markers of exposure, The framework (see Figure 1) with these markers of exposure and disease is a heuristic continuum that allows for portrayal of a series of biological events between exposure and resultant disease.[10] The actual steps in the continuum are arbitrary in terms of where they begin or end. Generally, the markers at the exposure end are believed to occur before the markers at the effects end. In the middle of the continuum are some biologic effects that could indicate a homeostatic response or that could be part of a causal pathway. At the effect end it is possible to identify alterations in structure or function that are preclinical.

Within each or between each of the steps in the continuum there are various host factors which can be identified as markers of susceptibility. Some of these can be acquired, some can be hereditary. Historically, epidemiologists have not utilized host factor characteristics to any great extent. Host factors have been generally represented by age, race and sex. Now it is becoming possible to utilize various phenotypes and genotypes as indications of susceptibility for appearance of disease or precursor conditios.[11]

Historically, the objective of epidemiology was to identify associations of an exposure with a disease. The "Achilles heel" of epidemiology is the identification of whether a person is exposed or not. When this is performed inaccurately, people are improperly classified. The hope is that certain biological markers will be available to allow us to reduce this kind of misclassification and to accurately identify who within a group of people has had an exposure.

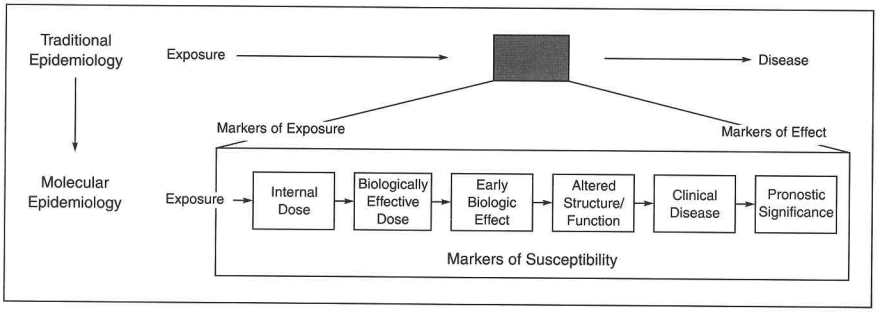

Figure 1. Enhancement of the traditional epidemiologic paradigm by the use of biological markers resulting in a molecular epidemiologic approach. In traditional epidemiology the mechanism of action is often a "black box", and associations between an exposure and disease are made by inference. In molecular epidemiology a continuum between an exposure and a disease is defined, and various markers are identified.

Similarly epidemiologic research in some parts has involved lumping of people who were heterogeneous in their disease categories. Using biomarkers could be a way of distinguishing them. Finally, one of the most exciting areas in contemporary research is the ability to assess such questions as: Who within a specific group might develop a disease? Consequently it is beginning to be possible to identify biomarkers that might allow us to discern within a population who is susceptible to a certain disease.[12]

There are also a great number of limitations that must be considered. The state-of-the-art is that may the biomarkers currently under discussion lack validation.[9] They have not been validated in the sense that they are proven useful beyond small-scale pilot studies. These early investigation often are conducted to determine the efficacy of an assay. There is little attention to characteristics of the study subjects. As research proceeds to larger and more diverse groups there is a need for collaboration with epidemiologists, statisticians, industrial hygienists and exposure assessors. This is usually not done. There are only a few examples of where they have actually been able to work together. It gets down to such questions as: Which discipline is going to take lead and where to publish the results.

To validate biological markers or ultimately to use them in population studies requires epidemiological expertise. The classic epidemiologic designs are cross-sectional, case-control and cohort-designs. Prior to conducting an epidemiologic study it is important to determine why and how a marker will be used.[13] What will it answer? what will it do better that just a questionnaire? what will it do better some sort of less labor intensive kind of effort? The lure of biomarker technologies needs to be weighed against their costs and their efficacy compared to traditional approaches. In assessing a health problem it is useful to think of a progression by which studies are performed. Cross-sectional studies are useful as a first cut. They can yield a lot of information but they have some problems because they have a concomitant ascertainment of both a dependent and independent variables. There are questions about temporality of what comes before what. But as a first cut, a small pilot cross sectional study can be quite useful.

The case-control study is a study that is quite amenable to the use of biological markers in a number of ways. Historically, often very heterogeneous types of cases have been lumped together as "cases". Using genetic and molecular techniques it is possible to distinguish disease subgroups. For example, the M344 probe identifies individuals with early stage bladder cancer that are likely to be the highest grade.[14] It may also be possible to identify a unique molecular fingerprint or spectrum that will indicate which carcinogen led to the mutations in tumor.[15] So there might be a heterogeneity in apparently homogeneous cases that was not recognized before the use of molecular genetic techniques.

One of the most powerful epidemiologic designs is the cohort study. Cohort studies usually involve following groups forward in time. This allows for resolution of temporal issues and the calculation of risk estimates. Biological markers can serve as dependent or independent variables in all of these epidemiologic study designs.

There is also the possibility of combining a variety of designs in a sort of hybrid format.[16] The classic one is where there is a cohort of people that is followed and a nested case-control study is conducted. This involves comparing all the cases of a particular disease and a sample of some of the people in the cohort who do not develop disease. The various epidemiologic designs can be applied to the biomarkers identified in Ihe continuum between exposure and disease, and the biomarkers can be used as dependent or independent variables in all epidemiologic designs to determine or not there is an outcome using one marker to predict another marker.

