



“Progress is impossible without change, and those who cannot change their minds cannot change anything.”
George Bernard Shaw (1856-1950)

“If you change the way you look at things, the things you look at change.”
Wayne Dyer (1940-2015)

Food for Thought ...

The Exposome – a New Approach for Risk Assessment

Fenna C. M. Sillé¹, Spyros Karakitsios², Andre Kleensang¹, Kirsten Koehler¹, Alexandra Maertens¹, Gary W. Miller³, Carsten Prasse¹, Lesliam Quiros-Alcala¹, Gurumurthy Ramachandran¹, Stephen M. Rappaport⁴, Ana M. Rule¹, Denis Sarigiannis^{2,5}, Lena Smirnova¹ and Thomas Hartung^{1,6}

¹Johns Hopkins University, Bloomberg School of Public Health and Whiting School of Engineering, Environmental Health and Engineering, Baltimore, MD, USA; ²Aristotle University of Thessaloniki, Center for Interdisciplinary Research and Innovation, HERACLES Research Center on the Exposome and Health, Thessaloniki, Greece; ³Columbia University Mailman School of Public Health, New York, NY, USA; ⁴University of California, School of Public Health, Berkeley, CA, USA; ⁵University School for Advanced Study (IUSS), Science, Technology and Society Department, Environmental Health Engineering, Pavia, Italy; ⁶University of Konstanz, Konstanz, Germany

Abstract

Complementing the human genome with an exposome reflects the increasingly obvious impact of environmental exposure, which far exceeds the role of genetics, on human health. Considering the complexity of exposures and, in addition, the reactions of the body to exposures – i.e., the exposome – reverses classical exposure science where the precise measurement of single or few exposures is associated with specific health or environmental effects. The complete description of an individual's exposome is impossible; even less so is that of a population. We can, however, cast a wider net by foregoing some rigor in assessment and compensating with the statistical power of rich datasets. The advent of omics technologies enables a relatively cheap, high-content description of the biological effects of substances, especially in tissues and biofluids. They can be combined with many other rich data-streams, creating big data of exposure and effect. Computational methods increasingly allow data integration, discerning the signal from the noise and formulating hypotheses of exposure-effect relationships. These can be followed up in a targeted way.

With a better exposure element in the risk equation, exposomics – new kid on the block of risk assessment – promises to identify novel exposure (interactions) and health/environment effect associations. This may also create opportunities to prioritize the more relevant chemicals for risk assessment, thereby lowering the burden on hazard assessment in an exposure-driven approach. Technological developments and synergies between approaches, quality assurance (ultimately as Good Exposome Practices), and the integration of mechanistic thinking will advance this approach.

1 Introduction

Claiming a new 'ome has become fashionable in the post-human genome era. Monya Baker (2013), in her piece in *Nature*, claimed the number to be in the thousands. She wrote: “*Botanist Hans Winkler had no idea what he was starting back in 1920, when he proposed the term ‘genome’ to refer to a set of chromosomes.*

Other 'omes existed even then, such as biome (collection of living things) and rhizome (system of roots), many of them based on the Greek suffix ‘-ome’ – meaning, roughly, ‘having the nature of.’ But it was the glamorization of ‘genome’ by megabuck initiatives such as the Human Genome Project that really set the trend in motion, says Alexa McCray, a linguist and medical informatician at Harvard Medical School in Boston, Massachusetts.

Received January 5, 2020;
© The Authors, 2019.

ALTEX 37(1), 003-023. doi:10.14573/altex.2001051

Correspondence: Thomas Hartung, MD PhD
Center for Alternatives to Animal Testing
Johns Hopkins Bloomberg School of Public Health
615 North Wolfe Street
W7032, Baltimore, MD 21205, USA
(thartun1@jhu.edu)

This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is appropriately cited.



