Incorporating Biomarkers into Cancer Epidemiology: A Matrix of Biomarker and Study Design Categories

Nathaniel Rothman, Walter F. Stewart, and Paul A. Schulte

Introduction

In 1987, the Committee on Biological Markers of the National Research Council/National Academy of Sciences proposed a classification scheme based in part on previous work by Perera and Weinstein (1) that broadly categorized biomarkers into markers of exposure, effect, and susceptibility (2). The emergence of this concept has motivated discussion on the study design implications of the use of biomarkers in epidemiological research (3–18). In this paper, we build on this foundation and provide an organizational framework for using different types of biomarkers in epidemiological studies and public health applications. Biomarkers associated with xenobiotic chemical carcinogens are the primary examples discussed. However, the study design implications of using biomarkers in observational studies are generally applicable to other exposures and diseases.

Biomarker Categories

Fig. 1 portrays the range of biomarker categories that reflect, directly or indirectly, the carcinogenic process due to xenobiotic exposures (2). This well described heuristic schema includes external exposure and biomarkers of internal dose, biologically effective dose, early biological effects, altered structure/function, and disease. Susceptibility, either acquired or inherited, is portrayed as potentially modifying the relationship between each step in the progression from exposure to disease. The schema in Fig. 1 is a powerful, organizing, albeit one-dimensional, framework for understanding the relationship between exposure, biomarkers, and the disease process. It does not, however, describe the progression of studies required to develop and validate a newly developed bioassay in human populations, nor does this schema provide insight into why, when, and how a particular biomarker can be used in a specific type of epidemiological study design. In order to explore this process, we begin with a discussion of each study category.

Study Design Categories

A series of different types of activities is usually carried out to develop and eventually apply a biomarker in etiological studies and in public health practice. These can be categorized as laboratory studies, transitional studies (where the biomarker is the outcome variable), etiological studies (where disease, or a valid surrogate, is the outcome variable), and public health applications (Table 1). Although biomarker research proceeds in an iterative fashion and is not strictly linear as conceptualized by these activities, this schema does serve to highlight the critical steps in the process.

Laboratory Studies

Basic research in biochemistry, molecular biology, and toxicology attempts to elucidate causative steps in carcinogenesis and modulators of that process with the use of in vitro and in vivo models. Many biomarkers that have been applied in human studies were identified initially as components or correlates of the tumorigenic process in animals (e.g., DNA adducts). More recently, molecular changes described in human tissues have become the basis of new assays (e.g., Refs. 19, 20). Increasingly, research efforts are focusing on developing biomarker assays for direct use in human studies.

Transitional Studies

Transitional studies as first defined by Hulka (10, 11) bridge the gap between the development of markers in the laboratory and their application in population-based studies. They optimize sample processing, evaluate assay accuracy and precision,
Matrix of Biomarkers and Study Designs

The first priority in evaluating a marker for use in epidemiological studies is to establish regulatory policy or serve as guidelines for clinical intervention. The following discussion applies primarily to markers of biologically effective dose and early biological effect.

At this stage of research, the biomarker has not been shown to predict an increase in risk of disease in humans. The marker may have some utility, however, if it has been shown to reflect biological changes considered relevant to cancer pathogenesis in animals exposed to human carcinogens (e.g., DNA adducts and cytogenetic damage) or if it has been shown previously to be elevated in human populations exposed to known carcinogens (e.g., DNA adducts and chromosomal aberrations). Under these conditions, biomarkers can be used to provide mechanistic insight into well established exposure-disease relationships, to supplement suggestive but inconclusive evidence of carcinogenicity of a chemical from epidemiological studies of disease, or to provide a preliminary evaluation of compounds introduced recently into the workplace or the general environment.

Applied studies are not capable, however, in and of themselves, of establishing or refuting a causal relationship between a given exposure or a given level of exposure and risk for developing disease. Until a marker is shown to predict disease risk, which can be established only by comparing risk of disease in individuals with and without the marker, results of studies using biomarkers as outcome measures are only suggestive; a biomarker may be overly sensitive (i.e., it may be responding to low levels of chemical exposures that are below the threshold of regulation).

While it is not our intention to introduce new jargon, we have found that dividing transitional studies into three broad categories (i.e., developmental, characterization, and applied studies) clarifies their distinctive research goals. In practice, elements of all three types of studies are often incorporated into a single field investigation.

Developmental Studies. When a promising new biomarker emerges from the laboratory, some very basic issues need to be resolved before considering its application in human studies. The first priority in evaluating a marker for use in epidemiological studies is to determine its reliability. As long as an assay is reliable, the ordering of subjects by the measure is preserved. Since this is all that is required for studying a marker-disease relationship, reliability and not accuracy is of initial importance. Reliability of laboratory assays can be assessed optimally through the analysis of blind replicate human samples representative of the range of values likely to be found in human populations. When, however, etiological studies are used to establish regulatory policy or serve as guidelines for clinical intervention, accuracy also becomes critical (although a "gold standard" is often lacking for new molecular markers). After determining whether assay reliability and accuracy are acceptable, it is then important to define the optimal conditions for collecting, processing, and storing biological specimens for eventual assay since variation in sample handling can sometimes introduce large variation in assay results, making the measures unsuitable for research.

