APPLICATION OF BIOLOGICAL MONITORING METHODS FOR CHEMICAL EXPOSURES IN OCCUPATIONAL HEALTH

D. Gayle DeBord, Ph.D., Dale Shoemaker, Ph.D.,
Clayton B'Hymer, Ph.D., and John Snawder, Ph.D. DABT

Introduction

Exposure assessment is a critical component of occupational health studies. Biomonitoring can be an essential tool to determine exposures and/or health effects resulting from those exposures. According to a National Research Council (NRC) report [NRC 2006], biomonitoring can provide an efficient means to measure exposures and, when used with other information derived from toxicological, epidemiologic or modeling studies, can estimate the absorption into the body (dose), the influence of an exposure to an individual and potential health risks. Large population studies that incorporate biomonitoring efforts can identify chemicals found in the environment (including the occupational environment) and monitor trends and the distribution of exposures in the general population [CDC 2014].

Environmental measurements involving various media (such as air, water, soil, food or surfaces) can be used to detect the presence of hazards and sometimes predict individual exposure; for example, air monitoring predicts effects when the lung is the target organ. However, the only way to evaluate which biological, physical or chemical agent has been absorbed into the body, the magnitude of the absorbed dose, and its contribution to total body burden, is through biomonitoring or measurement of biomarkers. The spectrum of exposures for which biomarkers can be useful includes the full range of occupational hazards from noise to stress to chemicals. While this document generally focuses on chemical exposures, the considerations of purposes, study design criteria, quality assurance, and the ethical and safety issues have general application. The guidance provided in this document is particularly relevant to urinary and blood biomonitoring and is generally relevant to other biomonitoring assays that measure response or susceptibility factors such as genetic toxicology tests or gene variants.

Applications of Biomonitoring

The benefits of biomonitoring (Table 1) range from the ability to measure actual body burden to the role of such measurements in providing informed risk communication. Biomonitoring can

enhance exposure assessment and provide information about health outcomes, depending upon which biomarker(s) is (are) selected.

Table 1. Benefits of Biomonitoring

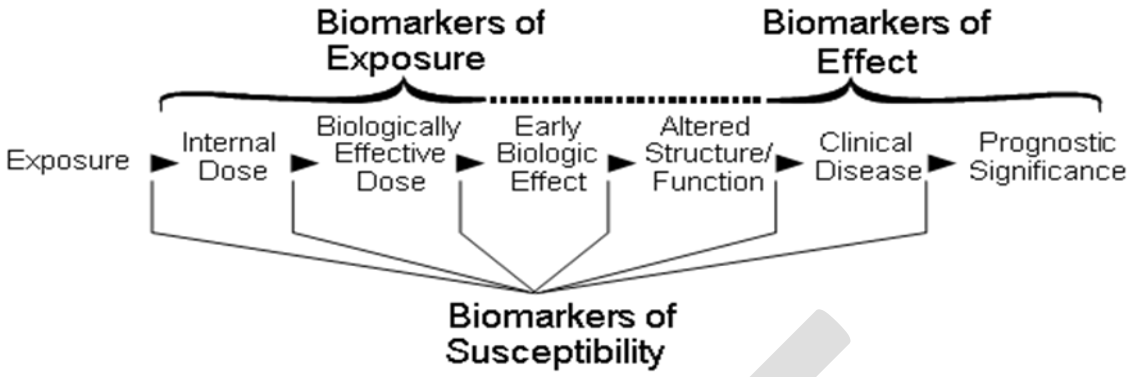
- Measures actual body burden
- Augments other exposure monitoring tools
- Captures all exposure routes, including dermal
- May detect unexpected exposures or unexpected routes of exposure
- Evaluates the effectiveness of control measures, including PPE
- May provide biomarkers of potential health risks
- Can be used to reconstruct exposures following acute or accidental events
- Enhances individual or group risk assessments
- Provides valuable information regarding risk communication including reassurance in cases where exposure is insignificant

[Adapted from Decker et al. 2013]

Categories of Biomarkers

Biomarkers have been defined by the National Academy of Sciences as measurable indicators in a biological system or organism, such as the presence of a chemical or its metabolite within biological specimens, measured alterations in structure or function or identifiable genetic variations (Figure 1) [NRC 1987]. Three categories of biomarkers have been identified: exposure, effect or response and susceptibility (Table 2) [NRC 1987]. While these categories can overlap, generally biomarkers of exposure provide a measure of body burden of a chemical or its metabolite. Susceptibility biomarkers estimate the impact that factors of the individual can have on exposure, uptake metabolism and/or repair.

Figure 1. Continuum from Exposure to Disease



[Adapted from NRC 1987 and Schulte and Perera 1993]

Biomarkers of Exposure

Occupational exposure routes are mainly by inhalation or through the skin [NIOSH 1994]. Route of exposure may affect the level of internal dose and the biologically effective dose. Biomarkers aggregate measurements of all exposure pathways. While this is often an advantage, such measurements cannot by themselves be used to determine the route(s) of exposure. Examples of biomarkers of exposure include blood lead [Sexton et al. 2004], urinary cadmium [Menke et al. 2009], and serum dioxin [Manh et al. 2014]. Biomarkers of exposure may or may not be reflective of future disease or health risk, but are indicative that an individual or population has been exposed. The level of a biomarker of exposure can change with exposure concentration or duration, route of exposure, the pharmacokinetics of the chemical that influences the distribution, metabolism and elimination, timing of sample collection, physiological variations in hydration status, and urinary flow [Aylward et al. 2012, 2014]. Biomarker level may also be modified by individual characteristics such as gene variants, exertion or even other exposures.

Table 2. The Three Types of Biomarkers

Type of Biomarker	Characteristics	Example
Exposure	Measurement that reflects absorption of a chemical into the body	Urine or blood concentration of agent
Effect	Measurable biochemical, physiological, or other alteration that can be	DNA mutation or cytogenetic change

	recognized that may or may not be a health impairment [WHO 2001]	
Susceptibility	Inherent or acquired sensitivities or resistance in response to specific exposures	Genetic polymorphisms in metabolic activation/deactivation enzymes

[Adapted from DeBord et al. 2015]

Biomarkers of Effect

Biomarkers of effect (also known as biomarkers of response) measure effects or responses in the body to an exposure. These changes may be early precursors of disease, specific clinical changes or markers for preclinical effects, but might not be related to any specific health outcome. Biomarkers of effect may result in a measurable change in a pathway, macromolecule (DNA, RNA, protein), or other internal substances. These changes may be reversible or permanent. An example of a biomarker of effect is the inhibition of blood acetyl cholinesterase (AChE) levels in the red blood cells (RBCs) of persons exposed to organophosphate (OP) pesticides [Richards et al. 1978]. Biomonitoring results of blood AChE levels can be used to determine exposure, potential health risks, and treatment implications [Richards et al. 1978]. A continuum exists across the biomarker spectrum. At low levels, a biomarker may be indicative of exposure or an effect but as the exposure continues or increases the level of the biomarkers will also increase leading to a response or action level that needs to be corrected before health is permanently impacted. While biomarkers of exposure tend to be specific for a particular chemical or agent, biomarkers of effect are less specific and may show the effect of multiple chemical or agent exposure. For example, measuring blood AChE levels provides evidence of exposure to OP pesticides, but does not identify the precise pesticide responsible for the decrease in blood enzyme. Similarly, carboxyhemoglobin levels increase in response to both carbon monoxide and dichloroethylene, which is metabolized internally to carbon monoxide.

Biomarkers of Susceptibility

Biomarkers of susceptibility indicate whether an individual may be at increased or decreased risk for developing a disease after an exposure has occurred. Susceptibility biomarkers may also identify individuals whose body burden may be increased or decreased relative to other individuals because of differences in absorption, metabolism, or other biological processes. Biomarkers of susceptibility could include other exposures from the environment, current health status, or genetic traits such as the activity of metabolizing enzymes. For example, the activity of CYP2E1, a cytochrome P450 enzyme involved in metabolism, can modify benzene

toxicity [McHale et al. 2012] and exposure to ethanol reduces methyl ethyl ketone metabolism [Liira et al. 1990].

Drivers for Biomonitoring in Occupational Health Studies

As described below, several drivers for utilization of biomarkers in occupational safety and health have been identified. These drivers include compliance, research, evaluation of intervention effectiveness, and risk assessment/management processes, which provide multiple benefits to workers.

Compliance

The Occupational Safety and Health Administration (OSHA) has three regulations that require biomonitoring in certain situations (Table 3). For benzene, 29 CFR 1910.1028 [OSHA 1980] requires urinary monitoring of phenol if an employee has been exposed to benzene in an unplanned release. The employee provides a urine specimen at the end of the work shift and another one 72 hours later. If the result of the 72-hour test is below 75 mg phenol/L urine, then no further testing is required. If the measurement of the 72-hour test is greater than 75 mg phenol/L urine, then additional medical surveillance is instituted at monthly intervals for 3 months.

Table 3. Chemicals for Which OSHA Requires Biomonitoring

Chemical	Biomarker (s)
Benzene	Phenol in urine
Cadmium	Cadmium in urine Beta-2 microglobulin in urine Cadmium in blood
Lead	Lead in blood Zinc protoporphyrin

[OSHA 1978, 1980, 1981]

Several biomonitoring tests are required by OSHA for employees exposed to cadmium at or above the action level for 30 or more days per year (or in a 12-month consecutive period) [OSHA 1981]. The action level is defined as an airborne concentration of cadmium of 2.5 $\mu g/m^3$ as calculated as an 8-hour time weighted average. Biomonitoring tests include: 1) cadmium in

urine (CdU), standardized to grams of creatinine (g/Cr), 2) beta-2 microglobulin in urine (β2-M), standardized to grams of creatinine (g/Cr), with pH specified and 3) cadmium in blood (CdB), standardized to liters of whole blood (lwb). β2-M is an effect modifier. Various additional actions are required depending upon the results and may include further monitoring, worksite assessment or even medical removal of the employee [OSHA 1981].

The third chemical for which OSHA requires biomonitoring is lead. OSHA 29 CFR 1910.1025 requires biomonitoring for all employees who are or may be exposed at or above 50 µg/m^3 in workplace air averaged over an 8-hour workday for workers in general industry. Monitoring of blood lead and zinc protoporphyrin levels is required every 6 months. The requirement changes to every two months if the blood lead level was at or above 40 µg/100 g of whole blood or to monthly if the blood lead levels were high enough that an employee was medically removed from his or her position [OSHA 1978].

The National Institute for Occupational Safety and Health (NIOSH) has no recommended exposure limits (RELs) based on biomonitoring. Other organizations have developed biomonitoring guidelines. The American Conference of Governmental Industrial Hygienists (ACGIH) publishes a list of biological exposure indices (BEIs®) for approximately 30 chemical agents [ACGIH 2015]. The German Deutsche Forschungsgemeinschaft (DFG), Health and Safety Executive in the UK, France's ANSES and the Swiss Accident Insurance Fund (Suva) also routinely publish lists of biological exposure limits [ANSES 2016; DFG 2006; HSE 2017; Suva 2013].