‘By virtue of that suffix, you are saying that you are part of a brand-new exciting science.’” (Baker, 2013). The human toxome, claimed earlier in this series of articles (Hartung and McBride, 2011), and subsequently an NIH Transformative Research Grant (Bouhifd et al., 2015a), was one of Baker’s examples addressed in a favorable way while poking fun at others. She concluded: *“Ideally, branding an area as an ‘ome helps to encourage big ideas, define research questions and inspire analytical approaches to tackle them.”*

Christian Wild in 2005 possibly made the wi(l)dest claim by suggesting the “exposome.” He even could have called it the “lifome” as it encompasses every impact of any exposure on the human body, e.g., chemicals, infections, radiation, as well as internal modulating factors such as the microbiome and metabolism. The goal of the exposome was to comprehensively measure these imprints by transcriptomics, epigenomics, proteomics, metabolomics, and others, primarily in biofluids such as blood and urine. The exposome can include well-established measures of exposure, such as traditional biomonitoring or environmental monitoring, but also includes untargeted discovery of unknown chemicals of biological importance. The exposome concept thus refers to the totality of environmental exposures from conception on, including external and internal components. Expanding on Wild (2005), Rappaport and Smith (2010) functionalized the exposome as the totality of chemicals that can be measured in blood, and in 2014, Rappaport defined the “blood exposome” (Rappaport et al., 2014). The field was furthered in a number of seminal articles (Vineis et al., 2009; Wild, 2009; Miller and Jones, 2014) and books (Miller, 2013; Dagnino and Macherone, 2019). This concept calls to mind a quote from Vietnamese Buddhist monk and peace activist Thich Nhat Hanh: *“You carry Mother Earth within you. She is not outside of you. Mother Earth is not just your environment.”*

The paper by Rappaport and Smith (2010) was critical in popularizing the exposome concept, which was furthered in a National Academies of Sciences report (NRC, 2010) and the National Institute of Environmental Health Sciences (NIEHS) Exposome Workshop in January 2015 (Dennis et al., 2017). Most of the discussion about an exposome, and most of this article, centers on human exposures and health effects. However, the concept can similarly be applied to environmental perturbations. In 2012, the National Research Council (NRC) of the US National Academy of Sciences defined the “eco-exposome” as *“the extension of exposure science from the point of contact between a stressor and receptor inward into the organism and outward to the general environment, including the ecosphere”* (NRC, 2012).

This article was very much inspired by the creation of the Johns Hopkins Exposome Collaborative¹ and its launch event on November 8, 2019. Members of the organization and guest speakers at the event were invited to collaborate on this piece.

2 Nature or nurture? – The debate has a new tool

“The exposome captures the essence of nurture; it is the summation and integration of external forces acting upon our genome throughout our lifespan” (Miller, 2013; Miller and Jones, 2014). Genetics can only explain a minor part of human diseases (Welter et al., 2014; Rappaport, 2016; Jones, 2016). Craig Venter, shortly after the completion of the human genome sequence, stated: *“We simply do not have enough genes for this idea of biological determinism to be right”* (McKie, 2001). After more than 2,000 genome-wide association studies (GWAS), relatively little disease variation has been explained (Visscher et al., 2012), and the median attributable risk of 28 diseases across pairs of monozygotic twins was found to be only 16% – indicating a rather small influence of entire genomes on disease risk (Rappaport, 2016). These results are consistent with the estimate that 70 to 90% of disease risks are environmental factors (Willett, 2002). This predominance of exposure in causing disease is not reflected in current research (Rappaport, 2016): For example, in PubMed citations (Feb. 6, 2016), “disease causes AND genetics” resulted in 566,685 hits, whereas “disease causes AND exposure” resulted in only 71,922. Miller (2013) summarized: *“It is odd that the majority of the discussion of Darwin occurs among geneticists when natural selection is primarily driven by the environment. The environment, viz., the exposome, is what is driving natural selection... if one removes the external forces acting upon our genome, evolution would grind to a halt.”* While less emphasized in research, public opinion is actually characterized by chemophobia, as aptly expressed by actress Sandra Bullock: *“If you can’t pronounce it, you probably shouldn’t be putting it in your body or in your environment.”*² Most people, of course, would have problems pronouncing the names of some vitamins and nutrients. Others, such as former vice president of the United States Dan Quayle (1947-), opine: *“It isn’t pollution that’s harming the environment. It’s the impurities in our air and water that are doing it.”*³ Sarcasm or ignorance? Who can tell?

The increasingly popular Developmental Origins of Health and Disease (DOHaD) hypothesis (Gluckmann and Hanson, 2006) suggests that exposures during critical and sensitive windows of susceptibility may induce epigenetic changes to prepare the embryo or fetus for extrauterine life; this stresses the relevance of the intrauterine exposome. Increasingly, a “life course” approach to understanding chronic diseases (Ben-Shlomo and Kuh, 2002) is being embraced. This indicates a strong impetus for exposure sciences. As Miller and Jones (2014) succinctly stated, *“We need an output of environmental exposures as tangible as the mutated gene.”* New approaches to identify environmental causes of diseases are fundamental to improving individual and population health (Baccarelli, 2019). Traditionally, this was attempted one chemical, one hazard/disease, and one population at the time and could take up to 50 years, even where there was

¹ <https://ehe.jhu.edu/research/the-exposome-collaborative/>

² Sandra Bullock Quotes. BrainyQuote.com, BrainyMedia Inc, 2019. https://www.brainyquote.com/quotes/sandra_bullock_579786 (accessed 28.12.2019).

