Biomonitoring in the Era of the Exposome

Kristine K. Dennis,1 Elizabeth Marder,1 David M. Balshaw,2 Yuxia Cui,2 Michael A. Lynes,3 Gary J. Patti,4,5 Stephen M. Rappaport,6 Daniel T. Shaughnessy,2 Martine Vrijheid,7 and Dana Boyd Barr1

1Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; 2Exposure, Response, and Technology Branch, Division of Extramural Research and Training, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina, USA; 3Department of Molecular and Cell Biology, College of Liberal Arts and Sciences, University of Connecticut, Storrs, Connecticut, USA; 4Department of Chemistry, Department of Medicine, Washington University, St. Louis, Missouri, USA; 5Department of Environmental Health Sciences, School of Public Health, University of California, Berkeley, Berkeley, California, USA; 7Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain

Introduction

More than ten years ago, shortly after the human genome was sequenced, Christopher Wild proposed an environmental complement to the genome in determining risk of disease, termed the exposome. He defined the exposome as the totality of exposures throughout the lifespan (Wild 2005)

Since the exposome was originally defined, research efforts have begun, leading to a revised working definition that may be summarized by the following elements. The exposome includes the cumulative measure of exposures to both chemical and nonchemical agents such as diet, stress, and sociobehavioral factors. It includes a series of quantitative and repeated metrics of exposures—both endogenous and exogenous—that describe, holistically, environmental influences or exposure over a lifetime (from conception to death). The exposome can include traditional measures of exposure (e.g., traditional biomonitoring, environmental monitoring) but also includes unintended discovery of unknown chemicals of biological importance (Miller and Jones 2014; Rappaport and Smith 2010; Wild 2005, 2012). Exposomic approaches go a step beyond traditional biomonitoring, aiming to capture all exposures that potentially affect health and disease.

As a cancer epidemiologist, Wild understood the importance of the environment to health and that current disease trends cannot be explained by genetics alone (Wild 2005). We are only beginning to understand the complexities of environmental exposures and their impacts on human health, whereas genetic influences on health have been extensively studied. At present, we have limited estimates of the impact of environmental exposures on health, and uncertainty exists even in those (Jones 2016; Rappaport 2016; Rappaport and Smith 2010). Biomonitoring serves as a key tool to define exposure–disease risks given the biological significance of internal exposure measurements. With the continued advancement of methods, biomonitoring strategies will be critical in achieving a comprehensive understanding of exposures that have personal and public health relevance. With full understanding of the complex interactions between genetics and environmental exposures, the mysteries of the etiology, trends, and prevention of many diseases can be solved.

In an effort to advance the framework for developing exposome approaches and characterization, a diverse group of scientists gathered at the National Institute of Environmental Health Sciences (NIEHS) Exposome Workshop in January 2015 to discuss the current state of the science and to provide recommendations to the environmental health sciences community on how to best advance exposome research. The state of the science along with the perspectives and recommendations of our working group, Biomonitoring for the Exposome, are described here.

Discussion

Traditional Biomonitoring

Exposure is commonly assessed by a spectrum of questionnaire data and ecological, environmental, or biological measurements.

Acknowledgments and Funding

The authors declare they have no actual or potential competing financial interests.

This work was supported by the National Institutes of Health (NIH), the National Institute of Environmental Health Sciences (NIEHS) and NIH grant P30 ES019776.

This manuscript is based upon the work of the Biomonitoring Working Group of the NIEHS Exposome Workshop held 14–15 January 2015 in Research Triangle Park, NC, USA.

This work was supported by the National Institutes of Health (NIH), the National Institute of Environmental Health Sciences (NIEHS) and NIH grant P30 ES019776.

Address correspondence to D.B. Barr, Department of Environmental Health, Rollins School of Public Health, 1518 Clifton Rd. NE, Mailstop: 1518:002-28B, Emory University, Atlanta, GA 30322 USA. Telephone: (404) 727-9065. E-mail: ddbarr@emory.edu

This manuscript is based upon the work of the Biomonitoring Working Group of the NIEHS Exposome Workshop held 14–15 January 2015 in Research Triangle Park, NC, USA.

This work was supported by the National Institutes of Health (NIH), the National Institute of Environmental Health Sciences (NIEHS) and NIH grant P30 ES019776.

K.K.D. and D.B.B are supported by Emory University’s Health and Exposome Research Center: Understanding Lifetime Exposures (HERCULES; grant P30 ES019776). M.A.L. has a patent that he shares with Ciencia, Inc. for a biomonitoring instrument that uses both grating-coupled surface plasmon resonance (GCSPR) imaging and grating-coupled surface plasmon–coupled emission (GCSPCE) imaging in a microarray format for the analysis of functional cell phenotyping. M.A.L. has consulted for Ciencia, Inc. in the past, but is not currently compensated as a consultant. M.A.L. also has had (and currently has) NIH/NIEHS support to develop this technology. G.J.P. is a scientific advisory board member for Cambridge Isotope Laboratories. All other authors declare they have no actual or potential competing financial interests.

