THE COLT LECTURE

Pergamon

Ann. occup. Hyg., Vol. 41, Supplement 1, pp. 1–6, 1997
British Occupational Hygiene Society
Published by Elsevier Science Ltd
Printed in Great Britain
0003-4878/97 $17.00 + 0.00
Inhaled Particles VIII

PII: S0003-4878(96)00125-1

MOLECULAR EPIDEMIOLOGY OF OCCUPATIONAL AND ENVIRONMENTAL LUNG DISEASE

P. A. Schulte

National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, OH 45226, U.S.A.

BIOMARKERS AND MOLECULAR EPIDEMIOLOGY

Advances in molecular biology and analytical chemistry are providing potentially useful tools for the epidemiologic study of occupational and environmental lung disease (NRC, 1989; Paoletti, 1995; Schulte, 1996). These tools are biological indicators, or biomarkers, of events that represent exposure to a xenobiotic, effects of exposure and susceptibility to effects of exposure. Taken together, these biomarkers can be used to describe a continuum of events following a xenobiotic exposure and proceeding through resultant diseases. Using a continuum concept expands the classic epidemiologic model that involves statistically inferring an association between an exposure and disease. This classic approach has yielded many useful findings for public and occupational health with regard to pulmonary diseases such as coal workers' pneumoconiosis, beryllium disease, asthma from wood dust and lung cancer from asbestos, cigarette smoking, bis-chloromethyl-ether and uranium. The epidemiologic assessment involving these lung diseases generally relied on statistical inference rather than mechanistic insight to associate exposure with disease; but that is not to say that investigators conducting such studies lacked biological hypotheses for such associations.

What differs from these previous investigations is the potential for identifying various events in a continuum between an exposure and disease (NRC, 1989). These events are represented by biomarkers. Biomarkers are generally biological measurements representing or correlating with an event, serving as the event, or predicting the event. Biomarkers can range from those on the macroscale, such as dysfunction, disease and death down to those on the microscale (such as the cell, gene or molecule). The middle of this range can include biomarkers such as various anatomic, physiologic and functional changes.

Biomarkers can serve as dependent or independent variables in classical epidemiological study designs such as cohort, case–control, cross-sectional and intervention studies. As a dependent variable, a biomarker of effect, [such as C8 oxidation of deoxyguanosine in coal workers (Schins, 1995) or 4-hydroxy-2-nonenal protein adducts in individuals exposed to ozone (Li *et al.*, 1996)] could, if validated, serve as surrogates for lung disease, or earlier stage events that could be associated with later disease, in an epidemiologic study. For independent variables, biomarkers of exposure can be used. For example, in addition to using a job title or a few

exposure measurements as an independent variable for exposure, it may be possible to use a biologic measure of cumulative dose for better defining biologically important exposure. Biomarkers may also be used to evaluate the factors that mediate between an exposure and disease. In epidemiology, these are called effect modifiers. For example, in people with the null GSTM1 allele, the risk of lung cancer given exposure to polycyclic aromatic hydrocarbons, was approximately two-fold greater than for those with the other prominent alleles given the same exposure level (Tang *et al.*, 1995).

These two uses of biomarkers, to define a continuum and to serve as variables and effect modifiers in epidemiologic research, have been termed as "molecular epidemiology" (Schulte and Perera, 1993). "Molecular epidemiology of lung disease" is a heuristic phrase that represents the natural confluence of powerful developments in basic biomedical sciences and the field-tested methods of epidemiology to study lung disease. Although molecular epidemiology can be viewed as an evolutionary step in epidemiology, a supplemental set of tools, or even a separate discipline, it does not represent a shift in the basic paradigm of epidemiology (Schulte, 1993). However, the term can be used to stand for the application of the whole range of biomarkers used in epidemologic investigations (Rothman *et al.*, 1995). These include molecular, genetic, cellular, histologic and physiologic markers. The literature discussing molecular epidemiology contains a difference of opinion about this term. The definition used in this paper emphasises epidemiology and treats the molecular descriptor as a way to define the variables to use in epidemiologic investigations. Others focus on the use of molecular tools to gain mechanistic insight but not necessarily to enhance understanding of epidemiologic associations. These molecular tools are most useful in epidemiology and risk assessment when the mechanistic information is used in conjunction with the established principles of epidemiology. Ultimately, the term "molecular epidemiology" should be seen as a signpost pointing to a more biological basis for epidemiology.

