Toward Greater Implementation of the Exposome Research Paradigm within Environmental Epidemiology

Jeanette A. Stingone,[1] Germaine M. Buck Louis,[2] Shoji F. Nakayama,[3] Roel C.H. Vermeulen,[4] Richard K. Kwok,[5] Yuxia Cui,[6] David M. Balshaw,[6] and Susan L. Teitelbaum[1]

[1]Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY 10029; email: jeanette.stingone@mssm.edu, susan.teitelbaum@mssm.edu

[2]Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, Maryland 20817; email: louisg@mail.nih.gov

[3]National Institute for Environmental Studies, Tsukuba 305-0053, Japan; email: fabre@nies.go.jp

[4]Institute for Risk Assessment Sciences, Environmental Epidemiology Division, Utrecht University, Utrecht 3584 CM, Netherlands; email: R.C.H.Vermeulen@uu.nl

[5]Epidemiology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709; email: richard.kwok@nih.gov

[6]Exposure, Response, and Technology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709; email: yuxia.cui@nih.gov, balshaw@niehs.nih.gov

Annu. Rev. Public Health 2017. 38:315–27

First published online as a Review in Advance on January 6, 2017

The *Annual Review of Public Health* is online at publhealth.annualreviews.org

https://doi.org/10.1146/annurev-publhealth-082516-012750

Keywords

exposome, environmental epidemiology, life-course epidemiology, multiple exposures

Abstract

Investigating a single environmental exposure in isolation does not reflect the actual human exposure circumstance nor does it capture the multifactorial etiology of health and disease. The exposome, defined as the totality of environmental exposures from conception onward, may advance our understanding of environmental contributors to disease by more fully assessing the multitude of human exposures across the life course. Implementation into studies of human health has been limited, in part owing to theoretical and practical challenges including a lack of infrastructure to support

comprehensive exposure assessment, difficulty in differentiating physiologic variation from environmentally induced changes, and the need for study designs and analytic methods that accommodate specific aspects of the exposome, such as high-dimensional exposure data and multiple windows of susceptibility. Recommendations for greater data sharing and coordination, methods development, and acknowledgment and minimization of multiple types of measurement error are offered to encourage researchers to embark on exposome research to promote the environmental health and well-being of all populations.

INTRODUCTION AND OBJECTIVES

In 2005, Christopher Wild championed the need for a comprehensive environmental exposure complement to the genome to refine epidemiologic findings and public health research (49). He referred to this new research paradigm as the "exposome" and defined it as that which "encompasses life-course environmental exposure (including lifestyle factors) from the prenatal period onwards" (49, p. 1848). Multiple researchers in fields ranging from perinatal and chronic disease epidemiology to exposure science have written commentaries and reviews that highlight the potential benefits of characterizing the exposome and promoting its integration into future studies (2, 6, 10, 24–26, 36, 38, 39, 45, 47, 49).

The exposome research paradigm is distinguished from traditional epidemiologic approaches by three characteristics: expanded and dynamic exposure assessment across multiple exposure domains (50); the integration of data on exposure and response across multiple scales of variation, including across populations, as well as over time and space; and the use of the resulting high-dimensional information on multiple exposure–response relationships for data-driven discovery (9). Researchers' views on the exposome are evolving, ranging from skepticism to enthusiasm despite challenges (10, 33, 36, 40). Additionally, researchers have differing views of the domains that should be included within the exposome research paradigm. Wild refined his original definition to explicitly include three domains of the exposome: internal (e.g., endogenous factors such as metabolism), specific external (e.g., environmental pollutants, chemical exposures, occupation), and general external (e.g., broader social, economic, and/or psychological factors such as socioeconomic status, mental stress, and climate) (50). Some have advocated for at least an initial focus on the internal exposome (37), whereas others have expanded the Wild definition to include measures of biological response to the environment (26).

This article adopts the view that there is merit to incorporating the exposome into environmental health research and applies a broad definition of the exposome that includes multiple domains of exposure as well as biological responses to those exposures. This definition includes multilevel exposures, for example the built environment of a neighborhood and the individual physical activity of its residents, biological markers of exposure and response, and exogenous stressors that may or may not have an associated biomarker. This view allows for flexibility in the implementation of the exposome within environmental health research because individual investigators can approach this paradigm from different perspectives in light of their targeted research goals. This broad view is inclusive of researchers who use the exposome paradigm to understand the exogenous exposures, including policy and social environments, that contribute to health disparities (19), and includes those who utilize untargeted approaches to identify metabolic profiles associated with known environmental exposures (13).

This review addresses the potential benefits and overarching challenges surrounding the incorporation of the exposome into epidemiologic research. Although it does not provide step-by-step

instructions for conducting exposome research, it summarizes the current approaches and initiatives aimed at implementing the exposome research paradigm within epidemiologic studies. In addition, it provides recommendations to address both scientific and practical barriers in order to move the field forward and encourage researchers to integrate the exposome within their own work.

