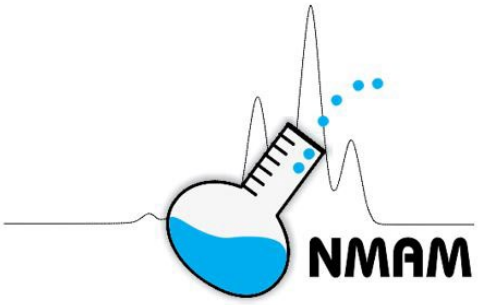

NIOSH Manual of Analytical Methods (NMAM), 5th Edition

Application of Biological Monitoring Methods for Chemical Exposures in Occupational Health

by D. Gayle DeBord, Ph.D., Dale Shoemaker, Ph.D., Clayton B'Hymer, Ph.D., and John Snawder, Ph.D., DABT, NIOSH (retired)

This chapter is a revision and expansion of Chapter F: Application of Biological Monitoring Methods published in the 4th edition of the NIOSH Manual of Analytical Methods (NMAM) [NIOSH 1998]. Additional information has been incorporated and terminology has been updated to reflect the most current practice. A lengthy section on biomonitoring analytical method validation has been added.

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Institute for Occupational Safety and Health

1 Introduction

Exposure assessment is a critical part of occupational health studies and risk mitigation activities. 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 the potential health risks. Large population studies that incorporate biomonitoring efforts can identify chemicals found in the environment (including the occupational environment) and monitor the trends and distribution of exposures in the general population [CDC 2021, 2022].

Environmental measurements involving various media (such as air, water, soil, food, or surfaces) can detect the presence of hazards and sometimes predict individual exposure; for example, air monitoring predicts effects when the lung is the target organ. However, to evaluate biological, physical, or chemical agents that have been absorbed into the body, the magnitude of the absorbed dose and its contribution to total body burden, the measurement of biomarkers (biomonitoring) is required. Exposures for which biomarkers can be useful include the full range of occupational hazards from noise to stress to chemicals. While this document generally focuses on chemical exposures, considerations pertaining to purpose (objectives), study design criteria, quality assurance, ethics, and safety have general application. The guidance provided in this document is particularly relevant to urine and blood biomonitoring and is generally relevant to other biomonitoring matrices and assays in which response or susceptibility factors such as genetic markers or gene variants are measured. This document generally focuses on the utility of biomonitoring as a research tool. Only a few examples that are cited in this chapter are relevant to occupational health practice.

2 Applications of biomonitoring

Biomonitoring has many benefits. Adapted from Decker et al. [2013], biomonitoring benefits can include the following:

- Measure actual body burden
- Augment other exposure monitoring tools
- Capture all exposure routes, including dermal
- Detect unexpected exposures or routes of exposure
- Evaluate the effectiveness of control measures, including personal protective equipment (PPE)
- Provide biomarkers of potential health risks
- Use to reconstruct exposures following acute or accidental events if appropriate biomarkers are available

- Enhance individual or group risk assessments
- Provide valuable information for risk communication

Biomonitoring can enhance exposure assessment and provide information about health outcomes, depending upon the selected biomarker(s).

3 Categories of biomarkers

The National Academy of Sciences has defined biomarkers as measurable indicators in a biological system or organism. Examples include 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]. Biomarkers can be classified into three categories: exposure, effect or response, and susceptibility (Table 1) [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 individual factors can have on exposure, uptake metabolism, and/or repair.

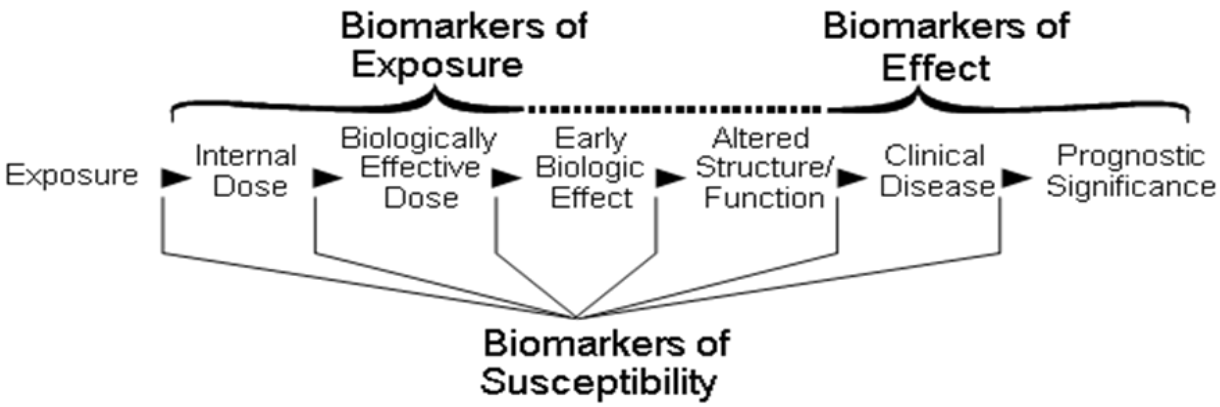

Figure 1 Continuum from exposure to disease. [Adapted from NRC 1987 and Schulte and Perera 1993]

Table 1. 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 chemical or metabolite
Effect	Biochemical, physiological, or other alteration that can be measured, which may or may not indicate a health impairment [WHO 2001]	DNA mutation or cytogenetic change
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]

a. Biomarkers of exposure

Occupational exposure occurs mainly by inhalation or through the skin [NIOSH 1998] and less often by ingestion. The route of exposure may affect the internal dose and the biologically effective dose, which is the dose that interacts with the target organ or tissue. Biomarker measurements aggregate exposure across all pathways and all agents. While this aggregation is often advantageous, biomarker measurements cannot be used by themselves to determine the route(s) of exposure and in certain cases, the specific exposure agent. 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 indicate that an individual or population has been exposed and absorbed an agent into the body. Exposure biomarkers may or may not reflect future disease or health risk. Exposure concentration, exposure duration, route of exposure, pharmacokinetics of the chemical as it affects the distribution, metabolism and elimination, timing of sample collection, physiological variations in hydration status, and urinary flow can influence biomarker levels [Aylward et al. 2012; Aylward et al. 2014]. Individual characteristics such as gene variants, exertion, and co-exposures can also modify biomarker levels through the modification of absorption or metabolism.

b. 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, or 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 biomolecules. These changes may be reversible or permanent. An example of a biomarker of effect is the decreased levels of the enzyme acetylcholinesterase (AChE) in 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 (or confirm) exposure and to assess potential health risks and treatment needs [Richards et al. 1978]. A continuum exists across the biomarker spectrum. At low levels, a biomarker may indicate an exposure or an effect. However, as the exposure continues or increases, the biomarker level also increases, leading to a response or action level that needs intervention 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 exposure to multiple chemicals or agents. For example, measuring blood AChE levels provides evidence of exposure to OP pesticides, but does not identify the specific pesticide responsible for the decrease in AChE. Similarly, carboxyhemoglobin levels increase after exposure to both carbon monoxide and methylene chloride, which is metabolized to carbon monoxide [Andersen et al. 1991].

c. 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 metabolic enzyme, can modify benzene toxicity [McHale et al. 2012], and exposure to ethanol reduces methyl ethyl ketone metabolism [Liira et al. 1990].