MOLECULAR EPIDEMIOLOGY AND LUNG DISEASE

Except for immunological markers, biomarkers (particularly cellular and molecular markers) have been infrequently used in the study of occupational and environmental lung disease. A TOXLINE (U.S. National Library of Medicine) title-field search was conducted of all the literature between January 1986 and July 1996 that could be identified with a search strategy that included various occupational and environmental lung diseases synonyms plus the terms "human and epidemiologic" plus "cellular", "molecular", "genetic", "DNA", or "biomarkers". Approximately 5221 citations (both human and epidemiologic) were identified, 70 of these were actually published investigations involving biomarkers in studies of occupational or environmental lung disease. When these 70 studies were assessed according to the established biomarker categories, the following percentage distribution was identified: internal dose, 9.4%; biologically effective dose, 20.2%; early biological effects, 25.5%; alternations of structure and function, 10.8%; clinical disease, 9.7%; and susceptibility, 20.2%. This distribution roughly indicates how often the markers were used and would probably change slightly with

different search strategies and classification logic. Nonetheless, the exercise provides a glimpse of how biological markers have been used to study occupational and environmental lung disease. In general, these studies provide no information about unknown causes of lung disease. However, they serve other important functions: they validate markers as indicators of exposure, effect, or susceptibility and they elucidate the causal pathway and biological plausibility of previously identified associations.

Two types of studies involved internal dose biomarkers: those measuring a xenobiotic in lung tissue indirectly by imaging and those using other biologic specimens such as blood to measure metabolites (e.g. of PAHs) or antibodies (e.g. of tuberculosis). Biomarkers of biologically effective dose were generally assessed in surrogate tissues such as peripheral blood lymphocytes. Most of these studies were exposure-selective, cross-sectional studies that assessed the relation of various DNA adducts with exposure. Occasionally, protein adducts were assessed. Case–control studies were also frequently used to study markers of biologically effective dose. At least one study compared cancer patients on the basis of smoking history and looked for marker–exposure associations based on tissue measurements. Many of the studies involving DNA adducts also assessed susceptibility factors such as various genotypes for p450 enzymes.

The category "early biologic effects" is the largest. However, the inclusion of a marker in this category is subjective and dependent on the state of the science. If a marker is a biologic effect that is known to be directly on the causal pathway to lung disease and is not reversible, it may be rightfully placed in the next category, "alteration of structure and function". At this time, for most markers, the extent to which a marker is a contributing causal factor in lung disease is not known so markers, such as oncogenes, tumour suppressor genes or proteins, are included as markers of early biologic effects. The most frequent marker found in this category is the p53 tumour suppressor gene. The studies involving altered structure and function are similarly subjective and are based on the extent to which a marker predicts disease. Many of the studies in this category, as well as in the clinical disease category, involved immunologic markers such as tumour necrosis factor alpha or IgE in asthma studies. In no cases were molecular changes used as dependent variables to represent disease in epidemiologic studies.