³ Dan Quayle Quotes. BrainyQuote.com, BrainyMedia Inc, 2019. https://www.brainyquote.com/quotes/dan_quayle_387491 (accessed 28.12.2019).

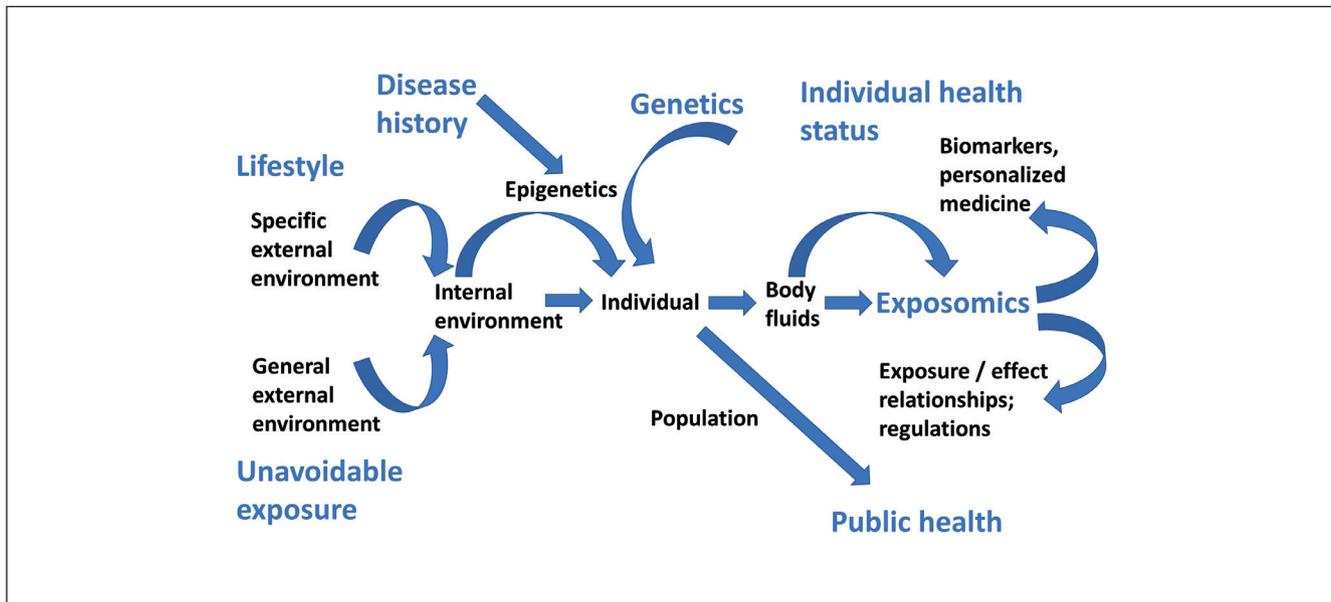


Fig. 1: The human exposome concept

The health status of an individual is imprinted by the specific external (lifestyle) and general external (unavoidable) environment together with the internal environment they influence. Through epigenetics, the individual's disease history as well as directly and indirectly genetics further impact, determining the exposome and reflecting the person's health status. On a population level, exposure/effect relationships can be deduced, supporting regulations to effect public health.

clear exposure and strong correlation of disease such as smoking and lung cancer. This approach misses all the interactions of mixtures, many confounders (individual differences, including disease history of individuals), and cannot typically be integrated easily with mechanistic understanding. Sure, epidemiologists have developed tools to identify covariables, but in each case only a *few likely* ones and plausible mechanisms of action are part of the Bradford-Hill criteria (Hartung et al., 2013) as supportive elements for causation.

