doi : http://dx.doi.org/10.5016/1806-8774.2006v8p1

ARBS Annual Review of Biomedical Sciences

pdf freely available at http://arbs.biblioteca.unesp.br

2006;8:1-8

# Population Monitoring for Assessment of Health Concerns Using Biomarkers

**William Au[1], Julian Preston[2]**

[1]Dept. of Preventive Medicine and Community Health, University of Texas Medical Branch, USA
[2]Environmental Carcinogenesis Division, US Environmental Protection Agency, USA

Received: 01 February 2006; accepted 06 March 2006

*Au W, Preston J. Population Monitoring for Assessment of Health Concerns Using Biomarkers. ARBS Annu Rev Biomed Sci 2006;8:1-8.* Exposure to environmental toxic substances is a significant cause of human health problems. In the modern world, exposure to small amount of these substances, either natural or man-made, is often unavoidable. The critical question to ask is, at what concentrations and exposure conditions are these substances hazardous to human health. One approach to address the question is to conduct population monitoring studies using reliable biomarkers that can be used to indicate excessive exposure, susceptibility to health consequences, and increased potential risk for health problems. In this review, the application of traditional and molecular biomarkers in population monitoring studies is systematically presented. A major emphasis in the review is to show how to design rigorous study protocols so that results from these studies will be appropriate for use in disease reduction programs. In addition, the benefits and limitations in the use of biomarkers are presented.



**Keywords:** population monitoring, biomarkers, chromosome aberrations, gene mutations, genetic susceptibility, health risk assessment

**[1]Correspondence**
William W Au, Dept. of Preventive Medicine and Community Health, 700 Harborside Drive, 2.102 Ewing Hall, The University of Texas Medical Branch, Galveston, TX 77555-1110, USA. E-mail: william.au@utmb.edu

# Table of Contents



# 1. Introduction

Traditional epidemiological studies are instrumental to our understanding of health hazards from exposure to environmental toxic substances. On the other hand, there are limitations to these investigations. For example, they generally do not provide information on dose response, mechanisms for disease induction and inter-individual differences in response to the exposure. To address these shortcomings, scientists began to incorporate analytical procedures into epidemiological investigations. The revised protocols became known as molecular epidemiological studies (Perera & Weinstein, 1982). Subsequently, the analytical procedures were expanded to include a variety of biomarkers that assess chemical metabolites, cellular response to exposure and biological consequences. In this review, the use of biomarkers in population monitoring studies is systematically presented.

# 2. Strategies for Designing Population Monitoring Studies

The underlying reason for conducting a population monitoring study is to investigate whether exposure to certain environmental hazardous substances at some particular level poses a health risk to the population. In these investigations, scientists often use biomarkers to make the linkage between exposure, induction of abnormal biological effects and risk for development of disease. When the studies are conducted appropriately, they can have tremendous impact on improved assessment of public health outcomes. Therefore, developing a vigorous strategy for the investigation is essential to the utility of the study (*e.g.* Au *et al*., 1998).

The general strategy for conducting a population monitoring study requires the fulfillment of the following criteria:

## 2.1. Characterization of the exposed population

Documentation of exposure conditions, health concerns, morbidity and/or mortality patterns of the population is useful because the information will be used to determine the study design and the impact of the study on health risk. Investigators will also find it useful to understand whether the motivation for study of an exposed population is based on the desire for improvement of health, for compensation, or for other purposes. The nature of the motivation may affect the extent of participation from the exposed populations or their willingness to provide unbiased information. The information will therefore determine if the study protocol should be adjusted to overcome the impact from different motivating factors. Regardless of the motivation, the study should be designed with a clearly-defined hypothesis.

## 2.2. Development of the study hypothesis

The information on the exposed population together with a careful research of the scientific literature will allow the scientists to formulate a good study hypothesis. The hypothesis should be testable with a well-defined scientific investigation. At this stage, the investigators will start to design the study. The investigators have to consider whether the study is built upon factual information and/or assumptions. An appropriately designed study is expected to confirm the fact and to test the assumptions adequately. The selection of specific biomarkers has to be considered and this topic will be addressed in a later section. The resources needed for the study such as the availability of personnel and financial and technical resources should be considered carefully. A frequently overlooked step in planning a study is the consideration of how to evaluate and interpret the results. Although it is convenient for investigators to consider this step after the data have been collected, the delay may cause serious problems, (*e.g.* the collected data may be inadequate for statistical evaluation and the data may be inappropriate to address the hypothesis).