Occupational Health Research and Surveillance

A major aim of biomarker research is to develop and validate biomarkers that reflect specific exposures or are quantitatively linked to adverse outcomes in humans to enable their use in risk prediction. Surveillance and research efforts can aid in hazard identification or monitoring of exposure trends over time. Biomonitoring can provide information on what has been absorbed or taken into the body and when used in conjunction with environmental exposure monitoring, may indicate whether environmental monitoring alone is sufficient if the purpose of biomonitoring is detecting exposures only. If a substance has a sufficiently long half-life in the body, biomonitoring can be used to estimate cumulative dose after repeated exposures and can help characterize the contribution from multiple exposure routes (e.g., inhalation and dermal). Sampling of environmental media focuses on a single route, while biomonitoring can be useful for assessing all routes of exposure.

Biomarker levels are affected by a number of different factors including the specimen matrix such as blood or urine, timing of the collection of the specimen, and the pharmacokinetic principles of absorption, distribution, metabolism and excretion (ADME) [Aylward et al. 2014]. Individual differences in physiological parameters such as age and weight may affect the distribution of a chemical just as genetic polymorphisms of metabolism enzymes can alter the concentration and pattern of metabolites. Differences in physical activity may affect the amount of a chemical that is absorbed. Clearance or excretion can be impacted by kidney function.

Biomonitoring may have special utility in assessing dermal exposure because sampling methods for skin may not be readily available for most chemicals and few, if any, reference standards are available [Ashley et al. 2011]. However, since biomarkers measure all routes of exposure, the amount attributable to skin absorption alone might not be determined [Decker et al. 2013] unless inhalation exposure is specifically precluded, say, by use of self-contained breathing apparatus (SCBA). Biomonitoring may also be useful in assessing the biological effects from breach of or improper use of personal protective equipment (PPE) and may allow for comparing exposures/doses associated with different work practices.

A few surveillance programs perform biomonitoring assays. One example is the Adult Blood Lead Epidemiology and Surveillance (ABLES) program established by NIOSH in 1987 to monitor occupational lead exposure [NIOSH 2014]. National blood lead level data are published in CDC's Morbidity and Mortality Weekly Report (MMWR) and elsewhere. CDC's National Biomonitoring Program is within the National Center for Environmental Health. The program periodically assesses a percentage of the U.S. population for exposure to a wide variety of chemicals. Four reports and a series of updated tables have been issued, and among its goals is the provision of U.S. population-based reference ranges for these chemicals. [CDC 2014, 2017].

NIOSH often uses biomonitoring in its occupational health research. In a study to assess war fighter exposures to JP-8 jet fuel at U.S. Air Force bases, a number of biomarkers of exposure were investigated [B'Hymer et al. 2012]. Three different biomarkers of exposure, S-benzylmercapturic acid (BMA), S-phenylmercapturic acid (PMA), and (2-methoxyethoxy)acetic acid (MEAA), were measured in urine to determine which biomarker accurately assessed exposures. Statistical analysis of the measured analytes demonstrated MEAA to be the most accurate or appropriate biomarker for JP-8 exposure using urinary concentrations, whether adjusted for creatinine or not. Correction of biomarker measurements is discussed in more detail later in this chapter.

Evaluation of Intervention Effectiveness

Measuring biomarker levels can be an effective means of evaluating interventions. Evaluating some interventions by measuring air levels may provide an indication of how well the control is working. However, for other interventions such as those used to control dermal exposure, it may not be readily apparent how effective those controls are without carrying out biomonitoring.

A study of paving workers exposed to polycyclic aromatic compounds (PACs) while working with hot-mix asphalt (HMA) used biomonitoring to evaluate the effectiveness of various interventions [McClean et al. 2012]. Air and hand-wash samples in that study were analyzed for the parent compounds corresponding to the urinary PAC metabolites. Interventions assessed included dermal protection, a powered air-purifying respirator (PAPR), biodiesel substitution (100% biodiesel provided to replace the diesel oil normally used by workers to clean tools and equipment), and reduced temperature of the HMA. Biodiesel substitution, dermal protection and PAPR use all decreased urinary PAC analytes from baseline data. Higher HMA temperatures were positively associated with urinary measures. Biodiesel substitution and lower HMA temperatures resulted in decreased PACs in air and on skin [Cavallari et al. 2012a,b].

Risk Assessment and Risk Management

Establishing a dose of concern and mode of action are two primary components of risk assessment [DeBord et al. 2015]. Environmental exposures can directly or indirectly cause alterations in a number of biological pathways that can be measured. For example, development of gene expression profiles helps identify genes, pathways or networks that are specific to the toxic end-point of interest [Brown and Botstein 1999]. Toxicogenomics biomarker data have also been used to demonstrate benchmark dose estimates [Thomas et al. 2007, 2011].

Determination of internal dose is important in risk assessment and provides highly relevant information that is more closely associated with disease response than external exposure estimates [Aylward and Hays 2011]. The capability of biomarkers to generate information that can be used for internal dose estimation and response markers will be important in their use in risk assessment.

Few studies have been performed that incorporate biomarkers of susceptibility into risk assessments. The advantage of such data has been discussed [Demchuk et al. 2007; Lohmueller et al. 2003; Scinicariello et al. 2010]. Thomas et al. [2007] has used toxicogenomic data to

better understand the mode of action and has shown the promise of such information in risk assessments.

General Considerations for Biomonitoring

Factors that Influence Biomonitoring

A worker exposed to an agent receives a dose of that agent only if it is absorbed into the body [NIOSH 1994]. Absorption can occur after dermal contact, inhalation, ingestion, or from a combination of those routes. The extent of absorption from an exposure and the rate of absorption depend on the properties of the chemical such as solubility in lipids and water and the route of exposure. Once absorbed, a chemical is distributed and partitions into various tissues due to tissue differences for such factors as perfusion, lipophilicity, pH, and permeability. Highly water-soluble chemicals may be distributed throughout the total body water, while more lipophilic substances may concentrate in the body fat or other lipid rich tissues, such as the brain.

The loss of a chemical compound from the body can loosely be defined as elimination, which depends on metabolism and excretion. Chemical compounds or their metabolites may be eliminated by numerous routes, including fecal excretion, urinary excretion, exhalation, perspiration, and lactation. A chemical compound can be excreted from the body without metabolism, in which case the parent compounds may be detectable in the urine, breath, or fecal material. In other cases, the chemical agent may be metabolized through oxidation, reduction, hydrolysis, or a combination of these processes, often followed by conjugation with an endogenous substrate. Conjugation of a chemical or metabolite is often a pathway for excretion. Conjugation reactions include glucuronidation, amino acid conjugation, acetylation, sulfate conjugation, and methylation [NIOSH 1994].

Metabolism, excretion, and the rates thereof, can be affected by age, diet, general health status, race, and other factors. In general, the metabolic products will be more water soluble than the parent chemical compounds. Where metabolism yields more than one product, the relative amounts of each and the parent-metabolite ratios are affected by an individual's general health status, diet, genetic makeup, degree of hydration, time after exposure, and other factors. The kidney is the major organ of excretion for some chemicals and is the primary route for water-soluble substances. These substances enter the urine by either glomerular filtration, tubular secretion, or sometimes both mechanisms [NIOSH 1994]. Excretion through the feces is an important route for many lipid soluble compounds, although urinary levels of these materials are also present.

Selection of an appropriate biomarker for an exposure requires sufficient knowledge of the distribution, metabolism, and excretion of the toxicant for selection of the proper compound to be determined, biological medium to be sampled, and time for obtaining a specimen. Often, most of the toxicological and pharmacological information available is from experimental animals and, thus, may not be directly applicable to humans. While addressing biomonitoring in emergency response, Decker et al. [2013] provided a list of considerations for conducting biomonitoring. The first consideration is the appropriateness of conducting biomonitoring, which includes whether biomonitoring could add information that is actionable and interpretable. The second consideration is the feasibility of conducting biomonitoring with respect to logistics and available methodology.

Monitoring Goals

Air monitoring (or workplace environmental monitoring) and biomonitoring have complementary goals and frequently are applied simultaneously in industrial hygiene investigations [NIOSH 1994]. Compared with biomonitoring, air monitoring offers advantages in certain situations. If the agent has acute toxic effects on the respiratory tract or the eyes, air monitoring is the logical tool for measuring the exposure [Hathaway and Proctor 2004]. For some chemicals, air monitoring can be conducted continuously, and thus, can detect peak exposures to potentially dangerous chemicals, especially in acute, rapidly evolving situations. Compared to air monitoring, biological monitoring offers a better estimate of the health risk in situations where routes of exposure other than inhalation are significant.

Toxicokinetics, which is the rate a chemical enters the body and the fate of the agent once in the body, will have a key role in the timing of the collection. The elimination half-life, or the time it takes for half of the agent to be eliminated, determines how long an agent or metabolite remains in the system. For example, PbB remains in the blood with a half-life of about a month while many organic solvents have a half-life of less than an hour. The amount of a chemical that accumulates in the body is dependent on its elimination half-life. Levels of rapidly disappearing biomarkers primarily reflect exposures during the previous several hours. On the other hand, biomarkers that disappear over the course of several weeks may reflect one, several, or numerous exposure incidents occurring anytime during a period of several weeks prior to the measurement. Some toxicants accumulate in one or several parts of the body and are in dynamic equilibrium with the sites of toxicity. In the case of polychlorinated biphenyls (PCBs), which accumulate in fatty tissue, the blood level of PCBs reflects the amount stored in the body [Brown and Lawton 1984; Zong et al. 2015]. When the site of critical action for a toxicant is known, the concentration of the biomarker at that site can be used as a measure of the biologically effective dose.

Biological Matrices

The most common matrices used for biomonitoring are exhaled air, blood, and urine [Landi and Caparosa 1997; NIOSH 1994]. Other matrices such as hair, nails, and feces also have been used. Monitoring exhaled air is limited to volatile chemicals. Exhaled air monitoring is not suitable for chemicals inhaled as aerosols, for gases and vapors that decompose upon contact with body fluids or tissues, or for chemicals that are highly insoluble in water [Fiserova-Bergerova et al. 1989].

The matrix to be sampled is a critical decision and will depend upon a number of elements including the timing of the sample and the chemical that is being measured. Monitoring for persistent chemicals in blood allows for the comparison with exposed populations with national reference ranges obtained by efforts such as NHANES and ability to follow elimination of the chemical compound over a long period of time [NRC 2006]. Most biomarkers present in the body can be found in the blood during some period of time after exposure [Needham et al. 2005]. A chemical in the blood is in dynamic equilibrium with various parts of the body: the site of entry, tissues in which the chemical is stored, and organs in which it is metabolized or from which it is excreted. Thus, the concentration of a biomarker in the blood may differ between regions of the circulatory system. This would be the case during pulmonary uptake or elimination of a solvent, which would cause differences in concentration between capillary blood (mainly arterial blood) and venous blood. Some considerations of monitoring chemicals in the blood are listed below.

1) The gross composition of blood is relatively constant between individuals and except for certain highly lipophilic compounds, it is not necessary to correct for volume or hydration differences such as you do for urine.
2) Obtaining specimens is straightforward and with proper care can be accomplished with relatively little risk of contamination. However, an important consideration is that obtaining blood specimens requires an invasive procedure and should be performed only by trained personnel [Taylor et al. 2004].
3) A hydrophobic chemical will persist in the body which results in a longer half-life of the agent in the body and may allow for accurate measurement of that agent years later [Calafat et al. 2015; NRC 2006]. The downside is that it may be difficult to determine the timing of exposure. Hydrophilic chemicals may be rapidly cleared from blood so sample collection may need to occur within minutes to hours of exposures [NRC 2006].