Instead of solving all of these associations one at the time, the exposome throws everything on one pile – first for the individual and then for the population. Each and every piece is incomplete: Who recalls the Sunday in 1986, when they sprayed a pesticide in the garden? The exposome reduces the pile, on the one hand, to what is measurable today, either the contaminant or the imprint it left. This reminds us somewhat of diagnosis in Traditional Chinese Medicine, focusing on pulse and tongue only, upon which relevant diseases are assumed to have an impact. This comparison is not meant to belittle this diagnostic approach, where a sophisticated pulse assessment takes up to 30 minutes (for a nice summary, see⁴), but to contrast to the Western medicine approach of measuring everything. Similarly, exposomics focuses on what can be measured in biofluids (blood or urine). This is still a lot, given the breadth of modern omics technologies.

3 The exposome – more than chemical exposures

When considering exposures, we typically think of environmental chemicals. The American comedy writer Robert Orben (1927-) pointedly said: “*There’s so much pollution in the air now that if it weren’t for our lungs there’d be no place to put it all.*”⁵ However, there is far more to consider about exposure. We are not 70 kg rats (Hartung, 2009) – already because approximately 33% of adults and 17% of children in the United States are currently estimated to be obese (Ogden et al., 2012), and 70 kg is a somewhat outdated average for the weight of a human. Other lifestyle (specific external) factors like smoking, diet, and recreational activities, may also play a role in human health and the exposome. Among the general external factors that we as individuals have little influence on are the (residential) environment, climate, socio-economic status and its consequences, as well as many others (summarized in Fig. 1).

Miller and Jones (2014) stressed “*The exposome must explicitly include how our bodies respond to environment pressures, including epigenetic changes and mutations, as well as the complex chemistry resulting from the biochemical reactions that sustain our lives.*” This prompted their redefinition (which we adopt for the purpose of this article; for other definitions see the glossary):

⁴ <http://www.itmonline.org/arts/pulse.htm>

⁵ Robert Orben Quotes. BrainyQuote.com, BrainyMedia Inc, 2019. https://www.brainyquote.com/quotes/robert_orben_104953 (accessed 28.12.2019).



Exposome: The cumulative measure of environmental influences and associated biological responses throughout the lifespan, including exposures from the environment, diet, behavior, and endogenous processes.

They see three distinct differences in this definition from Wild's original definition:

1. The concept of the cumulative biological responses (ongoing adaptations and maladaptations to external forces and chemicals), i.e., the body's response to these challenges.
2. The inclusion of behavior (in a very broad context, to include personal and volitional actions and those that result from family, community, or social units going beyond lifestyle to include the dynamic interaction with our surroundings, relationships, interactions, and physical and emotional stressors).
3. The addition of "endogenous processes" (i.e., the body's complex biochemical reactions)

This definition therefore recognizes the need to consider the totality of external and internal factors.

Beside the omics of biofluids, the exposome makes use of biomonitoring and biomarker data from omics or other techniques (internal), sensors for environmental or personal monitoring (specific external), geographic information systems (GIS) (general external), conventional methods such as survey instruments or job-exposure matrices (specific external), and reality mining from social networks or other sources (general external) (DeBord et al., 2016). The rapid development of material sciences and electronic technologies has enhanced our ability to measure and map specific exposures: These include, for example, sensors that measure a wide range of targeted environmental pollutants, monitors for physiological parameters (i.e., blood pressure, heart rate, stress, physical activity, sleep patterns), GPS devices for monitoring location, mobile phones for dietary assessments and social interactions, and high resolution atmospheric models for predicting ground-level pollutant levels as well as high resolution mass spectrometry and deep sequencing for adducts and metabolic and epigenetic imprints (Niedzwiecki et al., 2019).

4 The exposome – an unattainable vision?

One key technology for assessing the exposome is metabolomics, because it is the omics technology closest to a phenotype. Genes are not necessarily transcribed, mRNAs are not always translated, and some proteins are either not processed or not reflective of the phenotype. When metabolites change, the phenotype changes. Mass-spectrometry-based metabolomics is the key technology for characterizing chemical exposures based on their remnants and metabolites in the body. It should not take a historian (David Christian, 1946-) to explain the obvious: "*Living organisms are created by chemistry. We are huge packages of chemicals.*" However, given the inherent totality of an 'ome, even the metabolome falls short due to lost coverage of small molecules owing to factors such as sample preparation, choice of chromatography, ionization, measurement size window, etc. Consequently, only a fraction of the chemicals present in a sample is

measured. The goal of exposomics, therefore, is to cast a wide net when measuring classes of chemicals to form hypotheses regarding putative biomarkers for follow-up.