## 2.3. Selection of the cases (exposed individuals) in an objective and unbiased manner

This can be achieved by using predetermined criteria such as duration of exposure, proximity to the source of exposure, age ranges of the volunteers, and gender. The selection process should be open to all qualified individuals so that there will not be a selection bias and/or any under-representation of certain subgroups of individuals.

## 2.4 Selection of the cases with limited confounding factors

It is well known that previous or concurrent exposure to other hazardous substances will have an undesirable effect on the outcome of investigations. These other environmental factors may include occupational exposure to other toxic substances, use of toxic medications, smoking of cigarettes and living near hazardous waste sites. In addition, it is important to select healthy individuals because certain medical conditions will influence individual responses to the exposure.

## 2.5. Selection of matched controls

It should be emphasized that having appropriate controls is an essential component of the study. Ideally, the controls are identical to the cases in all aspects except the exposure conditions. However, it is an unrealistic expectation that the ideal conditions can be met in most population studies. In practice, the controls are selected based on the same criteria that are used for the cases, except for the exposure. Then, the two qualified populations are matched based on certain essential characteristics such as age, gender and ethnicity. If cigarette smokers cannot be eliminated from the cases, the controls need to be matched under this condition. Typically, the history of cigarette smoking habit is translated into pack-years for comparison from one individual to another. A pack-year is calculated by multiplying the number of cigarettes smoked per day by the number of smoking years. In general smoke-years can range from 20 to over 60 for smokers. The pack-years for controls are usually matched with that from cases with a range of +/- 5. Another approach is to subdivide the smokers into heavy, medium and light smokers, and conduct the matching with controls accordingly. A general approach for matching for age is +/- 5 years. For studies using small populations, *e.g.* around 25 cases and 25 controls, the matching can be done on a one-to-one basis. When the study populations are larger, frequency matching is more practical. In the latter approach, the matching is done based on the group characteristics rather than on individuals, (*e.g.* matching the mean age and pack-years for the two groups).

## 2.6. Collection of specimens for laboratory analysis

For population studies, the most readily available specimens are the body fluids, such as blood and urine. In addition, buccal smears, pap smears and sputum can also be used (Salama *et al.*, 1999). Other specimens will involve invasive procedures for collections therefore they are generally not appropriate for population investigations of the type considered in the present chapter.

## 2.7. Selection of biomarkers

For population monitoring studies, it is critical to have adequate exposure information. Without such information, it becomes impossible to establish a cause-effect association from the investigation. As indicated later, certain exposure biomarkers can be used for this purpose. The exposure biomarkers should compliment the selection of biomarkers that are indicative of biological effects from the exposure. The approach will produce data that will be essential for scientists to generate cause-effect association and to assess health risk.

## 2.8. Conducting the study and interpretation of results

The collection and analysis of samples from both the cases and controls should be conducted simultaneously. In addition, the samples should be coded so that the donors for the samples are unknown to the laboratory personnel. The use of these criteria will minimize negative impact to the study caused by potential temporal (seasonal), laboratory and personnel variations.

The collected data should be organized and analyzed as a whole using recommended statistical methods without any further selection. Assuming that all experiments for the study have been conducted according to acceptable procedures, the data should be reliable. The presence of outliers in the data set is expected. Their presence should be included in the report and explained rather than excluded to give what appears to be a better fit to the original hypothesis. It should be kept in mind that the hypothesis of the study may be erroneous and the study results that contradict the hypothesis may be correct. By ignoring the true meaning of the experimental data, a new and useful discovery may be missed.

# 3. Biomarkers for Population Studies

Several review papers have addressed the use of different biomarkers in population monitoring studies (Hulka *et al.*, 1990; Bonassi & Au, 2002; Albertini *et al.*, 2003). A summary of representative biomarkers is shown in Tables 1 and 2. The available biomarkers are grouped into several categories: biomarkers of exposure, susceptibility, early biological effects and health risk. A new interest is focused on developing biomarkers to provide a good understanding of the genome-based response to exposure (Table 2). Selection of specific informative biomarkers from such genome analysis can help in the prediction of adverse health outcomes. As mentioned earlier, a combination of appropriate biomarkers should be used so that a good understanding of the cause-effect relationship and a reliable assessment of health risk can be achieved.

Table 1. Traditional biomarkers for population monitoring studies[a].

| **Biomarkers that are useful for incication of** | | |
|---|---|---|
| Exposure[b] | Early Biological Effects | Health Risk |
| Chemical metabolites aberrations | Chromosome aberration | Chromosome aberration |
| DNA adducts | Micronuclei | Disease gene-specific expression |
| Protein adducts | HPRT gene mutation<br>Glycophorin A mutation<br>Sister chromatid exchanges<br>DNA strand breaks and repair | |

[a]The biomarkers under each category are listed in the order of usefulness, the highest on top.
[b]Data from environmental measurement of chemicals can be used in support of the biomarker data.