Characterization Studies. Biomarker characterization studies are designed generally to address questions about the presence or levels of a newly developed marker in human populations. In addition, they serve to identify factors that are confounders or effect modifiers of a marker (e.g., age, gender, and medications), which need to be measured and taken into account when using the marker in subsequent studies.

Applied Studies. Applied studies are investigations usually performed on healthy subjects exposed to particular xenobiotics (e.g., occupational exposures, smokers, etc.) where the biomarker is treated as the outcome variable. Most often, cross-sectional or short-term longitudinal designs (which may be experimental in some instances, e.g., controlled dietary modulation) are used. While this type of study has been commonly used in many classes of biomarkers in cancer epidemiology, the following discussion applies primarily to markers of biologically effective dose and early biological effect.

When a promising new biomarker emerges from the laboratory, some very basic issues need to be resolved before considering its application in human studies. The first priority in evaluating a marker for use in epidemiological studies is to determine its reliability. As long as an assay is reliable, the ordering of subjects by the measure is preserved. Since this is all that is required for studying a marker-disease relationship, reliability and not accuracy is of initial importance. Reliability of laboratory assays can be assessed optimally through the analysis of blind replicate human samples representative of the range of values likely to be found in human populations. When, however, etiological studies are used to establish regulatory policy or serve as guidelines for clinical intervention, accuracy also becomes critical (although a "gold standard" is often lacking for new molecular markers). After determining whether assay reliability and accuracy are acceptable, it is then important to define the optimal conditions for collecting, processing, and storing biological specimens for eventual assay since variation in sample handling can sometimes introduce large variation in assay results, making the measures unsuitable for research.

Characterization Studies. Biomarker characterization studies are designed generally to address questions about the presence or levels of a newly developed marker in human populations. In addition, they serve to identify factors that are confounders or effect modifiers of a marker (e.g., age, gender, and medications), which need to be measured and taken into account when using the marker in subsequent studies.

Applied Studies. Applied studies are investigations usually performed on healthy subjects exposed to particular xenobiotics (e.g., occupational exposures, smokers, etc.) where the biomarker is treated as the outcome variable. Most often, cross-sectional or short-term longitudinal designs (which may be experimental in some instances, e.g., controlled dietary modulation) are used. While this type of study has been commonly used in many classes of biomarkers in cancer epidemiology, the following discussion applies primarily to markers of biologically effective dose and early biological effect.

At this stage of research, the biomarker has not been shown to predict an increase in risk of disease in humans. The marker may have some utility, however, if it has been shown to reflect biological changes considered relevant to cancer pathogenesis in animals exposed to human carcinogens (e.g., DNA adducts and cytogenetic damage) or if it has been shown previously to be elevated in human populations exposed to known carcinogens (e.g., DNA adducts and chromosomal aberrations). Under these conditions, biomarkers can be used to provide mechanistic insight into well established exposure-disease relationships, to supplement suggestive but inconclusive evidence of carcinogenicity of a chemical from epidemiological studies of disease, or to provide a preliminary evaluation of compounds introduced recently into the workplace or the general environment.

Applied studies are not capable, however, in and of themselves, of establishing or refuting a causal relationship between a given exposure or a given level of exposure and risk for developing disease. Until a marker is shown to predict disease risk, which can be established only by comparing risk of disease in individuals with and without the marker, results of studies using biomarkers as outcome measures are only suggestive; a biomarker may be overly sensitive (i.e., it may be responding to low levels of chemical exposures that are below the threshold of regulation).

An exception is studies that evaluate the association between biomarkers and radiation and/or antineoplastic therapy in cancer patients.
threshold for disease risk) or insensitive, or reflect phenomena that are irrelevant to the disease process.

**Etiological Studies**

Etiological studies are distinct from transitional studies in that they involve either clinically ill subjects, asymptomatic subjects with early disease, or subjects positive for an intermediate process shown previously to be associated with increased risk of disease (e.g., colon adenomas and risk of colon cancer). These include ecological, case-control, case-case, prospective cohort (with nested variants), family, screening, and intervention studies. We focus our discussion on case-control, case-case, and prospective cohort studies, and then comment on the use of biomarkers as outcome measures in etiological studies.

**Case-Control Studies.** In case-control studies, the prevalence of an exposure or a biomarker is compared in cases and controls. This study design is used far more frequently than the prospective cohort study design because of its relatively greater efficiency and lower costs, and is logistically ideal for the study of rare diseases. As such, maximizing opportunities to integrate biomarkers creatively into case-control studies is important. Defining which biomarker categories can be used most effectively in this study design is critical to this process (12, 14) because some markers may be affected by disease itself, which raises complications of reverse causality.

**Case-Case Studies.** In a case-case study [Ref. 21; also described as a case series study (15) and a case-only study (16)], a series of subjects with the disease of interest are enrolled. Then, tumor characteristics (at the anatomic, cellular, chromosomal, or molecular level) of cases with exposure to a known or suspected carcinogen are compared with tumor characteristics of cases without a history of such exposures or suspected carcinogen. This approach was used in early studies of smoking and lung cancer histology (22) and has come into favor more recently to study morbid biomarker levels in these cases are then compared with morbid biomarker levels in nondiseased control group are more common in the population. A disadvantage of this strategy, however, is that subjects who are biomarker positive (“cases”) cannot be ascertained passively. Population-based screening or screening of selected exposed populations (e.g., specific occupations) is required.

