Use of the "Exposome" in the Practice of Epidemiology: A Primer on -Omic Technologies

D. Gayle DeBord[1], **Tania Carreón**[2], **Thomas J. Lentz**[3], **Paul J. Middendorf**[4], **Mark D. Hoover**[5], and **Paul A. Schulte**[3]

[1]National Institute for Occupational Safety and Health, Division of Applied Research and Technology, Cincinnati, OH

[2]National Institute for Occupational Safety and Health, Division of Surveillance, Hazard Evaluations, and Field Studies, Cincinnati, OH

[3]National Institute for Occupational Safety and Health, Education and Information Division, Cincinnati, OH

[4]National Institute for Occupational Safety and Health, Office of the Director, Atlanta, GA

[5]National Institute for Occupational Safety and Health, Respiratory Health Division, Morgantown, WV

Abstract

The exposome has been defined as the totality of exposure individuals experience over their lives and how those exposures affect health. Three domains of the exposome have been identified: internal, specific external and general external. Internal factors are those that are unique to the individual; and specific external factors include occupational exposures and lifestyle factors. The general external domain includes factors such as education level and financial status. Eliciting the exposome is daunting and at present not feasible and may never be fully realized. A variety of tools has been identified to measure the exposome. Biomarker measurements will be one of the major tools in exposomic studies. However, exposure data can also be obtained from other sources such as sensors, geographic information systems and conventional tools such as survey instruments. Proof of concept studies are being conducted that show the promise of the exposomic investigation and the integration of different kinds of data. The inherent value of exposomic data in epidemiologic studies is that they can provide greater understanding of the relationships among a broad range of chemical and other risk factors and diseases and ultimately lead to more effective and efficient prevention and control.

Keywords

Corresponding Author's Information: D. Gayle DeBord, Ph.D., National Institute for Occupational Safety and Health, Robert A Taft Laboratories, MS-R2, 1090 Tusculum Avenue, Cincinnati, Ohio 45226, ded4@cdc.gov, Phone – 513.841.4256.
All authors contributed equally to this article.

Disclaimer: The authors declare no conflict of interest.

WHAT IS THE EXPOSOME?

In 2005, Wild defined the exposome as the totality of exposure individuals experience from conception until death and its impact on chronic diseases (1) (a glossary of terms and definitions used in this article is provided in Appendix 1). Exposures can include toxicants in the general environment and in workplaces, diet, lifestyle choices and even socioeconomic status (Figure 1). People have unique characteristics that might make them more or less susceptible to stressors in their environment. A person's genetics, epigenetics, health status, and physiology, as well as changes in these personal components caused by previous exposures, can influence the effects of new or present exposures. For example, metabolic pathways can be disrupted that change susceptibility to the insult or to a disease.

The premise envisioned with the exposome concept was that the exposome is complementary to the genome and that an integrated understanding of the genome and the exposome would contribute to synergistically addressing chronic human health issues. The science of epidemiology is the primary means of understanding the exposome and its interaction with health status. Research suggests that environmental exposures have a much greater impact on health and disease than genetic factors alone (2). The inherent value of exposomic approaches and data in epidemiologic studies is to provide a greater understanding of the relationships among exposures and diseases and ultimately lead to prevention of chronic diseases. Epidemiologic research both can utilize exposomic data in health and disease research and it can be a means of understanding the exposome (Figure 2). The exposome concept was further refined by Wild (3) to include three broad domains: internal, specific external, and general external. Internal factors are those that are specific to the individual such as physiology, age, body morphology, and their genome. Specific external factors include diet and occupational and environmental exposures as well as physical, biological and physiological exposures. The third domain of general external factors includes broader social constructs such as home location, education level and socioeconomic status. Wild noted that the domains can be viewed as both overlapping and intertwining, and that it is sometimes difficult to place a particular exposure in one domain or another. For example, he observed that one can debate whether physical activity should be in the internal domain or in the specific external domain. A comprehensive and informative assessment of exposure can be achieved by combining aspects of the three domains in ways that can be used to guide the design, conduct, and interpretation of epidemiologic studies.

An individual's exposome is dynamic, which makes measuring the exposome challenging. Several critical life stages have been identified for which some exposures may have a greater impact with respect to future diseases. For example, the fetus or young child has rapidly growing cells and immature repair processes (4). Exposures at young ages may have significant influence over future health. Embryos or infants are rapidly maturing and may not have all the protective mechanisms in place to repair damage incurred from an exposure. For example, in the 1950s and 1960s some women were given diethylstilbesterol to prevent miscarriages (5). Their offspring, who were exposed *in utero*, have increased risk of reproductive tract cancers, decreased fertility, and difficult pregnancies (5, 6). The susceptibility of children and teenagers may differ from that of adults because they are still developing, have immature repair processes and have different hormone levels than adults.

As a result of these early-life exposures, the susceptibility to disease caused by later exposures may be increased. Additionally certain exposures can occur throughout different life-stages that are critically important as well (7). In utero exposures would mainly occur due to diet, pharmaceutical use or environmental/occupational exposures. Throughout their lifetime, individuals would have a steady-state exposure to some ambient agents such as allergens (7). Some xenobiotics may bioaccumulate resulting in a higher body burden with age. Occupational exposures would occur mainly during the working years; and as we age exposure to pharmaceuticals tends to increase.

Most epidemiologic studies get a snapshot look at exposures that affect health. Although the exposomic approach may allow for a broader view of exposure, the ability to measure past exposures to any great extent is limited; and measuring each agent to which a person may be exposed at any given time is not feasible at present or in the foreseeable future. Depending upon the study hypotheses, measurement of specific agents will be important along with a holistic exposure assessment approach that includes aspects of the three areas of the exposome described prior (3).

The challenges of measuring the complete exposome are daunting. Newer technologies such as omics, sensors, and geographic or spatial information are allowing for a more comprehensive understanding of the exposome. While the exposome is more likely to be useful in epidemiologic studies, it may also have other clinical or public health utility (8, 9). That utility could range from personalized medicine to improved risk assessment for regional exposures to chemicals, and will necessarily be dependent on the extent to which exposomic indicators are validated and to do that will require additional epidemiologic studies.

While the Wild definition of the exposome was developed to draw attention to the critical need for more complete environmental exposure assessment in epidemiological studies, others have broadened the definition of the exposome to include other factors such as behavior. Miller and Jones (10, p. 2) have defined the exposome as "the cumulative measure of environmental influences and associated biological responses throughout the lifespan, including exposures from the environment, diet, behavior, and endogenous processes". Their definition expands on that of Wild to consider cumulative biological responses as well as endogenous processes. Ultimately, health status is influenced by the interaction of an individual's environmental and genetic factors from conception to death (Figure 3). The explanation of both the exposome and the genome in a manner that can protect and promote health is important and there is need to pursue both.

WHAT HAS BEEN DONE IN THE FIELD?

Utilization of exposomic data in epidemiologic studies and surveillance efforts has occurred due to advancements in exposure science. Yet these advances also indicate the complexity of factors associated with total exposure assessment and the development of appropriate and effective risk management strategies. More importantly though, they help to determine what questions are most pertinent to performing comprehensive exposure assessments. Among the

questions to be considered, the growing scientific literature on this topic suggests the following (11):

- Which mixtures are most important from a public or occupational health perspective?

- What is the nature (i.e., duration, frequency, and timing) and magnitude (e.g., exposure concentration and dose) of relevant exposures for the population of interest?

- What is the mechanism (e.g., toxicokinetic or toxicodynamic) and consequence (e.g., additive, less than additive, more than additive) of the mixture's interactive effects on exposed populations?