Table 2. Genomic biomarkers for population monitoring studies[a].

| Biomarkers of Susceptibility by Genome-Wide Analysis | |
|---|---|
| Marker | Analysis |
| Variant Chemical metabolizing genes (e.g. CYP1A1, GSTM1) | Proteomic arrays |
| Variant DNA repair genes (e.g. XPD, XRCC1) | Gene expression arrays<br>DNA microarrays |

[a] Variations of genes are due to single nucleotide polymorphism, deletion, insertion, etc.

## 3.1. Biomarkers of susceptibility

A major challenge in population studies is the heterogeneous nature of the human population. A variety of factors contribute to the heterogeneity. Some of them are acquired in life such as aging, fitness, smoking of cigarettes (Au, 2001 and 2002); while others are genetically determined such as genetic variations for chemical metabolism and DNA repair (Au *et al.,* 2003; Norppa, 2003).

With the rapid development of molecular genetics, scientists are now able to investigate genetic variations as a contributing factor to the development of environmental disease (Table 2). The current focus is on inheritance of variant (polymorphic) chemical metabolizing and DNA repair genes that are key genetic susceptibility factors. Some investigators have combined the analysis of polymorphic genes with expression of biomarkers for health risk assessment (Au *et al.,* 2003; Norppa, 2003). For example, Harms *et al.* (2004) found significant increase of chromosome aberrations among cigarette smokers who had inherited the variant DNA repair (XPD) and the metabolizing genes (GSTM1) to increase their risk for lung cancer. It is generally quite difficult to appropriately account for the possible range of polymorphisms that can influence outcome. In fact, it is usually not necessary to do so, unless it is intended to better understand the underlying reasons for outliers of response.

## 3.2. Genome-based biomarkers

High throughput technologies are now available to screen for genome-based activities that can be used to help reduce uncertainty in the estimate of risk for disease following environmental exposures. As shown in Table 2, these techniques allow scientists to scan for DNA sequence variations (using DNA microarrays), gene expression (using gene expression arrays) and protein expression (using proteomic analysis). From these analyses, enormous amounts of data are generated that require extensive technology (bioinformatics and biostatistics) for interpretation. Although the technology is promising, it is not yet routinely used in population monitoring. However, it can quite reasonably be used at this time to provide information on the selection of informative biomarkers of response.

## 3.3. Biomarkers of adverse health outcomes

It is clear that the detection of abnormal expression of disease related genes (*e.g.* p53), is indicative of health risk. In addition, induction of chromosome aberrations has also been found to be predictive of health consequences. For example, prospective studies from the Nordic countries and from Italy have shown that normal populations having high chromosome aberrations had significantly increased cancer incidence and mortality than those with lower frequency of aberrations (Bonassi *et al*., 1995; Hagmar *et al.,* 1998). Another use of the chromosome assay is the determination of cancer-specific chromosome changes, *e.g.* leukemia-specific chromosome alterations in benzene-exposed workers (Zhang *et al.,* 2002).

## 3.4. Selection of biomarkers

As shown in Tables 1 and 2, different categories of biomarkers can be chosen for use in population monitoring studies. The selection of biomarkers is dependent upon the objective of the investigation. For

example, if the objective is to demonstrate exposure, biomarkers that are listed under the Biomarkers of Exposure category should be chosen. If the objective is to understand individual differences in response to an exposure, the assays under the Biomarkers of Exposure, Susceptibility and Early Biological Effects will most likely be chosen. Biomarkers of Health Risk will be used to address environmental health problems.

Irrespective of study objectives, other conditions should also be considered, for example, what is the stability/persistence of the biomarkers after the exposure? After an acute exposure, the expression of chemical metabolites, DNA/protein adducts and DNA strand breaks can be detected within a few hours but they would only remain detectable for up to a few days at most. On the other hand, the formation of chromosome aberrations could take up to one to two days and would remain detectable for months to years. Therefore, the selection of the biomarkers will dictate the time when biological specimens are collected for laboratory analyses. From chronic exposure, the biomarkers that can accumulate with time will be desirable, *e.g.* chromosome aberrations and HPRT gene mutations in lymphocytes. The selection of biomarkers is also dependent upon the use of different cell types for the investigation (Salama *et al.*, 1999). For example, the HPRT gene mutation assay is best performed using lymphocytes. The micronucleus assay can be used in a variety of cell types because the analysis itself does not depend upon the availability of dividing cells.