An intermediate marker must be correlated with disease risk in order to be used as a valid outcome measure. The criteria for validating intermediate markers have been discussed extensively by Schatzkin et al. (8) and include the sensitivity of the marker (i.e., the proportion of subjects who develop cancer who are positive for the biomarker), the relative risk of the association between the marker and disease, and a judgement about the extent to which the exposure-disease relationship is mediated through a process reflected directly or indirectly by the marker. Steenland et al. (14) have discussed the logistical challenges of the validation process and have noted that external exposure data, as well as biological samples, need to be collected in epidemiological studies in order to evaluate the comparative ability (and interrelationship) of these measures to predict disease in humans.

Several early prospective cohort studies collected blood samples but banked only the serum. These studies have made substantial contributions to understanding the role of nutrient intake and, to a lesser extent, infections in cancer causation. Unfortunately, large prospective cohort studies initiated in the 1980s and 1990s are, in general, banking most or all fractions of the peripheral blood sample to allow a far wider range of biological assays to be performed, particularly those that require DNA.

One limitation of many large cohort studies is that resources are often available to collect a biological sample at only one point in time. Although this is not a concern for DNA-based assays of inherited susceptibility markers (which remain constant in normal tissue), it poses some limitations for several other categories of markers, particularly for short-term markers of internal dose that reflect exposures which vary from day to day. When transient markers are used to classify individuals, misclassification is inevitable and is almost always nondifferential. As such, type II errors will be more likely as the degree of misclassiﬁcation increases (23, 24).

**Etiological Studies Using Validated Biomarkers as Outcome Measures.** The use of intermediate markers (e.g., biologically effective dose and early biological effect) as outcome measures in epidemiological research is likely to provide substantial research opportunities in the future since it allows risk factors for earlier stages of the disease process to be identiﬁed. From a logistical standpoint, subjects with the predisease markers are more common in the population. A disadvantage of this strategy, however, is that subjects who are biomarker positive (“cases”) cannot be ascertained passively. Population-based screening or screening of selected exposed populations (e.g., specific occupations) is required.

It should be noted, however, that the odds ratio from a case-case design will underestimate the odds ratio derived in a case-control design when the exposure of interest is associated with more than one tumor type. In this instance, the case-case design cannot be used to estimate the relative risk of disease from a specific exposure. A nondiseased control group is needed for this purpose.

**Prospective Cohort Studies.** In prospective studies, biological samples collected from exposed subjects while they are healthy either are analyzed at the time of collection or are banked for later analysis. The cohort is then followed forward in time; subjects who develop disease are identified, and premorbid biomarker levels in these cases are then compared with those levels in the controls. Often, a nested case-control approach is used. Samples from cases and only a sample of the controls are analyzed, which reduces the laboratory requirements and costs considerably. Although this study design is by far the most time-consuming and expensive type of observational epidemiological study, it is the only method available to test the biomarker-associated relative risk of cancer when the marker is transient or may be directly or indirectly affected by disease. An additional advantage of the cohort study is that it can be used to test the association between a marker and risk for multiple cancer sites, as well as noncancer outcomes, thus providing a more global understanding of its public health relevance.

Public Health Applications

Biomarkers can be used at the population level, where research results may be incorporated into policy making decisions, and at the individual level, where biomarkers may be used for screening purposes and ultimately in clinical practice (25).
Although a marker may not be useful for clinical screening, it may still have substantial utility as a research tool. Results from transitional and etiological studies can be used to enhance risk assessment at the population level. For example, the IARC recently concluded that there is only limited evidence in humans for the carcinogenicity of ethylene oxide (26) after considering available epidemiological studies of cancer. Its overall evaluation, however, was that “Ethylene oxide is carcinogenic to humans (Group 1)” (26), which was based in part on the observation that ethylene oxide is associated with a dose-related increase in hemoglobin adducts, chromosomal aberrations and sister chromosome exchanges in peripheral lymphocytes, and micronuclei in bone marrow cells of exposed workers (26).

A biomarker may have utility to screen populations at high risk of disease as part of a primary or secondary prevention effort. However, a substantial amount of information is required before a biomarker can be used for this purpose. In particular, the probability that an individual will develop cancer over a defined period given a constellation of biological and nonbiological risk factors must be estimated, along with a calculation of its uncertainty (27, 28). The enormous effort undertaken over the last few decades to evaluate serum lipid markers as screening tools for atherosclerotic heart disease risk is an instructive example of the extensive body of research that is required before a biomarker can be used effectively for screening and in clinical practice.

Incorporating Biomarkers into Epidemiological Studies
To facilitate the optimal use of biomarkers into epidemiological studies, we have constructed a two-dimensional matrix; study designs are presented along the vertical axis, and biomarker categories are presented along the horizontal axis (Fig. 2). The vertical axis of the matrix can be used to formulate a research program that develops a biomarker, explores its use in human populations, and applies it in cancer etiology studies. The horizontal axis of the matrix can be used during the planning phases of an epidemiological study to evaluate the potential contribution of each biomarker category.