Studies with biomarkers of susceptibility represent the second largest category. Generally, these involve phase 1 (e.g. CYP1A1) and phase II (GSTM1) metabolic polymorphisms. The existing knowledge about the causes of occupational and environmental lung disease has resulted from clinical, pathological and epidemiological investigations. These investigations did not involve molecular markers, but in a broader sense, they involved a range of biological markers at the histologic, physiologic and individual levels. The exceptions are immunopathology studies and assessments of cytotoxicity and cellular pathology. Thus, for diseases like coal workers' pneumoconiosis, silicosis, asbestosis and lung cancer from inhaled carcinogens, the major etiologic questions have been answered previously by classic epidemiologic methods. The contribution of molecular biomarkers in understanding these diseases probably involves issues of medical screening and surveillance (i.e. what early effects can be identified for therapeutic intervention); identification of those workers likely to progress to more serious disease (e.g.

progressive massive fibrosis in coal workers' pneumoconiosis) and contribution to risk assessments (e.g. what are the risks at low levels of exposure?).

FUTURE EFFORTS

In the future, molecular biomarkers might be useful in the area of new or undefined hazards. For example, many countries are wrestling with the question of potential health effects from manmade mineral fibres, such as refractive ceramic fibres. Are these fibres fibrogenic or carcinogenic? In some cases, these fibres have been produced for relatively short periods and with low airborne concentrations. The question of concern is what the health risk is at these concentrations. Biomarkers of effect such as those found for more widely studied fibres like asbestos, could be used to determine whether there are any exposure–response relationships in workers exposed to a new fibre of concern. A research study could be designed to compare workers exposed to asbestos with those exposed to refractive ceramic fibres for assays of identified biomarkers. Using appropriate controls and covariate adjustments, it would be possible to determine if the exposed workers were similar to each other or greater than the controls. This could be a signal to restrict exposures. If such a relationship was identified and risk could be inferred then precautions could be taken.

Similarly, with mixtures of substances that cause oxidative stress, understanding molecular-level events could provide useful surrogates for exposures or early effects. For example, it has been shown that 4-hydroxy-2-nonenal protein adducts may be associated with apoptosis and necrosis and may serve as a secondary toxic messenger for acute ozone injury from exposures to low concentrations. This marker could be used as a dependent variable in a population study.

Molecular markers may also be useful in distinguishing between etiological pathways. Consider for example, the mutation spectrum of the p53 tumour suppressor gene. Different carcinogens seem to cause different characteristic mutations. For example, lung cancers in cigarette smokers are different from lung cancers in nonsmokers. Cigarette smokers have more G:C to T:A transversions (Harris, 1996). The p53 mutation spectrum can be influenced by both the type and location of the promutagenic lesion and understanding these characteristics can provide etiological information. Understanding this kind of information in both humans and animals potentially increases our ability to assess cancer risks accurately. Cancer risk assessment has been forced to provide conservative risk estimates because in addition to gaps in classical exposure and disease data, limited knowledge exists about complex pathobiological processes during carcinogenesis, differences in metabolism of carcinogens, different DNA repair capacities, variable genomic stability among animal species, and variation among individuals with inherited susceptibilities (Sutter, 1995). Molecular biomarkers measured in animals and humans according to the established parallelogram approach used in genetic toxicology can provide useful data for low-dose assessments of risk (Sutter, 1995).

At this time, the lack of significant contributions of molecular markers to the epidemiologic study of occupational and environmental lung disease should not be viewed with concern about their utility. Most of the markers thus far identified

have been in the developmental/validation stage. The challenge now is to use categories of markers and assays that have been found with known carcinogens and toxicants to study and control substances that are suspected pulmonary hazards to fill in gaps in the data. Validation efforts are needed in the laboratory and the field before a marker can be used to assess a suspicious pulmonary toxicant. Nonetheless, as biomarkers are found to be useful, they need to be used proactively to identify potential hazards before widespread exposures occur. In the area of occupational and environmental lung disorders, biomarkers could be applied to topics such as: fibrosis and cancer from man-made mineral fibres, carcinogen risk assessment for diesel emissions, early effects from air pollution and exposure to oxidative stressors. Finally, more attention should be given to using molecular biomarkers to assess prevention and intervention efforts.