Received: 28 January 2016; Revised: 10 May 2016; Accepted: 21 June 2016; Published: 6 July 2016.

Note to readers with disabilities: EHP strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in EHP articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehponline@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.
Biological measures of exposure that determine an internalized dose are often preferred because they are usually more relevant to the health outcome studied. Traditional biological measurements, also called targeted analyses, measure a target chemical, metabolite, or reaction product in a biological medium such as urine or blood (see Appendix 1). These traditional biomonitoring measurements have become a key component of exposure assessment in many epidemiologic studies that attempt to link exposures to health outcomes.

Molecular epidemiology studies and regulatory agencies rely primarily on targeted analyses because of their current availability and historical use. Broad surveys such as the National Health and Nutrition Examination Study (NHANES) utilize these methods, allowing for quantification and longitudinal surveillance of known exposures across the U.S. population. NHANES data facilitate comparative identification of abnormal exposure levels in select population subsets. Major epidemiology studies such as those evaluating blood lead levels and mean IQ in children and prenatal pesticide exposures and neurological deficits in children and neurodegenerative disease in adults have linked significant health outcomes to specific exposures, informing opportunities for further mechanistic studies (Chin-Chan et al. 2015; Kaufman et al. 2014; Rosas and Eskenazi 2008). Other federal efforts in the United States include the National Biomonitoring Program (NBP) of the Division of Laboratory Sciences at the Centers for Disease Control and Prevention (CDC). The NBP produces a National Report on Human Exposure to Environmental Chemicals and regularly updates the NHANES biomonitoring data in that report (CDC 2009, 2015). Chemicals of potential concern such as arsenic, perchlorate, and environmental phenols, among others, continue to be added to NHANES, with the most recent report including data on > 250 chemicals. The CDC also provides grant funding to a variety of state laboratories to increase public health laboratory capacity for surveillance. Targeted analytical capabilities and worldwide use continue to expand through both public health and academic entities.

Historical use of biomonitoring. Traditional biomonitoring methods are well established for exposure assessment in epidemiology studies and in federal and state surveillance activities. Because of their historical use, they provide a number of strong advantages for exposure research (see Appendix 2). Biologically persistent chemicals are well-characterized with traditional methods, whereas short-lived chemicals are effectively measured only if the individual is undergoing continuous or continual exposures or if the timing of exposures is known. Chemicals such as phthalates, bisphenols, and parabens are well-characterized by targeted methods given their widespread use and presence in the environment. Often, chemicals of particular toxicological interest may be difficult to measure owing to barriers such as stability or presence in readily accessible biological matrices such as blood or urine. For example, short-lived chemicals such as various current-use pesticide and phthalate metabolites can only be detected in urine samples if exposure occurs within a few days of testing; therefore, continuous or longitudinal sample collection is necessary to capture exposure. For a selected group of 250–300 known persistent (~30–40%) and nonpersistent (~60–70%) chemicals, sample analysis provides exposure information for the chemical of concern within a specific window of exposure; reference data are available for most of these chemicals (CDC 2015).

The ~250 chemicals that are commonly measured in the United States are primarily driven by the CDC biomonitoring list of target analytes (CDC 2015). Most other programs also follow the CDC list because selection of these agents was informed by a public nomination process followed by expert ranking of the nominated chemicals (CDC 2012). An important caveat of this process is that the target list is partially based on ease of performance and compatibility with existing methods. Another concern is that some of the chemicals have little toxicological relevance or diminishing exposure across the population resulting from successful regulation of their release into the environment, or a combination of the two.

Biomonitoring methods. Although method development for traditional biomonitoring can be quite rigorous, this also translates into a slow and expensive process when developing analysis protocols for new chemicals of interest. These analyses often require relatively high volumes of sample, typically 0.5–1 mL for a single method (~10 mL urine and > 20 mL serum to measure the 250–300 currently biomonitored chemicals), which can be limiting for certain biospecimen types and age groups under study. For exposome research, these requirements restrict the number and types of chemicals that can be measured at any one time. Unknown or suspected chemicals of concern may not be measured or identifiable through targeted methods (see Appendix 2) (Rappaport et al. 2014); yet targeted analyses are valuable given the accuracy and depth at which a chemical of interest can be assessed. By coupling traditional biomonitoring methods with broader exposomic approaches, the benefits of both strategies can be fully realized.

Exposomic Approaches

An exposomic approach differs from traditional biomonitoring in that it can theoretically include all exposures of potential health significance, whether they are derived from exogenous sources (e.g., pollutants, diet, drugs) or endogenous sources (e.g., hormones, human and microbial metabolites) (Rappaport and Smith 2010; Rappaport et al. 2014). Because levels of chemicals in blood or other biospecimens reflect a wide range of exposures or the metabolic consequences of exposures, including psychosocial stress, other nonchemical stressors such as noise, and nutritional factors, exposomic biomonitoring offers an efficient means for characterizing individual exposure profiles. Incorporating the exposome paradigm into traditional biomonitoring approaches offers a means to improve exposure assessment in many ways (Wild 2012).