CURRENT STATE OF EXPOSOME RESEARCH WITHIN ENVIRONMENTAL EPIDEMIOLOGY

Rationale for the Exposome Research Paradigm

The exposome paradigm marks a significant departure from traditional approaches to environmental health research by promoting a more holistic view of environmental effects on human health. Traditional epidemiologic approaches often focus on a single exposure, or a class of exposures, and a single outcome or closely related outcomes. Interactions among multiple exposures are often investigated with a focus on an individual pollutant or a related class of toxicants and how its effects may be modified by other environmental, social, and/or genetic factors. The exposure and environmental health communities have already begun to acknowledge the limitations of a one-chemical-at-a-time approach and explore approaches to quantify the health effects of chemical mixtures (3). The exposome adds another level of complexity and is more similar to a mixture of mixtures because, by definition, it encompasses the totality of exposures (49). Owing to its expansive nature, characterizing the exposome requires multiple exposure assessment methodologies, including biological sample analyses and measurement of one's personal environment using sensors, wipes, simulation modeling, and interviews/questionnaires. As a result of this more holistic view, the exposome paradigm is more consistent with the dynamic nature of exposures experienced by individuals across their life course. The exposome encompasses exposure from before conception onward and through its application researchers can begin to investigate acute and cumulative exposures and their effects as well as to identify windows of susceptibility throughout the life course.

In addition to assessing multiple exposures (and multiple outcomes), the holistic view of the exposome promotes the inclusion of both longitudinal and hierarchical exposure measures within epidemiologic approaches. Hierarchical data structures arise when exposure measures are nested within each other. Examples include spatial hierarchy, where exposures can be measured at the individual, local, and community level, and familial hierarchy, where both parental and fetal measures are assessed. Hierarchical data can also arise from within individuals, such as when both exogenous (external) and endogenous (internal) measures of an exposure are assessed. There are existing epidemiologic methods to assess these types of multilevel exposures (10). However, expanded exposure assessment across the multiple domains (i.e., specific external, general external, and internal) of exposure that are the hallmarks of the exposome would greatly facilitate the investigation of multilevel or hierarchical data structures (38). Although the exposome does not necessarily possess an inherent hierarchical structure, the ability to incorporate hierarchical exposures within the exposome facilitates the statistical analysis of upstream contributors to environmental exposures such as climate change and allows for the investigation of exposures that lack internal biomarkers. It also allows researchers to investigate the interplay between the parental and offspring exposome during fetal and early childhood development.

Within this holistic framework, there is active discussion about how best to conceptualize and consider confounding in light of the totality of exposures being considered. If all environmental factors are coexposures within the exposome paradigm, then nonenvironmental factors remain

as potential confounders when assessing the combined effect of the totality of the environmental exposures (i.e., the exposome). The definition of environment used will then dictate which factors are nonenvironmental and thus potential confounders. From a practical standpoint, the statistical methodologies to estimate these combined effects are limited (40). Many researchers will continue to estimate effects between a single causative agent and outcomes within the exposome paradigm, which will require multiple confounding sets depending on the individual exposure of interest (38). In this case, the holistic nature of the exposome can improve researchers' ability to control for confounders and reduce unmeasured confounding by promoting expanded and improved data collection for a variety of environmental (i.e., nongenetic) coexposures (40).

The exposome affords epidemiologists an opportunity to utilize well-established epidemiologic methods aimed at minimizing measurement error and reducing bias, while looking for new associations between environmental exposures and health or disease end points. The complexity of the exposome underscores the need for continued vigilance and adherence to epidemiologic principles such as accurate assessment and adjustment for confounders, consideration of selection biases when reexamining existing populations or recruiting new study populations, and acknowledgment of measurement error and the potential impacts that such errors can have on analytic results. These can help ensure an accurate and reliable interpretation of the data that are emerging from exposome research. Findings from exposome-related analyses have the potential to generate signals that need to be interpreted and followed up using bioinformatics and in vivo and in vitro experiments. Promising signals or pathways will need subsequent validation in independent studies, similar to validation of genome-wide association study findings (18). Thus, the exposome represents a complement to, not a replacement for, the hypothesis-driven research that has successfully advanced the field of environmental health.

Current Approaches to Implement the Epidemiologic Method Relative to the Exposome

Successful implementation of the epidemiologic method relative to the exposome requires two components: the ability to accurately characterize the exposome and the ability to relate the complexity of the exposome to health-related end points. Many, but not all, of the current approaches seeking to examine health-related end points focus specifically on fetal and early childhood development owing to the close temporal association between exposure and effect within this window. Although the examples detailed below often examine the exposome within similar time windows, they illustrate the multiple, diverse avenues that researchers are taking to develop methods and incorporate the exposome into epidemiologic research.