The pharmacokinetics of the chemical is also a primary concern [Calafat et al. 2015]. The concentration of the biomarker in urine usually is correlated to its mean plasma level during the period the urine dwells in the bladder [Rosenberg et al. 1989]. In some instances, the urine concentration is affected by the amount of the biomarker stored in the kidneys, as with metals such as cadmium and chromium. The accuracy of the exposure estimate, using urine monitoring, depends upon the sampling strategy. The most influential factors are time of collection relative to the time of exposure, urine output, and specific biomarker characteristics. Measurements from 24-hour specimens are more representative than from spot samples and usually correlate better with intensity of exposure. However, collection, stabilization, and transportation of 24-hour specimens in the field are difficult and often not feasible.

Practical Considerations

Study Design

Attention needs to be paid to scientific rigor for utilizing biomarkers in occupational health studies. When developing biomarker methods, in most cases no gold standards exist for comparison or evaluation of results [NRC 2006]. Biomonitoring studies may be able to determine the extent of exposure, internal dose and, depending upon the biomarker, may be able to demonstrate risk from exposure. Because the level of most biomarkers tend to be measured in microgram or lower quantities, they may be subject to large variability due to confounders such as diet, lifestyle habits, and other environmental exposures. While environmental monitoring provides information on what is in the area, biomonitoring is individualized to a specific person, which may generate concern among participants. Health-related concerns come from measurements that may be outside of the normal or reference range should a reference range be known. Some individuals may have concern over any exposure especially if the results are not easily interpretable to what it means for their current or future health.

Population Selection

The selection of the population for biomonitoring is typically straightforward in occupational settings. Generally, an important concept to consider is inclusion of both male and female workers of all age ranges at the worksite. However, the population should be representative as a whole with respect to such parameters as age, sex, and risk classes [NRC 2006]. Confounders such as other exposures that might impact the level of the biomarker of interest should be assessed. Even when there is scientific rigor in selection of the population, a large degree of uncertainty will still exist because of random variations. The smaller the group, the greater the uncertainty due to inter-individual differences and laboratory variations [NRC 2006].

Statistical design is an important aspect of any study but is particularly important in biomonitoring studies in order to ensure that any effects seen are true and not the result of confounding variables [Gosho et al. 2012; NRC 2006]. One of the limitations of population studies is the lack of statistical power associated with many biomarker measurements either because of the small population size or because of the small change in biomarker levels [Gosho et al. 2012; Hunter 1997]. Description of the basic characteristics of each variable is needed. These include number of observations, mean, standard deviation, and confidence limits; all of these should help reveal which data are below the limits of detection, missing, miscoded, or outliers.

Selection of Biomonitoring Methods

The occupational health professional and the laboratory scientist should decide on appropriate methods so that the test results are interpretable and relevant to the exposure situation. Some methods already exist. For example, there are currently 50 BEIs that cover exposure to several hundred compounds (because of non-specificity) [ACGIH 2015]. Each of these have a documentation which includes many of the issues (e.g., sampling time, interferences) discussed below for the development of a new biomarker. However, if no determinant currently exists, a new one can be developed and validated. The goal of the biomonitoring should be consistent with the goal of the industrial hygiene investigation. Is the goal to measure exposure or a health effect related to the exposure or both? The method needs to be evaluated at least for the required sensitivity, specificity, biological relevance, and feasibility (Table 4) [NRC 2006]. Interferences from diet, drugs, alcohol, disease states, or other workplace chemicals or agents should be taken into account when interpreting the data. The method should also have a sufficiently low limit of detection to differentiate exposed from unexposed workers; and any effects due to sample matrix need to be assessed. In general, blood, serum, and urine specimens require different sample preparations and may require separate methodologies to eliminate matrix effects. Because of sample instability, some methods may be impractical or not feasible. The method should include guidelines for interpretation of the data. To minimize the risk of harm to workers, when two biomonitoring methods will provide the same information, the less invasive method should be used. For example, methods monitoring urine or exhaled breath are preferred over those monitoring blood if similar information can be obtained.

Table 4. Criteria for Biomarker Method Selection

- Sensitivity

- Specificity
- Biological Relevance
- Feasibility
- Pharmacokinetics

Sampling Strategy

Attention to specimen handling and collection is essential for obtaining quality data [NIOSH 1994]. The analytical laboratory should be consulted for standard operating procedures for specimen collection and storage instructions [NRC 2006]. Analytical methods should provide specific directions on the collection, storage, and transportation of specimens to the laboratory. Adherence to these directions is of the utmost importance to ensure sample integrity (Table 5). The method should include instructions for the timing of specimen collection, that is, whether specimens should be obtained during the work shift, at the end of the shift, or at some other time during the work week. The longer the half-life of the chemical agent, the less critical is the timing of the collection [Lauwerys and Hoet 1993]. Some toxicants will be normally present or have a background level; others will accumulate in the body over time such as a work week due to exposure and a biological half-life of greater than 12 hours. Therefore, it may be necessary to take a baseline measurement of the biomarker concentration prior to the start of the work week or work day to determine if the biomarker concentration increased over the time frame of investigation [Lauwerys and Hoet 1993]. Care should be taken not to contaminate the specimen with either chemicals or microbes, although the former is less a problem when metabolites are measured. The proper preservative (for urine or blood specimens) or anticoagulant for blood should be used, if appropriate. Stability of the biomarker is assured through proper storage and shipment of the specimen to the laboratory and proper storage by the laboratory.

Table 5. Considerations for Sampling

- Timing of specimen collection
- Biomarker baseline
- Contamination potential
- Proper preservatives
- Stability of the biomarker
- Logistics

Correction of Urinalysis Data for Dilution

The determination of biomarkers in individual urine samples can be confounded by urine dilution (or concentration if a worker is dehydrated), which can cause substantial variation. This variation can be attributed to fluid intake, temperature of the work environment, and the level of physical workload by the individual. In addition, the nature of a specific biomarker's excretion can be based on the mechanism of excretion which can be altered if the urine is either very concentrated or very dilute [Rosenburg et al. 1989]. The ranges of acceptability for creatinine and specific gravity are not well documented. However, the common practice to compensate for urine dilution is by adjusting the measured concentration to a normal value [Boeniger et al. 1993; Carrieri et al. 2001]. The most common types of normalization methods are based on creatinine levels, specific gravity of the urine specimen, and urine output. These methods have been described in detail [NIOSH 1994]. Urine osmolality has been suggested as a method of urine normalization [Yeh et al. 2015].

Creatinine adjustment is the most frequently used method to compensate for urine dilution. Creatinine is excreted by glomerular filtration at a relatively constant rate of 1.0-1.6 g/day. Urinary creatinine concentration can be determined by spectrometric or kinetic methods based on the Jaffe alkaline picrate reaction, enzymatic methods and methods based on mass spectrometry and liquid chromatography [Spenser 1986]. The adjusted value is expressed as the quantity of biomarker per unit quantity of creatinine. Adjustment for creatinine concentration, while correcting for dilution, introduces additional variation, which must be considered when the data are evaluated. Among the factors affecting the rate of creatinine excretion are the muscularity of the subject, physical activity, urine flow, time of day, diet, pregnancy, and disease [Barr et al. 2005; Boeniger et al. 1993]. Adjustment to the creatinine level is not appropriate for some compounds. For example, methanol is excreted from the kidney primarily by tubular secretion, a mechanism independent of creatinine's excretion by glomerular filtration. It has been reported that the mechanism of excretion of a biomarker can be altered if the urine is very concentrated or very dilute. Measurements of samples having creatinine concentration outside the range of 0.5 to 3 g/L are unreliable [Rosenberg et al. 1989]. Sauve et al. [2015] recently reported significant bias when using creatinine to normalize urinary biomarker levels.

Specific gravity (sp.g.) is used to adjust urine concentration by multiplying the measured concentration of the biomarker by the ratio of [(1.02-1)/(sp.g. -1)], where sp.g. is the specific gravity of the urine specimen and 1.02 is the assumed normal specific gravity value. In a study of the biomarker 1, 6-hexamethylene diamine (Gaines et al. 2010), it was determined that sp.g. adjustment was superior to creatinine adjustment. The mechanism of biomarker excretion can be altered if the urine is very concentrated or very dilute; and measurements of samples having

sp.g. outside the range of 1.010 to 1.030 have been reported as being unreliable (Rosenberg et al. 1989).

Urine osmolality is similar to urinary creatinine [Yeh et al. 2015]. The effect of sociodemographic factors such as age, gender, race, body mass index, chronic kidney disease status, daily total protein intake, plain water intake, and blood osmolality is not as great for urine osmolality as it is for creatinine.

Finally, urinary output has been used for a normalization method for biomarker levels. The measured concentration of the biomarker is multiplied by the ratio R/0.05, where R is the output for the sample in liters per hour (L/h). The urine output for the sample is computed from the volume (liters) of the sample and the time (hours) elapsed since the last voiding. The adjustment is to a mean output of 0.05 L/h. This method also has its possible limits and uncertainties from subject variation and hydration [NIOSH 1994].

Quality Assurance

Good quality data require the utilization of an effective quality assurance program. In 1992, regulations implementing the Clinical Laboratory Improvement Amendments (CLIA) of 1988 were published by the Health Care Finance Administration and the Public Health Service to ensure that analysis of human specimens was done accurately and under good quality control procedures (Table 6) [Laboratory requirements, 1992]. Any analysis of human specimens that can be used by a health care practitioner to assess the health of the individual or used in the diagnosis, prevention, or treatment of disease or impairment falls under the CLIA requirements.

Standard operating procedures should be developed for collecting, shipping and processing of biological specimens [NRC 2006]. In addition, the stability of the biomarker and range of its concentration in the population should be determined. Field blanks, spikes and duplicates should all be part of the quality assurance program. A field blank is a collection tube that can be empty or filled with water in some cases and run alongside specimens to determine if the collection tube is a source of contamination. Spikes, whether they are done in the field or in the analytical laboratory, are usually performed by putting a known amount of agent or its metabolite into the collection tube. Duplicates are created by splitting a specimen into two collection tubes. Tracking of specimens in the field after collection, shipping and receipt at the laboratory is necessary to ensure integrity of the samples. Data needed include date and time of collection, specimen number and description of the specimen (type of specimen, volume, etc.), shipping information (receipts, tracking numbers and inventory) [NRC 2006]. If shipping samples, DOT and IATA shipping regulations apply, IATA/ICAO Dangerous Goods by Air,

including DOT 49 CFR Part 172.704 – 172.804 and IATA/ICAO Shipping Infectious Substances and Diagnostic Specimens (GPO 2015; IATA 2015).