Exposure science is the discipline that studies the conditions for contact with toxicants, characterizing the quality and quantity of the toxicant from its sources to its transport and receipt by or interaction with the human body (12, 13). Regarding the evolution of exposure science (defined below) as the means for bridging the discipline of environmental science and environmental health science, Lioy (12) poses the following:

- *What does one do with such exposure information* (i.e., understanding variables that define contact with environmental stressors and the factors that influence the contact)*?*

- What role does exposure science play in situations beyond observational analyses and interpretation?

Addressing these questions through exposome-informed advances in exposure science and risk assessment methods will provide a foundation for improved tools for total exposure assessment and risk management.

In exploring implications for exposure science focusing on the exposome, Rappaport (2) advocates utilizing biomonitoring (e.g., blood sampling and other internal measures of dose) rather than focusing primarily on sampling exposures in food, water, and air. He also suggests the importance of better integration of these biomonitoring measures with environmental exposure measures to advance the field of exposure science and the understanding of the exposome as a means to characterize and control detrimental exposures. Pleil (14, p. 264) has suggested that biomarkers can be used to "assess the sustainability of the environmental conditions with respect to human health."

There are four conventions in the literature for characterizing environmental biomarkers and how they fit into categories of grouping schemes. These four conventions are origin, function, kinetics and medium (14). From the categories of biomarkers, a sequence of data management strategies might be applied to recognize patterns and statistics within the exposome which provide insight into the exposure-dose-response relationship involving environmental stressors.

Exposomic data may contribute to determining why some people will develop a disease while others with the same or greater exposure will not. A key factor in describing the exposome is the ability to accurately measure exposures and their effects on human health.

However, the evolution and maturity of the science of exposomics can be viewed as a practical extension of the principles of total exposure assessment that ultimately seeks to inform risk management strategies.

The U.S. Environmental Protection Agency has been at the forefront of research in total exposure assessment. Led by the Environmental Protection Agency's Office of Research and Development, the Cumulative Communities Research Program "focuses on exposure tools for advancing the science and understanding of cumulative risk to communities and individuals" (15, p. 353). There are multiple factors driving this approach for community (i.e., non-occupational) exposures, but most can be tied to the motivation that people want to know about the multiple stressors (e.g., pollutants) to which they are exposed, what the associated health risks are, and how these exposures and related risks can be prevented or reduced (16).

On a similar national-level surveillance effort, the French national occupational surveillance and prevention network (RNV3P) established a database for recording and tracking occupational health exposures and related adverse health effects (17, 18). A goal of the network and database is to provide better characterization of occupational disease-exposure relationships, thereby exploring a theoretical framework of the occupational exposome.

FUNDING INITIATIVES AND RESOURCES

National Institute of Environmental Health Sciences (NIEHS)

The National Institute of Environmental Health Sciences has long-supported the concept of the exposome, funding studies that are now being defined as exposomic in nature. The Institute recently established a strategic goal of transforming exposure science and has identified the exposome as a possible approach (19). They plan to advance characterizations of environmental exposure assessment at both the individual and population levels which will be accomplished through tools and technologies for multi-scale measurements.

Recently, the National Institute of Environmental Health Sciences funded the Health and Exposome Research Center: Understanding Lifetime Exposures (HERCULES) project at Emory University (http://emoryhercules.com/) with the stated goal of understanding lifetime exposures. The main aims of HERCULES are: 1) to provide greater access to exposome-related approaches such as systems biology, metabolomics, high throughput toxicology, spatial and temporal statistical models; 2) to facilitate communication of the importance of environmental factors in disease using exposome principles; and 3) to expedite translation of novel scientific findings to develop novel sustainability, prevention or treatment strategies in humans.

Human Early-Life Exposome (HELIX; http://www.projecthelix.eu/)—The HELIX project (20) is a European collaboration that has been established as a proof of concept study to characterize children's exposomes as they progress through early life. This project involves 13 partner institutions and will use data from six ongoing, prospective European birth cohorts of mothers and children living in Spain, France, the United Kingdom, Norway, Greece, and Lithuania. Traditional methods are being used for exposure assessment as well

as biomarker and omics measures to assess the exposomes. The project plans to measure external environmental exposures for food, water, air pollution, pesticides, noise, and ultraviolet radiation of up to 32,000 mother-child pairs and measure the growth, development, and health of the children, including birth outcomes, postnatal growth and body mass index, asthma and lung function, and neuro-development.

The intended outcome from this project is an exploration of the relationships between the early life exposome with omics markers and health in childhood. An important long-term goal is to estimate health impacts for the European population based on exposure levels and dose-response relationships developed from HELIX.

Health and Environment-wide Associations based on Large population Surveys (HEALS; http://www.heals-eu.eu/)—HEALS is a project funded by the 7th Framework Programme for the European Commission (21). The general objective of HEALS is to refine a methodology that integrates and applies analytical and computational tools for performing environment-wide association studies (EWAS) in support of European Union-wide environment and health assessments. HEALS interlinks activities that focus on aspects of individual exposure assessment to conventional and emerging environmental stressors and on the prediction of the associated health outcomes. The overall approach will be verified and refined in a series of population studies across Europe including twin cohorts, tackling different levels of environmental exposure, age windows of exposure, and socio-economic and genetic variability.

The external exposome will be estimated by integrating environmental, occupational and dietary data into exposure models. HEALS proposes to describe the internal exposome at the individual level by integrating omics derived data and biomonitoring data. The HEALS approach and tools will be tested by applying them in a number of population studies (including twins studies) across different exposure settings tackling key health endpoints for both children and the elderly. The overall population size involved in these studies to date is approximately 335,000 individuals covering different age, gender and socio-economic status groups. A high-level goal of HEALS is to develop scientific guidance on exposome-based risk assessment.

EXPOsOMICS (http://www.exposomicsproject.eu/)—EXPOsOMICS is a program developed by the European Union (22). It aims to develop a new approach to assess environmental exposures, primarily focusing on air pollution and water contaminants by developing a personal exposure monitoring (PEM) system (including sensors, smartphones, geo-referencing, satellites) to collect data on individuals' external exposome as well as analyzing biological samples (internal markers of external exposures) by using omics technologies. The program will search for relationships between external exposures measured using the personal exposure monitoring system, which has not previously been used in large scale studies, and profiles of chemicals, measured by omics, in the same individuals. Using omics techniques the collected exposure data can be linked to biochemical and molecular changes, and the results will help to improve understanding on how xenobiotics influence the risk of developing chronic diseases.

TOOLS USED IN THE STUDY OF THE EXPOSOME

Different sets of tools (Table 1) are needed to measure the three exposome domains as outlined by Wild (3). Specific exposures found in an individual's environment may be evaluated in a variety of ways, including biomarker-based metrics such as urinary metabolites of xenobiotics, and sensors (either personal or remote monitoring) to detect contaminants in the individual environment and to evaluate prior occupational/ environmental exposures. Personal monitoring might use sensors to determine physical activity and other exposures. Survey instruments or biomarkers can measure stress. Databases, geographic information systems, and surveys will be helpful to elucidate general external exposures such as educational level or urban or rural environments in the third domain identified by Wild (3).

Biomarker measurements will have a large role in developing the exposome, especially given the associated advantages of requiring a small sample size, providing high throughput, and relatively low-cost for obtaining a wealth of usable information.

If the concept of the exposome is to be realized, epidemiologists will need to incorporate exposure and health information from traditional sources as well as consider information from non-traditional sources. The National Research Council (2012) released a report that identified 21st Century techniques that would be important to assess exposure (23). The exposome is a paradigm shift from a single-exposure-to-disease concept to a recognition that health is impacted by multiple exposures. Different sources of exposure and health information that may be useful to capture are discussed below to provide epidemiologists with greater insight of the tools that may be available to them.

Omic technologies

Omic biomarkers are a class of biomarkers of current scientific interest (Table 2). The discovery and use of these biomarkers have increased rapidly, due to the advent of high-throughput technologies and innovations that include improved sample preparation, robotic sample-delivery systems, automated data processing, and use of multivariate statistical methods, with associated reductions in cost. Investigators have begun to use these biomarkers in larger-scale population studies of the exposome.