The three projects funded by the European Union's Framework Program 7 (22) represent a coordinated effort to advance this field through characterization of the exposome within multiple time windows, development of exposure assessment and analytic methods, and examination of exposome–response relationships related to a number of health end points. The Health and Environment-wide Associations based on Large population Surveys (HEALS; **http:/heals-eu.eu/**) will develop an integrated methodology for compiling and organizing exposure data across multiple domains and will relate these factors to predict health outcomes. These methods can then be applied to conduct European-wide health assessments related to the exposome, with the eventual goal of improved risk assessment. The EXPOsOMICS project leverages existing long-term European cohorts, and their stored biospecimens, to integrate external exposures from personal exposure monitoring technologies and population-based measures of exposure with internal measures resulting from -omics technologies (46). With a focus on air pollution and water contaminants, EXPOsOMICS will examine this segment of the exposome across studies

of populations at various points in the life course (46). Finally, the Human Early-Life Exposome (HELIX) project supplements existing data for 32,000 mother–child pairs from 6 European birth cohorts with new internal measures of exposure and biological response on a smaller subsample of 1,200 mother–child pairs (48). This approach effectively capitalizes on the existing study infrastructure and past data collection to facilitate the newer, exposome-related measurements, including personal exposure monitoring and analysis of molecular signatures within stored and new biological samples.

Within the United States, the National Institute of Environmental Health Sciences (NIEHS) has supported efforts to advance the technology, capacity, and infrastructure needed to implement the exposome within environmental health research (9). The Health and Exposome Research Center: Understanding Lifetime Exposures (HERCULES; **http://emoryhercules.com**) was funded to develop the expertise and necessary tools and methodologies to provide a more comprehensive assessment of exposures and to advance exposome-related research. More recently, the NIEHS has established an infrastructure to promote the incorporation of the exposome within child health research studies. The Children's Health Exposure Analysis Resource (CHEAR; **https://www.niehs.nih.gov/research/supported/exposure/chear/**) includes a network of laboratories with extensive analytic abilities for exposure assessment and measures of biologic response in a variety of biological samples. CHEAR also supports a data center that serves as a central data repository and provides methodological support for analysis of multiple environmental exposures, as well as develops novel statistical approaches for combining data across studies and analysis of high-dimensional exposure data.

Whereas these previous examples represent coordinated efforts to advance the incorporation of the exposome into epidemiologic research, other examples of current epidemiologic studies have integrated measures of the exposome within their initial study designs. One large-scale initiative is the Japan Environment and Children's Study (JECS) (23). The JECS has enrolled more than 100,000 expectant mothers and utilizes multiple exposure measurement methodologies, including frequent questionnaire application, extensive biospecimen analyses, simulation models, and ambient environmental measurements to prospectively follow children through age 13 years. Having data on multiple exposures within all the exposome exposure domains across time for a large sample of children will allow researchers to characterize the early-life and childhood exposomes within this population. The US National Institute of Child Health and Human Development (NICHD) Longitudinal Investigation of Fertility and the Environment (LIFE) study focuses on the preconception period and aims to characterize the maternal and paternal exposomes and assess their effects on fertility and birth outcomes (5). A recent publication from this study identified reductions in couples' fecundity associated with a number of paternal environmental exposures, including lead, selected phthalates, and polychlorinated biphenyls (4). Although these findings still reflect a chemical-class approach to analysis, the study's data availability on multiple parental exposures facilitates a later, agnostic-based, exposome approach to understanding fertility and time to pregnancy.

One example of an approach that examines the breadth of environmental exposures at a single time point related to disease is a data-driven analysis known as environment-wide association studies (EWAS). Using cross-sectional data from the National Health and Nutrition Examination Survey (NHANES), researchers examined 266 environmental factors collected at a single time point; they aimed to discover nongenetic contributors to type 2 diabetes by examining each factor within a statistical model and then simultaneously assessing the results across all models (30). The EWAS approach has since been applied to studies of preterm birth (32), communication impairments in nine-year-old children (42), and leukocyte telomere length (31). Individual studies have also focused on providing proof-of-principle data using agnostic techniques to identify environmental

signals. These studies often utilize well-defined exposure scenarios such as smoking to investigate whether specific biological signals can be traced in the biological system (16). These types of studies will result in best practices that can be implemented in future epidemiologic studies. Together these examples, from prospective cohorts to secondary analyses of cross-sectional data, illustrate that the exposome research paradigm can be implemented at different scales of epidemiologic research, depending on the research question and the availability and access to data and biospecimen resources.

CHALLENGES TO INCORPORATING THE EXPOSOME INTO EPIDEMIOLOGIC RESEARCH

Investigators have encountered three primary challenges to incorporating the exposome paradigm into epidemiologic research on a larger scale: (*a*) the lack of infrastructure to support the comprehensive exposure assessment activities that are critical to exposome research; (*b*) the difficulty in differentiating normal physiologic variation from environmentally induced changes in measures of biological response (33); and (*c*) the lack of analytical, bioinformatics, and statistical methodologies to process, integrate, and analyze high-dimensional data and the corresponding lack of cross-trained scientific investigators who can develop and implement these approaches.

First, owing to their complexity and the requirement for a large commitment of resources, exposome studies are well served by centralized support and coordination to ensure that potential exposure assessment strategies are rigorously evaluated. This approach promotes valid assessment of exposure and facilitates data sharing, replication, and validation of findings across studies. Encouraging the increased availability of tools and methodologies could especially benefit smaller efforts driven by investigators who could then apply the tools and approaches developed by larger-scale initiatives within their fields of research. Existing coordinated approaches of the European Initiative and the NIEHS CHEAR centers have already produced some discussion and development of analytic methods for exposomic data (1, 15). Similar investments in infrastructure and cross-study collaboration are necessary for life stages beyond pregnancy and birth, including support for linking the early-life exposome to adult disease, as suggested by the developmental origins of health and disease hypothesis (17).