Table 6. Considerations for Quality Assurance:
Key Components of the CLIA Quality Assurance Program

- Strict management of specimen collection, handling, storage, and transportation, thus ensuring sample integrity
- Thorough verification of a method by the laboratory before use on field specimens
- High level of analytical quality control
- Participation in proficiency testing programs, if available
- Documented instrument evaluation and maintenance programs
- Investigation of communication failures and complaints
- Documentation of performance and corrective measures

[NIOSH 1994]

Ethics

Ethical considerations are paramount when conducting biomonitoring. An Institutional Review Board (IRB) should review research studies before beginning the collection of specimens [NRC 2006]. Among other considerations, the method should be appropriate for the requirements of the investigation [NIOSH 1994], and the procedures should cause minimal harm to participants. Risk of using invasive methods must be outweighed by the benefits, informed consent from the participant is required, and the results should be kept confidential to the extent allowed by law [Records kept on individuals, 1974]. Should an employer initiate a biomonitoring program, the criteria described in this chapter should be considered (Table 7).

Table 7. Criteria for Implementing a Biomonitoring Program in the Workplace

- Specification of goals of the program
- Selection of validated biomarkers
- Establishment of clinical utility
- Acceptance by the population being studied (informed consent received)
- Establishment of link to exposure or disease

- Protection of privacy and confidentiality
- Notification of results to participant
- Process for addressing results and outliers

[Adapted from Schulte and DeBord 2000]

Laboratory and Field Safety

When dealing with human specimens, a biosafety program is essential [NIOSH 1994]. Pathogens such as hepatitis B and human immunodeficiency virus (HIV) may be present in blood, saliva, semen, and other body fluids. Pathogens can be transmitted by an accidental nick with a sharp object; exposure through open cuts, skin abrasions, and even dermatitis or acne; and indirectly through contact with a contaminated environmental surface. There are five major ways to reduce the potential for exposure to biological pathogens [CDC 1993].

1) Engineering controls that include mechanical or physical systems can be used to eliminate biological hazards. Examples include biosafety cabinets and self-sheathing needles.
2) Good housekeeping procedures, which involve cleanup of the work area, are essential to avoid contamination.
3) Employee work practices are essential to minimize exposure to pathogens, including good personal hygiene procedures and avoidance of needle recapping.
4) Personal protective equipment, such as gloves and masks, should be used.
5) Employees who have been identified as potential exposure candidates should be vaccinated for hepatitis B.

Employers are required by OSHA to implement an exposure control plan [OSHA 2001]. Standard precautions should be practiced with every biological specimen collected or received into the laboratory. It is not possible to know if a particular specimen contains pathogens; therefore, each sample should be treated as if contaminated.

Field personnel can be at risk for biological pathogens as well. They should follow standard precautions, such as using gloves during the collection of specimens and their processing for shipment. Field staff should be vaccinated for hepatitis B if they will be collecting blood. In addition, field staff should know the procedures in their blood borne pathogen policy in the event that they are exposed to blood.

Interpretation of Results

Measurement of a given biomarker at some level needs to be appropriately interpreted. The first consideration is whether the biomarker is above background levels. If above, this would

indicate that an exposure was occurring. The second consideration is does the measured level represent a risk. In industrial hygiene practice, increased biomarker concentrations can help identify the source of the exposure and determine the effectiveness of controls implemented to mitigate the exposure.

Two main approaches can be used to interpret biomonitoring results: descriptive and risk-based [NRC 2006]. The descriptive approach uses a statistical review of the data to establish ranges and comparisons of the individual results to subgroups within the cohort or to a control or reference group. Characterization of the half-life of the chemical agent to determine how long the exposure may have been going on, or whether this is a recent exposure, may also be determined [Aylward et al. 2012]. Risk-based approaches are more intensive [NRC 2006]. One approach to risk-based approaches can be used when exposure-response relationships have been demonstrated so that a measured exposure is known to cause an effect. One example would be increased acetylcholinesterase levels after exposure to OP pesticides. Additional approaches to risk-based interpretation include the use of forward or reverse dosimetry to interpret the measured biomarker concentrations. Reverse dosimetry can be used in conjunction with knowledge about the timing of exposure and pharmacokinetic information to estimate an external exposure resulting in the measured biomarker concentration, and this estimated exposure can be compared to exposure guidance values such as OELs, reference doses, or other risk-based guidance values (Clewell et al. 2008). Forward dosimetry estimates the concentration of a biomarker consistent with exposure at an exposure guidance value. This concentration can be used as a screening value to evaluation measured biomarker concentrations and interpret the exposure as well below, near, or above the risk-based exposure guidance value. This approach has been used extensively in both occupational and environmental biomonitoring contexts (Angerer et al. 2011; Lauwerys and Hoet 1993; Zidek et al. 2017).

Reference Levels

For proper interpretation of biomarker levels, biomarker measurements should be compared to biological action levels, if available. In the absence of published biomonitoring action levels, biomarker measurement levels indicating occupational exposure have been inferred by comparison with the normal background levels of the biomarker. Biomonitoring action levels vary in their derivation, some being from correlations with exposure, others with health effects. These should be used only when one has full understanding of their derivation.

Sources of biomonitoring action levels include:

a) Biological Exposure Indices (BEI®) adopted by the American Conference of Governmental Industrial Hygienists (ACGIH) [ACGIH 2015]
b) Biological Tolerance Values for Working Materials (BAT) published by the Deutsche Forschungsgemeinschaft's (DFG) Commission for the Investigation of Health Hazards of Chemical Compounds in the Work Area [DFG 2006]
c) Swiss Accident Insurance Fund (Suva) [Suva 2013]
d) Summary of Recommendations in Industrial Chemical Exposure. Guidelines for Biological Monitoring [Lauwerys and Hoet 1993]
e) OSHA standards [OSHA 1978, 1980, 1981]
f) Japan Society for Occupational Health. Recommendation of occupational exposure limits (2016–2017) [JSOH 2016]
g) Health and Safety Executive, UK. Biological monitoring guidance values [HSE 2017]
h) European Scientific Committee on Occupational Exposure Limits (SCOEL) [2014]
i) ANSES. Biological limit values for chemicals used in the workplace. [ANSES 2016]

Population-based sources for reference values:

a) German Human Biomonitoring Commission (HBM) values [Angerer et al. 2011; Schulz et al. 2011]
b) Centers for Disease Control and Prevention. National Report on Human Exposure to Environmental Chemicals [CDC 2017]
c) Biomonitoring Equivalents [Aylward et al. 2013; Exley et al. 2015; St-Amand et al. 2014; Zidek et al. 2017]

A limited but growing number of chemicals have values that can be used to interpret the measured biomarker concentrations in the context of health risk or external exposure guidance values. For those that do not, other approaches for interpretation can be used. In occupational health studies when biomarker data are available for exposed and unexposed populations that are otherwise similar, the upper limit of the range for the non-exposed population may serve as a reference level [NIOSH 1994]. When confronted with work-related exposures, levels of a biomarker significantly above that limit may suggest occupational exposure. For those biomarkers for which no measurable background level in unexposed humans has been demonstrated, the reference level is effectively the detection limit of the analytical method. In any case, levels of the biomarker above the reference level suggest there may have been an occupational exposure, but give no information on the potential health effect.

Variability

Biological monitoring data are subject to a number of sources of variability [Aylward et al. 2014; NIOSH 1994; Droz 1989] that need to be considered when interpreting results. Sources of variability [Aylward et al. 2014; NIOSH 1994] include:

1) Rates at which an agent is taken up by the body, metabolized, and excreted. These vary from person to person and are affected by the person's age, sex, and physical workload.

2) Route of exposure. For example, absorption through the lungs is much faster than absorption through the skin. Thus, the appearance and elimination of a biomarker will be slower if the agent enters through the skin. If the biomarker is rapidly excreted, the optimum timing for collection of biological samples will be different for the two routes of entry.

3) Fluctuation in environmental exposure. Such fluctuations will be tracked by the levels of rapidly eliminated biomarkers, those reflecting exposure of the immediately previous several hours.

4) Personal protective equipment worn and a person's work practices.

5) Existence of a biomarker in both a free and a conjugated form, the relative proportions of which can vary substantially from person to person. For example, aniline is present in urine as both the free amine and as acetanilide, its acetyl derivative [Greenberg and Lester 1947]. Some persons are genetically predisposed to excrete primarily free aniline; while others, primarily, acetanilide. In another example, the variability of metabolites of a number of solvents was evaluated [Valcke and Haddad 2015].

6) Concurrent exposure to several agents that compete for the same biotransformation sites in the body. This may lead to altered metabolism and excretion, which would change the relationship between exposure or health effect and the level of the biomarker [Ogata et al. 1993].

7) Concurrent exposure to several agents that are metabolized to the same biomarker. This frustrates the interpretation of the biological monitoring data. For example, trichloroacetic acid is a biomarker for trichloroethylene, 1,1,1-trichloroethane, and perchloroethylene.

8) Consumption of alcoholic beverages [Fiserova-Bergerova et al. 1993], since ethanol is metabolized by three pathways used for metabolism of other organic agents. After consumption of one drink, the ethanol concentration in the blood is about 1000 times higher than from a normal occupational exposure and may significantly affect the metabolism of industrial chemicals.

9) Medications [Rosenberg, 1994], health, and diet.

10) Smoking and other lifestyle factors.

Because of the variability of biomarkers, judgments on the exposure or health risk of workers frequently cannot be made based on a single determination [NIOSH 1994]. It may be necessary to conduct repeated samples.

Biomonitoring Analytical Method Validation

It is imperative that the bioanalytical methods used are well characterized, fully validated, and documented to a satisfactory standard in order to yield reliable results. Biomarker methods must also be validated for use in the field [Hunter et al. 2010; Schulte and DeBord 2000]. Field validation is not covered in this chapter. This document draws heavily from two primary sources: guidelines published by the US Food and Drug Administration (FDA) [FDA 2001; Viswanathan et al. 2007], which have been a biomonitoring standard since its publication in 2001; and a more recent set of guidelines from the European Medicines Agency [European Medicines Agency 2011]. Several other validation guidelines and research papers were also consulted and distilled into this final set of recommendations [Agilent Technologies 2010; Bader et al. 2010; Causon 1997; Green 1996; Hartmann et al. 1998; International Conference on Harmonisation 2005; Irish National Accreditation Board 2012; Linder and Wainer 1998; Mikkelsen and Cortón 2004; Peters and Maurer 2015; Theodorsson 2012; United Nations Office on Drugs and Crime 2009; Wieling et al. 1996]. These references demonstrate that while there is general agreement for some aspects of bioanalytical method validation, there are still many differences of opinions on terminology employed, on the number of parameters that should be evaluated, and on how best to evaluate many of those parameters. It should be noted that the driving force behind the majority of the literature guidelines has been the need to analyze for pharmaceuticals in biological matrices. Therefore, some adaptations of these guidelines have been taken into consideration for usage in occupational exposure assessment.

It is acknowledged that acceptance criteria that are wider than those defined in this guideline may be appropriate in special situations. Acceptance criteria should always be defined ahead of time and should be appropriate for the intended use of the method.

This section is not intended to:

1) Serve as a decision guide on whether biomonitoring is appropriate. Readers are referred to a decision matrix proposed by Decker et al. [2013].
2) Serve as a method development guide…method development steps should be completed before method validation occurs.

3) Serve to interpret any biomonitoring data.
4) Serve as a decision guide for the selection of the appropriate analytical technique and methodology to be employed.

While this section of the chapter addresses laboratory method validation, validation of the biomarker needs to be accomplished for the population in question. This validation process utilizes two steps, clinical validity and clinical utility [NIOSH 2010]. Clinical validity evaluates how well the test performs in the field such as the number of false positives or negatives. On the other hand, clinical utility evaluates how useful the information is.