4 Drivers for biomonitoring in occupational health studies

Several factors drive biomarker use in occupational safety and health. These drivers include risk management (including compliance), research and public health surveillance, evaluation of intervention effectiveness, and risk-assessment processes.

a. Risk management (including compliance)

The National Institute for Occupational Safety and Health (NIOSH) and other organizations recommend the use of medical surveillance as a component of a comprehensive risk management program to protect the health of workers [NIOSH 2012]. Frameworks for the use of biomarkers as medical surveillance tools, particularly when other sources of medical data are not readily available, have been published [Schulte 2005]; Medical surveillance can help identify sentinel adverse health effects among workers, suggesting failures in controlling exposures. Thus, help to identify the need for where improvements in workplace controls, such as engineering, or administrative controls or personal protective equipment are needed. Also, individual workers may benefit from the detection of disease in early stages when it may be more treatable with better clinical outcomes. Several examples of biomarkers that are used for medical surveillance purposes, including for compliance with established standards, are described below.

The Occupational Safety and Health Administration (OSHA) has three regulations that require biomonitoring in certain situations (Table 2). 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 milligram (mg) phenol/liter (L) urine, then no further testing is required. If the 72-hour test result is greater than 75 mg phenol/L urine, then additional medical surveillance is instituted at monthly intervals for 3 months.

Table 2. Chemicals requiring biomonitoring by OSHA

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]

OSHA requires several biomonitoring tests 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 an airborne cadmium concentration of 2.5 micrograms per cubic meter ($\mu g/m^3$) calculated as an 8-hour time-weighted average. Biomonitoring tests include cadmium in urine (CdU), standardized to grams of creatinine (g/Cr); beta-2 microglobulin in urine (β2-M), standardized to grams of creatinine (g/Cr), with pH specified; and cadmium in blood (CdB), standardized to liters of whole blood (lwb). β2-M is a marker of potential and actual kidney damage. 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 of all general industry employees who are or may be exposed to lead in workplace air at or above 50 $\mu g/m^3$ averaged over an 8-hour workday. Monitoring of blood lead and zinc protoporphyrin levels is required every 6 months. The requirement changes to every 2 months if the blood lead level is at or above 40 µg/100 g of whole blood or to monthly if the blood lead levels are high enough that an employee is medically removed from work [OSHA 1978].

NIOSH has no recommended exposure limits (RELs) based on biomonitoring. Other organizations have developed biomonitoring-based exposure limits. The American Conference of Governmental Industrial Hygienists (ACGIH) publishes a list of biological exposure indices (BEIs®) for approximately 30 chemical agents [ACGIH 2022]. The German Deutsche Forschungsgemeinschaft (DFG), Health and Safety Executive (HSE) in the UK, France's Agency for Food, Environmental and Occupational Health and Safety (ANSES), and the Swiss Accident Insurance Fund (Suva) also routinely publish lists of biological exposure limits [ANSES 2016; DFG 2015; HSE 2020; Suva 2019].

b. Research and public health 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 identifying hazards or monitoring exposure trends over time. Biomonitoring can provide information on what has been absorbed or taken up 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 only to detect exposure and not to give a measure of exposure intensity. 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 assesses exposure across all routes.

Biomarker levels are affected by a number of different factors. These include the specimen matrix, such as blood or urine; the timing of the specimen collection; and the pharmacokinetics of absorption, distribution, metabolism, and excretion (ADME) [Aylward et al. 2014]. Individual differences in physiological parameters, such as age, weight, and liver function, may affect the distribution of a chemical, and genetic polymorphisms of metabolic enzymes can alter the concentration and pattern of metabolites. Differences in physical activity may affect the amount of chemical absorbed. Kidney function can affect the clearance or excretion of a chemical or metabolite.

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, because biomarkers integrate exposure across all routes, it may not be possible to determine the amount of a measured biomarker attributable to skin absorption alone [Decker et al. 2013], unless inhalation exposure is precluded, for example, by use of an air-supplied respirator. Biomonitoring may also be useful in assessing the biological effects from the breach of or improper use of personal protective equipment (PPE) and may allow for comparing exposures/doses associated with different work practices and tasks.

A few surveillance programs include biomonitoring. One example is the Adult Blood Lead Epidemiology and Surveillance (ABLES) program established by NIOSH in 1987 to monitor occupational lead exposure [NIOSH 2021]. National blood lead level data are published in the Centers for Disease Control and Prevention's (CDC) Morbidity and Mortality Weekly Report (MMWR) and elsewhere. CDC also has a National Biomonitoring Program within the National Center for Environmental Health. This program periodically measures the amount of certain chemicals and metabolites in blood and urine samples collected from a representative sample of the U.S. population. CDC has issued several reports and a series of updated tables with the results of this biomonitoring program. An additional goal of this program is to provide U.S. population-based reference ranges for the monitored chemicals [CDC 2021, 2022].

NIOSH uses biomonitoring in its occupational health research. In a study to assess worker exposures to JP-8 jet fuel at U.S. Air Force bases, 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 most accurately assessed exposure [B'Hymer et al. 2012]. Statistical analysis of the measured analytes showed that the urinary concentration of MEAA, whether adjusted for creatinine or not, was the most accurate or appropriate biomarker for JP-8 exposure. Correction of biomarker measurements for urine dilution is discussed in more detail later in this chapter.

c. 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, the effectiveness of controls may not be readily apparent 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 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 were all associated with decreased urinary PAC concentrations compared with baseline concentrations. Higher HMA temperatures were positively associated with urinary PAC concentrations. Biodiesel substitution and lower HMA temperatures were associated with decreased PAC levels in air and on skin [Cavallari et al. 2012a,b].

d. Biomarker research for improving assessment of risk

Establishing a dose of concern and a mode of action are two primary components of risk assessment [DeBord et al. 2015]. Environmental exposures can directly or indirectly cause alterations in several 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 endpoint of interest [Brown and Botstein 1999]. Toxicogenomics biomarker data have also been used to demonstrate benchmark dose estimates [Thomas et al. 2011].

Determining 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 is important in their use in risk assessment.

Few studies have incorporated biomarkers of susceptibility into risk assessments. The advantages of susceptibility data have been discussed [Demchuk et al. 2007; Lohmueller et al. 2003; Schulte et al. 2015; Scinicariello et al. 2010]. These advantages include identification of the most genetically susceptible groups and opportunities for targeting preventative or therapeutic actions to high-risk populations.

5 General considerations for biomonitoring research studies

a. Factors that influence biomonitoring

Absorption can occur after dermal contact, inhalation, ingestion, or from a combination of these 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 because of 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 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 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 as it frequently makes a compound more water soluble so that it is more easily eliminated via the kidney. Conjugation reactions include glucuronidation, amino acid conjugation, acetylation, sulfate conjugation, and methylation [NIOSH 1998].

Metabolism, excretion, and the rates thereof, can be affected by age, diet, general health status, race, gender, and other factors. In general, 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. Substances enter the urine by either glomerular filtration, tubular secretion, or sometimes both mechanisms [NIOSH 1998]. 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 to select the proper compound to be determined, biological medium to be sampled, and time for obtaining a specimen. Often, most of the available toxicological and pharmacological information 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 was the appropriateness of conducting biomonitoring, which includes whether biomonitoring could add information that is actionable and interpretable. The second consideration was the feasibility of conducting biomonitoring with respect to logistics and available methodology.

b. Monitoring goals

Air monitoring (or workplace environmental monitoring) and biomonitoring have complementary goals and are frequently applied simultaneously in industrial hygiene investigations [NIOSH 1998]. Compared with biomonitoring, air monitoring offers advantages in certain situations. If the agent has acute toxic effects on the respiratory tract or eyes, air monitoring is the logical tool for measuring the exposure [Hathaway and Proctor 2004]. For some chemicals, direct-reading instruments can be used to detect peak air concentrations of potentially dangerous chemicals, especially in acute, rapidly evolving situations. Biological monitoring offers a better estimate of exposure in situations where routes of exposure other than inhalation are significant.