Untargeted analyses. With only a few hundred chemicals routinely assessable through targeted methods and with limitations for short-lived compounds, exposomic approaches are critical to understanding the thousands of chemicals people are exposed to daily through direct chemical exposures or consequences of exposure (e.g., cortisol levels due to stress or noise exposures) (CDC 2015). Through untargeted biomonitoring approaches such as high-resolution metabolomics (HRM), > 1,500 metabolites can be monitored with a relatively small amount of biological specimen (≤ 100 µL) and for the cost of a single traditional biomonitoring analysis of 8–10 target chemicals (Johnson et al. 2010; Jones 2016).

Untargeted analyses of small molecules or macromolecular adducts in blood, urine, or other matrices are well suited for exposome-wide association studies (EWAS), which compare profiles of hundreds or thousands of chemical features—analogous to ions with a given mass-to-charge ratio and a specified retention time in traditional biomonitoring—between diseased and healthy subjects (Rappaport 2012, 2016). Indeed, untargeted analyses performed using the current generation of liquid chromatography–high resolution mass spectrometers (LC-HRMs) can detect > 30,000 small-molecule features (Ivanisevic et al. 2013) and > 100 human serum albumin (HSA) adducts of reactive electrophilic chemicals (including reactive oxygen species) at the nucleophilic locus Cys34 (Grigoryan et al. 2012; Rappaport et al. 2012). Processing the rich sets of data from untargeted analyses of archived biospecimens offers a path for discovering health-imparing exposures that have thus far escaped scrutiny, a largely unrecognized benefit of exposomics. It is important to note that full annotation of molecular features is not required.
for case–control comparisons provided that LC-HRMS signatures are available (e.g., accurate mass, retention time, and MS/MS fragmentation). Archived biospecimens from well-designed cohort studies already exist. With continued advancement in untargeted analyses, there is potential to make significant advances in human health through uncovering unknown exposures (da Silva et al. 2015; Zhou et al. 2012).

**High-resolution metabolomics.** Although untargeted analyses encompass a wide range of the -omics techniques, HRM is a technique that is poised to advance exposomics research because of the breadth of coverage it offers of both endogenous and exogenous chemicals. At the present time, it is routine to detect tens of thousands of features with HRM, and this number will increase as the sensitivity of mass analyzers continues to improve. These features do not necessarily represent different chemical constituents but provide extensive data for evaluation of alterations in biological pathways (Mahieu et al. 2014). Extensive comparisons of the features of these various instruments are available elsewhere (Marshall and Hendrickson 2008).

With the additional advancements that have been made in bioinformatics methods to aid in feature extraction and data analysis, HRM has become an increasingly viable tool for broad exposome-level characterization (Jones 2016). Although features linked to human health will require chemical identification, the technology is in place for the feature extraction methods and annotation efforts that will increase the total number of chemicals that can be monitored (Soltow et al. 2013). Researchers are already demonstrating this expanded potential along with the ability to quantify chemicals under a high-resolution metabolomics platform (Go et al. 2015; Li et al. 2015). By definition, untargeted approaches are agnostic, allowing unknown or emerging exposures of concern (see Appendix 3) to be detected. These approaches are often hypothesis-generating and may require testing of newly discovered analytes/exposures in experimental models to confirm effects on biological responses.

**Detection of low-level xenobiotic exposures.** Persistent challenges exist with detecting chemicals present at low levels, defining reference values of “normal” exposure, and ultimately linking these measures to an exogenous source so intervention can occur. Because blood concentrations of xenobiotics (femtomolar to micromolar) tend to be much lower than those of chemicals derived from food, drugs, and endogenous sources (nanomolar to micromolar), untargeted analyses of xenobiotics are not as efficient and reliable as those of ingested and endogenous sources at detecting many exposures of interest (Rappaport et al. 2014). To determine the health impacts of these exposures, it will be necessary to develop semi-targeted or multiplexed methods that increase the signals of exogenous molecules relative to those of endogenous origin (Rappaport et al. 2014; Southam et al. 2014; Wei et al. 2010). Analyses of suspected chemicals of concern, also referred to as suspect screening, can be prioritized through measuring panels of chemicals with known biological effects but without identifying a specific hypothesis regarding the toxicological pathway. Untargeted and suspected chemical analyses both fall under exposomic biomonitoring and offer extraordinary potential for increased understanding of complex chemical exposures.