Second, as first suggested by Wild, the exposome is dynamic within and over time; thus, investigators must differentiate periods of random variation from periods of induced biologic response (50). This observation not only translates into a need for accurate and precise measurements for various populations in order to characterize typical biological variation, but also stresses the importance of capturing the temporal variation in the exposome that occurs throughout an individual's life course. Understanding this temporal variation is especially critical for time periods, such as development, puberty, and pregnancy, which are hypothesized to have specific patterns of greater biological variation and may be more susceptible to disruption by environmental exposures (8). Constructing a reference exposome is not feasible due to the dynamic nature of exposure across space and time. However, we can begin to characterize the temporal interrelatedness and variability of both exposures and responses throughout the life course.

Finally, as discussed above, investigations of the exposome are considerably different from traditional epidemiologic approaches. Practically, they require more data storage and more complex data management than do traditional studies. Analytically, statistical and other computational models are needed that can simultaneously accommodate heterogeneity from merging data across populations, longitudinal measurements, and large numbers of correlated exposures. Although genomic research also encounters a multitude of possible features coupled with their potential interactions, the longitudinal nature of environmental exposures is an added point of complexity for exposome research. This unique feature of the exposome needs to be considered in all phases

of research and will stimulate the development of analytic methods that are sensitive to these features. To develop and implement these novel approaches, current and upcoming epidemiologists will require training in complementary disciplines such as computer science, bioinformatics, and advanced biostatistics (27).

RECOMMENDATIONS FOR FUTURE RESEARCH AND SOCIETAL INVESTMENT

Catalog Existing Tools and Methodologies to Facilitate Exposome Research

As discussed above, researchers require easily accessible tools and approaches that can be used to characterize both exogenous and endogenous measures of exposure, including those exposures that do not have known corresponding biomarkers. The creation of a database to catalog and curate exposure assessment methods, ranging from questionnaires to assay protocols, along with their performance characteristics, would allow for wide implementation among the research community. This step would ensure that epidemiologists have access to the most appropriate exposure assessment tools, while fostering the development and testing of methods using collected data. Having an easily accessible inventory of exposure assessment tools would also facilitate the sharing of methods and data across studies. As part of the NIEHS exposome workshop in January 2015 (9), working groups began the process of cataloging available tools for assessing environmental exposures and associated biological responses within their respective fields (11, 12, 44). Thus, this recommendation can likely be implemented soon.

Similarly, establishing a standard methodology to examine and report findings from exposome-focused investigations would serve as a resource to investigators who are planning studies, as well as those who are at the publication stage. This research framework could emulate the STrengthening the Reporting of OBservational studies in Epidemiology–Molecular Epidemiology (STROBE-ME) guidelines (14). It could include guidance for describing biological specimen collection and analysis, as well as for incorporating guidelines for presenting agreement (or lack thereof) between external and internal measures of exposure and potential sources of bias and error throughout the exposure assessment process. Again, this recommendation can be initiated in the short term, and the research framework can continue to evolve as research within this field progresses.

Develop New Tools and Methods to Account for the Unique Aspects of the Exposome

As presented above, one of the primary challenges to interpreting exposome research is differentiating between periods of natural variation and environmentally induced variation in measures of the exposome and then relating these measurements to health outcomes. The key to solving this challenge is two-pronged: First, define patterns of variation for exposures, particularly during critical and sensitive windows of human reproduction and development, and then design studies that examine the effect of changes in the exposome throughout the life course. Similar to how we use NHANES to understand background levels of contaminants, defining variation in critical and sensitive windows is an important step in allowing researchers to place into its appropriate context the variation observed as a result of environmental exposures.

The second step is to create study designs that facilitate comparisons in the exposome over time and in response to environmental exposures. Transgenerational, prospective life-course studies with repeated follow-up, including multiple periods of biological sampling, are one way to explore changes in the exposome that occur from preconception on throughout the life course. However,