Toxicokinetics, which is the rate at which a chemical enters the body and the fate of the chemical once in the body, has a key role in the timing of sample collection. The elimination half-life, or the time it takes to eliminate half of the chemical from the body, determines how long a chemical or metabolite remains in the system. For example, lead in blood has 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, as well as any on-going exposure and the intensity of that exposure. Levels of rapidly disappearing biomarkers (short half-lives) 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.

c. Biological matrices

The most common matrices used for biomonitoring are exhaled air, blood, and urine [Landi and Caporaso 1997; NIOSH 1998]. Other matrices such as hair, nails, saliva, and feces also have been used.

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 enables exposed populations to be compared with national reference ranges obtained by efforts such as CDC's Biomonitoring Program and the elimination of the chemical compound to be followed over a long period of time [NRC 2006]. Most biomarkers present in the body can be found in the blood for 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, the tissues in which the chemical is stored, and the 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:

- 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 for urine.
- 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].
- A hydrophobic chemical will persist in the body, which results in a longer half-life and may allow for accurate measurement of the chemical 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 toxicokinetics of the chemical is also a primary concern [Calafat et al. 2015]. The concentration of the biomarker in urine is usually 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 often more representative than from spot samples and are usually a better estimate of exposure. If the exposure is relatively constant or the half-life is long, then the concentration in a spot sample may reasonably predict the 24-hour concentration; however, if variability is high or the half-life is short (a few hours), then a spot sample may not be very predictive. However, collection, stabilization, and transportation of 24-hour specimens in the field are difficult and often not feasible.

6 Practical considerations for biomonitoring research studies

a. Study design

Attention needs to be paid to scientific rigor when using 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 levels of most biomarkers are usually measured in microgram or lower quantities, concentrations may be subject to large variability because of influences such as diet, lifestyle habits, and other environmental exposures. While environmental monitoring provides information on what is in the area, biomonitoring is person-specific, which may generate concern among participants. Health-related concerns may arise from measurements that are outside of the normal or reference range should a normal or reference range be known. Some individuals may have concern over any exposure especially if the results are not easily interpretable in terms of their current or future health.

b. Population selection

The selection of the population for biomonitoring is typically straightforward in occupational settings. Generally, an important concept to consider is the inclusion of

both male and female workers of all ages at the worksite. However, if possible, 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 selecting the population, a large degree of uncertainty will still exist because of random variations. The smaller the group, the greater the uncertainty because of inter-individual differences and laboratory variations [NRC 2006].

Statistical design is an important aspect of any study but is particularly important in biomonitoring studies 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 some population studies is the lack of statistical power associated with many biomarker measurements because of small population sizes or small changes in biomarker levels [Gosho et al. 2012; Hunter 1997]. A description of the basic characteristics of each biomarker variable is needed. These characteristics 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.

c. Selection of biomonitoring methods

The occupational health professional and the analytical 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, 50 BEIs currently cover exposure to several hundred compounds (because of non-specificity) [ACGIH 2022]. Each of these BEIs has documentation that includes many of the issues (e.g., sampling time, interferences) discussed below for the development of a new biomarker. However, if no method 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, and biological relevance [NRC 2006]. Toxicokinetics and feasibility should also be considered. Interferences from diet, drugs, alcohol, disease states, or other chemicals or agents should be considered when interpreting the data. The method should also have a sufficiently low limit of detection to differentiate exposed from unexposed workers, and any effects because of the 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 not be practical or feasible. The method should include guidelines for interpreting the data. To minimize the risk of harm or discomfort to workers, when two biomonitoring methods will provide the same information, the less invasive method

should be used. For example, methods using urine or exhaled breath are preferred over those using blood if similar information can be obtained.

d. Sampling strategy

Attention to specimen handling and collection is essential for obtaining quality data [NIOSH 1998]. Consider these factors when collecting samples: the timing for the specimen collection, a baseline for the biomarker, possibilities for contamination, the need for preservatives, the stability of the biomarker, and sampling logistics. The analytical laboratory should be consulted for standard operating procedures for specimen collection and storage [NRC 2006]. Analytical methods should provide specific directions on the collection, storage, and transport of specimens to the laboratory. Adherence to these directions is of the utmost importance to ensure sample integrity. 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, although investigators may choose a different timing than recommended in the method depending on the purpose of the biomonitoring. The longer the half-life of the chemical, the less critical is the timing of the collection for exposure biomarkers [Lauwerys and Hoet 1993]. Some toxicants will be normally present or have a background level in the body; others will accumulate in the body over time, such as over a work week, based on factors like exposure and biological half-lives. Therefore, it may be necessary to take a baseline measurement of the biomarker concentration prior to the start of the work week or workday to determine if the biomarker concentration increased over the time frame of investigation [Lauwerys and Hoet 1993]. Recent trends in changing work schedules and work shifts may have implications for biomonitoring and should be considered when developing a sampling strategy. 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.

e. Correction of urinalysis data for dilution

The interpretation of biomarkers in individual urine samples can be affected by urine dilution (or concentration if a worker is dehydrated). The hydration status of the worker can cause substantial variation in analyte concentration. This variation can be attributed to fluid intake, temperature of the work environment, and the individual's physical workload. In addition, the excretion mechanism of a specific biomarker can be altered if the urine is either very concentrated or very dilute [Rosenburg et al. 1989]. The ranges of acceptability for urine dilution are not well documented. However, the common practice

to compensate for urine dilution is to adjust the measured concentration to a normalized value [Boeniger et al. 1993; Carrieri et al. 2001]. The most common normalization methods are based on creatinine levels, specific gravity, and urine output. These methods have been described in detail [NIOSH 1998]. Urine osmolality has also been suggested as a normalization method [Yeh et al. 2015].

Creatinine adjustment is the most frequently used method to normalize 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 other methods based on mass spectrometry and liquid chromatography [Spencer 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 [Sauve et al. 2015]. Among the factors affecting the rate of creatinine excretion are the muscularity of the individual, physical activity, urine flow, time of day, diet, pregnancy, age, 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. 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] reported significant bias when using creatinine to normalize urinary biomarker levels.

Specific gravity (sp.g.) is used to normalize an analyte's concentration in urine by multiplying the measured concentration of the biomarker by the ratio of [(sp.g.ref-1)/(sp.g.spec –1)], where sp.g.spec is the specific gravity of the urine specimen and sp.g.ref is the average reference specific gravity of human urine [Cone et al. 2009; Goldberger et al. 1995; Levine and Fahy 1945; WADA 2016]. The reference value of 1.020 is often used, especially in the United States, but a value of 1.015 was used in a study of Bangladeshi females [Miller et al. 2004].

The choice of which of these two normalization methods to use, if any, should be influenced by the chemical of interest in a study. Often there is little difference, but there are exceptions. 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. Measurements of samples having sp.g. outside the range 1.010–1.030 have been reported as being unreliable [Rosenberg et al. 1989].