This guidance generally applies to bioanalytical procedures such as gas chromatography (GC) and high-performance liquid chromatography (LC) regardless of the type of detector utilized (e.g., ultraviolet, electron capture, flame ionization, atomic emission, mass spectrometric, etc.). These guidelines are applicable for quantitative determination of analytes in biological matrices such as blood, serum, plasma, urine, saliva, etc. The recommendations can be adjusted or modified depending on the specific type of analytical method used and matrix involved. Because of some of the unique and inherent characteristics of ligand-binding assays or immunoassays, this set of guidelines may not be appropriate. The FDA guidelines do discuss immunoassays more comprehensively [FDA 2001]. While the general validation principles can and should be applied, the specific guidelines and limits described herein may not be applicable. For more in-depth discussion of the reasons for this and adjusted recommendations, consult the guidelines from the FDA and the European Medicines Agency [European Medicines Agency 2011; FDA 2001; Viswanathan et al. 2007].

Validation Parameters

Method validation should include determination of the following parameters: selectivity, carryover, calibration curve, limits of detection and quantitation (LOD and LOQ), accuracy, precision, stability, matrix effects, recovery, and robustness (Table 8). Each of these parameters will be defined and discussed in more detail. These parameters should be reported when validating the method.

Selectivity

Selectivity is the ability of the method to measure unequivocally (quantify) and to differentiate the analyte(s) in the presence of other components which may be expected to be present in the sample. These might include metabolites, parent compound, impurities, degradation products, matrix components, etc. Sometimes the term specificity is used interchangeably with selectivity, although specificity has the implication that a method produces a response for a

single analyte. Since that is rarely the case, the term selectivity will be used here. Multianalyte methods should, of course, be able to differentiate all of the analytes of interest from one another and from the matrix. Likewise, if an internal standard (or multiple internal standards) is employed, the method must be able to differentiate them as well.

Table 8. Biomonitoring Analytical Method Validation Parameters	
Selectivity	Recovery
Carryover	Robustness
Calibration Curve	Limit of Detection/Limit of Quantitation
Accuracy	Acceptance Criteria for Analytical Run
Precision and Bias	Acceptance Criteria for Quality Control Samples
Stability	Acceptance Criteria for Calibration Standards
Matrix Effects	

Selectivity should be investigated in ten (or more) individual sources of the appropriate blank matrix. The use of fewer sources is acceptable in case of rare or difficult to obtain matrices. While interferences should ideally be kept to a minimum, it is recognized that some biomarkers are produced endogenously while others are present due to background exposures. These types of methods can still be invaluable in differentiating between exposed and non-exposed individuals and should therefore not be dismissed because of measurable background levels, especially if those levels are known and consistent.

Investigation of suspected interferences may be carried out by spiking these compounds at realistic concentrations into blank matrices or matrices spiked with low levels of the analyte(s) of interest. It is recognized that suspected interferences may be difficult to obtain or predict. It is common to disregard an interference if it is less than 20% of the LOQ for the analyte and 5% for any internal standard [EMA 2011].

Carryover

Carryover should be addressed and minimized during method development. Carryover should be assessed during validation by injecting one or more blank samples after a high concentration sample or a high calibration standard. If carryover is unavoidable, specific measures should be provided in the method to prevent or correct for carryover so that it does not affect accuracy and precision of the samples.

This could include the injection of matrix blanks or solvents blanks after certain samples. Randomization of samples should be avoided in this case as it may interfere with the assessment of carryover. While "there is no standard acceptable magnitude of carryover for a passing analytical run" [Viswanathan et al. 2007], a good measure would be to use the same guideline as that of the blank and that the carryover in the blank sample following the high concentration standard should be less than 20% of the LOQ for the analyte and 5% for any internal standard [EMA 2011].

Calibration Curve

The calibration or standard curve is the relationship between the instrument response with regard to known concentrations of the analyte. Each analyte studied in the method should have a separate calibration curve. The calibration standards should be prepared in the same matrix as that of the intended samples by spiking the blank matrix with known concentrations of the analyte. A sufficient number of standards should be used to adequately define the relationship between concentration and response. This number will be a function of the anticipated range of values and the nature of the analyte/response relationship. The simplest possible model that adequately describes the concentration-response relationship should be used. A calibration curve should consist of a blank sample (a matrix sample processed without analyte and without internal standard), a zero sample (a matrix sample processed without analyte and with internal standard), and at least six non-zero standards. The blank and zero samples are not used when calculating the calibration curve parameters.

More calibration standards may be required to adequately describe a higher-function (non-linear) model. Selection of weighting and use of complex regression equations should be justified. The calibrators may be injected using single or replicate samples.

Calibration curve parameters (slope, intercept, and correlation coefficient) should be reported, but are themselves not sufficient to determine the validity of the curve. The back calculated concentrations of the calibration standards should be within ±15% of the nominal value, except for the LOQ for which it should be within ±20%. At least 75% of the calibration standards should meet this criterion. Excluding the standards should not change the calibration model used. A minimum of three valid calibration curves should be reported during the validation study. The range of the calibration curve determines the range of the method (more on this in the next section.)

Limit of Detection and Limit of Quantitation

The limit of detection (LOD); also referred to as the method detection limit (MDL) is the lowest concentration of an analyte that the bioanalytical procedure can reliably differentiate from background noise [FDA 2001]. The LOD should not be used for quantitative measurement but may be useful for semi-quantitative or qualitative determinations [Tiwari and Tiwari 2010]. Several literature methods for determining the LOD can be utilized [Bader et al. 2010; NIOSH 1995]. In brief, this approach analyzes five or more low-level standards (prepared in the appropriate matrix) that range from less than the expected LOD to no greater than 10 times the expected LOD. These standard responses are plotted and the regression equation and standard error (s_y) are calculated. The LOD is then 3 s_y/slope. Alternatively, analyze at least 10 standards near the anticipated LOD and calculate the standard deviation [ASTM 2013]. Report LOD (MDL) as 3 times the standard deviation of the blank signal (correcting for background). LLOQ is concomitantly taken as 10 times the standard deviation of the blanks.

The lower limit of quantitation (LLOQ) is the lowest concentration of analyte in a sample that can be quantified reliably with acceptable accuracy and precision. By definition, the calibration standards should include the lower and upper limits of quantitation (L/U LOQ) as values should not be extrapolated beyond the range of the calibration curve. The accuracy and precision criteria at each end of the curve have been found to be acceptable. In addition, the analyte signal of the LLOQ should be at least ten times s_y/slope (or 3.33×LOD).

The LOD, LLOQ, and ULOQ should all be reported in the validation report. They should also be reported when samples are analyzed. When reporting biomonitoring results, report results below the LOD as "Not Detected (ND)" and report results between the LOD and LLOQ numerically and enclosed in parentheses to denote the greater imprecision of these results.

Accuracy

Accuracy is the closeness of the determined value obtained by the method to the true or accepted value of the analyte [FDA 2001]. The term trueness is sometimes used for this value. Accuracy is a function of bias and precision [NIOSH 1995] and should be assessed on samples spiked with known amounts of the analytes. It is recommended that these samples should be spiked from a separately prepared stock solution of the analyte than that used to make up the calibration standards.

Accuracy should be measured using a minimum of three samples per level at a minimum of three levels. The low level should be within 3 times the LLOQ, the medium level near the center of the calibration curve, and the high level within 75% of the ULOQ [NIOSH 1995]. The mean concentration should be within 15% of the nominal value for the upper two concentration

levels and within 20% of the nominal value for the low level [FDA 2001]. More levels, more replicates, or both are recommended during method validation.

Accuracy should be estimated between runs by analyzing at least three concentration levels from at least three runs analyzed on at least two different days. The mean concentration should be within 15% of the nominal value for the upper two concentration levels and within 20% of the nominal value for the low level [FDA 2001].

Precision

Precision is the closeness of agreement of a series of individual measures of multiple aliquots of a single, homogeneous sample. Precision is usually expressed as the coefficient of variation (CV) or the relative standard deviation (RSD) which are different terms for the same thing. Generally, the same data and runs used to determine accuracy are used to determine precision. Precision is often divided into three separate, but important parts:

1) A short-term precision sometimes called repeatability or within-run precision. Repeatability expresses the precision under the same operating conditions over a short interval, usually a single run or a 24-hour window.
2) Intermediate precision, sometimes called between-run, between-day, or inter-assay precision, measures variables within a single laboratory. These variables could include different days, analysts, equipment, etc. It is not considered necessary to study these effects individually, and the extent to which intermediate precision should be established is dictated by the intended use of the method.
3) Long-term precision, or reproducibility, expresses the precision between labs (collaborative study, round-robin, inter-laboratory trial). Once again, the extent to which reproducibility should be established is dictated by the intended use of the method. One has to use caution as many authors use the term reproducibility for within-laboratory studies at the level of intermediate precision, which can lead to confusion [ASTM 2014].

Biological monitoring methods should always be evaluated for repeatability and for selected areas of intermediate precision. The precision between days is nearly always an important parameter. Checking the precision while changing lots or manufacturers of solvent, different chromatography columns or extraction cartridges, for example, is also encouraged while changing analysts or instrumentation is often less important and sometimes impractical or impossible. For methods that are published in the NIOSH Manual of Analytical Methods (NMAM), the process independent laboratory analysis aims to measure reproducibility by

having at least one independent laboratory analyze a series of blind samples. Consensus standards often require testing by multiple laboratories (e.g., ASTM) for full validation.

Precision should be measured using a minimum of five samples per level at a minimum of three levels. As previously mentioned, these samples and levels will usually be the same to measure accuracy. The within-run precision (repeatability, as measured by CV or RSD) should be within 15% for the upper two concentration levels and within 20% for the low level. More levels, more replicates, or both are acceptable. Intermediate precision (between days, etc.) follows the same guidelines, a precision of no greater than 15% RSD at the upper levels and no greater than 20% at the lower level.

Stability

An evaluation of stability should be carried out to ensure that no step taken during preparation, analysis, and storage of samples affects the concentration of the analyte in the matrix. Stability in a biological fluid is a function of the storage conditions, the chemical properties of the analyte(s), the matrix, and the container system. The stability of an analyte in a particular matrix and container system is relevant only to that condition and should not be extrapolated to other matrices and container systems. Conditions used in stability experiments should be the same or similar to those used for the actual study samples. Such conditions might include sample matrix, anticoagulants, preservatives, container materials, storage conditions, and analytical conditions.

Stability determinations should use a set of samples from a freshly prepared stock solution of the analyte spiked into the appropriate matrix. Stability should be evaluated using low (within three times the LLOQ) and high (near the ULOQ) concentrations with at least three replicates for each level under each stability condition evaluated. One set of samples is analyzed immediately after preparation and the other set after the applied storage condition. The stability samples are analyzed against a calibration curve from freshly spiked calibration standards and the obtained mean concentrations are compared to the nominal concentrations. The mean concentration at each level should be within 20% of the nominal concentration.