One of the greatest challenges with these studies is applying omics technologies to generate meaningful results. Epidemiologists must strive to understand the principles of omics and determine when it is appropriate to include biomarkers identified using these technologies. In addition, no single omics approach will suffice to characterize the exposome, and integration of omics outcomes and other sources of exposure information will be needed to deepen our understanding of the causes of disease. For example, metabolomics studies have shown links between gut flora, diet and cardiovascular disease (24, 25). Omics technologies have shown utility in determining toxicity and mode of action in risk assessments and assessing the health impact of an exposure using analysis of variance (ANOVA) approaches together with pair-wise comparisons between dose groups and the corresponding control (26). Thus, what might be considered an advance in exposomics can equally be considered a natural combining and extension of traditional tools.

Molecular Epidemiology

Many of the techniques in assessing exposure and the exposome involve molecular epidemiologic studies. Molecular epidemiology is the use of biological markers (exposure, effect, susceptibility) in epidemiologic research (27, 28). In assessing the exposome, biomarkers of exposure may be the most useful. The development of adductomics, which measure the full complement of protein adducts might be useful in improving exposure assessment in epidemiologic studies because of the ability to reflect extended exposure (28, 29). As some omics applications mature and systems biology becomes more incorporated in molecular epidemiology, the understanding of the exposome will be increased, and it should be possible to more broadly explore exposure-disease relationships, to study effect modifiers, and to obtain insight into temporal and multi-level factors in health and disease.

Sensor Technologies

Remote sensing is a key innovation in exposure science and is defined as the measurement of some property of a phenomenon, object, or material that is not in direct physical contact with the population being studied (23). The use of sensors for not only remote monitoring but personal monitoring is growing. Sensors are now used to measure clinical parameters such as blood pressure and glucose levels, and new sensors are being developed to measure biomarkers, such as portable adhesive sweat analyzers (30). The devices—with remote-sensing-based spatial referencing technologies and modeling—enable the continuous, real-time, real-world assessment of exposures that vary according to an individual's location, activity, and lifestyle. The use of smart phones and tablet technology is only likely to grow and provide enhanced opportunities to collect exposure information (23, 31). Apple has created ResearchKit (https://developer.apple.com/researchkit/) a framework for Apple platforms that can be used in medical research of diseases such as Parkinson's, asthma, diabetes, heart disease and breast cancer.

In another example used in a current exposomic study, the ExpoApp, a mobile application created specifically for the HELIX project, uses a Global Positioning System (GPS) and a built-in accelerometer to track a person's location and measure physical activity every 10 seconds (32). Participants in the HELIX project will wear ExpoApp-enabled smartphones for a week, along with air pollution and ultraviolet radiation monitors, and the data will be used to calculate the amount of air inhaled and an individual's exposure to air pollutants. A challenge is that individuals with access to information from such applications may alter their behavior and thereby inject biases into the study data.

Geographic Information Systems (GIS)

Geographic Information Systems are informative for tracking the acquisition, editing, analysis, storage, and visualization of geographic data (33). Geographic Information Systems allow for mapping a variety of data, for example environmental, topographical or health-related to understand trends and patterns. Information from Geographic Information Systems would fall into and provide clarity to exposures in the general external domain of the three domains of the exposome described by Wild (3). Geographic Information Systems have been reported to enhance exposure assessment in epidemiologic studies, because such systems can provide information on broad environmental contaminant levels maps or to

define a population (34). For example, the Research Center on Health Disparities, Equity, and the Exposome (http://rchdee.uthsc.edu/) maintains the Public Health Exposome Data Information System, a longitudinal, 30-year data set that integrates over 20,000 environmental and health data records (9). The database contains data indicators of the atmosphere, climate, water, land cover and land use that have been geocoded. In addition, a considerable amount of other data are geocoded and included in this database, derived from ground stations and field samples of emissions, heavy metals, toxic dump and storage sites, and Brownfields collected by the Environmental Protection Agency with other federal, state, and local agencies.

Conventional measurements

Survey instruments gather information for which there is no way to quantitatively measure exposures and have long been used in epidemiologic studies. These tools may be important in documenting retrospective exposures, which currently is not feasible with other exposomic tools. Modeling the data to improve exposure matrices will be an important aspect of using exposomic data.

Exposome informatics

With the advances in molecular medicine and development of omics technologies, the field of biomedical informatics has evolved as a discipline. Genome-wide association studies (GWAS) and studies combining genomic and phenotypic data have required the development of new biostatistical methods for quality control, imputation, and analysis issues such as multiple hypothesis testing. Developments in exposome science have revealed the need for evaluating the interrelationships among phenotype, genotype, and exposure data. As shown in the different approaches discussed here, these data will be heterogeneous, wide-ranging, and massive; will frequently involve time series; and will require high velocity processing (35). Some efforts have been made for the analysis of EWAS, but new research is required to develop approaches for data management, analysis and visualization.

Reality mining

The term "reality mining" refers to the analysis of behavioral and self-reported data extracted from social networks and other portable devices' applications (36). By continually logging and time-stamping information about a subject's activity, location, and proximity to other users, it is currently possible to identify patterns in the data and translate them into maps of social relationships. By definition of the exposome, all exposures are to be taken into account. Logistically, that is impossible but by using social media networks or citizen scientists' efforts, such as HabitatMap (http://habitatmap.org/), additional exposure information could be obtained. This relational information could have much broader implications including the improvement of existing computational models of exposure, disease status of an individual, and disease spread (37). A recent report demonstrated how the mobile application, Yelp, was used to identify foodborne illnesses in New York City restaurants (38).

Environment-wide association studies

In an EWAS employing techniques partially adapted from GWAS, Patel et al. (39, 40) used chemical, clinical and questionnaire data from the National Health and Nutrition Examination Survey (NHANES) cohorts to evaluate associations between multiple environmental factors and type 2 diabetes mellitus and serum lipid levels, respectively. The study followed two methodological steps analogous to those in a GWAS. First, the authors considered a panel of environmental assays, and identified those significantly associated or correlated with diabetes or serum lipids while controlling for multiple confounders. Second, they validated the associations by testing significant findings in other NHANES independent cohorts. The authors identified environmental factors, including select chemicals, corroborating earlier findings. Additionally, Patel et al. (41) screened for gene-environment interactions by integrating results from GWAS and EWAS. Properly designed, EWAS studies can lead to discovery of biomarkers for exposure and disease and establish a molecular basis for the cause of environmental diseases (42). Patel et al. (43) described the "exposome globe" [e.g. Figure 4], which is a visual depiction of the network of replicated correlations between individual exposures of the exposome. The exposome globe allows visualization of clusters of exposure.

Although omics technologies are at the forefront in studies of the exposome and associated biomarkers, omic-based measurements do not always reflect exposure and may instead be products of normal cellular function (3). Proteomics, transcriptomics and metabolomics are a few of the omics technologies that have shown great promise for exposomic studies. As our understanding of basic biologic pathways grows, perturbations in the pathways as measured by omics result in improved interpretation with respect to how health is affected. Other types of exposure information of value to EWAS include databases that contain exposure information or population demographics data that can impact exposures. Use of remote sensors to gather environmental data may help ascertain exposures to the population as a whole. Personal monitoring using sensors to determine exposures, physiological factors, and geographic location is beginning to be used in environmental health studies to assess the "total" exposure of study participants.

STUDY DESIGN IN EXPOSOMIC STUDIES

Ideally, the best approach in exposomics would be to use a longitudinal (prospective) cohort study design, as it allows follow up of individuals and repeated sampling as well as monitoring during windows of increased sensitivity. However, neither the biomarkers nor the environmental data could be recorded continuously and it would be necessary to carry out a series of cross-sectional investigations of the study population.