Urine osmolality is a measure of the number of dissolved particles per unit of water in the urine [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 to normalize biomarker levels. The measured concentration of the biomarker is multiplied by the ratio R/0.05, where R is the urine 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, which is an average for adults [Rosenberg et al. 1989]. This method also has possible limitations and uncertainties from individual variation and hydration [NIOSH 1998].

f. Quality assurance

An effective quality assurance program is needed to ensure good quality data. In 1992, the Health Care Finance Administration and the Public Health Service published regulations implementing the Clinical Laboratory Improvement Amendments (CLIA) of 1988 to ensure that the analysis of human specimens was done accurately and under good quality control procedures [42 CFR 493 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. The CLIA Quality Assurance Program includes these key components [NIOSH 1998]:

- 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

Standard operating procedures should be developed for collecting, shipping, and processing of biological specimens [NRC 2006]. In addition, the stability of the biomarker and its concentration range 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 container that can be empty or filled with uncontaminated water and analyzed alongside specimens to determine if the collection container is a source of contamination. Spikes, whether done in the field or in the analytical laboratory, are usually performed by adding a known amount of chemical or its metabolite into the collection container. Duplicates are created by splitting a specimen into two collection containers. Tracking of specimens in the field after collection, during shipping, and after receipt at the laboratory is necessary to ensure sample integrity. The data needed include date and time of collection, specimen number and description of the specimen (type of specimen, volume, etc.), and shipping information (receipts, tracking numbers, and inventory) [NRC 2006]. If shipping samples, Department of Transportation (DOT) and International Air Transport Association (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 2021; IATA 2022].

g. Laboratory and field safety

When dealing with human specimens, a biosafety program is essential [NIOSH 1998]. 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, including dermatitis or acne; and indirectly through contact with a contaminated environmental surface. There are four major ways to reduce the potential for exposure to biological pathogens [CDC 2020]:

- Primary barriers (safety equipment): Examples include biosafety cabinets, enclosed containers and self-sheathing needles.
- Personal protective equipment (PPE): Examples include gloves, coats, safety glasses, and face shields.
- Secondary barriers (facility design and construction): Examples include ventilation strategies, decontamination systems, facility design, and vaccine usage.
- Facility practices and procedures: Examples include training and biosafety programs.

OSHA requires employers to implement an exposure control plan when employees may be exposed to bloodborne pathogens [OSHA, no date; OSHA, 1992]. Standard precautions should be practiced with every biological specimen collected or received in the laboratory. It is not possible to know if a particular specimen contains pathogens; therefore, each sample should be treated as if contaminated with pathogens.

Field personnel can be at risk for exposure to biological pathogens as well. They should follow standard precautions, such as using gloves during specimen collection and 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 organization's bloodborne pathogen policy in the event they are exposed to blood.

h. Ethics

Ethical considerations are paramount when conducting biomonitoring. An Institutional Review Board (IRB) should review research studies before collecting specimens [NRC 2006]. Among other considerations, the sampling method should be appropriate for the requirements of the investigation [NIOSH 1998], and the procedures should cause minimal harm to participants. The benefits of using invasive methods must outweigh the risks. Informed consent from the participant is required and the results should be kept confidential to the extent allowed by law [5 CFR 552a 1974]. Should an employer initiate a biomonitoring program, these criteria should be considered [Adapted from Schulte and DeBord 2000]:

- Specification of goals of the program
- Selection of validated biomarkers
- Establishment of clinical utility, if applicable
- 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 participants
- Process for addressing results and outliers

i. Interpretation of results

The measured concentration of a given biomarker 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 or had occurred. The second consideration is whether the measured level represents 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

the exposure is recent, may also be relevant [Aylward et al. 2012]. Risk-based approaches are more intensive [NRC 2006]. A risk-based approach could be used when exposure-response relationships have been demonstrated so that a measured exposure is known to cause an effect. One example would be decreased AChE 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. 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 evaluate measured biomarker concentrations and to interpret the exposure as below, near, or above the risk-based exposure guidance value. The forward dosimetry approach has been used extensively in both occupational and environmental biomonitoring contexts [Angerer et al. 2011; Lauwerys and Hoet 1993; Zidek et al. 2017].

j. Reference levels

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

Sources of biological action levels:

- Biological Exposure Indices (BEI®) adopted by the American Conference of Governmental Industrial Hygienists (ACGIH) [ACGIH 2022]
- 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 2015]
- Swiss Accident Insurance Fund (Suva) [Suva 2019]
- Summary of Recommendations in Industrial Chemical Exposure. Guidelines for Biological Monitoring [Lauwerys and Hoet 1993]
- OSHA standards [OSHA 1978, 1980, 1981]

- Japan Society for Occupational Health. Recommendation of occupational exposure limits (2016–2017) [JSOH 2016]
- Health and Safety Executive, UK. Biological monitoring guidance values [HSE 2020]
- European Scientific Committee on Occupational Exposure Limits (SCOEL) [2014]
- ANSES. Biological limit values for chemicals used in the workplace. [ANSES 2019]

Population-based sources for reference values:

- German Human Biomonitoring Commission (HBM) values [Angerer et al. 2011; Apel et al. 2017; Schulz et al. 2011]
- Centers for Disease Control and Prevention. National Report on Human Exposure to Environmental Chemicals [CDC 2021]
- 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 nonexposed population may serve as a reference level [NIOSH 1998]. For those biomarkers for which no measurable background level in nonoccupationally exposed 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 an occupational exposure but give no information on the potential health effect. The extent to which biomarkers can be used is dependent upon a number of critical factors: (1) adequate investment in validation, (2) obtaining international agreement on exposure guidelines, (3) exploring the utility of biomarkers in regulation, (4) applying biomarkers to critical occupational safety and health questions, (5) developing the exposome, (6) utilizing biomarkers to address emerging occupational health issues, and (7) continuing to address the ethical and social justice issues related to biomarkers. [Schulte and Hauser 2012].

k. Variability

Biological monitoring data are subject to several sources of variability [Aylward et al. 2014; Droz 1989; NIOSH 1998] that need to be considered when interpreting results. Sources of variability [Aylward et al. 2014; NIOSH 1998] include the following:

- Rates at which an agent is taken up by the body, metabolized, and excreted. These rates vary from person to person and are affected by the person's age, sex, and physical workload.
- Route of exposure. For example, absorption through the lungs is typically 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 collecting biological samples will be different for the two routes of entry.
- Fluctuations in environmental exposure. Such fluctuations will be tracked by the levels of rapidly eliminated biomarkers, i.e., those reflecting exposure in the previous several hours.
- Personal protective equipment worn, and a person's work practices.
- 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.
- Concurrent exposure to several chemicals that compete for the same biotransformation pathways in the body. This may lead to altered metabolism and excretion, which could change the relationship between exposure or health effect and the level of the biomarker [Ogata et al. 1993]. In one example, the variability of metabolites of several solvents was evaluated [Valcke and Haddad 2015]. Modelling evidence indicates the impact of multiple exposures depends upon chemical concentrations, biochemical properties, subpopulations, and internal dose metrics considered.
- Concurrent exposure to several chemicals that are metabolized to the same biomarker. This complicates the interpretation of the biological monitoring data. For example, trichloroacetic acid is a biomarker for trichloroethylene, 1,1,1-trichloroethane, and perchloroethylene.
- Consumption of alcoholic beverages [Fiserova-Bergerova et al. 1993]. Ethanol is metabolized by three pathways used for metabolism of other organic chemicals. After consuming one drink, the ethanol concentration in the blood is about 1000 times higher than from a typical occupational exposure and may significantly affect the metabolism of industrial chemicals.
- Medications [Rosenberg 1994], health, and diet.
- Smoking and other lifestyle factors.