We now describe in detail some of the main applications of biomarkers in human studies of cancer etiology, with a primary focus on transitional and etiological studies. We discuss the potential uses and limitations of each biomarker category in these studies and describe the role that each study design plays in the iterative process of cancer research, where laboratory, transitional, and etiological studies serve to extend, challenge, and ultimately reinforce each other as they reveal patterns of cancer causation. If desired, readers can focus on particular cells of interest using the cell numbers in Fig. 2. The application of biomarkers in genetic epidemiology studies of familial cancers, screening and intervention trials, risk assessment and clinical prevention, and early cancer diagnosis and treatment is outside the scope of this paper.

### Markers of Internal Dose
Biomarker assays of internal dose are measurements of a parent compound or its metabolite in an accessible biological matrix, such as serum or urine. In some instances, they may complement or serve as an alternative to questionnaire and environmental exposure data.

#### Transitional Studies (Fig. 2; Cell 1)
Transitional studies determine whether markers of internal dose are detectable in human populations, evaluate the exposure-marker relationship,
effect modifiers of that relationship (e.g., nutrition and demographic variables), and measure marker kinetics. Both observational (e.g., smokers and occupationally exposed populations) and experimental studies (e.g., dietary manipulation) have been used to evaluate marker characteristics.

There are several applications of internal dose markers in transitional studies. They can determine whether a compound present in a particular environment is absorbed and excreted, which can help to establish the biological plausibility of associations described in etiological studies (e.g., measurement of aflatoxin metabolites in urine of exposed subjects; Ref. 29). They can serve as internal dosimeters to evaluate biomarkers further downstream in the exposure-disease continuum (e.g., adducts and cytogenetic markers). In addition, they can be used to validate other sources of exposure data (e.g., questionnaire and environmental monitoring data), even if the markers themselves are not used in subsequent etiological studies (9).

**Case-Control Studies (Fig. 2; Cell 2).** It is reasonable to assume that in most instances cumulative exposure is the single best predictor of cancer risk. In general, questionnaires are used to obtain exposure data. In some instances, however, a measure of internal dose may be more suitable or may be a valid surrogate of cumulative exposure, even when blood or biopsy material is obtained from cases and controls at the time of the study. The ideal biomarker should persist over time (e.g., fat-soluble substances such as 1,1,1-trichloro-2,2-bis(p-chlorophenyl)ethane (DDT) metabolites, which have been measured in case-control studies of breast cancer (30)) and not be affected by disease status. Many xenobiotic compounds, however, are metabolized to excretable substances that have relatively short half-lives (e.g., aflatoxin metabolites; Ref. 29).

In general, short-term markers have limited use in case-control studies. For example, the measure of aflatoxin metabolites in urine would have little value in a case-control study of breast cancer (30). In this instance, even a short-term internal dose marker might be expected to classify the long-term exposure status of subjects more accurately than questionnaire data. For example, a nested case-control study conducted in Shanghai, China, demonstrated that aflatoxin exposure, assessed by measuring several aflatoxin metabolites (and an N7-guanine aflatoxin adduct) in banked urine samples, was associated with an increased risk of HCC, while aflatoxin exposure assessed by questionnaire was not associated with elevated risk (31).

**Markers of Biologically Effective Dose**

Macromolecular adducts, which are considered markers of biologically effective dose, integrate both external exposure and a spectrum of processes that activate or detoxify procarcinogens and repair DNA damage (1). They can be used to evaluate the ability of genotoxic compounds to form adducts in exposed humans and may have utility as exposure markers in etiological studies.

Theoretically, the association between carcinogen-DNA adduct levels and disease risk should be stronger than the association between external exposure (or internal dose) and disease. Practically, however, obtaining a sample from the target tissue is possible for only a few sites (e.g., hematopoietic and lymphatic systems, bladder, cervix, and oral cavity). While more readily accessible surrogates have been developed (e.g., peripheral WBC DNA adducts), their correlation with DNA adducts measured in relevant tissues should be assessed.

**Transitional Studies (Fig. 2; Cell 5).** Transitional studies determine whether markers of biologically effective dose are detectable in the general or specific target populations, assess marker kinetics, evaluate the dose-response relationship between the exposure and the marker [e.g., PAH exposure and adduct formation (32, 33)], and assess potential effect modifiers of that relationship.

Since DNA adduct formation is thought to play a central role in early stages of tumorigenesis for many genotoxic agents (34, 35) and elevated adduct levels have been found in human populations exposed to several known or suspected human carcinogens (35), it is reasonable to assume that a population with elevated DNA adduct levels may be at higher risk of...
Matrix of Biomarkers and Study Designs

306

developing cancer. To date, however, only one study has actually demonstrated that DNA adducts measured in any biological matrix are associated with excess cancer risk in humans (31, 36). Until there is a greater body of evidence that supports the ability of measured macromolecular adducts to predict excess cancer risk and characterizes that risk quantitatively, the interpretation of studies that use adduct biomarkers as outcome measures will be somewhat uncertain.

**Case-Control Studies** (Fig. 2; Cell 6). There is only a limited role for the use of adduct biomarkers to assess exposure in case-control studies since most available markers have relatively short half-lives and are therefore subject to the same limitations as many internal dose markers. For example, peripheral WBC PAH-DNA adducts probably reflect exposure that has occurred over several weeks to months (Refs. 37, 38; although adducts in longer-lived mononuclear cells may reflect more distant exposure), while hemoglobin adducts reflect exposure that has occurred over the last 120 days, the life span of RBC.