Candidate exposures vs. agnostic approaches

Two strategies have been identified for characterizing the exposome (2). One is a 'bottom-up' strategy in which all the exposures in a person's exposome are measured at set time points. Although this approach would have the advantage of identifying important exposures in the air, water, or diet, it is not currently feasible and would miss essential components of the internal chemical environment due to such factors as gender, obesity, inflammation, and

stress (44). At the other extreme is to measure a combination of omic endpoints and legacy biomarkers in repeated blood specimens. This strategy has been referred to as 'top-down exposomics' (2). This data-driven (or "agnostic") approach lacks specific hypotheses (2). According to Rappaport (2), the exposome would consist of a profile of exogenous and endogenous exposures, but would not pinpoint their source. Since it is currently not feasible to measure all chemicals in the blood, it has been proposed (44, 45) to focus initially on the most prominent classes of toxicants known to cause disease, namely, reactive electrophiles, endocrine (hormone) disruptors, modulators of immune responses, agents that bind to cellular receptors, and metals. Exposures to these agents can be monitored in the blood either by direct measurement or by looking for their effects on physiological processes (such as receptor-based signaling). These measurements could help generate signatures or profiles for these exposures in the blood. These profiles could be used to help identify key exposures associated with a disease by comparing the profiles between cases of that disease and controls, preferably from longitudinal studies. Additionally, once important profiles of biomarkers have been identified the sources of exposure could be determined and methods to reduce the exposures could be identified. Therefore, discovery-driven research and hypothesis-driven research should be considered complementary and synergistic.

CHALLENGES IN STUDY DESIGN

While many epidemiologic studies measure multiple variables, the multiplicity of variables in exposomic studies can be daunting (46, 47) and may require approaches that are different from traditional epidemiology. From the epidemiology point of view, the following challenges have been identified in exposome studies:

Reverse causality

For an exposure to be a cause, the exposure must precede the outcome. Reverse causation is a situation in which the outcome precedes and causes the exposure instead of the other way around. This could occur, for example, if a person moved or changed his/her address as a result of a condition in the domicile that was making him or her sick. It is of particular concern in retrospective and cross-sectional studies. In prospective cohort studies, since exposure is determined in advance of disease onset, the probability of reverse causation is greatly diminished (1). EWAS study designs have been developed that can reduce the possibility of reverse causation (48).

Testing multiple variables for associations

Statistical inference problems with the use of omics methodologies involve the simultaneous test of thousands of null hypotheses. The multitude of comparisons made in these studies will result in both false positive (Type 1 errors) and, if the correction for multiple comparisons is overly conservative or power is inadequate, false negative (Type 2 errors) results (49).

Univariate models consider each predictive variable separately, and their association with the outcome of interest is tested using the same statistical model. The Family Wise Error Rate (FWER) and the False Discovery Rate (FDR) have been used to characterize the number of

associations that could falsely be declared statistically significant (50, 51). The Family Wise Error Rate is the probability of making one or more false discoveries, or type I errors, among all the hypotheses. Bonferroni and other correction methods are commonly used; however, these often produce exceedingly conservative thresholds (50).

In an exploratory approach ('top-down exposomics', described above), it may be preferable to use a less conservative correction strategy. The False Discovery Rate is the expected proportion of errors among all associations declared statistically significant. To control for the False Discovery Rate, one must define the proportion of positive findings that are allowed to be false, usually 5%. Methods have been developed for both approaches and are described in detail elsewhere (50).

Correlation among variables

While univariate methods are useful for uncovering simple relationships between predictors and responses, they are also likely to overlook relationships involving combinations of factors. Different multivariate methods have been developed for this purpose. Multivariate analyses aim to summarize the information contained in large datasets into a few synthetic variables [the principal components] that capture the latent structure of the data. Because these methods effectively reduce the number of dimensions necessary to represent the data, they are often referred to as dimension-reduction methods. These methods have been described in detail elsewhere (50). Due to the density of the data in exposomic studies, identification of independent associations will be challenging. Patel and Ioannidis (46, 47) have proposed some agnostic approaches to aid in the analysis of the large number of variables found in exposomic studies. Smith et al. (52) proposed the use of genetic traits and Mendelian randomization techniques to study highly confounded risk factors and disease causation. Such a technique may be useful for exposome studies to investigate the effects of modifiable risk factors of diseases that are too heavily confounded to be studied by conventional approaches.

Variability over time and between subjects

Variability over time and between subjects is associated with a multitude of intrinsic and extrinsic factors, some known and some unknown. Unlike the genome, omics endpoints are dynamic and likely to show variability in different cells and tissues, and throughout the life of an individual (53). Panel studies in small population samples have been proposed to measure the effect of short-term variability on exposure and omics biomarkers, on individual behaviors (physical activity, mobility, time activity), and on personal and indoor exposures (20).

Variability of exposure data

For exposures with a short biological half-life and little constancy in the underlying exposure behavior, temporal variability may be particularly high. For such exposures, intra-individual compared with inter-individual variability is known to be high, and only repeat measurements over time provide improved exposure estimates (20). Studies to measure daily repeat biomarkers of non-persistent chemicals (phthalates, phenols, organophosphate

pesticides) in urine have been proposed to characterize intra- and inter-individual variability in these urine biomarkers, and where possible, correct for the uncertainties in a larger cohort.

Analytical measurement error

With the development of stable and cost-effective high-throughput platforms, large amounts of experimental data are generated. However, data obtained from these platforms are highly sensitive to experimental conditions, and can therefore include considerable noise in the form of measurement error. Statistical methods should be able to estimate technical nuisance variation (50).

Measurements

Many challenges in measuring the exposome exist. The sheer diversity and variability of exposures individuals experience throughout their lives is impossible to measure. Exposures constantly change in a person's environment over time (54). Environmental contaminants appear and disappear and an individual may make changes in their lifestyle choices that can increase or decrease exposures. Additionally, the relevance of past exposures remains unknown and the inability to measure past exposures hampers the exposome (54). Certain life stages have been identified as being more susceptible to environmental exposures (i.e., lifecycle "windows of susceptibility") so relevancy across time may also be an issue.

Biomarkers, one of the primary ways to measure the exposome, have numerous challenges. Their relevance to exposure or their predictive value in non-target tissues may not have been established. The full validation of the biomarker both in the laboratory and the population has not always been performed making the interpretation or relevance of the measurement difficult. Omics technologies have multiple advantages that make them well-suited for exposomic studies but also have limitations in interpretation and validation for exposures and effect of exposures. The management of large datasets is still a challenge for these technologies; and these technologies also require major investments in equipment and expertise.

Other techniques that may be important to measure the exposome also have limitations that include the uncertainty about relevance of the measurement, variability over time, and the inability to measure past exposures. While environmental monitoring (personal or remote), geospatial information, and reality mining may tell us what exposures a person may have had, the actual dose that an individual internalized and relevance of that exposure cannot always be determined. The inherent problem of connecting these external exposures to internal biomarker measurements has always been a challenge and is further exacerbated with exposomic research. Newer approaches such as those used by Patel et al. may be helpful in this regard (55). Better models that integrate exposure data from multiple sources will be more useful for determining health impact. Additionally, surveys or questionnaires can be improved to assess current or past exposures in greater depth and accuracy to facilitate research findings.

MULTI-LEVEL ANALYSIS

Since the exposome can involve various types of data, such as biological, economic, behavioral, and social data, there will be challenges to model data from these areas and from individual, group and ecologic levels in regards to their role as determinants of disease. Difficulties in inference from ecologic data can impede epidemiologic research concerning the effects of the exposure on individual-level health behaviors and disease risks. Multi-level analysis seeks to explain relations involving both individual-level and aggregate level variables (56). Population and group-level factors may modify the relationship between individual-level risk factors and risk of disease (57). There are various approaches available to epidemiologists to perform multi-level modeling that utilize random effects models and generalized estimating equations. Multi-level models can be implemented using SAS Proc MIXED and SAS Proc GLIMMIX (58, 59).