Successful applications of molecular markers will not be rapid if a passive approach is taken by the occupational and environmental communities. The mapping of the human DNA (and various other genomes) is yielding a vast amount of raw information that (unless interpreted) will not provide tangible tools for research and control. Occupational and environmental researchers and risk assessors need to increase their knowledge and access to the raw data. Computer-driven "bioinformatics" coupled with the ability to rapidly obtain partial sequences of expressed genes portend revolutionary changes in scientific research and the practice of occupational medicine and in the collation of data for useful regulatory and prevention purposes (Rodbell, 1996). Assuming the irreversibility of gene patterns indicates hazard to the cell and possibly to the human body, dose–response relationships, under appropriate cell/tissue culture and exposure conditions can be accurately determined. The important difference from other toxicological approaches is that the method involves the total expressed gene pattern. Epidemiologists and industrial hygienists need to envision investigations that involve the laboratory screening of chemicals followed by field research (Rodbell, 1996).

More research is also needed to assess the relationship of biomarkers in blood to pathologic changes in lungs. Linkage of these two specimens with each other and with lung imaging technology should enhance the utility of biomarkers.

Finally, investigators need to integrate molecular epidemiologic approaches with other approaches for etiologic, intervention, and clinical research on lung diseases. Such integration can be accomplished in part by encouraging banking of biological specimens when possible.

The literature review showed that only a small number of biomarkers were used in epidemiological studies of lung disease. Generally, clusters of studies used the same markers (e.g. p53). This small number is good in one sense, since widespread use of unvalidated markers is unwarranted. Nonetheless, a large and diverse array of biomarkers have been identified for various lung diseases and their causal pathways. More of these markers need to be field tested and used in epidemiologic studies.

The use of molecular biological markers to study and control occupational and environmental lung disease can raise various ethical, legal and social issues. Researchers need to anticipate these issues and the impacts on participants in studies and make efforts to protect them.

REFERENCES

Harris, C. (1996) p53 tumour suppressor gene: at the crossroads of molecular carcinogenesis, molecular epidemiology and cancer risk assessment. *Environ. Health Perspectives* **104**, 435–39.

Li, Li, Hamilton, R. F., Jr, Kirichenko, A. and Holian, A. (1996) 4-hydroxy-2-nonenal as a mediator for oxidative injury. *Toxicol. appl. Pharmacol.* **139**, 135–143.

National Research Council (1989) *Biological Markers in Pulmonary Toxicology*. National Academy Press, Washington, D.C.

Paoletti, P. (1995) Application of biomarkers in population studies for respiratory non-malignant diseases. *Toxicol.* **101**, 99–105.

Rodbell, M. (1996) Bioinformatics: An emerging means of assessing environmental health. *Environ. Health Perspectives* **104**, 136.

Rothman, N., Stewart, W. F. and Schulte, P. A. (1995) Incorporating biomarkers into cancer epidemiology: a matrix of biomarker and study design categories. *Cancer Epid. Biomarkers Prev.* **4**, 301–311.

Schins, R. P. F., Schilderman, A. E. L. and Borm, P. S. A. (1995) Oxidative DNA damage in peripheral blood lymphocytes in coal workers. *Int. Arch. occup. Environ. Health* **67**, 153–157.

Schulte, P. A. (1993) A conceptual and historical framework for molecular epidemiology. In *Molecular Epidemiology: Principles and Practices* (Edited by P. A. Schulte and F. P. Perera), pp. 3–44. Academic Press, San Diego.

Schulte, P. A. (1996) Use of biomarkers to investigate occupational and environmental lung disorders. *Chest* **109**, 95–123.

Sutter, T. R. (1995) Molecular and cellular approaches to extrapolation for risk assessment. *Environ. Health Perspective* **103**, 386–89.

Tang, D. L., Chiamprasert, S., Santella, R. M. and Perera, F. P. (1995) Molecular epidemiology of lung cancer: carcinogen—DNA adducts, GSTM1 and risk. [Abstract] Annual Meeting of the American Association for Research, Toronto.