Biological monitoring methods should be evaluated for sample stability. For NMAM methods, the specific stability test will be indicated. The following stability tests are recommended:

a. Short-term stability. The stability samples are kept at room temperature (or sample processing temperature if different) for a minimum of 4 to 24 hours and then analyzed. The length of time should be equal to or exceed the expected time that the samples will be maintained at that temperature during the study or course of the analysis method. In the

case of chromatographic analysis, a cooled autosampler may be required to improve sample stability. Stability should be verified at room temperature and in cooled conditions such as found with an autosampler. Light sensitivity should also be evaluated, and it should be determined whether the samples require tinted glassware or sample vials to extend stability. In the case of chromatographic analysis, which may require long run times, stability should be evaluated to a minimum of 72 hours (or appropriate if the solutions would normally be left standing for longer periods) should a system failure require re-injection of samples stored in an autosampler.

b. Long-term stability. The stability samples are kept stored (most likely in a freezer but it could be in a refrigerator in some circumstances) for at least the same duration as the study samples and preferably longer. Data points may be collected at intermediate times through use of either replicate aliquots or from re-analysis of the same samples provided sufficient volume is originally supplied. The samples should be stored under the same conditions as the field/study samples, so if different temperatures will be used (some samples at -20 ^{o}C and some at -70 ^{o}C, for example), then studies at both temperatures should be performed.

c. Freeze and thaw stability. The stability samples are stored and frozen at the intended storage temperature for 24 hours and then thawed at room or processing temperature. When the samples are completely thawed, they should be refrozen under the same conditions for 12 to 24 hours. A minimum of three freeze-thaw cycles should be performed but more are required if the field/study samples are expected to go through a higher number of freeze-thaw cycles. Analysis of the stability samples after each freeze-thaw cycle is not required nor advised. Analysis at the completion of three (or more) cycles is sufficient to show stability unless determining the maximum number of cycles over which there is analyte(s) stability is desired.

d. Stock solution and working solutions stability. It is not necessary to study the stability of every concentration of working solutions. The stability of the analyte(s) stock solution(s), at least one concentration of the analyte(s) working solutions and the internal standard(s) solution(s) should be studied. The solutions should be evaluated at room temperature for at least six hours (or appropriate if the solutions would normally be stored or left standing out for longer periods). If the stock solutions are refrigerated or frozen during storage or usage, the stability should be documented by comparison with freshly prepared solutions. Aprotic solvents such as acetonitrile may have to be considered as a major component for

a stock standard solution and lower temperatures (refrigeration) utilized for long-term storage.

e. Post-preparative stability. The stability of the sample after processing should also be examined. If the samples are stored after processing (either in a dry extract or in the injection phase), then stability samples should be treated the same, looking at both the analyte(s) and the internal standard(s). Stability of the processed sample should also be assessed on-instrument or in the autosampler at those times and temperatures, taking into account the anticipated run time for the batch size.

f. Reinjection stability. The reinjection stability and reproducibility should be evaluated to determine if all or parts of an analytical run could be reanalyzed in case of instrument failure. This parameter only pertains to particular techniques like chromatographic or spectrometric analysis.

Matrix Effects

Matrix effects should be investigated and are especially important for methods that employ mass-spectrometric (MS) detection. The matrix effect in MS is typically due to the suppression or enhancement of ionization of analytes by the presence of matrix components in the biological specimens [Smeraglia et al. 2002; Trufelli et al. 2011]. For each analyte and internal standard, the matrix factor (MF) should be calculated using at least six lots of blank matrices from individual donors. Pooled matrices should not be used. If the matrix is difficult to obtain, less than six different lots may be used, but this should be justified, documented, and matrix effects still investigated.

The MF is calculated by taking the ratio of the peak analyte response (usually peak area for chromatographic and spectrometric analysis) in the presence of matrix (measured by analyzing blank matrix spiked with analyte(s) after extraction) to the peak response in the absence of matrix (a pure solution of the analyte(s)). An MF of 1 signifies no matrix effect. An MS value of less than 1 suggests ionization suppression while a value of greater than 1 may be due to ionization enhancement. For a method utilizing internal standards, the MF of the internal standard(s) (IS(s)) is also calculated; then the IS-normalized MF is calculated by dividing the MF of the analyte by the MF of the internal standard.

An MF or IS-normalized MF of 1 is not necessary for a reliable bioanalytical assay. However, highly variable MF in individual subjects would introduce variability and would be one of the causes of a lack of reproducibility in the analysis. The MF or IS-normalized MF should be

examined at two concentration levels: low (within three times the LLOQ) and high (near the ULOQ) for each of the six individual lots of the matrix. The variability in matrix factors (as measured by CV or RSD) should be less than 15% for both of the concentration levels. Higher variability may suggest the need for an internal standard or the need to change to a different internal standard. For example, in MS, it is recommended to use an isotopically-labeled IS that better adequately mimics the analyte and thereby reduces the variability [Trufelli et al. 2011].

Recovery

The recovery is the measured amount of a spiked quantity of the analyte divided by the theoretical value for that analyte in the sample analyzed; this ratio is a measure of the recovery for a quality control sample. The determination of recovery of a bioanalytical method is done by adding a known amount of the analyte (or internal standard) to the matrix and analyzing by the method. The method response obtained is compared to the theoretical concentration of pure authentic standard, expressed in percent. Absolute recoveries can be difficult to obtain for methods that include a derivatization step, as the derivatives are often not available as reference substances.

While it is often desirable to obtain recovery as close to 100% as possible in order to maximize the accuracy of a method, acceptable precision and bias can often be obtained from methods with moderate recoveries. [Dolan 2015; FDA 2001; UNODC 2009]. Extraction recovery is generally an issue investigated during the analytical method development. While the absolute value for recovery is less important and not required to meet any certain threshold, the recovery should be consistent, precise, and reproducible. Recovery values can be calculated from the same data and runs used in the accuracy and precision parts of the validation. This involves analysis of five replicates at three concentration levels. The recovery should be reproducible (better than 20% as measured by CV or RSD) for each of the concentration levels.

Robustness/Ruggedness

Ruggedness or robustness is a measure of the susceptibility of a method to small changes that might occur during routine analysis. Validation of a method must not necessarily include ruggedness testing, but it can be very helpful during the method development phase. Problems that may occur during validation are often detected in advance and warnings can be included in the methodology about what parameters must be more tightly controlled in order to achieve the desired levels of precision and accuracy.

While not every parameter of every type of method can be listed here, it is important for the analyst to consider every step of the method and strive to include ruggedness testing in each of

them. Some typical examples of variations include variations in pH (of samples, of mobile phases, of extraction solvents), variations in mobile phase composition, and variations in columns (GC, LC, and extraction columns/cartridges - different lots or different suppliers). Other factors can include variations in temperature (room, oven, column, processing, thawing, autosampler, etc.), variations in flow rate (both in extraction and analysis), as well as various volumes and times used during the entire bioanalytical method.

Acceptance criteria for an analytical run

An analytical run consists of the following components:

1) A blank matrix (a processed matrix sample without analyte and without internal standard)
2) A zero sample (a processed matrix sample with internal standard)
3) A minimum of six non-zero calibration levels
4) A minimum of three levels of quality control (QC) samples in duplicate or a minimum of 5% of the total number of unknown samples, whichever is higher. Add QC samples in multiples of two when needed.
5) The field/study/unknown samples.

All samples should be processed as one single batch of samples. Calibration standards and QC samples should be spiked independently using separately prepared stock solutions. A single batch is comprised of samples that are handled at the same time, by the same analyst, under homogeneous conditions.

The standard curve samples, blanks, QCs, and unknowns can be arranged as considered appropriate within the run and support detection of drift over the course of the run. Acceptance criteria should be pre-established in the protocol, in the study plan, or in a standard operating procedure. In cases where a whole run consists of more than one batch, acceptance criteria should be applied to the whole run and to the individual batches.

Accuracy acceptance criteria for calibration standards [EMA 2011; FDA 2001]

The back-calculated concentrations of the calibration standards should be within ±15% of the nominal value, except for the LLOQ for which it should be within ±20%. A minimum of 75% of the calibration standards should be within these limits for the analytical run to qualify. If one of the calibration standards does not meet these criteria, this standard should be discarded, provided the calibration model does not change, and the calibration curve without this standard should be re-evaluated and regression analysis performed.

If the discarded calibration standard is the LLOQ or the ULOQ, then the next calibration standard becomes the LLOQ or ULOQ, and the calibration range is lower for that run. The revised calibration range should cover all concentrations of QC samples.

Accuracy acceptance criteria for QC samples [EMA 2011; FDA 2001]

The accuracy values of the QC samples should be within ±15% for the upper concentration (50-100% ULOQ) and within ±20% for the lower concentration (100-200% LLOQ). At least 67% of the QC samples should be within these limits for a run to be accepted. Furthermore, at least 50% of the QC samples at each concentration should comply with this criterion.

In case where one or both of these criteria are not fulfilled, the analytical run should be rejected. The samples can be re-extracted and analyzed, corrective action taken on the instrumentation, or other investigation of the cause of the failure analyzed and corrected. In the case of the simultaneous determination of several analytes, there will be a calibration curve for each analyte in the method. If an analytical run is acceptable for one analyte but has to be rejected for another analyte, the data for the accepted analyte can be used.

Conclusions

This chapter describes biomonitoring as a valuable tool for demonstration of compliance with exposure limits, occupational health research and surveillance, evaluation of intervention effectiveness, and risk assessment and risk management. A major aim of biomarker research is to develop and validate biomarkers that reflect specific exposures or are quantitatively linked to adverse outcomes in humans to enable their use in risk prediction. The NRC identified several strategies to use biomarkers of effect to extrapolate dose and evaluate dose response [NRC 2007]. Biomonitoring can be used not only to identify exposures but also to evaluate trends of exposure over time.

Some general considerations for biomonitoring have been noted, such as the goals of biomonitoring for the study or program being initiated. The biological matrices of interest, whether blood, urine, exhaled breath or some other specimen, will have an integral role in the selection of which biomarker is most pertinent for the specified goals of the study. If the goal is to understand whether exposure occurred or not, it may be possible to measure the chemical agent or its metabolite in exhaled breath or urine. If the goal is to determine a health risk, then measurement of a biomarker of effect may be more appropriate.

Appropriate study design and statistical analyses will aid in ensuring that the study has scientific rigor and will optimize uncertainty. Written standard operating procedures and a quality

assurance program will help to improve the data quality and provide interpretable results. Adherence to field and laboratory safety procedures are imperative to protect staff from exposure to blood-borne pathogens. Ethical considerations should be at the forefront of any biomonitoring investigation.

Biomarkers need to be validated in the laboratory and for the population for which they will be used. The steps for laboratory validation of the analytical methods are covered in this chapter. Equally important is the population validation of biomonitoring methods, which is not in the scope of this chapter [Schulte and DeBord 2000]. Population validation determines the utility of the method in various population groups. It entails understanding interpersonal variability according to demographic and behavioral characteristics, determining the underlying prevalence of the marker, and identifying the optimal handling and logistical considerations [Schulte and DeBord 2000; Schulte and Perera 1997].

The potential of biomonitoring in occupational health is tempered by limitations in study design, interpretation of results, communication of results, and ethical uses. The ability of advances in existing and emerging technologies to develop new biomarker methods exceeds our practical ability to evaluate and validate all of them. The real challenges for occupational health professionals are to decide which measurement methods may be of value, to understand what information the biomarker measurements are actually providing, and, finally, to determine appropriate actions based on that information.