FUTURE PROSPECTS

Wild (1) originally developed the concept of the exposome to draw attention to the need for comprehensive exposure assessment in epidemiologic studies. Lichtenstein et al. (60) estimated that the attributable risk from the genome for chronic disease was only somewhere between 10–30%, and the environment was responsible for the other 70–90%. Measuring the exposome has a critical role in understanding chronic disease formation and progression. The exposome provides improved exposure assessments by the integration of different types of exposure information, but can the utility of this concept be expanded beyond exposure assessment?

Improved exposure assessments can feed into the systems-biology approach to evaluate how exposures disrupt normal biological processes. This type of approach could provide information used in delineating the mode of action of the response or toxicity, and more accurately inform risk assessments and risk management. Early biological effects using a systems-biology approach and computational toxicology efforts offer great promise for improving risk assessments (26). The biomarker techniques such as the omics approaches have shown promise and can provide information on mode of action and dose-response relationships (61). As these techniques evolve, estimation of internal dose and response markers will be a critical test of these new technologies for application in risk assessment strategies (26). However, it has been reported that many research "findings" are false or overinflated. Improving the accuracy of research findings will require better powered studies, diminishing bias, and improvement in our understanding of R values (62, 63). Juarez (9) has explored the public health exposome as a way to integrate environmental contextual data with measures of health outcomes. The exposome is a new way to conduct health disparities research to understand the social and environmental factors and their effect on health. Others have suggested that the exposome could have a major role in clinical care or as a basis for transformative research (8) (64). The benefits of the exposome are many for both individual and population health research. The exposome can aid in identifying modes of action of stressors, identifying unknown exposures and improving our understanding of disease. This can lead to better risk assessments, better translation of science into practice and ultimately to disease prevention.

The concept of the exposome is a paradigm shift. While the concept is daunting, in particular measuring all exposures an individual has in a lifetime and predicting health impact, studies are being conducted using exposomic principles. Epidemiologic studies can be improved by incorporation of exposomic principles, thus improving measured associations with many health outcomes and conditions.

Acknowledgments

Appendix 1

Environment- or Exposome-wide assessment studies (EWAS)	Studies that collect multiple kinds of exposure data from multiple sources which are then related to health effects
Exposome	Concept describing the totality of exposure experienced by an individual during their life and the health impact of those exposures (1)
Exposome informatics	Data management framework to deal with large multi-scale data sources in exposome studies (32)
Genome	The genetic material of an organism, encoded in DNA and including both the genes and the non-coding sequences of the DNA (68)
Genome-wide association studies (GWAS)	Studies that evaluate markers across the genome to elucidate associations with diseases
Geographic Information Systems (GIS)	Systems that manage geographic data (33)
Metabolome	Sum of all low molecular weight metabolites present in a biological sample (69)
Omics technologies	The collective characterization of components and measurement of molecules from a biological field of study, which involves a large scale data acquisition system that can be used to measure biological states or responses. Examples include the genome (DNA), transcriptome (RNA), and proteome (proteins)
Phenome	All of the phenotypes of a cell, tissue, or organism
Proteome	The full set of proteins encoded by a genome (68)
Reality mining	Analysis of behavioral and self-reported data extracted from social networks and other portable devices' applications (36)
Sensors	Devices that measures an event or change. There are many diverse types of sensors available
Total exposure	All exposures experienced by an individual
Transcriptome	Complete set of RNA transcripts produced by the genome at any one time (67)

Glossary

EWAS   Environment-wide association studies

GWAS   Genome-wide association studies

| **HEALS** | Health and Environment-wide Associations based on Large population Surveys |
| **HELIX** | Human Early-Life Exposome |
| **NHANES** | National Health and Nutrition Examination Survey |

References

1. Wild CP. Complementing the genome with an "exposome": The outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol Biomarkers Prev. 2005; 14(8):1847–1850. [PubMed: 16103423]

2. Rappaport SM. Implications of the exposome for exposure science. J Expo Sci Environ Epidemiol. 2011; 21(1):5–9. [PubMed: 21081972]

3. Wild CP. The exposome: from concept to utility. Int J Epidemiol. 2012; 41(1):24–32. [PubMed: 22296988]

4. Needham LL, Calafat AM, Barr DB. Assessing developmental toxicant exposures via biomonitoring. Basic Clin Pharmacol Toxicol. 2008; 102(2):100–108. [PubMed: 18226062]

5. Goodman A, Schorge J, Greene MF. The long-term effects of in utero exposures–the DES story. N Engl J Med. 2011; 364(22):2083–2084. [PubMed: 21506735]

6. Herbst AL, Ulfelder H, Poskanzer DC. Adenocarcinoma of the vagina. Association of maternal stilbestrol therapy with tumor appearance in young women. N Engl J Med. 1971; 284(15):878–881. [PubMed: 5549830]

7. Athersuch TJ. The role of metabolomics in characterizing the human exposome. Bioanalysis. 2012; 4(18):2207–2212. [PubMed: 23046263]

8. Li-Pook-Than J, Snyder M. iPOP goes the world: integrated personalized Omics profiling and the road toward improved health care. Chem Biol. 2013; 20(5):660–666. [PubMed: 23706632]

9. Juarez P, Hood D. Sequencing the public health genome. J Health Care Poor Underserved. 2013; 24(1 Suppl):114–120. [PubMed: 23395949]

10. Miller G, Jones D. The nature of nurture: refining the definition of the exposome. Toxicol Sci. 2014; 137(1):1–2. [PubMed: 24213143]

11. Sexton K, Hattis D. Assessing cumulative health risks from exposure to environmental mixtures - three fundamental questions. Environ Health Perspect. 2007; 115(5):825–832. [PubMed: 17520074]

12. Lioy PJ. Exposure science: a view of the past and milestones for the future. Environ Health Perspect. 2010; 118(8):1081–1090. [PubMed: 20308034]

13. Cohen Hubal EA, Barr DB, Koch HM, et al. The promise of exposure science. J Expo Sci Environ Epidemiol. 2011; 21(2):121–122. [PubMed: 21224897]

14. Pleil JD. Categorizing biomarkers of the human exposome and developing metrics for assessing environmental sustainability. J Toxicol Environ Health B. 2012; 15(4):264–280.

15. Zartarian VG, Schultz BD. The EPA's human exposure research program for assessing cumulative risk in communities. J Expo Sci Environ Epidemiol. 2010; 20(4):351–358. [PubMed: 19367326]

16. Williams P, Dotson G, Maier A. Cumulative risk assessment (CRA): transforming the way we assess health risks. Environ Sci Technol. 2012; 46(20):10868–10874. [PubMed: 22938698]

17. Bonneterre V, Faisandier L, Bicout D, et al. Programmed health surveillance and detection of emerging diseases in occupational health: contribution of the French national occupational disease surveillance and prevention network (RNV3P). Occup Environ Med. 2010; 67(3):178–186. [PubMed: 19776024]

18. Faisandier L, Bonneterre V, De Gaudemaris R, et al. Occupational exposome: A network-based approach for characterizing Occupational Health Problems. J Biomed Inform. 2011; 44(4):545–552. [PubMed: 21362498]

19. National Institute of Environmental Health Sciences. Exposure Biology and the Exposome. http://www.niehs.nih.gov/research/supported/dert/programs/exposure/index.cfm). (Updated 10/05/2015. Accessed 10/30/2015)

20. Vrijheid M, Slama R, Robinson O, et al. The human early-life exposome (HELIX): project rationale and design. Environ Health Perspect. 2014; 122(6):535–544. [PubMed: 24610234]

21. Adekoya N, Myers JR. Fatal harmful substances or environmental exposures in agriculture, 1992 to 1996. J Occup Environ Med. 1996; 41(8):699–705. [PubMed: 10457514]

22. Adgate JL, Kukowski A, Stroebel C. Pesticide storage and use patterns in Minnesota households with children. J Expo Anal Environ Epidemiol. 2000; 10(2):159–167. [PubMed: 10791597]

23. National Research Council. Exposure Science in the 21st Century: A Vision and a Strategy. Washington, DC: National Academies Press; 2012.