**Hybrid approaches.** Various terms are used to describe hybrid approaches, including suspect screening or semi-targeted analyses. Because both targeted and untargeted approaches have beneficial attributes as well as drawbacks, using a hybrid exposomics approach may enable us to exploit advantages while minimizing the limitations of each technique. One of the obvious limitations of a targeted approach is its inability to provide exposure information on a wide array of chemicals. However, targeted analysis can typically provide validated and quality-assured detection and quantification at very low concentrations that may not be available using an untargeted approach until HRM and the necessary bioinformatic data extraction techniques mature. As mentioned above, the development of these quantitative techniques for HRM is underway with advancements in instrumentog (Go et al. 2015; Marshall and Hendrickson 2008). Furthermore, the generic extraction methods used in untargeted analysis may not be able to capture all of the chemicals of interest (e.g., limited extraction of nonpolar chemicals using a typically polar solvent extraction), whereas more specialized extractions can specifically target chemical classes.

**Semi-targeted analysis.** Semi-targeted analysis can utilize various approaches including a two-step strategy: discovery using metabolomics followed by a more fully quantitative targeted measurement. Another potential approach would involve a known or measured chemical exposure in individuals for which metabolomic measurements could also be made. For instance, untargeted metabolomic analysis of each group would then allow for a search for new exposure biomarkers and unique metabolic pathway perturbations to help elucidate the effect mechanism.

Traditionally, targeted analysis data have been used for risk assessment purposes, so shifting solely to a newer platform may take some time. The hybrid approach can be useful for both exposomic analysis and informing targeted analysis approaches. For example, a targeted chemical concentration can be used as an “outcome” for metabolome-wide association studies (e.g., evaluating biochemical alterations relative to targeted chemical concentrations), or a metabolomic analysis can help identify important chemicals that need to be rigorously quantified for health or risk assessments. Of course, each of the two approaches stands on its own, and they have done so for many decades. By combining the two, however, we have a much more powerful approach to understanding chemical exposures, biological alterations, and disease.

**Overarching Issues**

**Matrix selection.** Whether using a traditional biomonitoring or an exposomic approach, careful attention must be given to which matrices can be practically collected and which matrices are relevant for assessing chemical exposures. The matrices available for collection during different life stages and a nonexhaustive list of the chemicals that are appreciably present in these matrices have been reviewed elsewhere (Barr et al. 2005). Typically, the least-invasive matrix in which the chemicals appreciably collect, such as blood and urine, is the preferred matrix.

Although most analysis of exposure is performed with urine or blood samples as a consequence of the ease with which they can be collected, there are other sample types that have begun to be explored for their value in exposome interrogation. For example, saliva, which can be collected from school-age children and adults, is a problematic matrix to collect from infants and toddlers because of choking dangers associated with the collection devices and the inability of young children to actively secrete it. Even if the matrix, in this case saliva, can be noninvasively collected, the target chemical or suite of potential chemicals may not enter the matrix for a variety of reasons, including protein binding of chemicals that prevents their secretion into saliva (Lu et al. 1998). In addition, saliva is nonsterile; therefore, contributions of the oral microbiome can influence the composition of the analytes to be measured. Buccal and nasal swabs have also been used to assess the biological consequences of external exposures. In those sample types, DNA, mRNA, and their adducts have been the principal focus to date (Beane et al. 2011; Spira et al. 2004; Zhang et al. 2010), but these samples (as well as fecal samples) are also compromised by the presence of a strong microbial community that can influence the composition of the exposome constituents.

Other biological samples (e.g., selected blood cells, sweat, teeth, nails) can include information about recent historical exposures in their composition. The use of alternative
samples as historical measures of exposure may become important in future studies. Teeth are a matrix that has demonstrated particular promise for characterizing prenatal exposures to metals and to some organic chemicals because of their defined growth patterns (Andra et al. 2015). We can use the “record” of prior exposures recorded in hair, deciduous teeth, or molecular “fingerprints” in other samples to provide historical measures of certain exposures (Arora et al. 2012; Hu et al. 2007); however, validation of the time represented in the exposure history may be laborious.

There are limitations to these sample sets because external deposits of specific chemicals can make the interpretation of measured levels in these samples different from those measured in blood, for example. In addition, standardized protocols and reference standards are lacking for many alternative matrices, making standardization of results across studies difficult.

An important consideration when choosing samples for exposome-type research is the anticipated presence of the particular chemical(s) in the harvested samples. Because chemicals may display unanticipated pharmacodynamics and biotransformation, it may ultimately be essential that multiple sample types are collected from each individual in the effort to define the exposome. Blood circulates throughout the body, so there is an advantage to its assessment because it has been exposed to the myriad of routes by which an environmental chemical may enter the body. However, some analytes are known to specifically accumulate in particular tissues; thus, a broad-spectrum assessment of multiple patient samples will provide the best insights into exposures.

Analytical considerations for matrix effects. In addition to the relevant matrices that can be collected, we must consider the alterations in response that may be obtained in analytic systems related to other components of the matrix. Such matrix effects can enhance analytic signals or work to suppress signals (Panuwet et al. 2016). In fact, each individual sample will exert its own matrix effects that can make quantification difficult, particularly in mass spectrometry–based methods. Mass spectrometers are inherently sensitive to matrix effects such that the analytical signal of a given concentration can vary over orders of magnitude if appropriate internal standards for normalizing the mass spectral signal have not been used (Baker et al. 2005). In particular, these variations could present challenges when attempting to quantify features in untargeted analysis approaches.