Because of the variability in biomarker concentration, judgments on the exposure or health risk of workers frequently cannot be made based on a single sample [NIOSH 1998]. It may be necessary to collect multiple or repeated samples.

7 Biomonitoring analytical method validation

It is imperative that bioanalytical methods are well characterized, fully validated, and documented to a satisfactory standard to yield reliable results. Bioanalytical methods must also be validated for use in the field [Hunter et al. 2010; Schulte and DeBord 2000], which is not covered in this chapter. This chapter draws heavily from two primary sources: guidelines published by the U.S. Food and Drug Administration (FDA) [FDA 2018; Viswanathan et al. 2007], which have been a biomonitoring standard since their original publication in 2001; and a set of guidelines from the European Medicines Agency [EMA 2011]. Several other validation guidelines and research papers were also consulted and distilled into this final set of recommendations [Bader et al. 2012; Causon 1997; Green 1996; Hartmann et al. 1998; ICH 2005; Irish National Accreditation Board 2019; Linder and Wainer 1998; Mikkelsen and Cortón 2004; Peters and Maurer 2015; Theodorsson 2012; UNODC 2009; Wieling et al. 1996]. These references demonstrate that while there is general agreement for some aspects of bioanalytical method validation, many differences of opinions remain on the terminology employed, the number of parameters that should be evaluated, and how best to evaluate many of those parameters. The driving force behind most of the literature guidelines has been the need to analyze for pharmaceuticals in biological matrices. Therefore, some adaptations of the above guidelines have been taken into consideration for use in occupational exposure assessment.

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

This section is not intended to serve as a decision guide on the following:

- Whether biomonitoring is appropriate. Readers are referred to a decision matrix proposed by Decker et al. [2013].
- Method development. Method development steps should be completed before method validation occurs.
- Interpreting biomonitoring data.
- Selecting the appropriate analytical technique and methodology.

While this section addresses laboratory method validation, validation of the biomarker needs to be accomplished for the population of interest. 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, while 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 modified depending on the analytical method used and the matrix involved. Because of some unique and inherent characteristics of ligand-binding assays or immunoassays, this set of guidelines may not be appropriate for such assays. The FDA guidelines discuss immunoassays more comprehensively [FDA 2018]. While the general validation principles can and should be applied, the specific guidelines and limits described herein may not be applicable for immunoassays or for all bioanalyses cases. For more in-depth discussion of the reasons for this and about the adjusted recommendations, consult the guidelines from the FDA and the European Medicines Agency [EMA 2011; FDA 2018; Viswanathan et al. 2007].

a. Validation parameters

Method validation should include determining the following parameters: selectivity, carryover, calibration curve, bias and precision, accuracy, stability, matrix effects, recovery, robustness, limits of detection and quantitation (LOD and LOQ), and acceptance criteria for an analytical run, quality control samples, and calibration standards. Each of these parameters will be defined and discussed in more detail. These parameters should be reported when validating the method. Immunoassays have slightly different terminology than chemistry methods with respect to validation parameters and terminology [Andreasson et al. 2015]. To date, no immunoassays are represented in the NMAM and this chapter has focused on chemistry methods.

1) Selectivity

Selectivity is the ability of the method to measure unequivocally and to differentiate the analyte(s) in the presence of other components that may be expected to be present in the sample. These other components might include metabolites, parent compound(s), impurities, degradation products, matrix components, etc. Sometimes the term specificity is used interchangeably with selectivity, although specificity implies that a method produces a response for a single analyte. Because that is rarely the case, the term selectivity will be used here. Multianalyte methods should be able to differentiate 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 it as well.

Selectivity should be investigated in ten (or more) individual sources of the appropriate blank matrix. The use of fewer sources is acceptable for rare or difficult to obtain matrices. Investigating suspected interferences may be done by spiking these compounds at realistic concentrations into blank matrices or matrices spiked with low levels of the analyte(s) of interest. 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 less than 5% for any internal standard signal [EMA 2011].

While interferences should ideally be kept to a minimum, it is recognized that some biomarkers are produced endogenously while others are present because of background (nonoccupational) exposures. These types of methods can still be invaluable to differentiate between occupationally exposed and nonoccupationally exposed individuals and should, therefore, not be dismissed because of measurable background levels, especially if those levels are known and consistent.

2) **Carryover**

Carryover should be addressed and minimized during method development. "Carryover" is a type of sample contamination. Carryover causes a target analyte response to be generated in later measurements, or runs, where the sample does not actually contain the analyte or contains less than indicated by the intensity of the response. 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 solvent blanks after certain samples. Randomizing samples should be avoided in this case as it may interfere with assessing carryover. While "there is no standard acceptable magnitude of carryover for a passing analytical analysis" [Viswanathan et al. 2007], a good measure would be to use the same guideline as that of the blank, i.e., carryover in the blank sample following the high concentration standard should be less than 20% of the LOQ for the analyte and less than 5% for any internal standard signal [EMA 2011].

3) **Calibration curve**

The calibration or standard curve is the relationship between the instrument response and known concentrations of the analyte. Each analyte studied in the method should have a separate calibration curve. Calibration standards should be prepared in the same matrix as 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. Use the simplest model that adequately describes the concentration-response relationship. 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 nonzero standards. Usually, the blank and zero samples are not used when calculating the calibration curve parameters, but there may be methods where use of the zero sample is appropriate.

More calibration standards may be required to adequately describe a higher-function (nonlinear) 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 nonzero 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 Accuracy section).

4) Bias and precision

Bias is a systematic deviation of the method average or measured value from an accepted value [ASTM 2020a]. 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:

- 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.
- Intermediate precision, sometimes called between-run, between-day, or inter-assay precision, captures 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.

- Long-term precision, or reproducibility, expresses the precision between labs (collaborative study, round-robin, inter-laboratory trial) [ASTM 2022]. The extent to which reproducibility should be established is dictated by the intended use of the method. One must use caution as many authors use the term reproducibility for within-laboratory studies at the level of intermediate precision, which can lead to confusion.

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 after changing lots or manufacturers of solvent, e.g., different chromatography columns or extraction cartridges, is also encouraged, while changing analysts or instrumentation is often less important and sometimes impractical or impossible. For methods that are published in NMAM, the process of an 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 used 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, with acceptable precision of no greater than 15% RSD at the upper levels and no greater than 20% at the lower level.

5) **Accuracy**

Accuracy is the closeness of the determined value obtained by the method to the true or accepted value of the analyte [FDA 2018]. 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. These samples should be spiked from a stock solution of the analyte prepared separately from that used to make up the calibration standards.

Accuracy should be determined 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 2018]. 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, and 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 2018].

6) Stability

An evaluation of stability should be conducted to ensure that steps taken during sample preparation, analysis, and storage, affecting the concentration of the analyte in the matrix, are evaluated so that parameters can be set around the sampling and analysis. 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 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(s) after the applied storage condition(s). 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:

- 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. 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 amber 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 an appropriate time period if solutions would normally be left standing for longer periods) should a system failure require delayed injection of samples stored in an autosampler.

- Long-term stability. The stability samples are stored (most likely in a freezer but it could be in a refrigerator in some circumstances) for at least the same duration as planned for 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 planned for the field/study samples, so if different storage temperatures will be used (some samples at –20 °C and some at –70 °C, for example), then studies at both temperatures should be performed.

- 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.

- 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 an appropriate time period 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.