24. Tang W, Wang Z, Levison B, et al. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. N Engl J Med. 2013; 368(17):1575–1584. [PubMed: 23614584]

25. Wang Z, Klipfell E, Bennett BJ, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature. 2011; 472(7341):57–63. [PubMed: 21475195]

26. DeBord DG, Burgoon L, Edwards SW, et al. Systems Biology and Biomarkers of Early Effects for Occupational Exposure Limit Setting. J Occup Environ Hyg. [Available online ahead of print July 1, 2015].

27. Schulte, PA.; Perera, F., editors. Molecular epidemiology: principles and practices. San Diego, CA: Academic Press; 1993.

28. Rothman, N.; Hainaut, P.; Schulte, P., et al. Molecular Epidemiology: Principles and Practices. Lyon, France: International Agency for Research on Cancer; 2011.

29. Rappaport SM, Li H, Grigoryan H, et al. Adductomics: Characterizing exposures to reactive electrophiles. Toxicol Lett. 2012; 213(1):83–90. [PubMed: 21501670]

30. Ahlbom A, Navier IL, Norell S, et al. Nonoccupational risk indicators for astrocytomas in adults. Am J Epidemiol. 1986; 124(2):334–337. [PubMed: 3728449]

31. Howard JH, DeBord DG, Hoover MD. The future of industrial hygiene. Synergist. 2014; 25(8):7.

32. Potera C. The HELIX Project: tracking the exposome in real time. Environ Health Perspect. 2014; 122(6):A169. [PubMed: 24892972]

33. Longley, PA.; Goodchild, MF.; Maguire, DJ., et al. Geographic Information Systems and Science. 2nd. New York, NY: Wiley; 2005.

34. Nuckols JR, Ward MH, Jarup L. Using geographic information systems for exposure assessment in environmental epidemiology studies. Environ Health Perspect. 2004; 112(9):1007–1015. [PubMed: 15198921]

35. Martin Sanchez F, Gray K, Bellazzi R, et al. Exposome informatics: considerations for the design of future biomedical research information systems. J Am Med Inform Assoc. 2014; 21(3):386–390. [PubMed: 24186958]

36. Eagle N, Pentland A. Reality mining: sensing complex social systems. Personal Ubiquitous Comput. 2006; 10(4):255–268.

37. Madan A, Cebrian M, Lazer D, et al. Social Sensing for Epidemiological Behavior Change. Ubicomp 2010: Proceedings of the 2010 Acm Conference on Ubiquitous Computing. 2010:291–300.

38. Harrison C, Jorder M, Stern H, et al. Using online reviews by restaurant patrons to identify unreported cases of foodborne illness - New York City. MMWR Morb Mortal Wkly Rep. 2014; 63(20):441–445. [PubMed: 24848215]

39. Patel C, Bhattacharya J, Butte A. An Environment-Wide Association Study (EWAS) on type 2 diabetes mellitus. PLoS One. 2010; 5(5):E10746. [PubMed: 20505766]

40. Patel C, Cullen M, Ioannidis J, et al. Systematic evaluation of environmental factors: persistent pollutants and nutrients correlated with serum lipid levels. Int J Epidemiol. 2012; 41(3):828–843. [PubMed: 22421054]

41. Patel C, Chen R, Kodama K, et al. Systematic identification of interaction effects between genome- and environment-wide associations in type 2 diabetes mellitus. Hum Genet. 2013; 132(5):495–508. [PubMed: 23334806]

42. Rappaport SM. Biomarkers intersect with the exposome. Biomarkers. 2012; 17(6):483–489. [PubMed: 22672124]

43. Patel C, Manrai A. Development of exposome globes to map out environment-wide associations. Pac Symp Biocomput. 2015; 20:231–242. [PubMed: 25592584]

44. Rappaport SM, Smith MT. Epidemiology. Environment and disease risks. Science. 2010; 330(6003):460–461. [PubMed: 20966241]

45. Rappaport S, Barupal D, Wishart D, et al. The blood exposome and its role in discovering causes and disease. Environ Health Perspect. 2014; 122(8):769–774. [PubMed: 24659601]

46. Patel C, Ioannidis J. Studying the elusive environment in large scale. JAMA. 2014; 311(21):2173–2174. [PubMed: 24893084]

47. Patel C, Ioannidis J. Placing epidemiological results in context of multiplicity and typical correlations of exposures. J Epidemiol Community Health. 2014; 68:1096–1100. [PubMed: 24923805]

48. Patel C, Rehkopf D, Leppert J, et al. Systematic evaluation of environmental and behavioural factors associated with all-cause mortality in the United States National Health and Nutrition Examination Survey. Int J Epidemiol. 2014; 42(6):1795–1810. [PubMed: 24345851]

49. Kelsey, JL.; Whittemore, AS.; Evans, AS., et al. Methods in Observational Epidemiology. New York, NY: Oxford University Press; 1996.

50. Chadeau-Hyam M, Campanella G, Jombart T, et al. Deciphering the complex: Methodological overview of statistical models to derive OMICS-based biomarkers. Environ Mol Mutagen. 2013; 54(7):542–557. [PubMed: 23918146]

51. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B. 1995; 57(1):289–300.

52. Smith GD, Lawlor DA, Harbord R, et al. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. PLoS Med. 2007; 4(12): 1985–1992.

53. Mias G, Snyder M. Personal genomes, quantitative dynamic omics and personalized medicine. Quant Biol. 2013; 1(1):71–90. [PubMed: 25798291]

54. Wild CP, Scalbert A, Herceg Z. Measuring the exposome: a powerful basis for evaluating environmental exposures and cancer risk. Environ Mol Mutagen. 2013; 54(7):480–499. [PubMed: 23681765]

55. Patel C, Ioannidis J, Cullen M, et al. Systematic assessement of the correlation of household income with infectious, biochemical, physiological factors in the United States, 1999–2006. Am J Epidemiol. 2014; 181(3):171–179. [PubMed: 25589242]

56. Korff MV, Koepsell T, Curry S, et al. Multi-level analysis in epidemiologic research on health behaviors and outcomes. Am J Epidemiol. 1992; 135(10):1077–1082. [PubMed: 1632420]

57. Diez-Rouz A, Aiello A. Multilevel analysis of infectious diseases. J Infect Dis. 2005; 1(191 Suppl 1):S25–33.

58. Bingenheimer J, Raudenbush S. Statistical and substantive inferences in public health: issues in the application of multilevel models. Annu Rev Publ Health. 2004; 25:53–77.