The use of biomarkers in studies of human populations requires that two questions be considered. What does the marker mean and what are you going to tell the subjects? These are also corollary issues such as, the impact of labeling a person as part of a subgroup that has a high frequency of a given marker. This may have impact on a person's credit rating, impact on insurability, and an impact on his/her ability to get a job. Many molecular biomarkers can be very powerfully misperceived in the population. There are many examples of this and it is a great concern. It is quite important that when designing larger scale studies using biologic markers, that first of all, they be validate (that is, there meaning in relation to diseases is know) so investigators have a little more understanding when they start to interpret the results and communicate them. Finally, one of the great hopes is that biological markers will be useful in assessing risks. Biological markers will be most useful in two areas of risk assessment: quantifying the interindividual variability in people and making risk assessments targeted to specific groups.[12,17,18]

Asscientistassessphenomenaatthemolecular and genetic level, they are observing more heterogeneity than was ever dreamed of. Some of this variability can be quantified thus overcoming one of the weaknesses of quantitative risk assessment which is the inability to pay adequate attention to variability. It is becoming possible to factor variability into assessment of risks.

Secondly, in terms of making risk assessments that are targeted to appropriate groups in the population it is possible to use various markers to identify subpopulations based on a frequency of a marker characteristic, who might be at heightened or diminished risk, and the possibly make policy decisions that will deal with these subpopulations.

In general, from an environmental health and epidemiologic point of view those are the strengths and limitations of biological markers. In summary, epidemiologists should not to use biomarkers in field studies until there has been extensive validation of them? Unfortunately, there is not much funding to do that kind of research. It is difficult to obtain grants to conduct validation studies, moreover, it is difficult to motivate scientists who are in the "cutting-edge" of research to go back and perform more mundane efforts. It is not the kind of work that is being fostered by the granting authorities. Therefore, although there are many scientists developing new markers and new assays, epidemiologists are unable to use them because they are not validated for population research. Until we have that kind of information we are not going to be able to adequately use and interpret these markers in public health studies.

References

1. **National Research Council. Biological markers in environmental health research.** Environ **Health Perspective** 74:1-19,1987.
2. **Schulte PA, Perera FP, and Rothman N. Molecular Epidemiology. Encyclopedia of molecular biology and molecular medicine. New York. pp 258-262,1996.**
3. **National Research Council. Biologic Markers in Reproductive Toxicology. Washington,** D.C. **National Academy Press** 1989.
4. **National Research Council. Biologic Markers in Pulmonary Toxicology Washington** D.C. **National Academy Press** 1989.
5. **National Research Council. Biologic Markers in ImmunolotoxicologyWashingtonD.C. NationalAcademy Press, 1992.**
6. **Snow J. On the mode of communication of cholera (2nd** Ed) Cnurch II, **Lonoon** Reproduced in "Snow an cholera". Commonwealth Fund. **New York. 1936.** Reprinted by **Hafners New York, 1965.**

7. **Phillips** DH, Hemminki, **Alhoen** A, Hewer A and Grover PL. Monitoring occupational exposure to carcinogens: Detection by 32P-postlabeling of aromatic DNA adducts in white blood cells from foundry workers. Mutat Re 204:531-41,1988.
8. Shulte PA, Boeniger M, Walter JT, Schober SE, Peira MA, Gulati DK, **Wojciechowski** J, Garza A, Froelich R, **Strauss G,** Halperin WE, **Herrick R, Griffith** J. Biological markers of hospital workers exposed to low levels of ethylene oxide. Mutat Res 278:237-41,1992.
9. Shulte PA, Perera FP. Validation In Shulte P and Perera FP eds. Molecular Epidemiology: Principles and Practices. San Diego Academic Press, 1993 pp 79-107.
10. Shulte PA. A conceptual framework for the validation and use of bioiogical markers. Env Research 48:129-144, 1989.
11. Daly AK, Cholerton S, **Gregory** W, **Idle** JR. Metabolic poiymorphisms. Phram Ther 57:129-160, 1993.
12. **Shields** PG, Harris CC. Molecular epidemiology and the genetics of environmental cancer. JAMA 246:681-689.
13. Rothman N. Epilogue. In PA Schulte and FP Perem (eds). Molecular Epidemiology: Principles and Practices. San Diego, Academic Press. 1993 pp 565-569.
14. Rao JY, Hemstreet GP, **Hurst** RE, Bonner RB, Jones PL, Min KW and Fradet Y. Alterations in phenotypic biochemical markers in bladder epithelium during tumorigenesis. Proc. Nat. Acad. Sci., USA (In press).
15. **Hollstein** M, Sidransky D, Vogelstein **B,** Harris CC. p53 mutations in human cancer. Science 253:49-53, 1991.
16. Schulte P, Rothman N, Schottenfels D. Design considerations in molecular epidemiology. In PA Schulte and FP Perera. Molecular Epidemiology: Princlples and practices Academic Press San Diego, 1992 pp 159-198.
17. Schulte P. A conceptual and historical framework for molecular epidemiology. In PA Schulte and FP Perera. Molecular Epidemiology: Principles and Practices. San Diego, Academic Press, 1993 pp 3-44.
18. Perera **FP.** The potential usefulness of biological markers in risk assessment. Environ Health Perspect 76:141-145, 1987.