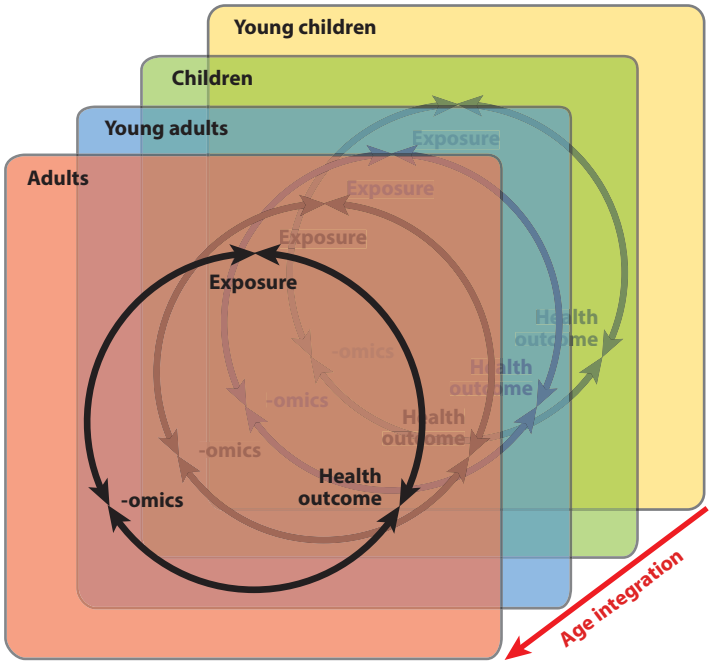

Figure 1

Framework for methodologies that integrate across individual studies to assess the exposome. Each square represents a cohort at a specific life stage. Black arrows illustrate the potential bidirectional relationships between measures of exposure assessment, -omics technologies, and health outcomes assessed within each cohort. Integrating across the ages represented by the different life-stage cohorts will allow investigators to explore the exposome across the life course.

they require lifetimes for completion and an extensive commitment of resources, even when building on existing studies. Additionally, relevant exposures may change during these expansive periods of time. As an alternative, we recommend the development of methods that account for timing of exposure within multiple critical and sensitive windows by merging across different cohorts that have focused on shorter windows. This approach is being investigated by the EXPOsOMICS research group (46). As shown in **Figure 1**, exposure is linked to health end points using advanced exposure science technologies including -omics within each critical life stage, and then integrated across life stages for a life-course view. These methods would require the ability to account for the heterogeneity in study populations as well as the potential differences in exposure assessment and the presence of coexposures. Developing these methods will require an investment of resources and a commitment to increase data sharing and consortium building.

In addition to developing new epidemiologic study designs and approaches for the exposome, the fields of exposure science and biostatistics will also need to develop approaches that can account for the increased complexity of the exposome. New technologies such as sensors that enable individualized measures of external exposure and advanced modeling techniques that better estimate population-level exposures are two examples of the types of improvements in exposure science that will benefit exposome research and minimize measurement error. The Pediatric Research using Integrated Sensor Monitoring Systems (PRISMS; **https://www.nibib.nih.gov/research-funding/prisms**) is a program to develop technologies such as smart phone applications and wearable and home-based sensors to improve exposure assessment in

studies of childhood asthma. It serves as a model of the types of initiatives that are needed for other diseases and stages of life.

New statistical methods are needed specifically to handle the high-dimensional data that will be produced by exposome studies. These methods need to accommodate the increased levels of heterogeneity from combining data from potentially different study populations, as well as the greater amount of uncertainty that arises from examining multiple exposures. However, these methods must continue to account for confounding factors and provide interpretable estimates of disease risk that can be used in risk assessment and health communication. Similarly, we will require better data management tools and increased available storage for the products of exposome-related research (27).

All these new tools and methodologies, whether they are exposure assessment techniques, bioassays, or new approaches to statistical modeling, need to be rigorously assessed in populations with gender, racial, socioeconomic, and geographic diversity. Researchers have suggested that populations with greater susceptibility or higher levels of environmental exposures, such as migrants and low socioeconomic groups, should be a priority for exposome research (41). The exposome has great potential to improve our understanding of health disparities but only if measures and methods are assessed and adapted for various health states in diverse populations. Doing so may require additional outreach to communities that are typically underrepresented in epidemiologic research to both improve recruitment into prospective studies and identify existing biospecimen repositories that represent diverse populations.

Address the Multiple Sources and Types of Uncertainty Present within Exposome Research

The multiple exposures and responses within exposome research can be measured by a variety of methods, ranging from self-administered questionnaires to laboratory assays. Each approach has its own potential for error or misclassification, which when considered in the context of the exposome contributes to the overall uncertainty of the research study. Although uncertainty in the characterization of the exposome should not stop progress, its implications on the obtained results should be thoughtfully considered and reported by investigators. A growing body of empirical research has focused on uncertainty and measurement error for single exposures in epidemiology (7). However, it is not a straightforward process to translate this work to multiple exposures and across the life course.

We caution against using exposome research to rule out exposures of interest, owing to the potential for multiple sources of nondifferential error or misclassification. However, even when exposure misclassification is hypothesized to be nondifferential, additional conditions need to be met in order to conclude that the bias would be toward the null (20). Additionally, the potential for false discovery can be heightened in the context of highly correlated exposures within the exposome, particularly when these exposures are measured with different levels of error (51). As discussed by Zidek et al. (51), simulation studies have illustrated that the association between an outcome and a causal exposure with high measurement error can fail to be observed and instead manifest as an association between the outcome and a highly correlated surrogate of the causal exposure measured with less error. This situation is especially problematic when relying solely on statistical significance to denote associations of interest. As a result, the roles of measurement error and collinearity need to be considered when positive associations are observed in exposome research. As in traditional epidemiologic studies, these findings need to be followed up with additional targeted studies. In the long term, we encourage the development of statistical methodologies that can combine measurement error across multiple sources within the exposome.