59. Witte JS, Greenland S, Kim L-L, et al. Multilevel Modeling in Epidemiology with GLIMMIX. Epidemiology. 2000; 11(6):684–688. [PubMed: 11055630]

60. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med. 2000; 343(2):78–85. [PubMed: 10891514]

61. Thomas R, HJ Clewell I, Allen B, et al. Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. Toxicol Sci. 2011; 120(1):194–205. [PubMed: 21097997]

62. Ioannidis J. Why most published research findings are false. PLoS Med. 2005; 2(8):E124. [PubMed: 16060722]

63. Ioannidis J. Why most discovered true associations are inflated. Epidemiology. 2008; 19(5):640–648. [PubMed: 18633328]

64. Louis G, Sundaram R. Exposome: time for transformative research. Stat Med. 2012; 21(22):2569–2575.

65. Herceg Z, Vaissiere T. Epigenetic mechanisms and cancer: an interface between the environment and the genome. Epigenetics. 2011; 6(7):804–819. [PubMed: 21758002]

66. Aiken RD. Quality-of-life issues in patients with malignant gliomas. Semin Oncol. 1994; 21:273–275. [PubMed: 8153670]

67. Horgan RP, Kenny LC. 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. Obstet Gynecol. 2011; 13(3):189–195.

68. Ferlini A, Scotton C, Novelli G. Biomarkers in Rare Diseases. Public Health Genom. 2013; 16(6):313–321.

69. Wishart DS, Jewison T, Guo AC, et al. HMDB 3.0–The Human Metabolome Database in 2013. Nucleic Acids Res. 2013; 41(Database issue):D801–807. [PubMed: 23161693]
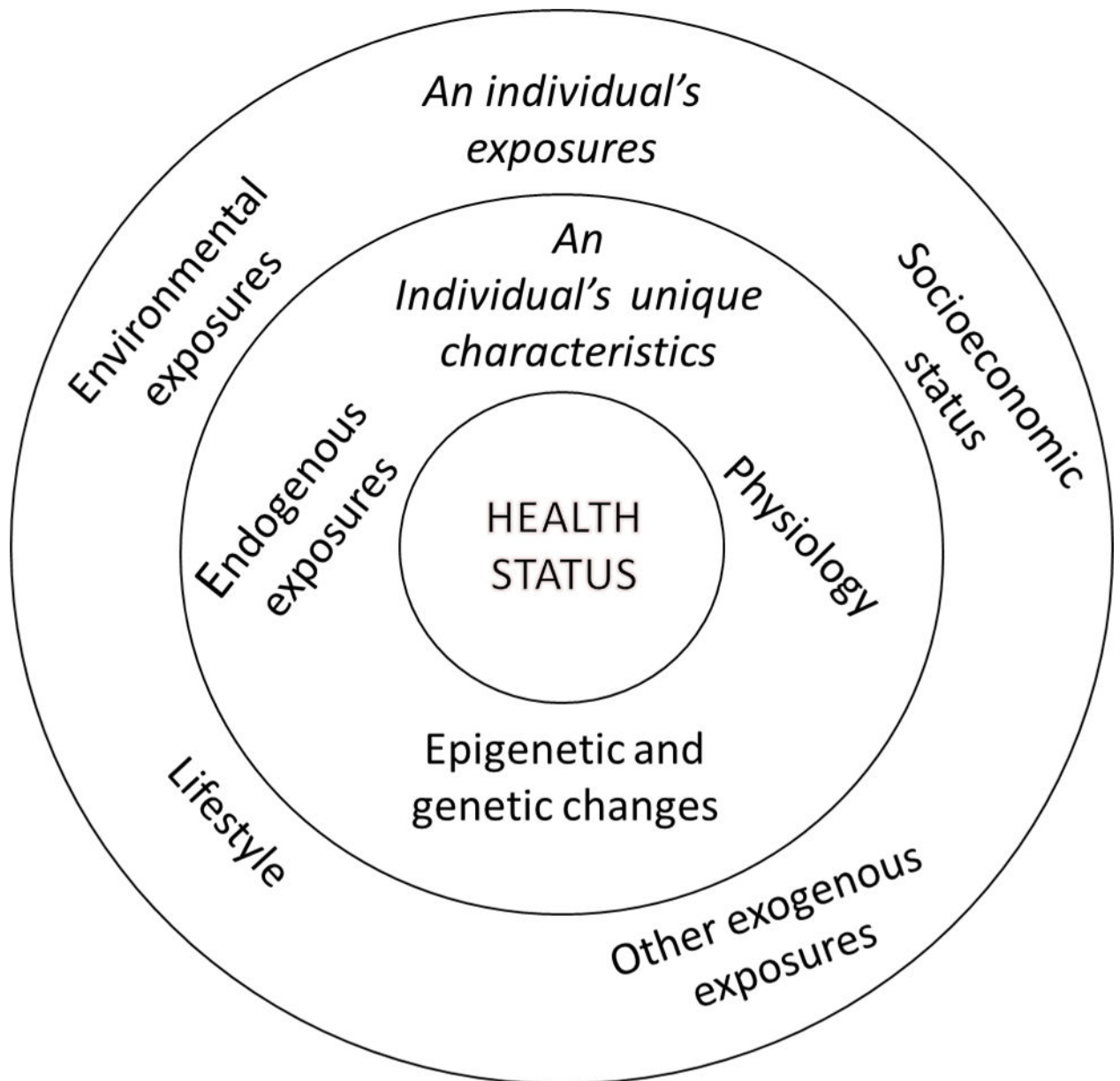
Figure 1.
Concept of the exposome. An individual has many sources of exposure. How those exposures are modulated and impact health depends on the individual's unique characteristics.
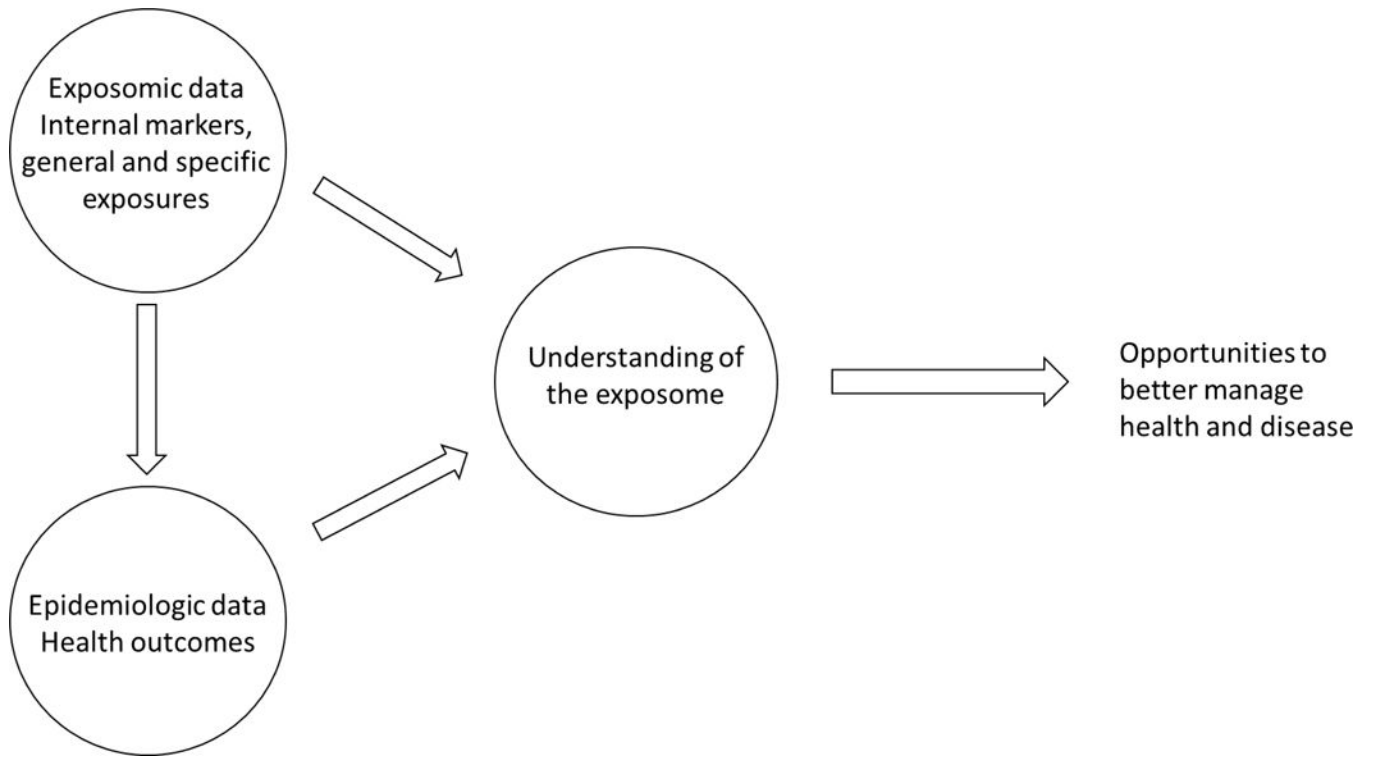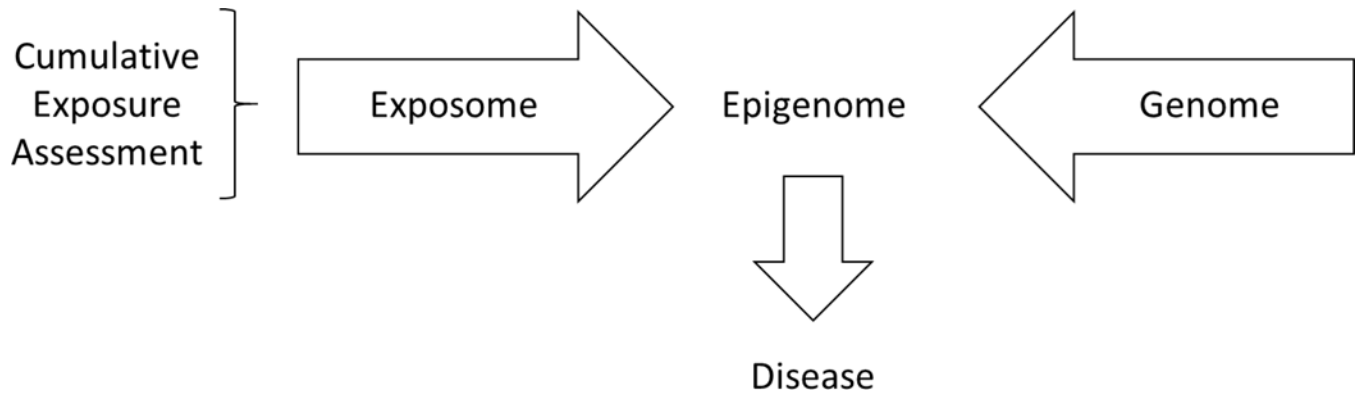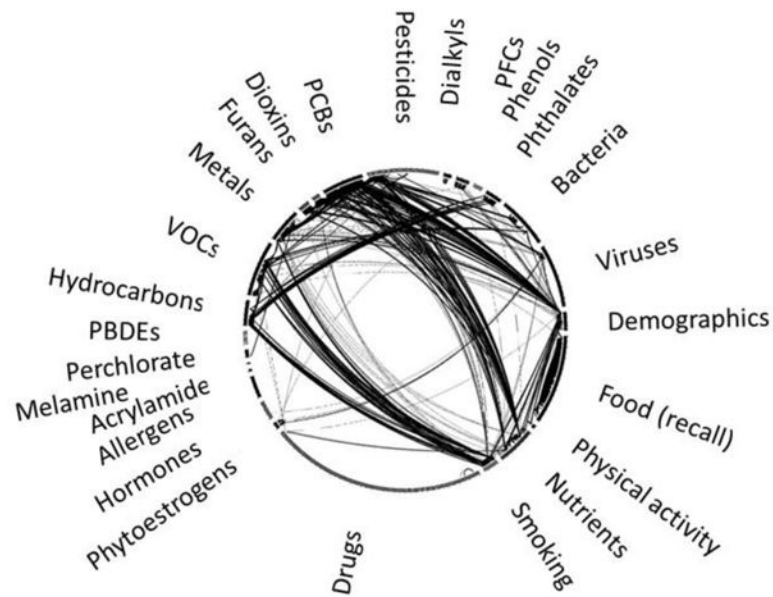
Figure 2.
Exposome-informed epidemiologic research