Rappaport, in his foreword to Dagnino and Macherone (2019), wrote: "*The vast chemical space and dynamic range of the blood exposome emphasize the importance of untargeted methods for discovering causes of disease. In fact, state-of-the-art liquid chromatography-mass spectrometry can now detect thousands of small molecules and proteins in a few microliters of archived blood from prospective cohorts. Likewise, preprocessing methods for metabolomics and proteomics are rapidly evolving to filter and normalize untargeted proteomics and metabolomics data that are appropriate for statistical analyses. Thus we can now conduct exposome-wide association studies (EWAS) as complements to GWAS in epidemiologic studies (Rappaport, 2012) and thereby identify causal 'signatures and fingerprints' of exposure that fueled Wild's speculation in 2005.*" Keeping in line with the pathway perturbation concept, not all perturbations at the level of gene expression should be taken into account, because many would result in adaptive, not toxic, responses (Smirnova et al., 2015). A reliable way to identify the perturbations that lead to toxic responses is to couple gene expression-based molecular response pathways with the prevalent pathways identified from bioinformatic analysis of metabolite profiles. Metabolomics data provide a closer link to potential phenotypic/clinical observations. Joint transcriptomic and metabolomic data analysis allow us to identify, among the perturbed pathways, those more likely to be associated with adverse outcomes (Hartung and McBride, 2011; Sarigiannis et al., 2018a). Thus, the main weakness of transcriptomics, namely, the failure to distinguish between adaptive and toxic responses, can be overcome. Metabolomics is developing rapidly and is increasingly entering toxicology (Bouhifd et al., 2013; Ramirez et al., 2013), including *in vitro* toxicology (Ramirez et al., 2018). This massive rise in the coverage of large and small molecules in a given sample increases the demand for bioinformatics and cheminformatics from data processing tools to comprehensive databases, and statistical and computational tools for modeling metabolic networks, with the inherent pitfalls illustrated previously: We showed on the one hand that using a data analysis approach based exclusively on pathway annotations has the potential to miss much that is of interest; in the case of misidentified metabolites, it can produce perturbed pathways that are statistically significant yet uninformative for the biological sample at hand. On the other hand, a targeted approach, by narrowing focus and minimizing – but not eliminating – misidentifications, renders the likelihood of a spurious pathway much smaller. The limited number of metabolites, however, also makes statistical significance harder to achieve (Maertens et al., 2017).

With so much information being amassed, new types of metabolomic databases are emerging (Go, 2010) that are not only storing, managing, and analyzing metabolomic data, but also serve as "*gateways to the vast information space of metabolism in living systems.*" The mushrooming of metabolomics has driven the development of a number of freely and commercially available databases that provide information on the chemical structures,

physicochemical and pharmacological properties, spectral profiles, experimental workflows, and biological functions of metabolites. These comprehensive, well-annotated, and user-friendly databases (Go, 2010) are able to store and manage the vast amounts of disparate and complex data with relevant associated annotations and feature user-friendly interfaces with a range of tools that can be tailored according to a specific project. However, we are still missing the equivalent of a gene ontology⁶, i.e., a metabolite ontology connecting each metabolite to its biological role. In short, a great deal of data infrastructure is still needed before metabolomics can begin to compete with transcriptomics. On the other hand, the lessons learned from transcriptomics about the importance of publicly available datasets, data curation, and ontologies can help point the way.

5 Epigenetics as the bridge between genetics and the environment is key to exposomics

Epigenetics arbitrates between genes and the environment. Epigenetic mechanisms control spatial and temporal transcriptomic activity. They include (1) DNA methylation, (2) nine types of histone modifications (acetylation, methylation of lysines, methylation of arginines, phosphorylation, ubiquitylation, sumoylation, ADP ribosylation, deamination and proline isomerization), (3) epitranscriptome (100+ chemically distinct modifications of ribonucleosides), and (4) non-coding RNAs (microRNAs (miRNAs) and long intervening noncoding RNAs (lincRNAs)) that regulate cell-specific gene expression by influencing DNA-protein interactions and chromatin structure (Torres-Padilla et al., 2007; Mikkelsen et al., 2007; van der Heijden et al., 2008; Lee et al., 2012; Smith et al., 2012; Cantone and Fisher, 2013). Dynamic epigenetic programming is regulated during early development; methylation patterns, for example, are then mitotically inherited in somatic cells, thus preserving cell identity of differentiated cells (Smith et al., 2012; Li et al., 2013; Cantone and Fisher, 2013). Emerging data suggest non-traditional roles for CpG and non-CpG (i.e., CpT, CpA, or CpC) methylation (Ziller et al., 2011) that may play a particularly important role in the response to the environment. In addition, disruption of imprinted genes' tissue-specific expression patterns can lead to a variety of behavioral phenotypes (reviewed in Plasschaert and Bartolomei, 2014). Epigenetic regulation is heritable but can be influenced by environmental stimuli, *in utero* circumstances, and aging. As botanist Luther Burbank (1849-1926) noted, "*Heredity is nothing but stored environment.*"⁷ How true for the exposome and epigenetics!