Macromolecular adducts, however, do pose one advantage over markers of internal dose; they may function as a phenotypic measure of cancer susceptibility by evaluating the tendency to activate and detoxify carcinogens and repair DNA for subjects exposed to the compound of concern at the time of biological sampling. For example, two studies of PAH-DNA adducts in peripheral WBC of lung cancer patients and healthy controls have suggested that given similar current smoking patterns, cases have a greater tendency to form adducts than do controls (39, 40). The tendency to form adducts, taking into account current exposure level, could be modeled as an effect modifier of the relationship between lifelong cigarette exposure (assessed by questionnaire) and disease.

**Case-Case Studies** (Fig. 2; Cell 7). Markers of biologically effective dose may complement exposure assessment in case-case studies. They can be measured in blood or urine samples taken at the time of diagnosis or, potentially, in normal tissue adjacent to tumor in pathological materials (e.g., PAH-DNA adducts in lung; Refs. 41, 42). As noted previously, the use of these markers is subject to the constraints of marker kinetics and potential disease influence.

**Cohort Studies** (Fig. 2; Cell 8). Markers of biologically effective dose may be useful in prospective cohort studies since the problem of reverse causality is minimized. When a macromolecular adduct is correlated highly with external exposure [e.g., aflatoxin B₁ exposure and N⁷-guanine aflatoxin adducts in urine (29)], it can be used to address the primary question of whether that exposure is associated with excess cancer risk, directly and/or through interactions with other exposures [e.g., risk of HCC due to both aflatoxin B₁ and hepatitis B exposure (31, 36)]. As noted previously, the utility of an exposure biomarker is based on its ability to classify subjects into long-term exposure categories relative to other available measures of exposure.

Since adduct formation in humans is often variable, adducts may reflect information about individual susceptibility, as well as exposure (1). In this instance, the association between adduct levels and cancer risk then becomes a test of whether this observed variation is relevant. If high quality data are available on external exposure at the time samples were collected since adducts in peripheral blood and urine probably reflect only recent exposure (29, 37, 38).

When reliable data are not available about external exposure at the time the biological sample is collected, the adduct-associated relative risk for cancer can be calculated, but the relevance of interindividual variation in adduct formation cannot be tested directly because the population cannot be stratified on exposure. Further, if this variation reflects processes that are unrelated to cancer risk, the strength of the exposure-disease relationship (assessed via adduct formation) may be weakened unknowingly. If, however, a measure of internal dose is available, then the internal dose-associated relative risk for cancer can be compared to the adduct-associated relative risk (31).

**Markers of Early Biological Effects and Altered Structure/Function**

Markers of EBE and ASF represent processes that are intermediate between exposure and disease. Early biological effect markers reflect cyto genetic damage and somatic cell mutation frequency, while ASF markers reflect precancerous alterations in cell cycle control, measured at the morphological (e.g., hyperproliferation) or at the molecular (e.g., altered oncogene expression) level. While some cytogenetic markers can be studied at the target site (e.g., micronuclei in exfoliated bladder cells and oral/nasal mucosa), most cytogenetic and somatic cell mutation assays are performed on peripheral blood cells; as with macromolecular adducts, the correlation between markers in peripheral blood cells and relevant events occurring at target sites may vary by exposure category, assay type, marker kinetics, and target site.

**Transitional Studies** (Fig. 2; Cells 9, 13). Transitional studies determine whether newly developed EBE or ASF markers are detectable in the general population or in particular subgroups. These studies, performed on subjects with exposures of interest, can potentially determine early biological effects of the exposure and provide insight into possible disease mechanisms. As in vitro, animal, and human evidence accumulates that a given marker reflects carcinogen-mediated damage, there is usually a greater tendency to consider the marker as a surrogate for disease. However, until the marker-associated relative risk for disease is known, the marker cannot be used to establish directly an association between a given exposure and cancer risk.

**Case-Control Studies** (Fig. 2; Cells 10, 14). Markers of EBE and ASF may have use in case-control studies if the markers are not altered by the presence of disease, are relatively stable, and are evaluated on patients who have not yet been treated (particularly by radiation therapy or chemotherapy). Certain persistent markers, such as stable chromosomal translocations, potentially integrate both exposure and biological processes that modify exposure-mediated chromosomal damage. The interpretation of markers used in this setting is enhanced when data are available on external exposure. Marker levels and external exposures can be evaluated independently for their association with excess cancer risk and then combined in a stratified analysis. In this instance, the research goal is to determine whether, within a given level of exposure, subjects with higher marker levels have an increased cancer risk. In general, such investigations should be regarded as hypothesis generating, to be followed up in prospective studies if feasible.

**Case-Case Studies** (Fig. 2; Cells 11, 15). Markers of EBE and ASF that persist over time and are unaffected by disease status can be compared in subgroups of cases defined by exposures...
and/or tumor characteristics under the conditions described above.