Promote Data Sharing and Cross-Discipline Collaboration

To develop the methodologies needed to advance exposomic research, exposure scientists, statisticians, and data scientists need access to epidemiologic data, particularly stored biological specimen resources for diverse populations. The National Institutes of Health (NIH) has a long-standing policy to promote data sharing among its intramural (29) and extramural research (28), although there are obstacles and considerations, particularly when using data derived from human subjects. In 2014, the Roundtable on Environmental Health Sciences of the National Academies of Sciences, Engineering, and Medicine held a workshop on the principles of and obstacles to sharing data from environmental research. As part of this work, researchers discussed ways to increase data sharing, including conducting quality control prior to sharing, developing common language and standards, and providing incentives to share beyond the minimal amount of data (34). Efforts aimed at improving data sharing in the environmental health community would facilitate the development of tools and methods needed to advance exposome research. In addition, epidemiologic researchers both within and external to the environmental health community need to collaborate with researchers who are engaged in exposome methods development to ensure that these approaches are consistent with core epidemiologic principles of study design and analysis.

Consider Causal Inference in the Context of the Exposome

Just as statistical approaches need to be adapted to accommodate unique aspects of the exposome, the approach to causal inference may change with the increased complexity of the exposome. Epidemiology draws on the replication of results and/or the use of meta-analyses to create a consensus out of individual observational research studies (43). The question arises about whether replication is possible if the exposome varies in composition across different populations. Additionally, the exposome is composed of highly correlated exposures, both known and unknown. Within this context, the causal agent and the avenues for intervention may not be clear. There are examples in public health, such as the use of tobacco and exposure to diesel engine exhaust, where an intervention aimed at prevention was implemented without isolating a single chemical responsible for the adverse health effect (21, 35). However, it is possible that targeted studies of individual chemicals/exposures, which incorporate knowledge of sources, toxicology, and other scientific information, can inform public health in a way that might not be possible with the untargeted approaches of the exposome.

As causal analysis methods applicable to observational research continue to advance, investigators incorporating the multiple exposure domains of the exposome will have both the data and the tools for more comprehensive analysis and interpretation of findings from a causal perspective. For example, mediation analyses, which can increase our understanding of the mechanistic path from exposure to disease, will be greatly facilitated by having comprehensive data on exposures, including those with and without biomarkers and associated biological responses for study populations. We may be able to rely more on these investigations of mechanistic pathways to infer causality once these diverse data are more widely available as the exposome research paradigm is implemented.

CONCLUSION

The goal of incorporating the exposome into epidemiologic research is to improve our ability to uncover the environmental contributors to health and disease. The discovery science enabled by the exposome presents a complement to the hypothesis-driven investigations of modern epidemiology, and both are needed to move the field of environmental epidemiology forward. As discussed in this review, many challenges have hindered the incorporation of the exposome on a larger scale

within environmental health research. Infrastructure and coordinated support for the new technologies and approaches needed to measure multiple domains of the exposome within populations have been limited. Characterizing the complexity of the exposome is a daunting task, and many epidemiologists have not been trained to manage, store, or analyze the high-dimensional data that result from exposome investigations. In addition, analytic methods that can account for the complexity of multiple, highly correlated exposures measured at multiple time points have been limited.

Over the past ten years, however, the field has experienced some progression from initial discussions of the paradigm to empirical examples of how the exposome could be implemented, including proof-of-principle studies showing that environmental signals can be traced within the body (16). Investigators have now initiated efforts, such as the HELIX (48) and JECS (23) studies, to design epidemiologic research that capitalizes on new technologies and seeks to characterize the early-life exposome and relate it to health end points. The infrastructure available to support these types of studies has also expanded. To continue this progression and successfully conduct research within the exposome paradigm, we need to adapt current methodologies and create new approaches to handle the unique requirements of assessing the exposome. These include the high dimensionality of data, the dynamic nature of exposure and response throughout the life course, the need for greater and broader data sharing, and the demand for transdisciplinary research teams composed of investigators with diverse skill sets. These recommendations are offered to encourage researchers to embark on exposome research and to further develop the needed approaches and analytic methods in our collective efforts to promote the environmental health and well-being of all populations.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work began as part of a January 2015 workshop on the exposome hosted by the NIH and the National Institute of Environmental Health Sciences. This work was supported in part by NIH grants P30ES023515 and U2CES026555. J.A.S. was supported by the Elizabeth Mascia Scholarship from the Mount Sinai Children's Environmental Health Center. We thank M. Chadeau-Hyam for producing **Figure 1**.