Goodman (2017) recently discussed epigenetics in toxicology, noting that toxicologists have embraced epigenetic regulation of gene expression, as evidenced by their increasing focus on discerning potential epigenetic mechanisms by which chemical and physical agents might cause toxicity. There appears to be a rela-

tion between the epigenome and disease (Jirtle and Skinner, 2007; Waterland and Michels, 2007), and epigenome-wide association studies (Rakyan et al., 2011) are being carried out. Smirnova et al. (2012) summarize epigenetics in a broad variety of toxicological studies. There is, however, no consensus about how this might be systematically incorporated into risk assessment.

Because they can be dynamically regulated while being heritable, epigenetic marks serve at the interface of genes and the environment and have become likely candidates for a mechanistic link between exposure and disease susceptibility. In one of our previous articles in this series, we introduced the term "epigenetic scar" or memory (Smirnova et al., 2015), defined as "*long-term changes in epigenetics induced by stressors, which affect regulation of gene expression.*" This imprinting can lead to disease manifestations, but also can be protective by leveraging the homeostasis of the cell, tissue, organ or organism in response to changes in the exposome. An important aspect of the robustness of an organism's epigenetic response to exposure is redundancy.

In particular, miRNAs have some interesting properties that predestine them as biomarkers (Smirnova and Maertens, 2019):

- miRNA genes are very often encoded on several chromosomes.
- One miRNA may regulate hundreds of mRNAs and one mRNA can be regulated by several miRNAs.
- Often, one mRNA has several binding sites for different miRNAs.
- 60% of mRNAs are miRNA targets.
- A relatively small number of miRNAs, e.g., *Homo sapiens*: ~2578
- Can be measured sensitively by PCR amplification
- Conserved through the phyla, i.e., easy to translate between species
- Fine-tuning of mRNA network through positive and negative feedback loops
- miRNA-targets more likely to be affected by chemicals than non-miRNA-targets
- miRNAs are more stable than mRNAs, allowing analysis in archived formalin-fixed paraffin-embedded (FFPE) blocks, biofluids, archived and degraded samples.
- Circulating (secreted) miRNAs in biofluids lend themselves as biomarkers for organ toxicity and injury.
- Organ and developmental stage-specific expression eases interpretation.

From one perspective, epigenetic mechanisms appear to buffer "noisy" gene expression to maintain homeostasis in response to environmental stress. From another, epigenetic marks can be inherent targets of the exposure. Potential mechanisms for environment-induced DNA methylation changes include increased/decreased expression or inhibition of DNA methyltransferases (Duncan et al., 2013; Schaevitz et al., 2014) and toxicant-induced reactive oxygen species leading to demethylation of DNA (Coulter et al., 2013). Changes in the exposome may alter

⁶ <http://geneontology.org>

⁷ https://www.brainyquote.com/quotes/luther_burbank_277483



miRNA expression, which may change the fine-tuning of gene expression and lead to adverse outcomes in the long run. As elegantly shown by Fraga et al. (2005), differences in the epigenetic signatures of monozygotic twins were indistinguishable at the age of three, and drifted apart later in life, suggesting exposure contribution to the epigenome. Another study demonstrated influences of the environment on the epigenome using the yellow agouti (*Ay*) mouse model, in which coat color variations depend on epigenetic marks established early in development; this model has been used to investigate the impacts of exposure on the fetal epigenome (Feil and Fraga, 2011).

Plusquin et al. (2019) note that “*Genomic DNA isolated from blood cells or other pertinent tissues is being expansively exploited for the discovery of biomarkers of effect and exposure. Technology to measure epigenetic marks on a genomic scale complemented with novel tools for data-analysis have recently been developed and continue to be enhanced.*” Increasingly, they are therefore included in exposomics studies – especially miRNAs as biomarkers for environmental exposure (recently systematically reviewed by Vrijens et al., 2015). In any case, addressing the epigenetic response to perturbations in the exposome should be a key to biomarker identification.