Sample collection and storage. Collection and storage procedures are particularly important considerations for internal exposure measurements. Failures in the proper collection and storage of specimens can result in lost sample integrity, samples that are not suitable for analysis, and contamination/degradation of important chemicals. Because of the sensitivity of some methods such as HRM, biospecimens must be carefully collected and well-maintained. Specific attention to freeze-thaw cycles, potential contamination risks, and collection protocols is needed to ensure that the data extracted from each sample are accurate. It is nearly impossible to control for every preanalytic challenge in sample collection and storage for an untargeted analysis, which is one reason that both targeted and untargeted analyses are quite complementary. In addition, both targeted and untargeted approaches can only measure a limited amount of the exogenous and endogenous chemicals that are present in our bodies. The types and number of chemicals within us that are measurable largely depend upon the matrix selected and the method used [Children’s Health Exposure Analysis Resource (CHEAR) 2016a, 2016b].

Variability of exposures. Temporal variability. Temporal, spatial, and genetic variability and variability in biological distribution of chemicals are important elements to characterize in internal exposure studies. It is important to understand if a single sample obtained at a given life stage represents average exposure over time (e.g., a blood sample for dichlorodiphenyldichloroethylene (DDE) measurements during adulthood obtained during a time of much physiologic change, such as pregnancy), or if peak exposures during a critical window are more important to consider. For short-lived chemicals, new technologies and approaches that facilitate collection of real-time data, high-dimensional analyses, and uncovering biological response markers of transient exposures offer strategies for capturing historically difficult measurements (Dennis et al. 2016).

Spatial variability. In addition, it is important to understand how temporal variability may vary over geographic areas and in different exposure scenarios. In this respect, exposure assessment can become very complex. Multiple samples within a population are generally preferred over a single sample so that both temporal and spatial variability can be assessed; however, the collection of multiple samples is often cost-prohibitive and can be an undue burden on participants. To appropriately interpret internal exposure data in the context of risk or health outcome, it is imperative to ascertain the degree of variability in space and time.

Pharmacodynamic variability. Ideally, we would have information on variability in pharmacodynamics to potentially evaluate resulting exposure data (e.g., whether a given chemical distributes to tissues differently among individuals). Most of the pharmacodynamic information we have for specific chemicals is derived from animal studies, and these processes may not be the same in humans. In addition to exposure and pharmacokinetic variability, laboratory and sampling variability should also be assessed and, if possible, teased apart from true intra-person variability.

Fit-for-purpose use. The “fit-for-purpose” concept has gained popularity in traditional biomonitoring (Lee et al. 2006). This concept addresses the balance between overall cost of analysis and the degree of analytical rigor required to use the internal exposure measure results for a given purpose. In instances where legal implications exist or regulatory decisions are to be made, maximum analytical rigor is required. But for exploratory studies and for many epidemiologic studies, statistical power derived from a larger number of samples, but with sufficient precision to detect differences, is often preferred. In these cases, relaxation of analytical rigor may translate into lower costs that, in turn, could enable the number of samples analyzed to increase. Furthermore, in untargeted approaches, authentic standards are not always necessary to evaluate a chemical’s relationship to disease or alterations in biomolecular concentrations. In addition, many “add-on” studies use samples collected for different analyses for which the sample collection/storage may represent more precision, thus not warranting the increased cost of strict analytical rigor.

For each given study or study question, it is important to consider the analysis and the criteria that are necessary to meet the study objectives. For example, if a study seeks to control for smoking but needs validation of the questionnaire, a low-resolution method such as an immunoassay for molecular indications of smoking may be most suitable for the study; this would maximize the money available for other needs in the study. Many times, substantial resources are dedicated to perfecting an analytic method rather than using a portion of those funds to determine which measurements are actually critical to answering a research question. The issue of balance in analytic rigor and cost needs to be addressed in each study.

Extant data also represent a “fit-for-purpose” approach. Extant data were often collected to answer a certain set of research questions; thus, they are not always applicable to a different study question. However, extant data do represent a source for generating hypotheses that can be further tested using prospective, longitudinal studies. For example, NHANES data offer a resource to evaluate the extent of U.S. population exposures to particular chemicals and can serve as a tool for the exposure component of risk assessment. Although the data are cross-sectional, they serve as a useful hypothesis-generating resource.
Unknown analytes. Characterizing unknown analytes remains a major challenge for understanding the exposome. Research efforts should prioritize the development of methods to determine relevant exposures and to identify sources of specific chemical signatures. By linking shifts in the microbiome, the metabolome, the proteome, and so forth to unknown analytes, we can start to determine the profiles of unknown toxicant exposures and their consequences. Additionally, biomonitoring techniques that can assess changes in cellular composition or in the developmental capacity of cells may indicate risks for later health conditions such as cancer and neurodegenerative diseases. Even if the identity of an analyte is unknown, linking unknown exposures to potential disease consequences creates further support for the investment of resources necessary to understanding cumulative lifetime exposures.