Disclaimer

Mention of any company or product does not constitute endorsement by NIOSH. In addition, citations to websites external to NIOSH do not constitute NIOSH endorsement of the sponsoring organizations or their programs or products. Furthermore, NIOSH is not responsible for the content of these websites. All web addresses referenced in this document were accessible as of the publication date.

References

Agilent Technologies [2010]. Validation of analytical methods: A primer. http://www.chem.agilent.com/Library/primers/Public/5990-5140EN.pdf.

ACGIH [2015]. TLVs and BEIs based on the documentation of the Threshold Limit Values for chemical substances and physical agents and Biological Exposure Indices. Cincinnati, Ohio: American Conference of Governmental Industrial Hygienists. www.acgih.org

Angerer J, Aylward LL, Hays SM, Heinzow B, Wilhelm M [2011]. Human biomonitoring assessment values: approaches and data requirements. Int J Hyg Env Health *214*:348-360.

ANSES [2016]. Biological limit values for chemicals used in the workplace. Maisons-Alfort: French Agency for Food, Environmental and Occupational Health and Safety, https://www.anses.fr/en/content/biological-limit-values-chemicals-used-workplace

Ashley K, Brisson MJ, White KT [2011]. Review of standards for surface and dermal sampling. In: Brisson MJ, Ashley K, eds. Surface and dermal sampling (ASTM STP 1533). West Conshohocken, PA: ASTM International, pp. 3-16.

ASTM International [2013]. ASTM D6785: Determination of lead in workplace air using flame or graphite furnace atomic absorption spectrometry. ASTM International: West Conshohocken, PA.

ASTM International [2014]. ASTM E691: Standard practice for conducting an interlaboratory study to determine the precision of a test method. ASTM International: West Conshohocken, PA.

Aylward LL, Hays SM [2011]. Biomonitoring-based risk assessment for hexabromocyclododecane (HBCD). Int J Hyg Environ Health *214*(3):179–187.

Aylward LL, Hays SM, Smolders R, Koch HM, Cocker J, Jones K, Warren N , Levy L, Bevan R [2014]. Sources of variability in biomarker concentrations. J Toxicol Environ Health B *17*(1):45-61.

Aylward LL, Kirman CR, Adgate JL, McKenzie LM, Hays SM [2012]. Interpreting variability in population biomonitoring data: role of elimination kinetics. J Exp Sci Environ Epidemiol *22*:398-404.

Aylward LL, Kirman CR, Schoeny R, Portier CJ, Hays SM [2013]. Evaluation of biomonitoring data from the CDC National Exposure Report in a risk assessment context: perspectives across chemicals. Environ Health Perspect *121*:287-294.

Bader M, Barr D, Göen T, Schaller KH, Scherer G, Angerer J [2010]. Reliability criteria for analytical methods. In: The MAK-collection for occupational health and safety; Part IV:

biomonitoring methods. Vol. 12. Weinheim, Germany: WILEY-VCH Verlag GmbH & Co. KGaA, pp. 55-101.

Barr DB, Wilder LC, Caudill SP, Gonzalez AJ, Needham LL, Pirkle JL [2005]. Urinary creatinine concentrations in the U.S. population: implications for urinary biologic monitoring measurements. Environ Health Perspect *113*:192–200.

B'Hymer C, Mathias P, Krieg E, Cheever KL, Toennis CA, Clark JC, Kesner JS, Gibson RL, Butler MA [2012]. Evaluation and comparison of urinary metabolic biomarkers of exposure for the jet fuel JP-8. J Toxicol Environ Health *75*:661-672.

Boeniger MF, Lowry LK, Rosenberg J [1993]. Interpretation of urine results used to assess chemical exposure with emphasis on creatinine adjustments: a review. Am Ind Hyg Assoc J *54*:615-627.

Brown JF, Lawton RW [1984]. Polychlorinated biphenyl (PCB) partitioning between adipose tissue and serum. Bull Environ Contam Toxicol *33*(3):277-80.

Brown PO, Botstein D [1999]. Exploring the new world of the genome with DNA microarrays. Nat Genet *21*(1 Suppl):33–37.

Calafat AM, Longnecker MP, Koch HP, Swan SH, Hauser R, Goldman LR, Lanphear BP, Rudel RA, Engel SM, Whyatt RM, Wolff MS [2015]. Optimal exposure biomarkers for nonpersistent chemicals in environmental epidemiology. Environ Health Perspect *123*(7):A166-168.

Carrieri M, Tevisan A, Bartolucci GB [2001]. Adjustment to concentration dilution of spot urine samples: correlation between specific gravity and creatinine. Int Arch Occup Environ Health *74*:63–67.

Causon R [1997]. Validation of chromatographic methods in biomedical analysis: viewpoint and discussion. J Chromatogr B *689*:175-180.

Cavallari JM, Osborn LV, Snawder JE, Kriech AJ, Olsen LD, Herrick RF, McClean MD [2012a]. Predictors of airborne exposures to polycyclic aromatic compounds and total organic matter among hot-mix asphalt paving workers and influence of work conditions and practices. Ann Occup Hyg *56*(2):138-147.

Cavallari JM, Osborn LV, Snawder JE, Kriech AJ, Olsen LD, Herrick RF, McClean MD [2012b]. Predictors of dermal exposures to polycyclic aromatic compounds among hot-mix asphalt paving workers. Ann Occup Hyg *56*(2):125-137.

CDC [1993]. Biosafety in microbiological and biomedical laboratories, 3rd ed., Richmond JY, McKinney RW, eds. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, HHS Publication No. (CDC) 93-8395.

CDC [2014]. National biomonitoring program. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, http://www.cdc.gov/biomonitoring/about.html.

CDC [2017]. National report on human exposure to environmental chemicals. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, https://www.cdc.gov/exposurereport/index.html.

Clewell HJ, Tan YM, Campbell JL, Andersen ME [2008]. Quantitative interpretation of human biomonitoring data. Toxicol Appl Pharmacol *231*:122-133.

DeBord DG, Burgoon L, Edwards S, Haber LT, Kanitz MH, Kuempel E, Thomas RS, Yucesoy B [2015]. Systems biology and early effects dose-response for OEL setting. J Occup Environ Hyg *12*(Suppl 1):S41-S54.

Decker JA, DeBord DG, Bernard B, Dotson GS, Halpin J, Hines CJ, Kiefer M, Myers K, Page E, Schulte P, Snawder J [2013]. Recommendations for biomonitoring of emergency responders: focus on occupational health investigations and occupational health research. Mil Med *178*(1):68-75.

Demchuk E, Yucesoy B, Johnson V, Andrew M, Weston A, Germolec D, De Rosa C, Luster M [2007]. A statistical model for assessing genetic susceptibility as a risk factor in multifactorial diseases: lessons from occupational asthma. Environ Health Perspect *115*(2):231–234.

DFG [2006]. Bonn, Germany: Deutsche Forschungsgemeinschaft, http://media.wiley.com/product_data/excerpt/76/35273147/3527314776.pdf.

Dolan J [2015]. HPLC solutions #88: matrix effects, http://www.sepscience.com/Techniques/LC/Articles/217-/HPLC-Solutions-88-Matrix-Effects.

Droz PO [1989]. Biological monitoring I: sources of variability in human response to chemical exposure. Appl Ind Hyg *4*(1):F20-F24.

EMA [2011]. Guideline on bioanalytical method validation. London: European Medicines Agency, http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/08/WC500109686.pdf.

European Scientific Committee on Occupational Exposure Limits (SCOEL) [2014]. List of recommended health-based biological limit values (BLVs) and biological guidance values (BGVs), http://ec.europa.eu/social/main.jsp?catId=148&intPageId=684&langId=en.

Exley K, Aerts D, Biot P, Castelyn L, Kolossa-Gehring M, Schwedler G, Castaño A, Angerer J, Koch HM, Esteban M, Schindler BK, Schoeters G, Hond ED, Horvat M, Bloemen L, Knudsen LF, Joas R, Joas A, Sepai O [2015]. Pilot study testing a European human biomonitoring framework for biomarkers of chemical exposure in children and their mothers: experiences in the UK. Environ Sci Pollut Res *22*:15821-15834.

FDA [2001]. Guidance for industry: bioanalytical method validation. Silver Spring, MD: U.S. Department of Health and Human Services, Food and Drug Administration, http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070107.pdf.

Fiserova-Bergerova V [1993]. Biological monitoring VIII: interference of alcoholic beverage consumption with biological monitoring of occupational exposure of industrial chemicals. Appl Occup Environ Hyg *8*(9):757-760.

Fiserova-Bergerova V, Lowry LK, Rosenberg J [1989]. Biological monitoring II: measurements in exhaled air. Appl Ind Hyg *4*(2):F10-F13.

Gaines LGT, Fent KW, Flack SL, Thomasen JM, Ball LM, Zhou H, Whittaker SG, Nylander-French LA [2010]. Effect of creatinine and specific gravity normalization on urinary biomarker 1, 6-hexamethylene diamine. J Environ Monit *3*:591–599.

Gosho M, Nagashima K, Sato Y [2012]. Study designs and statistical analyses for biomarker research. Sensors *12*:8966-8986.

Government Publishing Office (GPO) [2012]. Transportation – Hazardous materials regulation. http://www.ecfr.gov/cgi-bin/text-idx?SID=c75175814d4d7f6397795d585538964d&mc=true&node=pt49.2.172&rgn=div5 Accessed September 23, 2015.

Green JM [1996]. A practical guide to analytical method validation. Anal Chem *68*:305A-309A.

Greenberg LA, Lester D [1947]. The metabolic fate of acetanilid and other aniline derivatives: the role of p-aminophenol in the production of methemoglobinemia after acetanilid. J Pharmacol Expo Ther *90*(2):150-153.

Hartmann C, Smeyers-Verbeke J, Massart DL, McDowall RD [1998]. Validation of bioanalytical chromatographic methods. J Pharm Biomed Anal *17*:193-218.

Hathaway GJ, Proctor NH [2004]. Chemical hazards of the workplace, 5th ed. New York: Wiley.

HSE [2017]. Biological monitoring guidance values. Liverpool, UK: Health and Safety Executive, http://www.hsl.gov.uk/online-ordering/analytical-services-and-assays/biological-monitoring/bm-guidance-values.

Hunter DJ [1997]. Methodologic issues in the use of biological markers in cancer epidemiology: cohort studies. In Toniolo P, Boffetta P, Shuker DEG, Rothman N, Hulka B, Pearce N, eds. Application of biomarkers of cancer epidemiology. IARC Scientific Publication No 142. Oxford, England: Oxford University Press.

Hunter DJ, Losina E, Guermazi A, Burstein D, Lassere MN, Kraus V [2010]. A pathway and approach to biomarker validation and qualification for osteoarthritis clinical trials. Curr Drug Targets *11*(5): 536–545.

International Conference on Harmonisation (ICH) [2005]. Validation of analytical procedures: text and methodology. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1__Guideline.pdf.

IATA [2015]. Dangerous goods. Montreal, Quebec, Canada: International Air Transport Association, http://www.iata.org/whatwedo/cargo/dgr/pages/index.aspx.

Irish National Accreditation Board [2012]. Guide to method validation for quantitative analysis in chemical testing laboratories, http://inab.ie/media/PS15.pdf.