**Cohort Studies** (Fig. 2; Cells 12, 16). Prospective cohort studies are the optimal observational design to incorporate EBE and ASF markers (subject to the feasibility of collecting the appropriate tissue) since the problem of reverse causality is minimized. These markers may provide mechanistic insight into the exposure-disease relationship, can be tested for their potential to function as disease surrogates in subsequent studies, and may be shown to have potential as screening tools for primary and secondary prevention.

In the absence of data on external exposure, the goal of these studies is relatively straightforward: to determine whether elevated marker levels are associated with overall and site-specific excess risk of cancer. As noted previously, if data are available on potentially important external exposures (e.g., from questionnaires, medical records, and environmental monitoring), these studies can then explore the more refined question of whether, given the same level of exposure to a particular agent, subjects with higher biomarker levels have an elevated cancer risk (14).

As an example, chromosomal aberrations in peripheral lymphocytes were evaluated recently for their ability to predict excess cancer risk in two small cohort studies that analyzed samples collected from individuals tested at various clinics over a 20-year period (43, 44). The study by Hagmar et al. (43), carried out in four Nordic countries, found an elevated cancer risk associated with the highest tertile of chromosomal aberration frequency for all sites combined, while the study by Bonassi et al. (44), carried out in Italy, found significantly elevated risk for all cancer types and for respiratory tract tumors and lymphatic/hematopoietic tumors. These studies provide support for the hypothesis that chromosomal aberrations measured in peripheral lymphocytes reflect (or are correlated with) relevant, pathogenic, exposures and/or processes and suggest that they may be useful as an outcome measure in subsequent studies. In contrast, Hagmar et al. (43) found that sister chromatid exchange and micronuclei formation were not associated with excess cancer risk (although there was limited power to test the latter marker). As additional years of follow-up are accumulated by this cohort, it will be interesting to follow the relative risk of cancer associated with each marker.

The studies by Hagmar et al. (43) and Bonassi et al. (44) are an important first step in the process of exploring the relevance of EBE biomarkers for disease risk in humans and will undoubtedly be extended by studying candidate markers in ongoing large prospective cohort studies. However, the applicability of particular assays will depend on the method that was used to process and store peripheral blood samples. For example, many currently available cytogenetic and somatic cell mutation assays require cultured lymphocytes. Although most if not all recent prospective studies have stored the major blood fractions, only a few have cryopreserved lymphocytes (which maintains their viability) so that they can be cultured in the future. Fortunately, recently developed PCR-based somatic cell mutation assays (19, 45) can be used in studies that have stored nonviable WBC since these assays require only DNA for analysis.

**Markers of Susceptibility**

There has been increasing interest in evaluating the impact of genetic susceptibility in cancer epidemiology. In this section, we focus on susceptibility genes that are common in the population and are generally considered polymorphisms (i.e., the minor allele frequency is >0.01), are probably associated with relative risks under 10 and as such do not exhibit familial patterns of inheritance, and may interact with particular exposures (46–47). Under these circumstances, which apply to a wide spectrum of potential susceptibility genes, the case-control study (known as an association study in genetic epidemiology parlance) is an effective approach to evaluating these factors (46–49). A discussion of methods used in genetic epidemiology studies of familial cancer, which evaluate gene mutations generally present at low frequencies and associated with high cancer risk, is outside the scope of this paper (for reviews see Refs. 18 and 48).

Markers of susceptibility include polymorphisms in genes responsible for chemical activation or detoxification, DNA repair, and genomic stability, among other processes. Some categories of susceptibility genes may interact with very specific types of chemical exposures (e.g., cytochrome P-450 enzyme subtypes and Phase II-conjugating enzymes), while others may confer more general susceptibility. Markers can be measured at the DNA level (if the genetic basis of a polymorphic phenotype has been identified) or the phenotype level (e.g., drug probes of hepatic enzyme activity, DNA repair measured in peripheral lymphocytes (50), in vitro susceptibility of lymphocytes for drug-induced chromosomal damage (51)).

**Transitional Studies** (Fig. 2; Cell 17). Transitional studies play an important role in the study of genetic susceptibility markers in that they survey the prevalence of particular alleles in specific racial subgroups. Further, they can evaluate the correlation between genotype and phenotype assays (52), estimate the likelihood and impact of allele misclassification (53), and evaluate potential induction effects (54). This information can be used in designing subsequent etiological studies by estimating sample size needs and by identifying exposures (e.g., enzyme inducers such as alcohol and certain drugs) that must be collected via questionnaire.

In addition, transitional studies can potentially evaluate the biological plausibility of gene-environment interactions observed in etiological studies. For example, the *GSTM1* gene encodes the cytosolic enzyme glutathione S-transferase M1, which conjugates activated aflatoxin metabolites. A case-control study in a geographic area with high exposure to aflatoxin demonstrated recently that subjects with the *GSTM1* null genotype (who lack a functional enzyme) were at increased risk of HCC (55). To support the biological plausibility of the case-control results, these investigators showed that the same genotype was associated with higher levels of aflatoxin albumin adducts in healthy subjects currently exposed to aflatoxin (55). Similarly, Vincis et al. (56) have shown that among subjects who smoke cigarettes, the slow (versus rapid) NAT2 acetylation phenotype is associated with a greater tendency to form 4-amino-1-naphthyl hemoglobin adducts, particularly at lower levels of cigarette use. This supports the observation from several case-control studies that slow acetylators in the general population are at increased risk for bladder cancer (57). The underlying assumption of this study design is that surrogates (e.g., protein adducts in peripheral blood) reflect similar activation and detoxification processes occurring at the target site (i.e., DNA adduct formation).