- 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, considering the anticipated run time for the batch size.

- 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.

7) **Matrix effects**

Matrix effects are especially important for methods that employ mass-spectrometric (MS) detection and should be investigated. The matrix effect in MS is typically because of 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, which is 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, and a value of greater than 1 may be due to ionization enhancement. For a method utilizing internal standards (IS), the MF of the 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 samples 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 matrix lots. The variability in matrix factors (as measured by CV or RSD) should be less than 15% for both 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 adequately mimics the analyte and thereby reduces the variability [Trufelli et al. 2011].

8) Recovery

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 analyte (or internal standard) to the matrix followed by analysis using the method. The method response 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.

The recovery of an analyte should be optimized to ensure efficient and reproducible extraction. While it is often desirable to obtain recovery as close to 100% as possible to maximize the accuracy and sensitivity of a method, recovery need not be 100% as long as the recovery is consistent and reproducible enough to obtain acceptable precision and bias [FDA 2018; UNODC 2009]. Extraction recovery is generally an issue investigated during the analytical method development. The recovery needs to be consistent, precise, and reproducible, while the absolute value for recovery is less important, and is not required to meet any certain threshold and should be fit for purpose. 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.

9) Robustness/ruggedness

Robustness or ruggedness is a measure of the susceptibility of a method to small changes that might occur during routine analysis. Validation of a method need 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 initial ruggedness testing. Warnings can be included in the methodology about parameters that must be tightly controlled 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 step. Some typical examples of variations include pH (of samples, mobile phases, extraction solvents), mobile phase composition, and columns (GC, LC, and extraction columns/cartridges with different lots or suppliers). Other sources can include variations in temperature (room, oven, column, processing, thawing, autosampler, etc.), flow rate (both in extraction and analysis), and various volumes and times used during the entire bioanalytical method.

10) Limit of detection and limit of quantification

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 2018]. 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. 2012; NIOSH 1995]. In brief, one method involves the analysis of 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 (sy) are calculated. The LOD is then 3 sy/slope. Alternatively, analyze at least 10 standards near the anticipated LOD and calculate the standard deviation [ASTM 2020b]. The LOD (MDL) is reported as 3 times the standard deviation of the blank signal (correcting for background). The lower limit of quantitation (LLOQ) is concomitantly taken as 10 times the standard deviation of the blanks.

The LLOQ is the lowest concentration of analyte in a sample that can be quantified reliably with acceptable accuracy and precision. 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 within the defined criteria. In addition, the analyte signal of the LLOQ should be at least 10 times sy/slope (or 3.33×LOD).

The LOD, LLOQ, and ULOQ should be reported in the validation report. These 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 enclose in parentheses to denote the greater imprecision of these results. It is also common to use one less significant figure when reporting results in this area.

11) Acceptance criteria for an analytical run

An analytical run consists of the following components:

- A matrix blank (a processed matrix sample without analyte and without internal standard)
- A zero sample (a processed matrix sample with internal standard)
- A minimum of six nonzero calibration levels
- 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 additional samples are needed
- Field/study/unknown samples

All samples should be processed as a single batch. Calibration standards and QC samples should be spiked independently using separately prepared stock solutions. A single batch is comprised of samples 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 to support detection of drift or carryover 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.

12) Accuracy acceptance criteria for QC samples [EMA 2011; FDA 2018]

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 the case where the criteria are not fulfilled, the analytical run should be rejected. The samples can be re-extracted and analyzed, corrective action can be taken on the instrumentation, or other cause(s) of the failure investigated, 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 is rejected for another analyte, the data for the accepted analyte can be used.

13) Accuracy acceptance criteria for calibration standards [EMA 2011; FDA 2018]

The back-calculated concentrations of the calibration standards should be within ±15% of the nominal value, except for the LLOQ, which 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, the 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 narrower for that run. The revised calibration range should cover all concentrations of QC samples.

8 Conclusions

This chapter describes biomonitoring as a valuable tool for conducting research and public health surveillance, evaluating intervention effectiveness, and improving risk assessments. 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 to evaluate dose response [NRC 2007]. Biomonitoring can be used not only to identify exposures but also to evaluate exposure trends over time. While several international agencies have developed biological exposure levels for some chemicals, OSHA has only mandated three tests for occupational compliance.

Some general considerations for biomonitoring research 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 the biomarker most pertinent for the study. If the goal is to understand whether exposure occurred or not, it may be possible to measure the chemical 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 in occupational research studies will aid in ensuring that studies of biomarkers have scientific rigor and will minimize uncertainty. Written standard operating procedures and a quality assurance program will help improve data quality and provide interpretable results. Adherence to field and laboratory safety procedures are imperative to protect staff from exposure to bloodborne 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 issues. 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 biomarker measurements are providing, and, finally, to determine appropriate actions based on that information.

Disclaimer

Mention of any company or product does not constitute endorsement by the National Institute for Occupational Safety and Health (NIOSH), Centers for Disease Control and Prevention (CDC). 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.

9 References

ACGIH [2022]. 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, https://www.acgih.org/tlv-bei-guidelines/policies-procedures-presentations/overview.

Andersen ME, Clewell HJ III, Gargas ML, MacNaughton MG, Reitz RH, Nolan RJ, McKenna MJ [1991]. Physiologically based pharmacokinetic modeling with dichloromethane, its metabolite, carbon monoxide, and blood carboxyhemoglobin in rats and humans. Toxicol Appl Pharmacol *108*(1):14–27.

Andreasson U, Perret-Liaudet A, van Waalwijk van Doorn LJ, Blennow K, Chiasserini D, Engelborghs S, Fladby T, Genc S, Kruse N, Kuiperij HB, Kulic L, Lewczuk P, Mollenhauer B, Mroczko B, Parnetti L, Vanmechelen E, Verbeek MM, Winblad B , Zetterberg H, Koel-Simmelink M, Teunissen CE [2015]. A practical guide to immunoassay method validation. Front Neurol *6*:179, https://doi.org/10.3389%2Ffneur.2015.00179.

Angerer J, Aylward LL, Hays SM, Heinzow B, Wilhelm M [2011]. Human biomonitoring assessment values: approaches and data requirements. Int J Hyg Environ Health *214*(5):348–360, https://doi.org/10.1016/j.ijheh.2011.06.002.

Apel P, Angerer J, Wilhelm M, Kolossa-Gehring M [2017]. New HBM values for emerging substances, inventory of reference and HBM values in force, and working principles of the German Human Biomonitoring Commission. Int J Hyg Environ Health *220*(2, Pt A):152–166, https://doi.org/10.1016/j.ijheh.2011.06.002.

ANSES [2019]. 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. J ASTM Int *8*(6):9, http://doi.org/10.1520/JAI103678.

ASTM [2020a]. ASTM D1356-20 Standard terminology relating to sampling and analysis of atmospheres. ASTM International: West Conshohocken, PA: American Society for Testing and Materials.

ASTM [2020b]. ASTM D6785-20 Determination of lead in workplace air using flame or graphite furnace atomic absorption spectrometry. ASTM International: West Conshohocken, PA: American Society for Testing and Materials.

ASTM [2022]. ASTM E691-22 Standard practice for conducting an interlaboratory study to determine the precision of a test method. ASTM International: West Conshohocken, PA: American Society for Testing and Materials.

Aylward LL, Hays SM [2011]. Biomonitoring-based risk assessment for hexabromocyclododecane (HBCD). Int J Hyg Environ Health *214*(3):179–187, https://doi.org/10.1016/j.ijheh.2011.02.002.