Japan Society for Occupational Health (JSOH) [2016]. Recommendation of occupational exposure limits (2016–2017). J Occup Health *58*:489-518.

Laboratory requirements. Clinical laboratory improvement amendments of 1988; Final rule. 42 CFR 493 (1992).

Landi MT, Caparosa N [1997]. Sample collection, processing and storage. In: Toniolo P, Boffetta P, Shuker DEG, Rothman N, Hulka B, Pearce N, eds. Application of biomarkers of cancer epidemiology. IARC Scientific Publication No 142. Oxford, England: Oxford University Press.

Lauwerys RR, Hoet P [1993]. Industrial chemical exposure. Guidelines for biological monitoring, Ch. 1. Ann Arbor, MI: Lewis Publishers.

Liira J, Riihimaki V, Engstrom K [1990]. Effects of ethanol on the kinetics of inhaled methyl ethyl ketone in man. Br J Ind Med *47*:325-330.

Linder W, Wainer IW [1998]. Requirements for initial assay validation and publication in J. Chromatography B (editorial). J Chromatogr B *707*:1-2.

Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN [2003]. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet *33*(2):177–182.

Manh HD, Kido T, Okamoto R, XianLiang S, Anh LT, Supratman S, Maruzeni S, Nishijo M, Nakagawa H, Honma S, Nakano T, Takasuga T, Nhu DD, Hung NN, Son LK [2014]. Serum dioxin levels in Vietnamese men more than 40 years after herbicide spraying. Environ Sci Technol *48*(6):3496-3503.

McClean MD, Osborn LV, Snawder JE, Olsen LD, Kriech AJ, Sjodin A, Li Z, Smith JP, Sammons DL, Herrick RF, Cavallari JM [2012]. Using urinary biomarkers of polycyclic aromatic compound exposure to guide exposure-reduction strategies among asphalt paving workers. Ann Occup Hyg *56*(9):1013-1024.

McHale CM, Zhang L, Smith MT [2012]. Current understanding of the mechanism of benzene-induced leukemia in humans: implications for risk assessment. Carcinogenesis *33*(2):240-252.

Menke A, Muntner P, Silbergeld EK, Platz EA, Guallar E [2009]. Cadmium levels in urine and mortality among U.S. adults. Environ Health Perspect *11*(2):190-196.

Mikkelsen SR, Cortón E [2004]. Validation of new bioanalytical methods. In: Bioanalytical chemistry. Hoboken, NJ: John Wiley & Sons, Inc. pp. 322-345.

Needham LL, Ozkaynak H, Whyatt RM, Barr DB, Wang RY, Naeher L, Akland G, Bahadori T, Bradman A, Fortmann R, Liu LJ, Morandi M, O'Rourke MK, Thomas K, Quackenboss J, Ryan PB, Zartarian V [2005]. Exposure assessment in the national Children's Study: introduction. Environ Health Perspect *113*(8):1076-1082.

NIOSH [1994]. Application of biological monitoring methods: Chapter F. By Teass AW, Biagini RE, DeBord DG, Hull RD. In: Eller PM, Cassinelli ME, eds. NIOSH manual of analytical methods. 4th ed. Cincinnati, OH: U.S. Department of Health and Human Services, Centers for Disease

Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 94-113.

NIOSH [1995]. Guidelines for air sampling and analytical method development and evaluation. By Kennedy ER, Fischbach TJ, Song R, Eller PM, Shulman SA. Cincinnati, OH: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 95-117.

NIOSH [2010]. Genetics in the workplace: implications for occupational safety and health. By DeBord DG, Schulte PA, Butler MA, McCanlies E, Reutman S, Ruder A, Schill A, Schubauer-Berigan M, Schuler C, Weston A. Cincinnati, OH: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 2010-101.

NIOSH [2014]. Adult blood lead epidemiology and surveillance (ABLES). Cincinnati, OH: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, http://www.cdc.gov/niosh/topics/ables/default.html.

NRC (National Research Council) [1987]. Biological markers in environmental health research. Environ Health Perspect *74*:3–9.

NRC (National Research Council) [2006]. Human biomonitoring for environmental chemicals. Washington, DC: The National Academies Press.

NRC (National Research Council) [2007]. Toxicity testing in the 21st century: a vision and a strategy. Washington, DC: The National Academies Press.

Ogata M, Fiserova-Bergerova V, Droz PO [1993]. Biological monitoring VII: occupational exposures to mixture of industrial chemicals. Appl Occup Environ Hyg *8*(7):609-617.

OSHA [1978]. Lead. Washington, DC: U.S. Department of Labor, Occupational Safety and Health Administration, https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10030.

OSHA [1980]. Benzene. Washington, DC: U.S. Department of Labor, Occupational Safety and Health Administration, https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10042.

OSHA [1981]. Cadmium. Washington, DC: U.S. Department of Labor, Occupational Safety and Health Administration, https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10035.

OSHA [2001]. Bloodborne pathogens and needlestick prevention. Washington, DC: U.S. Department of Labor, Occupational Safety and Health Administration, https://www.osha.gov/SLTC/bloodbornepathogens/index.html.

Peters FT, Maurer HH [2015]. Review: bioanalytical method validation – How, how much and why? http://www.gtfch.org/cms/images/stories/media/tk/tk68_3/Peters.pdf.

Richards DM, Kraus JF, Kurtz P, Borhani NO, Mull R, Winterlin W, Kilgore WW [1978]. A controlled field trial of physiological responses to organophosphate residues in farm workers. J Environ Pathol Toxicol *2*(2):493-512.

Rosenberg J [1994]. Biological monitoring IX: concomitant exposure to medications and industrial chemicals. Appl Occup Environ Hyg *9*(5):341-345.

Rosenberg J, Fiserova-Bergerova V, Lowry LK [1989]. Biological monitoring IV: measurements in urine. Appl Ind Hyg *4*(4):F16-F21.

Sauve J, Levesque M, Huard M, Drolet D, Lavoue J, Tardif R, Truchon G [2015]. Creatinine and specific gravity normalization in biological monitoring of occupational exposures. J Occup Environ Hyg *12*(2): 123–129.

Schulte PA, DeBord DG [2000]. Public health assessment of genetic information in the occupational setting. In: Khoury M, Burke W, eds. Genetics and public health. New York: Oxford Press.

Schulte PA, Perera FP [1993]. Validation. In: Schulte PA, Perera FP, eds. Molecular epidemiology: principles and practices. San Diego, CA: Academic Press.

Schulte PA, Perera FP [1997]. Transitional studies. In Toniolo P, Boffetta P, Shuker DEG, Rothman N, Hulka B, Pearce N, eds. Application of biomarkers of cancer epidemiology. IARC Scientific Publication No 142. Oxford, England: Oxford University Press.

Schulz C, Wilhelm M, Heudorf U, Kolossa-Gehring M [2011]. Update of the reference and HBM values derived by the German Human Biomonitoring Commission. Int J Hyg Environ Health *215*:26–35.

Scinicariello F, Yesupriya A, Chang MH, Fowler B [2010]. Modification by ALAD of the association between blood lead and blood pressure in the U.S. population: results from the third National Health and Nutrition Examination Survey. Environ Health Perspect *118*(2):259–264.

Sexton K, Needham LL, Pirkle JL [2004]. Human biomonitoring of environmental chemicals: measuring chemicals in human tissue is the "gold standard" for assessing the people's exposure to pollution. Am Sci *92*(1):38-45.

Smeraglia J, Baldrey SF, Watson D [2002]. Matrix effects and selectivity issues in LC-MS-MS. Chromatographia *55*(Supplement):S95-S99.

Spencer K [1986]. Analytical reviews in clinical biochemistry: the estimation of creatinine. Ann Clin Biochem *23*:1-25.

St Amand A, Werry K, Aylward LL, Hayes SM, Nong A [2011]. Screening of population level biomonitoring data from the Canadian Health Measures Survey in a risk-based context. Toxicol Lett *231*(2):126-134.

Swiss Accident Insurance Fund (SUVA) [2013]. Grenzwerte am arbeitsplatz 2015 MAK-werte, BAT-werte, grenzwerte für physikalische einwirkungen. Biomonitoring. http://www.suva.ch/english/factsheet-grenzwerte.pdf.

Taylor L, Jones RL, Ashley K, Deddens JA, Kwan L [2004]. Comparison of capillary earlobe and venous blood monitoring for occupational lead surveillance. J Lab Clin Med *143*:217-224.

Theodorsson E [2012]. Validation and verification of measurement methods in clinical chemistry. Bioanalysis *4*(3):305-320.

Thomas RS, Allen BC, Nong A, Yang L, Bermudez E, Clewel lll HJ, Andersen ME [2007]. A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. Toxicol Sci *98*(1):240-248.

Thomas RS, Clewell III HJ, Allen BC, Wesselkamper SC, Wang NCY, Lambert JC, Hess-Wilson JK, Zhao QJ, Andersen ME [2011]. Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. Toxicol Sci *120*(1):194–205.

Tiwari G, Tiwari R [2010]. Bioanalytical method validation: an updated review. Pharm Methods *1*(1): 25–38.

Trufelli H, Palma P, Famiglini G, Cappiello A [2011]. An overview of matrix effects in liquid chromatography-mass spectrometry. Mass Spectrom Rev *30*:491-509.

UNODC [2009]. Guidance for the validation of analytical methodology and calibration of equipment used for testing of illicit drugs in seized materials and biological specimens. Vienna, Austria: United Nations Office on Drugs and Crime, http://www.unodc.org/documents/scientific/validation_E.pdf.

Valcke M, Haddad S [2015]. Assessing human variability in kinetics for exposures to multiple environmental chemicals: a physiologically based pharmacokinetic modeling case study with dichloromethane, benzene, toluene, ethylbenzene, and m-xylene. J Toxicol Environ Health A *78*(7):409-31.

Viswanathan CT, Bansal S, Booth B, DeStefano AJ, Rose MJ, Sailstad J, Shah VP, Skelly JP, Swann PG, Weiner R [2007]. Workshop/conference report – quantitative bioanalytical methods validation and implementation: best practices for chromatographic and ligand binding assays. Am Assoc Pharm Sci J *9*(1):E30-E42.

Wieling J, Hendriks G, Tamminga WJ, Hempenius J, Mensink CK, Oosterhuis B, Jonkman JHG [1996]. Rational experimental design for bioanalytical methods validation. Illustration using an assay method for total captopril in plasma. J Chromatogr A *730*:381-394.

WHO [2001]. Environmental health criteria 222: biomarkers in risk assessment: validity and validation. Geneva, Switzerland: World Health Organization.

Yeh H-C, Lin Y-S, Kuo C-C, Weidemann D, Weaver V, Fadrowski J, Neu A, Navas-Acien A [2015]. Urine osmolality in the US population: implications for environmental biomonitoring. Environ Res *136*:482–490.

Zidek A, Macey K, MacKinnon L, Patel M, Poddalgoda D, Zhang Y [2017]. A review of human biomonitoring data used in regulatory risk assessment under Canada's Chemicals Management Program. Int J Hyg Environ Health *220*(2):167-178.

Zong G, Grandjean P, Wu H, Sun Q [2015]. Circulating persistent organic pollutants and body fat distribution: evidence from NHANES 1999-2004. Obesity *23*(9):1903-1910.