**Case-Control Studies** (Fig. 2; Cell 18). The case-control design is highly efficient for examining the role of genetic susceptibility markers, particularly when high quality exposure data are available from questionnaires or environmental monitoring. Some studies have suggested that there may be subtle
exposure-gene interactions, where particular alleles are associated with greater risk of cancer at lower versus higher levels (58) of exposure to carcinogens, or vice versa (59). These interactions are likely to be missed if adequate attention is not given to exposure assessment. There is a critical need for more valid and reliable methods of obtaining data on dietary, environmental, and occupational exposure to carcinogens by questionnaire (60, 61), given that there will be many opportunities to study exposure-gene interactions in future studies.

The following examples are drawn from the study of polymorphisms in enzymes that oxidize or conjugate chemical carcinogens. When a phenotype is fixed and determined by a known genetic allele, DNA-based assays can be used to evaluate cancer risk associated with a given polymorphism [e.g., GSTMI (62, 63)] since disease status does not affect genomic DNA. The interaction of particular alleles and environmental exposures can then be assessed. Techniques developed recently that use buccal swabs to obtain oral epithelial cells for PCR-based genomic DNA analysis are promising cost-efficient and feasible strategy for case-control studies. They are ideally suited, in particular, to collecting samples on population-based subjects.

When an enzyme is inducible and there is a genetically conferred tendency toward inducibility that has been identified [e.g., CYP1A1 (64)], DNA-based assays can be used to measure susceptibility of a subject for induction. If the inducers of the enzyme are known (e.g., cigarettes), then exposure to these inducers can be estimated by questionnaire and modeled with the presence or absence of the inducibility polymorphism to develop an integrated measure of enzyme activity.

When the genetic basis of a measured phenotype is unknown, a phenotype assay must be used [e.g., CYP1A2 (65)]. Phenotype assays can be applied usually to patients with early, asymptomatic disease. The application of phenotype assays to subjects with advanced disease may, however, have more limitations; disease status may alter assay results directly, and, if the phenotype is inducible, the current pattern of exposure to enzyme inducers may not be representative of the usual exposure pattern.

**Case-Case Studies (Fig. 2; Cell 19).** The use of case-case studies to evaluate genetic markers has been discussed recently by Piegorsch et al. (16). These studies can be used to compare the prevalence of a given allele among cases with different patterns of exposure since a polymorphism that modifies a specific chemical-tumor (or tumor subtype) relationship may be present in a higher proportion of cases exposed to that agent compared to cases without that exposure. Any source of genomic DNA can be used for analysis. Even if the study is limited to the analysis of archived, formalin-fixed, paraffin-embedded tumor samples, the presence of some alleles may be measurable in normal tissue adjacent to the tumor.

Case-case studies assume that the probability of a healthy person being exposed to a particular compound or level of a compound is independent of the probability that they have the polymorphism under study (16). More explicitly, this assumes that genetic predisposition to the disease is not related to greater risk for acute or subchronic toxic reactions to the exposure. While this is a reasonable assumption for the majority of chemical exposures and genetic polymorphisms described to date, it should be tested, if possible, on a sample of healthy subjects who are exposed to various levels of the chemical.

Finally, case-case studies can be an efficient first step to determine whether a particular chemical exposure-genetic interaction is present in a case-control study, particularly with a large number of subjects. The prevalence of a given allele can be evaluated and compared among a subsample of cases with high, middle, and low level exposure. If the frequency of the allele in each group of cases is similar, it may not be worthwhile to analyze DNA from remaining cases and all controls. It may, however, still be useful to study a subsample of controls to determine whether the allele is associated with increased risk of disease independent of exposure.

**Cohort Studies (Fig. 2; Cell 20).** If WBC or blood clots from serum tubes have been stored in prospective cohort studies, DNA can be extracted and analyzed for genetic polymorphisms. Clearly, since genotype assays can be performed in case-control studies, they are not a rationale for performing costly and time-consuming prospective studies. If, however, DNA samples are already available, it may be worthwhile to evaluate relevant genetic polymorphisms, since most analyses require only nanogram quantities of DNA and the interaction between particular genetic polymorphisms and exposures ascertained prospectively can be explored. In addition, a cohort study can test the association between a given polymorphism and risk for multiple diseases.

Genotype assays potentially can be used in retrospective cohort studies if tumor samples from cases have been stored and can be retrieved. In this instance, the distribution of genetic alleles among cases with and without the exposure of interest can be compared. In addition, a control group of living subjects matched to cases on ethnicity, age, and sex could be assembled and used for comparison, assuming that the genetic make-up of the target population had not changed over recent years.

**Markers of Disease**

Markers of disease include tumor characteristics such as tumor histology and the presence and nature of oncogene and tumor suppressor gene mutations. Although these mutations are causally associated with the cancer process itself, their characteristics may have etiological relevance (66).

**Transitional Studies (Fig. 2; Cell 21).** Transitional studies can be used to evaluate the reliability of tumor marker analysis and compare various analytical techniques to each other.