# 4. Confounding Factors

Two factors are the frequent cause for generating non-reproducible data. They are small sample size and inadequately matched exposed and control populations. The problem with small sample size can be minimized by conducting a sample size power calculation. For example, the calculation will indicate how many exposed and control subjects are needed to determine the difference between the two populations if the difference is 50%. Matching of populations requires some effort, *e.g.* having good knowledge of the background information regarding the exposed and the potential control populations. The knowledge includes the availability of individuals for selecting the exposed and control subjects, age distribution, ethnic composition and exposure conditions. Since the control population has little incentive for participation in the study, the investigators need to consider what constitute an appropriate compensation to encourage their participation.

# 5. Interpretation of Results

As discussed above, the design of a study in which a biomarker or biomarkers are to be used to predict if an exposure has occurred, how large an exposure has occurred and/or what is the magnitude of any attendant risk to the exposed populations is critical as regards the ability to provide a valid interpretation of the results. An appropriately designed study should be one that follows the requirements of an epidemiological study, albeit a relatively small one (Preston, 1999; Bonassi & Au, 2002). This requires, as noted previously, the selection of an appropriate biomarker for the intended purpose of the study (see Tables 1 and 2), significant detail on exposure received, and account being taken in exposed and control groups of known confounders of response. In addition, the size of the exposed and control groups has to be large enough for statistical definition between the groups to be attainable. It is inappropriate to conduct a study and then try and establish if it meets these design requirements.

The general format for interpreting the results is to conduct comparisons of the frequency of a particular biomarker in a putative exposed group with the frequency of that same biomarker in a control group that has been matched for known confounders of response (*e.g.* age, sex, smoking status, diet). Standard statistical tests for group comparisons (*e.g.* Student's t) can be employed for this purpose. It is considerably more informative to have available groups of individuals in the exposed population who received different defined exposures. The ability to demonstrate a dose-response relationship for a biomarker allows for extrapolation to exposure levels below those at which increases in response can be measured using biomarkers.

For biomarkers of exposure such as DNA adducts, micronuclei or sister chromatid exchanges, the conclusions that can be drawn from a well-conducted study are: l) an exposure to a genotoxic agent has occurred; and 2) groups receiving different levels of exposure can be identified. This information can be useful for adverse health outcome follow up and for identifying areas where steps need to be taken to reduce exposure.

For biomarkers of response (*e.g.*, ones that can potentially predict cancer outcomes at a group level) such as chromosome aberrations, specific chromosomal alterations (*e.g.* reciprocal translocations) and cancer-related mutations, the conclusions that can be drawn are related to cancer risk estimation. For example, dose-response curves for informative biomarkers can be used to estimate tumor response curve shape at doses below those at which tumors can be observed. In this regard the biomarker serves as a surrogate for tumors. The use of data in this predictive way is consistent with the U. S. Environmental Protection Agency's Draft Final Guidelines for Carcinogen Risk Assessment (U.S. EPA, 2003) and with the ILSI framework for Human Cancer Risk Assessment (Meek *et al.*, 2003). There has been a certain degree of validation of this approach from the studies of Hagmar and colleagues (Hagmar *et al.*, 1999) and Bonassi and colleagues (Bonassi *et al.*, 1995). These studies demonstrated a relationship between increased chromosome aberration frequency and increased likelihood of cancer. However, the analysis could not demonstrate a relationship between exposure to industrial chemicals and increased chromosome aberration frequency in a particular group.

It should be noted that the use of biomarkers in peripheral lymphocytes for estimating exposures to ionizing radiation is a very well-established method for dose assessment. The ability to produce standard references curves in vitro for the same cell type against which in vivo frequencies can be compared for dose calculation provides a significant advantage (Tucker, 2001; Camparoto *et al.*, 2003).

The general interpretations described here, it needs to be stressed, are those that can be drawn from a well-controlled, well designed study. Unfortunately, the published literature is rife with studies that do not fit these desired characteristics, largely because of size of study groups, inadequate account being taken of known or potential confounders, and insufficient information on exposures (for dose- response calculations). This deficiency has, in fact, led to the misinterpretation, in a number of instances, of the effects of chemicals to humans at environmental exposure levels. Improved experimental design for all such biomonitoring studies is needed (See Bonassi & Au, 2002, for additional information).