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, https://doi.org/10.1080/10937404.2013.864250.

Aylward LL, Kirman CR, Adgate JL, McKenzie LM, Hays SM [2012]. Interpreting variability in population biomonitoring data: role of elimination kinetics. J Expo Sci Environ Epidemiol *22*(4):398–404, http://doi.org/10.1038/jes.2012.35.

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*(3):287–294, https://doi.org/10.1289/ehp.1205740.

Bader M, Barr D, Göen T, Schaller KH, Scherer G, Angerer J [2012]. Reliability criteria for analytical methods. In: The MAK-collection for occupational health and safety Part IV: biomonitoring methods (2010). Vol. 12. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, https://doi.org/10.1002/3527600418.bireliabe0012.

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*(2):192–200, https://doi.org/10.1289/ehp.7337.

B'Hymer C, Krieg E Jr., 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*(11):661–672, https://doi.org/10.1080/15287394.2012.688483.

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*(10):615–627, https://doi.org/10.1080/15298669391355134.

Brown JF Jr., Lawton RW [1984]. Polychlorinated biphenyl (PCB) partitioning between adipose tissue and serum. Bull Environ Contam Toxicol *33*:277–280, https://doi.org/10.1007/bf01625543.

Brown PO, Botstein D [1999]. Exploring the new world of the genome with DNA microarrays. Nat Genet *21*(1):33–37, https://doi.org/10.1038/4462.

Calafat AM, Longnecker MP, Koch HP, Swan SH, Hauser R, Goldman LR, Lanphear BP, Rudel RA, Engel SM, Teitelbaum SL, Whyatt RM, Wolff MS [2015]. Optimal exposure biomarkers for nonpersistent chemicals in environmental epidemiology. Environ Health Perspect *123*(7):A166–A168, https://doi.org/10.1289/ehp.1510041.

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*(1):63–67, https://doi.org/10.1007/s004200000190.

Causon R [1997]. Validation of chromatographic methods in biomedical analysis. Viewpoint and discussion. J Chromatogr B Biomed Sci Appl *689*(1):175–180, https://doi.org/10.1016/s0378-4347(96)00297-6.

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, https://doi.org/10.1093/annhyg/mer088.

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, https://doi.org/10.1093/annhyg/mer108.

CDC [2020]. Biosafety in microbiological and biomedical laboratories, 6th ed., Meechan PJ, Potts J, eds. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institutes of Health, https://www.cdc.gov/labs/pdf/CDC-BiosafetyMicrobiologicalBiomedicalLaboratories-2020-P.pdf.

CDC [2021]. 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 [2022]. 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*(1):122–133, https://doi.org/10.1016/j.taap.2008.04.021.

Cone EJ, Caplan YH, Moser F, Robert T, Shelby MK, Black DL [2009]. Normalization of urinary drug concentrations with specific gravity and creatinine. J Anal Toxicol *33*(1):1–7, https://doi.org/10.1093/jat/33.1.1.

DeBord DG, Burgoon L, Edwards SW, 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, https://doi.org/10.1080/15459624.2015.1060324.

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, https://doi.org/10.7205/milmed-d-12-00173.

Demchuk E, Yucesoy B, Johnson VJ, Andrew M, Weston A, Germolec D, De Rosa CT, 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, https://doi.org/10.1289/ehp.8870.

DFG [2015]. List of MAK and BAT values 2015. Report 51. Permanent Senate Commission for the Investigation of Health Hazards of Chemical Compounds in the Work Area. Bonn, Germany: Deutsche Forschungsgemeinschaft (DFG), https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527695539.oth1.

Droz PO [1989]. Biological monitoring I: sources of variability in human response to chemical exposure. Appl Ind Hyg *4*(1):F20–F24, https://doi.org/10.1080/08828032.1989.10389872.

EMA [2011]. Guideline on bioanalytical method validation. Committee for Medicinal Products for Human Use (CHMP). London: European Medicines Agency, http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/08/WC500109686.pdf.

Exley K, Aerts D, Biot P, Casteleyn 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 Int *22*(20):15821–15834, https://doi.org/10.1007/s11356-015-4772-4.

FDA [2018]. Bioanalytical method validation: guidance for industry. Silver Spring, MD: U.S. Department of Health and Human Services, Food and Drug Administration, https://www.fda.gov/media/70858/download.

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, https://doi.org/10.1080/1047322X.1993.10388193.

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 *12*(3):591–599, https://doi.org/10.1039/b921073c.

Goldberger BA, Loewenthal B, Darwin WD, Cone EJ [1995]. Intrasubject variation of creatinine and specific gravity measurements in consecutive urine specimens of heroin users. Clin Chem *41*(1):116–117.

Gosho M, Nagashima K, Sato Y [2012]. Study designs and statistical analyses for biomarker research. Sensors *12*(7):8966–8986, https://doi.org/10.3390/s120708966.

GPO [2021]. Transportation – Hazardous materials regulation. Title 49 Subtitle B Chapter I Subchapter C Part 172. Government Publishing Office, https://www.ecfr.gov/current/title-49/subtitle-B/chapter-I/subchapter-C/part-172.

Green JM [1996]. A practical guide to analytical method validation. Anal Chem *68*(9):305A–309A, https://doi.org/10.1021/ac961912f.

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*(2):193–218, https://doi.org/10.1016/s0731-7085(97)00198-2.

Hathaway GJ, Proctor NH, eds. [2004]. Proctor and Hughes' chemical hazards of the workplace, 5th ed. Hoboken, NJ: John Wiley & Sons, Inc.

HSE [2020]. Biological monitoring guidance values. In: EH40/2005 Workplace exposure limits, 4th edition. Liverpool, UK: Health and Safety Executive.

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 Sci Publ (142):39–46.

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, https://doi.org/10.2174%2F138945010791011947.

IATA [2022]. Dangerous goods. Montreal, Canada: International Air Transport Association, http://www.iata.org/whatwedo/cargo/dgr/pages/index.aspx.

ICH [2005]. Validation of analytical procedures: text and methodology. Q2R1. International Conference on Harmonisation, https://database.ich.org/sites/default/files/Q2%28R1%29%20Guideline.pdf.

Irish National Accreditation Board [2019]. Guide to method validation for quantitative analysis in chemical testing laboratories (ISO 17025). PS15. Dublin, Ireland: Health and Safety Authority, https://www.inab.ie/inab-documents/mandatory-policy/guide-to-method-validation-for-quantitative-analysis-in-chemical-testing-laboratories-iso-17025-.pdf.

JSOH [2016]. Recommendation of occupational exposure limits (2016–2017). Japan Society for Occupational Health. J Occup Health *58*(5):489–518, https://doi.org/10.1539%2Fjoh.ROEL2016.

Landi MT, Caporaso 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 Sci Publ (142):223–236

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

Levine L, Fahy JP [1945]. Evaluation of urinary lead determinations. J Ind Hyg Toxicol *27*:217–223.

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*(5):325–330, https://doi.org/10.1136/oem.47.5.325.

Linder W, Wainer IW [1998]. Requirements for initial assay validation and publication in J. Chromatography B. Editorial. J Chromatogr B Biomed Sci Appl *707*(1–2):1–2, https://pubmed.ncbi.nlm.nih.gov/9613926/.

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, https://doi.org/10.1038/ng1071.

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, https://doi.org/10.1021/es404853h.

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, https://doi.org/10.1093/annhyg/mes058.