Annotation of spectra for unknown chemicals can be quite time-consuming and therefore only completed on a select number of features. Limitations regarding chemical annotation will best be overcome through a concerted effort across many research groups to identify, catalogue, and disseminate information related to newly identified small molecules. Additionally, continued focus on bioinformatics techniques to extract information about chemical features of importance will allow semi-targeted approaches to be used for unknown and low-abundance chemicals.

The omics technologies all have potential for discovering unknown analytes. Through ongoing advancements in mass spectrometry, low-abundance chemicals can be targeted and characterized. With comprehensive coverage of the metabolome, reference metabolic profiles combined with health outcome data would provide a baseline for identification of unknown analytes with health relevance. Through a concerted effort across laboratories, identification and cataloguing unknown analytes will become a tangible task for advancing the exposome.

Overcoming Gaps and Barriers to Exposome Research

Several data gaps or barriers exist in both targeted and untargeted analyses. For untargeted analyses, the ability to identify and quantify low-abundance analytes—most environmental chemicals—is still immature. Untargeted approaches may need new, more sensitive mass spectrometric approaches or chemoselective probes to improve the detection of low-abundance chemicals. We reemphasize that analytic standards are not required for discovery of new and relevant biomarkers; they become necessary only when a new biomarker is identified and needs to be validated.

Although many biomonitoring resources are available through public health and academic laboratories, few laboratories exist with the capacity to measure a wide array of “known” toxicants, particularly in nonstandard matrices (i.e., matrices other than blood and urine) (see Appendix 4). Having access to such capacity is particularly important for new investigators, who may not have established relationships with such laboratories. Additionally, accurate and reproducible measures across laboratories remain a challenge. The CHEAR initiative, led by the National Institute of Environmental Health Sciences, represents a unique opportunity to provide a standardized laboratory network with access to targeted and untargeted analyses of biospecimens and so may serve to fill these gaps (NIEHS 2015).

Datasets. The application of untargeted metabolomics to identify environmental exposures correlated with human health has its own unique challenges. The largest reference databases for metabolomics are the Metabolite and Tandem MS Database (METLIN) and the Human Metabolome Database (HMDB) (Tautenhahn et al. 2012; Wishart et al. 2009). To date, METLIN and HMDB have largely focused on naturally occurring metabolites. To our knowledge, the number of compounds in METLIN and HMDB that may be potentially relevant to exposure studies has not yet been carefully assessed. The number of databases available for metabolomics continues to expand and has unique utility depending on the research question. A more expansive discussion of metabolomics database resources is available (Go 2010). To facilitate large-scale exposome studies, the field may benefit from having a database or from having database search functionalities specifically dedicated to environmental exposure chemicals. As discussed above, discovery experiments are typically most successful when a small subset of features can be targeted for structural identification. Thus, databases and repositories curating information on the human exposome would provide powerful mechanisms for prioritizing features of interest to environmental health scientists.

Bioinformatic Approaches

Although bioinformatics were covered under the scope of the Biostatistics and Informatics Workgroup at the NIEHS Exposome Workshop, it is worthwhile to mention a few bioinformatic needs that are specific to the development of exposomic biomonitoring approaches. As highlighted throughout this article, characterizing the complexities of the exposome requires use of...
Biomonitoring in the era of the exposome

broad coverage techniques to link internal biochemical perturbations to external exposures. Bioinformatic requirements for these types of data analyses are substantial, yet they offer a high return on investment. Through pathway analysis and data extraction algorithms, biological pathway perturbations can provide great insight into broad disease processes. Additionally, detection of low-level xenobiotic and unknown chemicals can be greatly enhanced through bioinformatic techniques. Further development of bioinformatic tools and data storage and handling will be key to advancing our understanding of the health impact of complex exposures.

**Implementing the Exposome**

External exposures and the actual body burden of said exposures can be quite variable. There is much to be learned about combining external and internal measures to maximize understanding of exposure and how to mitigate exposures that have negative health consequences. Coupling technologies and utilizing real-time monitoring tools can increase our overall understanding of exposures spatially and temporally. Exposome studies in Europe such as The Human Early-life Exposome (HELIX); Health and Environment-wide Association Studies based on Large population Surveys (HEALS); and EXPOsOMICS have started to demonstrate specific approaches for capturing this type of information [Community Research and Development Information Service (CORDIS) 2014, 2015; Vrijheid et al. 2014].