6 Exposomics gives a spin to traditional exposure biomonitoring, overcoming some of its limitations

Biomonitoring, in the closer sense of chemicals (as discussed in NRC, 2006), involves the quantitative measurement of the accumulation of chemicals in tissues:

Biomonitoring: The measurement of the body burden of toxic chemical compounds, elements, or their metabolites, in biological substances.

Recently, with the continuous improvement of measurement technologies, such studies have grown in number of studies, number of chemicals measured, and number of people monitored. This is met with some skepticism, as our ability to measure expands faster than our ability to interpret, and again, in contrast, emphasizes the hypothesis-generating focus of exposomics. For example, Jurgens et al. (2008) noted that “*Biomonitoring reveals the amount of a chemical in an individual’s body, but such knowledge is largely meaningless unless we know at what level in body fluids or tissues health consequences do and don’t occur.*” Similarly, the National Academy of Sciences stated: “*The ability to generate new biomonitoring data often exceeds the ability to evaluate whether and how a chemical measured in an individual or population may cause a health risk or to evaluate its sources and pathways for exposure.*” Clearly, exposomic investigations must be followed up with specific hypotheses to provide the mechanistic understanding linking biomonitoring data with disease pathways and morbidity.

The exposome offers new opportunities for interpreting biomonitoring data as shown by Pino et al. (2017). Still, biomarkers have some disadvantages, as noted by Náráy and Kudász⁸ of the Hungarian Institute of Occupational Health (in their case for monitoring workers). They note that biomarkers:

- “*are usually unable to specify the source of the exposure (occupational or non-occupational);*
- *may not be sufficiently specific to a particular chemical;*
- *are not suitable for identification of workplace contaminations in general;*
- *may be interfered by other chemicals in the biological medium (e.g. medications);*
- *are not useful at all for the assessment/monitoring of acute and/or local toxic effects (e.g. irritation);*
- *and the provision of samples for biomonitoring may be a burden for workers (e.g. blood samples).”*

This shows that a specific use area for biomonitoring or exposomics might come with specific additional considerations. Exposomics contrasts with traditional biomonitoring, which is typically limited to a select group of known chemicals (usually in the tens and rarely lower hundreds) (Buekers et al., 2018). Some prominent disadvantages are that such studies (e.g., the National Health and Nutrition Examination Survey, NHANES⁹) do not take continuous repeated measures over time, thereby limiting detection of short-lived chemicals and suspected chemicals of concern, which are less likely to be captured (Dennis et al., 2017). The focus is usually on the relatively exact measurement of individual chemicals, which is time-intensive for development and validation, and expensive and time-consuming for execution, as well as often limited to few laboratories with such capabilities. Multiple methods are required for a large suite of chemicals, and typically 0.5-2 mL of blood or other biospecimens are needed per analyzed chemical. Additionally, chemicals added for monitoring are not always the most important ones from a toxicologic perspective.

Biomonitoring increasingly incorporates biomarker measurement. The concepts of exposome and biomarkers overlap closely (Rappaport, 2012).

Biomarker: A characteristic that can be objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention (Puntmann, 2009)

Biomarker discovery has grown dramatically in the quest for biomarkers of disease in drug development and diagnostics. Rappaport (2012) distinguishes between biomarkers of environmental exposure (causal pathway) and biomarkers of disease (reactive pathway). Compared to biomonitoring, which assesses the presence of a chemical, biomarkers have a proven predictive value. While biomonitoring would measure a given chemical in any tissue or body fluid, a biomarker would be one that correlates with

⁸ [https://oshwiki.eu/wiki/Biological_monitoring_\(biomonitoring\)#Advantages_and_disadvantages_of_biological_monitoring](https://oshwiki.eu/wiki/Biological_monitoring_(biomonitoring)#Advantages_and_disadvantages_of_biological_monitoring)

⁹ <https://www.cdc.gov/nchs/nhanes/index.htm>

the extent of exposure or its effect (though by including more biomarkers, the difference becomes increasingly blurred). The World Health Organization (WHO) defined in 1993¹⁰: “‘Biomarkers’, are recognized as providing data linking exposure to a chemical with internal dose and outcome and as relevant to the process of risk assessment.” Over the last decades, the validation and acceptance of biomarkers has been formalized (Goodsaid and Mattes, 2013). The FDA-NIH Biomarker Working Group has made the BEST (Biomarkers, EndpointS, and other Tools) Resource available¹¹ for further reference.