# 6. Conclusion

Population monitoring studies using appropriate biomarkers have provided valuable information for the better understanding of the potential health risks from exposure to environmental toxic substances. These include the understanding of exposure dose and conditions that are hazardous to health, mechanisms for induction of disease, inter-individual variations in response to exposure, approaches for better risk assessment procedures and applications to disease reduction programs. As these assays improve in sensitivity and precision, it is possible that the data can be used not only for disease reduction strategies at the population level but at the individual level as well. Consequently, the data may be useful for monitoring of efficacy in disease intervention activities. Therefore, opportunities for conducting biomarker monitoring studies that are relevant and that have health applications are likely to increase quite considerably.

# 7. Acknowledgment and Disclaimer

RJP wishes to thank Drs Andrew Kligerman and Hal Zenick for their helpful comments on this chapter. This document has been reviewed in accordance with the U.S. Environmental Protection Agency policy and approved for publication although it does not necessarily reflect EPA policy.

# 8. References

Albertini R, Clewell H, Himmelstein MW, Morinello, E, Olin S, Preston J, Scarano L, Smith MT, Swenberg J, Tice R, Travis C. The use of non-tumor data in cancer risk assessment: reflections on butadiene, vinyl chloride, and benzene. Regul Tox Pharm 2003;37:105-32.

Au WW. Life style factors and acquired susceptibility to environmental disease. Int J Hyg Env Health 2001;204:17–22.

Au WW, Cajas-Salazar N, Salama AS. Factors contributing to discrepancies in population monitoring studies. Mutat Res 1998;400:467-78.

Au WW. Susceptibility of children to environmental toxic substances. Int J Hyg Environ Health 2002;205:1-3.

Au WW, Sierra-Torres CH, Tyring SK. Acquired and genetic susceptibility to cervical cancer. Mutat Res 2003;544:361-4.

Bonassi S, Au WW. Biomarkers in molecular epidemiology studies for health risk prediction. Mutat Res 2002;511:73-86.

Bonassi S, Abbondandolo A, Camurri L, Dal Prá A, De Ferrari M, Degrassi F, Forni A, Lamberti L, Lando L, Padovani M, Sbrana I, Vecchio D, Puntoni R. Are chromosome aberrations in circulating lymphocytes predicitve of future cancer onset in humans? Prelimiary results of an Italian cohort study. Genet Cytogenet 1995;79:133-5.

Camparoto MZ, Ramalho AT, Natarajan AT, Curado MP, Sakamoto Hojo, ET. Translocation analysis by FISH-painting method for retrospective does reconstruction in individuals exposed to ionizing radiation 10 years after exposure. Mutat Res 2003;530:1-7.

Hagmar L, Brøgger A, Hansteen IL, Heim S, Högstedt B, Knudsen L, Lambert B, Linnainmaa K, Mitelman F, Nordenson I, Reuterwall C, Salomaa S, Skerfving S, Sorsa M. Chromosomal aberrations in lymphocytes predict human cancer: a report from the European Study Group on Cytogenetic Biomarkers and Health (ESCH). Cancer Res 1998;58:4117-21.

Harms C, Salama SA, Sierra-Torres CH, Cajas-Salazar N, Au WW. Polymorphisms in DNA repair genes, chromosome aberrations and lung cancer. Environ Mol Mutag *in press*.

Hulka BS, Wilcosky TC, Griffith JD. Biological Markers in Epidemiology: Oxford University Press, New York, 1990.

Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKeeman LD, Longfellow DG, Pastoor T, Seed J, Patton DE. A framework for human relevance analysis of information on carcinogenic modes of action. Crit Rev Toxicol 2003;33:591-653.

Norppa H. Genetic susceptibility, biomarker response, and cancer. Mutat Res 2003;544:339-48.

Perera FP, Weinstein IB. Molecular epidemiology and carcinogen-DNA adduct detection: new approaches to studies of human cancer causation. J Chronic Dis 1982;35:581-600.

Preston, RJ. Cytogenetic effects of ethylene oxide, with an emphasis on population monitoring. Crit Rev Toxicol 1999;19:263-82.

Salama SA, Serrana M, Au WW.Biomonitoring using accessible human cells for exposure and health risk assessment. Mutat Res 1999;436:99-112.

Tucker JD. FISH cytogenetics and the future of radiation biodosimetry. Radiat Prot Dosimetry 2001;97:55-60.

Data for this paper were retrieved from the U.S. EPA (Environmental Protection Agency). Draft Final Guidelines for Carcinogen Risk Assessment, 2003, EPA/630/P-03/001A. World Wide Web (URL:www.epa.gov/ncea/raf/cancer 2003.htm.) (March, 2006).

Zhang L, Eastmond DA, Smith MT. The nature of chromosomal aberrations detected in humans exposed to benzene. Crit Rev Toxicol 2002;32:1-42.