Similarly, Emory University’s NIEHS-funded Human Exposome Research Center: Understanding Lifetime Exposures (HERCULES) has developed infrastructure that has supported several environmental health studies using hybrid biomonitoring approaches (Go et al. 2014, 2015; Jones 2016; Zhang et al. 2014). HELIX also uses a hybrid approach for data collection. HELIX specifically focuses on cohorts of mother-child pairs to better understand which developmental periods may be particularly vulnerable to environmental exposures (Vrijheid et al. 2014). Along with personal external exposure monitoring strategies, traditional biomonitoring techniques have been combined with untargeted omics analyses (e.g., metabolomics, proteomics, transcriptomics, epigenomics) with a particular focus on repeat sampling to capture nonpersistent biomarkers. By performing omics–exposure and omics–health association studies, researchers aim to uncover biologically meaningful omics signatures. The HELIX design is one example of a current approach that integrates traditional and nontraditional techniques to better understand the exposome. Although HELIX offers one initial study structure for understanding the exposome, continued emphasis for exposomic approaches should be placed on developing techniques to measure nonpersistent chemicals that do not place undue burdens on study participants or significant financial constraints on the research study.

**Recommendations**

The following recommendations are suggested for approaching internal exposure assessment for exposome research:

**Recommendation 1: Encourage secondary analyses of samples collected for traditional targeted chemical studies.** High-quality samples (i.e., samples that have been collected and stored properly) from longitudinal epidemiology studies should be used for untargeted analysis and alternative measurement techniques. For this aim to be successful, it is critical that methods for sample collection and storage be standardized. Investment should be made in maintaining established cohorts and in developing protocols that optimize stabilization of samples for storage (e.g., does one analyte stabilize actually destabilize other analytes of interest? Would adding a known xenobiotic act as a standard for normalization? Should multiple small aliquots be stored at the time of collection to facilitate different analytical needs?).

**Recommendation 2: Evaluate and use standardized measurement platforms with measurement harmonization.** A general prototype platform or reference samples should be established under which different technologies can be tested. By establishing this platform, researchers can have a standardized way of demonstrating capacity with

<table>
<thead>
<tr>
<th>Appendix 2. Key advantages and disadvantages of traditional biomonitoring for determination of exposure.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traditional biomonitoring for determination of exposure</strong></td>
</tr>
<tr>
<td><strong>Advantages</strong></td>
</tr>
<tr>
<td>• Well-established and reliable methods for both long-lived (biologically persistent) chemicals and short-lived chemicals with continuous exposures</td>
</tr>
<tr>
<td>• Highly selective methods</td>
</tr>
<tr>
<td>• Provides accurate and precise measurements of biologically persistent chemicals</td>
</tr>
<tr>
<td>• Often targets known chemicals of toxicologic importance</td>
</tr>
<tr>
<td>• Reference data exist for most chemicals</td>
</tr>
<tr>
<td>• Targeted approach allows specific hypotheses of well-documented chemicals to be studied</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Appendix 3. Key advantages and disadvantages of exposomic approaches for determination of exposure.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exposomic approaches for determination of exposure</strong></td>
</tr>
<tr>
<td><strong>Advantages</strong></td>
</tr>
<tr>
<td>• Agnostic approaches are encouraged for detection of emerging exposures of concern</td>
</tr>
<tr>
<td>• Techniques (and development of techniques) promote identification of unknown/existing exposures of concern</td>
</tr>
<tr>
<td>• Links exogenous exposures to internal biochemical perturbations</td>
</tr>
<tr>
<td>• A large number of features can be detected (~10,000) for the cost of a single traditional biomonitoring analysis</td>
</tr>
<tr>
<td>• Includes biomolecular reaction products (e.g., protein adducts, DNA adducts) for which traditional biomonitoring measurements are often lacking or cumbersome</td>
</tr>
<tr>
<td>• Requires a small amount of biologic specimen (~100 μL or less) for full-suite analysis</td>
</tr>
<tr>
<td>• Enables detection of “features” that are linked to exposure or disease for further confirmation</td>
</tr>
<tr>
<td>• Encourages techniques to capture short-lived chemicals</td>
</tr>
<tr>
<td>• Aims to measure biologically meaningful lifetime exposures, both exogenous and endogenous, of health relevance</td>
</tr>
<tr>
<td>• Includes lifetime exposures but does not place undue burdens on study participants</td>
</tr>
</tbody>
</table>
new approaches, which would allow efficient integration of effective methods into research protocols. One approach would be to use samples from NHANES or from a similarly well-characterized data set as a “challenge” or “quality control” set for new and emerging technologies. Moreover, development of or participation in multi-lab proficiency testing programs will ensure harmonization of data across studies.