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, https://doi.org/10.1093/carcin/bgr297.

Menke A, Muntner P, Silbergeld EK, Platz EA, Guallar E [2009]. Cadmium levels in urine and mortality among U.S. adults. Environ Health Perspect *117*(2):190–196, https://doi.org/10.1289/ehp.11236.

Mikkelsen SR, Cortón E [2004]. Validation of new bioanalytical methods. In: Bioanalytical chemistry. Hoboken, NJ: John Wiley & Sons, Inc., https://doi.org/10.1002/0471623628.ch16.

Miller RC, Brindle E, Holman DJ, Shofer J, Klein NA, Soules MR, O'Connor KA [2004]. Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations. Clin Chem *50*(5):924–932, https://doi.org/10.1373/clinchem.2004.032292.

Needham LL, Ozkaynak H, Whyatt RM, Barr DB, Wang RY, Naeher L, Akland G, Bahadori T, Bradman A, Fortmann R, Liu LJS, 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, https://doi.org/10.1289/ehp.7613.

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, https://www.cdc.gov/niosh/docs/95-117/default.html.

NIOSH [1998]. Application of biological monitoring methods: Chapter F. By Teass AW, Biagini RE, DeBord DG, Hull RD. In: Eller PM, Cassinelli ME, O'Connor PF, 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. 98-119, https://www.cdc.gov/niosh/docs/2003-154/pdfs/chapter-f.pdf.

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, https://www.cdc.gov/niosh/docs/2010-101/default.html.

NIOSH [2012]. Medical surveillance for health care workers exposed to hazardous drugs. 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. 2013-103, http://www.cdc.gov/niosh/docs/wp-solutions/2013-103/pdfs/2013-103.pdf.

NIOSH [2021]. 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 [1987]. Biological markers in environmental health research. Committee on Biological Markers of the National Research Council. Environ Health Perspect *74*:3–9, https://doi.org/10.1289/ehp.74-1474499.

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

NRC [2007]. Toxicity testing in the 21st century: a vision and a strategy. National Research Council, 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, https://doi.org/10.1080/1047322x.1993.10388168.

OSHA [no date]. 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. Date accessed: July 21, 2022.

OSHA [1978]. Lead. Occupational Health and Safety Standard 1910.1025. 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. Occupational Health and Safety Standard 1910.1028. 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. Occupational Health and Safety Standard 1910.1027. 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 [1992]. Bloodborne pathogens. Occupational Health and Safety Standard 1910.1030. Washington, DC: U.S. Department of Labor, Occupational Safety and Health Administration, https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.1030.

Peters FT, Maurer HH [2015]. Review: bioanalytical method validation—How, how much and why? https://orion.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, https://doi.org/10.1080/1047322X.1994.10388324.

Rosenberg J, Fiserova-Bergerova V, Lowry LK [1989]. Biological monitoring IV: measurements in urine. Appl Ind Hyg *4*(4):F16–F21, https://doi.org/10.1080/08828032.1989.10390349.

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, https://doi.org/10.1080/15459624.2014.955179.

Schulte PA [2005]. The use of biomarkers in surveillance, medical screening, and intervention. Mutat Res, Fundam Mol Mech Mutagen *592*(1–2):155–63, https://doi.org/10.1016/j.mrfmmm.2005.06.019.

Schulte PA, DeBord DG [2000]. Public health assessment of genetic information in the occupational setting. In: Khoury M, Burke W, Thomson E, eds. Genetics and public health in the 21st century. New York: Oxford Press, https://doi.org/10.1093/acprof:oso/9780195128307.001.0001.

Schulte PA, Hauser JE [2012]. The use of biomarkers in occupational health research, practice, and policy. Toxicol Lett *213*(1):91–99, https://doi.org/10.1016/j.toxlet.2011.03.027.

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 Sci Publ (142):19–29.

Schulte PA, Whittaker C, Curran CP [2015]. Considerations for using genetic and epigenetic information in occupational health risk assessment

and standard setting. J Occup Environ Hyg *12*(Suppl 1):S69–S81, https://doi.org/10.1080%2F15459624.2015.1060323.

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*(1):26–35, https://doi.org/10.1016/j.ijheh.2011.06.007.

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, https://doi.org/10.1289/ehp.0900866.

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

Sexton K, Needham LL, Pirkle JL [2004]. Human biomonitoring of environmental chemicals: measuring chemicals in human tissues is the "gold standard" for assessing the people's exposure to pollution. Am Sci *92*(1):38–45, http://www.jstor.org/stable/27858331.

Smeraglia J, Baldrey SF, Watson D [2002]. Matrix effects and selectivity issues in LC-MS-MS. Chromatographia *55*:S95–S99, http://doi.org/10.1007/BF02493363.

Spencer K [1986]. Analytical reviews in clinical biochemistry: the estimation of creatinine. Ann Clin Biochem *23*(Part 1):1–25, https://doi.org/10.1177/000456328602300101.

St-Amand A, Werry K, Aylward LL, Hayes SM, Nong A [2014]. Screening of population level biomonitoring data from the Canadian Health Measures Survey in a risk-based context. Toxicol Lett *231*(2):126–134, https://doi.org/10.1016/j.toxlet.2014.10.019.

SUVA [2019]. Limit values at the workplace. MAK/BAT values (explanations) physical influences, physical loads. Swiss Accident Insurance Fund (SUVA) , https://www.suva.ch/de-CH/material/Richtlinien-Gesetzestexte/erlaeuterungen-zu-den-grenzwerten, [in German].

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*(4):217–224, https://doi.org/10.1016/j.lab.2003.12.011.

Theodorsson E [2012]. Validation and verification of measurement methods in clinical chemistry. Bioanalysis *4*(3):305–320, https://doi.org/10.4155/bio.11.311.

Thomas RS, Clewell HJ III, 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, https://doi.org/10.1093/toxsci/kfq355.

Tiwari G, Tiwari R [2010]. Bioanalytical method validation: an updated review. Pharm Methods *1*(1):25–38, https://doi.org/10.4103%2F2229-4708.72226.

Trufelli H, Palma P, Famiglini G, Cappiello A [2011]. An overview of matrix effects in liquid chromatography-mass spectrometry. Mass Spectrom Rev *30*(3):491–509, https://doi.org/10.1002/mas.20298.

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):409–431, https://doi.org/10.1080/15287394.2014.971477.

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. AAPS J *9*(1):E30–E42, http://doi.org/10.1208/aapsj0901004.

WADA [2016]. Endogenous anabolic androgenic steroids. Measurement and reporting. World Anti-Doping Agency Technical Document (WADA) TD2016EAAS:1–16, https://www.wada-ama.org/sites/default/files/resources/files/wada-td2016eaas-eaas-measurement-and-reporting-en.pdf.

WHO [2001]. Environmental Health Criteria 222. Biomarkers in risk assessment: validity and validation. Geneva, Switzerland: World Health Organization, https://inchem.org/documents/ehc/ehc/ehc222.htm.

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*(1–2):381–394, https://doi.org/10.1016/0021-9673(96)00006-4.

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, https://doi.org/10.1016%2Fj.envres.2014.09.009.

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, https://doi.org/10.1016/j.ijheh.2016.10.007.

Zong G, Grandjean P, Wu H, Sun Q [2015]. Circulating persistent organic pollutants and body fat distribution: evidence from NHANES 1999–2004. Obesity (Silver Spring) *23*(9):1903–1910, https://doi.org/10.1002/oby.21161.