LITERATURE CITED

1. Agier L, Portengen L, Chadeau-Hyam M, Basagana X, Giorgis-Allemand L, et al. 2016. A systematic comparison of linear regression-based statistical methods to assess exposome-health associations. *Environ. Health Perspect.* 124:1848–56
2. Andra SS, Austin C, Arora M. 2016. The tooth exposome in children's health research. *Curr. Opin. Pediatr.* 28:221–27
3. Braun JM, Gennings C, Hauser R, Webster TF. 2016. What can epidemiological studies tell us about the impact of chemical mixtures on human health? *Environ. Health Perspect.* 124:A6–9
4. Buck Louis GM, Barr DB, Kannan K, Chen Z, Kim S, Sundaram R. 2016. Paternal exposures to environmental chemicals and time-to-pregnancy: overview of results from the LIFE study. *Andrology* 4:639–47
5. Buck Louis GM, Sundaram R, Schisterman EF, Sweeney AM, Lynch CD, et al. 2013. Persistent environmental pollutants and couple fecundity: the LIFE study. *Environ. Health Perspect.* 121:231–36
6. Buck Louis GM, Yeung E, Sundaram R, Laughon SK, Zhang CL. 2013. The exposome - exciting opportunities for discoveries in reproductive and perinatal epidemiology. *Paediatr. Perinat. Epidemiol.* 27:229–36

7. Burns CJ, Wright JM, Pierson JB, Bateson TF, Burstyn I, et al. 2014. Evaluating uncertainty to strengthen epidemiologic data for use in human health risk assessments. *Environ. Health Perspect.* 122:1160–65
8. Cohen Hubal EA, de Wet T, Du Toit L, Firestone MP, Ruchirawat M, et al. 2014. Identifying important life stages for monitoring and assessing risks from exposures to environmental contaminants: results of a World Health Organization review. *Regul. Toxicol. Pharmacol.* 69:113–24
9. Cui Y, Balshaw DM, Kwok RK, Thompson CL, Collman GW, Birnbaum LS. 2016. The exposome: embracing the complexity for discovery in environmental health. *Environ. Health Perspect.* 124:A137–40
10. DeBord DG, Carreon T, Lentz TJ, Middendorf PJ, Hoover MD, Schulte PA. 2016. Use of the "exposome" in the practice of epidemiology: a primer on -omic technologies. *Am. J. Epidemiol.* 184:302–14
11. Dennis KK, Auerbach SS, Balshaw DM, Cui Y, Fallin MD, et al. 2016. The importance of the biological impact of exposure to the concept of the exposome. *Environ. Health Perspect.* 124:1504–10
12. Dennis KK, Marder E, Balshaw DM, Cui Y, Lynes MA, et al. 2016. Biomonitoring in the era of the exposome. *Environ. Health Perspect.* **https://doi.org/10.1289/EHP474**
13. Ellis JK, Athersuch TJ, Thomas LD, Teichert F, Perez-Trujillo M, et al. 2012. Metabolic profiling detects early effects of environmental and lifestyle exposure to cadmium in a human population. *BMC Med.* 10:61
14. Gallo V, Egger M, McCormack V, Farmer PB, Ioannidis JPA, et al. 2011. STrengthening the Reporting of OBservational studies in Epidemiology-Molecular Epidemiology (STROBE-ME): an extension of the STROBE statement. *PLOS Med.* 8:e1001117
15. Gennings C, Stingone J, Pajak A, Teitelbaum S. 2016. Statistical considerations for combining data across studies. In *Abstr. 2016 Conf. Int. Soc. Environ. Epidemiol. (ISEE)*, Abstr. 3146. Research Triangle Park, NC: Environ. Health Perspect. **http://dx.doi.org/10.1289/ehp.isee2016**
16. Guida F, Campanella G, Sandanger T, Lund E, Vermeulen R, et al. 2014. Identification of short-term, long-term and lifelong DNA methylation markers of exposure to tobacco smoke: evidence from EPIC and NOWAC studies. *Occup. Environ. Med.* 71(Suppl. 1):A29–30
17. Heindel JJ, Balbus J, Birnbaum L, Brune-Drisse MN, Grandjean P, et al. 2015. Developmental origins of health and disease: integrating environmental influences. *Endocrinology* 156:3416–21
18. Ioannidis JPA, Thomas G, Daly MJ. 2009. Validating, augmenting and refining genome-wide association signals. *Nat. Rev. Genet.* 10:318–29
19. Juarez PD, Matthews-Juarez P, Hood DB, Im W, Levine RS, et al. 2014. The public health exposome: a population-based, exposure science approach to health disparities research. *Int. J. Environ. Res. Public Health* 11:12866–95
20. Jurek AM, Greenland S, Maldonado G. 2008. How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null? *Int. J. Epidemiol.* 37:382–85
21. Kagawa J. 2002. Health effects of diesel exhaust emissions—a mixture of air pollutants of worldwide concern. *Toxicology* 181–182:349–53
22. Karjalainen T. 2014. *Bridging the Science-Policy Gap: EU-Funded Research for Better Environmental Health*. Luxembourg: Eur. Comm. (EU) Bookshop. **http://bookshop.europa.eu/en/bridging-the-science-policy-gap-pbKI0114345/**
23. Kawamoto T, Nitta H, Murata K, Toda E, Tsukamoto N, et al. 2014. Rationale and study design of the Japan Environment and Children's Study (JECS). *BMC Public Health* 14:25
24. Lewis RM, Demmelmair H, Gaillard R, Godfrey KM, Hauguel-de Mouzon S, et al. 2013. The placental exposome: placental determinants of fetal adiposity and postnatal body composition. *Ann. Nutr. Metab.* 63:208–15
25. Lioy PJ, Rappaport SM. 2011. Exposure science and the exposome: an opportunity for coherence in the environmental health sciences. *Environ. Health Perspect.* 119:A466–67
26. Miller GW, Jones DP. 2014. The nature of nurture: refining the definition of the exposome. *Toxicol. Sci.* 137:1–2
27. Mooney SJ, Westreich DJ, El-Sayed AM. 2015. Commentary: epidemiology in the era of big data. *Epidemiology* 26:390–94
28. NIH (Natl. Inst. Health). 2003. *Final NIH statement on sharing research data*. NIH NOT-OD-03-032, Feb. 26. **https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html**
29. NIH (Natl. Inst. Health). 2015. *Intramural Research Program Human Data Sharing (HDS) policy*. NIH 3016, July 31. **https://oma1.od.nih.gov/manualchapters/intramural/3016/**