Figure 3.
An example of an exposome and genome interaction

Figure 4. Example of an overall exposome correlation globe (43)
575 exposures are grouped by a priori defined environmental health categories and displayed in in the globe. Line thickness is proportional to size of the absolute value of correlation coefficient. Adapted and used with permission from Pacific Symposium on Biocomputing. PCBs (polychlorinated biphenyls), PFCs (polyfluorinated compounds) and VOCs (volatile organic compounds.

Author Manuscript

Table 1

Potential Tools to Measure the Exposome

Tool	Domain
Biomonitoring and biomarker data from omics or other techniques	Internal
Sensors for environmental or personal monitoring	Specific external
Geographic Information Systems	General external
Conventional methods such as survey instruments or job-exposure matrices	Specific external
Reality mining from social networks or other sources	General external

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Examples of omics technologies that may be useful for the exposome

Epigenetics	The study of the totality of all heritable changes in gene expression and chromatin organization that are independent of the DNA sequence itself and that can be inherited in a stable manner over cell divisions (65).	The epigenome is comprised of a number of post-translational modifications such as DNA methylation, histone modification and specifically positioned nucleosomes. The epigenome can be dynamic, influenced by environmental factors and extracellular stimuli, and change in response to these factors and is instrumental in developing the phenome (54, 66). It has been suggested that the epigenome may function as an interface between environmental factors and the genome, and that its deregulation by environmental stressors is likely to disrupt different cellular processes and contribute to disease risk (65).
Transcriptomics	The study of the sequence of an RNA mirrors the sequence of the DNA from which it was transcribed. By analyzing the transcriptome, researchers can determine when and where each gene is actively expressed at any given moment. Gene expression microarrays measure packaged mRNA (mRNA with the introns spliced out) as a summary of gene activity (67).	The transcriptome is the complete set of RNA transcripts produced by the genome at any one time. Studies of transcriptome profiles may provide insights into how they are affected by development, disease, or environmental factors (54). Some studies have shown that the transcriptome is highly responsive to environmental exposures. For example, specific environmental exposures were shown to alter the expression of as much as 30% of the transcriptome in specific blood cells, which may also affect the phenome.
Proteomics	The study of the full set of proteins encoded by a genome is known as proteomics and involves the identification, characterization and quantitation of expressed proteins in biological samples (68).	Protein profiles may be useful to understand the underlying cause of disease and may or may not be useful in identifying exposures. Protein profiles can be useful for describing the phenome and may be helpful in delineating health outcome.
Metabolomics	The study of low molecular weight metabolites present in a biological sample (69).	The metabolome is the sum of all low molecular weight metabolites present in a biological sample. Metabolic profiles obtained using metabolomic techniques constitute the measurable part of the metabolic phenotypes (or metabotypes). These metabotypes show large interindividual differences and are characteristic of an individual at a particular time of his/her life; moreover, differences can be seen at the population level when comparing populations from different regions (54). Metabotypes are both genetically and environmentally determined, and it is possible to establish metabolic profiles from samples collected under different conditions (e.g., exposures, treatments, or physiologic states) and evaluate the effect of other environmental exposures on these profiles (54).
Adductomics	The study of macromolecular adducts in the context of an entire genome (29).	DNA adducts are compounds that bind to DNA, causing damage and mutations. These mutations can result in cancer and birth defects in multicellular organisms. The science of adductomics seeks to identify all DNA adducts and the target sequence of each adduct. It is possible to measure adducts of electrophiles that result from reactions with DNA, with glutathione, and with blood proteins such as hemoglobin and human serum albumin (29).
Genomics	The study of genes or their function (68).	While the exposome is complementary to the genome, some identification of genes that may impact exposure or effect of exposure may be useful. For example, genetic variants that code for metabolizing enzymes may increase or decrease the level of internal dose.