**Case-Control Studies (Fig. 2; Cell 22).** Case-control studies can use tumor markers to separate cancer cases into more homogenous categories and then determine whether particular categories have a stronger association with external exposure. For example, when Taylor et al. (67) compared the occupational histories of leukemia patients with mutations in any ras oncogene and population-based controls, they found an excess risk for having worked in an a priori high-risk occupation, which was not detected when all leukemia patients (both with and without ras mutations) were analyzed together as one group.

**Case-Case Studies (Fig. 2; Cell 23).** In case-case studies, tumor characteristics of cases with exposure to a known or suspected carcinogen are compared with tumor characteristics of cases without a history of such exposures (Ref. 21). For example, in the case-control study of leukemia and occupation described above (67), results similar to the case-control analysis were obtained when the occupational histories of cases with and without ras mutations were compared directly to each other. A similar approach has been used to study the association of

---

8 R. Sinha, N. Rothman, E. Brown, M. Knize, S. Rossi, D. Rhodes, O. Levander, and J. Felton. Heterocyclic amine content in red meat cooked by various techniques to different degrees of doneness, manuscript in preparation.
specific chemical exposures and the presence and pattern of mutations in the \( p53 \) gene [e.g., aflatoxin exposure and codon 249 mutations in HCC (68), and cigarette use and G:C to T:A transversions, consistent with PAH exposure, in non-small cell lung cancer (69)].

**Cohort Studies (Fig. 2; Cell 24).** The association of external exposure and risk among cases with similar tumor markers can be evaluated in prospective and retrospective cohort studies, if tumor samples can be retrieved.

**Discussion**

We believe that presenting biomarker categories and study designs as a two-dimensional matrix provides an organizational framework for reviewing the literature about a particular exposure-disease relationship and evaluating the potential application of biomarkers in new studies. Specifically, the matrix can be used in two ways. By using the vertical study design axis, one can summarize the development and use of a biomarker via laboratory, transitional, and etiological studies and public health applications. For example, Table 2 depicts this process for a bioassay that measures the \( N7 \)-guanine aflatoxin DNA adduct assay, a marker of biologically effective dose [reviewed by Groopman et al. (70)].

Alternatively, when designing an epidemiological study, the horizontal biomarker axis can be used to systematically consider whether biomarkers from each category can (or should) be incorporated into a study and to compare the utility of a given marker in different types of studies. For example, Fig. 3 evaluates the use of biomarkers from each category in a case-control study and a prospective cohort study evaluating the association between cigarette smoking and lung cancer. In the case-control study, exposure assessment relies almost exclusively on questionnaire data. Inherited susceptibility markers and tumor markers have the greatest application in this study, while other markers have a more limited role. In contrast, the cohort study (using a nested design for efficiency) can evaluate the association of each marker category with disease, as well as the relationship between sequential markers. Markers reflecting later stages in lung tumorigenesis can be evaluated in samples from subjects who develop cancer within a relatively short period of time after enrollment (i.e., several years), markers reflecting earlier stages in carcinogenesis can be studied in

---

**Table 2 Application of aflatoxin (N\(^7\)-guanine) DNA adduct assay, a marker of biologically effective dose**

<table>
<thead>
<tr>
<th>Study design</th>
<th>Study objective*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laboratory studies</td>
<td>Develop assay for ( N7 )-guanine adduct in urine</td>
</tr>
<tr>
<td>Transitional studies</td>
<td>Optimize processing urine samples for long-term storage</td>
</tr>
<tr>
<td>Developmental</td>
<td>Detect adducts in Chinese and East African populations</td>
</tr>
<tr>
<td>Characterization</td>
<td>Correlate adducts with dietary exposure</td>
</tr>
<tr>
<td>Applied</td>
<td>Evaluate aflatoxin exposure and risk of hepatocellular cancer</td>
</tr>
<tr>
<td>Etiological studies</td>
<td>Determine if oltipraz therapy alters adduct levels</td>
</tr>
<tr>
<td>Risk assessment</td>
<td>Potential to screen exposed populations and estimate disease risk at the group level</td>
</tr>
<tr>
<td>Clinical prevention</td>
<td>Potential to evaluate and modify risk among individuals</td>
</tr>
</tbody>
</table>

*Reviewed in Groopman et al. (70).
samples from subjects who develop cancer in subsequent years, and DNA-based markers of inherited genetic susceptibility can be evaluated in all cases diagnosed since they do not vary over time in healthy tissue. The efficiency of these various goals will, however, depend on the size and age structure of the cohort.

Conclusions
We have found this matrix valuable for reviewing the literature and for designing our own studies. In addition, after several years of use, we have found it to be an effective instructional tool in the classroom. We hope that it will help others consider why, when, and how to use biomarkers in their own research.

Acknowledgments
We thank Drs. Douglas Bell, Neil Caporaso, John Groopman, Richard Hayes, Robert Hoover, Miriam Foirer, John Potter, Mark Schiffman, Maryn Smith, Paul Strickland, Jack Taylor, and Sholom Wacholder for helpful discussions.

References


Incorporating biomarkers into cancer epidemiology: a matrix of biomarker and study design categories.

N Rothman, W F Stewart and P A Schulte


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/4/4/301

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.