30. Patel CJ, Bhattacharya J, Butte AJ. 2010. An environment-wide association study (EWAS) on type 2 diabetes mellitus. *PLOS ONE* 5:e10746
31. Patel CJ, Manrai AK, Corona E, Kohane IS. 2016. Systematic correlation of environmental exposure and physiological and self-reported behaviour factors with leukocyte telomere length. *Int. J. Epidemiol.* **https://doi.org/10.1093/ije/dyw043**
32. Patel CJ, Yang T, Hu Z, Wen Q, Sung J, et al. 2014. Investigation of maternal environmental exposures in association with self-reported preterm birth. *Reprod. Toxicol.* 45:1–7
33. Peters A, Hoek G, Katsouyanni K. 2012. Understanding the link between environmental exposures and health: Does the exposome promise too much? *J. Epidemiol. Community Health* 66:103–5
34. Pool R, Rusch E, Roundtable Environ. Health Sci. 2016. *Principles and Obstacles for Sharing Data from Environmental Health Research: Workshop Summary*. Washington, DC: Natl. Acad. Press
35. Proctor RN. 2012. The history of the discovery of the cigarette–lung cancer link: evidentiary traditions, corporate denial, global toll. *Tob. Control* 21:87–91
36. Rappaport SM. 2011. Implications of the exposome for exposure science. *J. Expo. Sci. Environ. Epidemiol.* 21:5–9
37. Rappaport SM, Smith MT. 2010. Environment and disease risks. *Science* 330:460–61
38. Robinson O, Vrijheid M. 2015. The pregnancy exposome. *Curr. Environ. Health Rep.* 2:204–13
39. Rogler G, Vavricka S. 2015. Exposome in IBD: recent insights in environmental factors that influence the onset and course of IBD. *Inflamm. Bowel Dis.* 21:400–8
40. Siroux V, Agier L, Slama R. 2016. The exposome concept: a challenge and a potential driver for environmental health research. *Eur. Respir. Rev.* 25:124–29
41. Smith MT, de la Rosa R, Daniels SI. 2015. Using exposomics to assess cumulative risks and promote health. *Environ. Mol. Mutagen.* 56:715–23
42. Steer CD, Bolton P, Golding J. 2015. Preconception and prenatal environmental factors associated with communication impairments in 9 year old children using an exposome-wide approach. *PLOS ONE* 10:e0118701
43. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, et al. 2000. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 283:2008–12
44. Turner MC, Nieuwenhuijsen M, Anderson K, Balshaw DM, Cui Y, et al. Assessing the exposome with external measures: commentary on the state of the science and research recommendations. *Annu. Rev. Public Health* 38:215–39
45. Vineis P. 2015. Exposomics: Mathematics meets biology. *Mutagenesis* 30:719–22
46. Vineis P, Chadeau-Hyam M, Gmuender H, Gulliver J, Herceg Z, et al. 2016. The exposome in practice: design of the EXPOsOMICS project. *Int. J. Hyg. Environ. Health*. In press. **http://dx.doi.org/10.1016/j.ijheh.2016.08.001**
47. Vrijheid M. 2014. The exposome: a new paradigm to study the impact of environment on health. *Thorax* 69:876–78
48. Vrijheid M, Slama R, Robinson O, Chatzi L, Coen M, et al. 2014. The human early-life exposome (HELIX): project rationale and design. *Environ. Health Perspect.* 122:535–44
49. Wild CP. 2005. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol. Biomark. Prev.* 14:1847–50
50. Wild CP. 2012. The exposome: from concept to utility. *Int. J. Epidemiol.* 41:24–32
51. Zidek JV, Wong H, Le ND, Burnett R. 1996. Causality, measurement error and multicollinearity in epidemiology. *Environmetrics* 7:441–51

TABLE OF CONTENTS FOR VOLUME 1:

Annual Review of Public Health

Volume 38, 2017

Contents

Epidemiology and Biostatistics

Social Environment and Behavior

Environmental and Occupational Health

Public Health Practice and Policy

Health Services

Indexes

Errata

An online log of corrections to *Annual Review of Public Health* articles may be found at http://www.annualreviews.org/errata/publhealth