**Recommendation 3: Use existing resources and databases to obtain information on current exposures that may be important.** Significant efforts have been made to expand databases such as the HMDB,

### Appendix 4. Biomonitoring resources.

<table>
<thead>
<tr>
<th>Category</th>
<th>Resource/Location</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Targeted</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Laboratory Response Network for Chemical Threats Laboratory for Exposure Assessment and Development for Environmental Research (LEADER), Emory University</td>
<td><a href="http://emergency.cdc.gov/lrn/chemical.asp">http://emergency.cdc.gov/lrn/chemical.asp</a></td>
<td><a href="http://www.leadslaboratory.org">www.leadslaboratory.org</a></td>
</tr>
<tr>
<td>Clinical Pharmacology Analytical Services, University of Minnesota Biomarker Core, Center for Tobacco Control Research and Education, University of California, San Francisco</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analytical Chemistry Core, Superfund Research Center, Duke University</td>
<td><a href="http://sites.nicholas.duke.edu/superfund/core/analytical-chemistry-core/">http://sites.nicholas.duke.edu/superfund/core/analytical-chemistry-core/</a></td>
<td><a href="https://chearprogram.org/about/labnetwork">https://chearprogram.org/about/labnetwork</a></td>
</tr>
<tr>
<td><strong>Untargeted</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wishart Research Group, University of Alberta Berkeley Center for Exposure Biology, University of California, Berkeley Clinical Biomarkers Lab, Emory University</td>
<td></td>
<td><a href="http://clinicalmetabolomics.org/">http://clinicalmetabolomics.org/</a></td>
</tr>
<tr>
<td>Mayo Clinic Metabolomics Resource Core, Rochester, MN</td>
<td><a href="http://www.mayo.edu/research/centers-programs/metabolomics-resource-core/overview">http://www.mayo.edu/research/centers-programs/metabolomics-resource-core/overview</a></td>
<td><a href="https://chearprogram.org/about/labnetwork">https://chearprogram.org/about/labnetwork</a></td>
</tr>
<tr>
<td><strong>Funding/biomonitoring support</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDC-funded state biomonitoring grants in 2009 and 2014 (CA, NY, WA, MA, NJ, VA, UT, AZ, CO, NM)</td>
<td><a href="http://www.cdc.gov/biomonitoring/state_grants.html">http://www.cdc.gov/biomonitoring/state_grants.html</a></td>
<td></td>
</tr>
<tr>
<td><strong>International biomonitoring labs and programs</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Biomonitoring in the era of the exposome

Kyoto Encyclopedia of Genes and Genomes (KEGG) human metabolic pathways, and METLIN databases (Kanehsa and Goto 2000; Kanehsa 2002; Smith et al. 2005; Wishart et al. 2009, 2013). Mining these well-developed resources in conjunction with new data analyses will enable a more comprehensive exposure characterization.

**Recommendation 4: Provide guidance for the use of existing databases and develop tools to allow searches across multiple databases.** To facilitate researchers’ integrating exposomic approaches into their studies, resources regarding existing databases should be streamlined. Integration of existing databases such as the HMDB, LIPID MAPS Structure Database and METLIN or search options that can readily work across these resources would enhance their utility for exposome research (LIPID MAPS 2015; Smith et al. 2005; Wishart et al. 2009, 2013).

**Recommendation 5: Foster and facilitate dialogue concerning the utility of targeted, untargeted approaches with traditional biomonitoring techniques.** Discussions should focus around the development of semi-targeted or multiplexing strategies (Wei et al. 2010). Specific discussions should emphasize approaches for capturing short-lived chemicals while minimizing undue financial and participant burdens. Through generating discussion regarding established methods, researchers can have a structured dialogue concerning the utility of targeted, untargeted, and hybrid methods.

**Recommendation 6: Develop chemistries that may enable the detection of low-abundance chemicals and to enable differentiation of endogenous molecules from exogenous molecules.** Through methods such as multiplexing, interfering chemicals can be removed to allow detection of low-level environmental chemicals that are often difficult to detect because of higher-abundance endogenous chemicals from food, drugs, and normal metabolic processes (Rappaport et al. 2014). Investments in the development of semi-targeting or multiplexing strategies should be a high priority.

**Recommendation 7: Develop bioinformatics techniques to enhance detection of unknown chemicals using untargeted methods.** With continued efforts such as ExpoCast, untargeted analysis can be combined with advanced bioinformatics techniques to help prioritize risk assessment, to determine which exposures often co-occur, and to establish markers of disease risk (Dennis et al. 2016; Johnson et al. 2015; Rager et al. 2016; Yu et al. 2013; Wambough et al. 2013).

**Recommendation 8: Encourage development of pharmacokinetic models.** Through building simulated human response models, researchers would be able to incorporate kinetic and dynamic variability to inform interpretation of biomonitoring data.

**Conclusions**

Measurable long-term improvements to human health are attainable through working towards a holistic understanding of environmental influences. In order to assess the exposome, traditional biomonitoring should be coupled with untargeted discovery of unknown chemicals of biological importance. It is critical to note that the advances described here, including those still in early stages of development, require commitment of scientific resources and energy to bring such approaches to fruition. Continued discussion and integration of approaches will be necessary to address the inherent complexity of the exposome. Broad characterization and understanding of internal exposures and their consequences are achievable under the exposome paradigm through combining emerging technologies and untargeted approaches with traditional biomonitoring techniques.

**References**