Notably, the biomarker concept also has been applied to *in vitro* research (Blaauboer et al., 2012), in part with the hope that biomarkers identified to reflect a mechanism might translate to biomarkers in humans. This may especially hold true for extracellular (secreted or by cell lysis) substances that might also be found in biofluids.

Exposomics uses omics and other data sources to discover biomarkers of both exposure and disease by casting a wide net and accepting that many biomarkers may actually be patterns of signals instead of single compounds. The goal is ultimately not to measure a myriad of signals in each individual (patient), but to reduce the information to a biomarker equivalent. The mechanistic foundation of the biomarker concept as a “marker of biology” is advantageous and an important outcome of the data-driven exposome approach for chemical risk assessment (Steckling et al., 2018).

An additional advantage of the exposome concept is the integration of toxicokinetics for linking exposure (and exposure biomarkers) with internal dosimetry. Starting from exposure biomarker data (Sarigiannis et al., 2016), we are able to reconstruct external exposure, estimate the internal dose in the target tissue that results in the observed biological perturbation, and compare this to a relevant *in vitro* identified biological pathway activating dose (BPAD) (Judson et al., 2011).

7 Advantages of exposomics over other approaches of exposure assessment

The comprehensive coverage of various exposures over time and in combinations, as well as the diversity of individuals considered, is the key strength (and weakness) of the approach. Beyond this, the already quoted NIEHS workshop (Dennis et al., 2017) saw the following advantages of an exposome approach:

- Agnostic approaches are encouraged for detection of emerging exposures of concern.
- Techniques (and development of techniques) promote identification of unknown/emerging exposures of concern.
- Links exogenous exposures to internal biochemical perturbations
- A large number of features can be detected (> 10,000) for the cost of a single traditional biomonitoring analysis.

- Includes biomolecular reaction products (e.g., protein adducts, DNA adducts) for which traditional biomonitoring measurements are often lacking or cumbersome
- Requires a small amount of biologic specimen (~100 µL or less) for full-suite analysis
- Enables detection of “features” that are linked to exposure or disease for further confirmation
- Encourages techniques to capture short-lived chemicals
- Aims to measure biologically meaningful lifetime exposures, both exogenous and endogenous, of health relevance.

8 Challenges for exposomics

Exposome studies also come with their own set of challenges. Some see the lack of stringent *a priori* hypotheses, the multiple testing problem, and the resulting probability of false-positive findings as problems reflecting the traditional view in epidemiology (Wacholder et al., 2004) that the lack of a highly curated *a priori* hypothesis impacts robustness of findings (as they depend not only on the strength of the observed data, but also on the strength of the *a priori* hypothesis). This view, however, misses the hypothesis-generating character of the exposome. The goal is to select from among the myriad variables those worthy of follow-up. Dealing with multiple testing already at this stage is problematic.

The challenge of exposure variability

The variability of exposure data (e.g., exposures with short biological half-lives and little constancy in the underlying exposure behavior) creates obvious difficulties in interpretation. While the genome is essentially static, the exposome varies throughout life, presenting challenges for epidemiologic studies due to temporal, spatial, and genetic variability in both exposures and individuals’ responses to the exposures. The variability in biological distribution of chemicals in the body (the toxicokinetic variability), i.e., how chemicals are distributed and metabolized differently by individuals (Tsaïoun et al., 2016), impacts internal exposure studies. It is also important to know whether a single sample obtained at a given life stage represents the relevant exposure over time or if exposures during a critical window are more important. Even when exposome studies include lifetime exposures, they often do not place enough emphasis on defining and measuring windows of susceptibility (e.g., *in utero*) to accurately capture the most biologically important exposures. Here, it should be noted that the HEALS project in the EU has identified ten critical windows of susceptibility during the human life course, providing a comprehensive guide to designing future exposome studies properly.

The measurement challenge

In 2014, Rappaport defined the “blood exposome” (Rappaport et al., 2014), showing that chemicals derived from foods, drugs, and endogenous sources covered the same dynamic range

¹⁰ <https://apps.who.int/iris/bitstream/handle/10665/39037/9241571551-eng.pdf?sequence=1&isAllowed=y>

¹¹ <https://www.ncbi.nlm.nih.gov/books/NBK326791/>



of blood concentrations (pM–mM), while those from pollution sources were present at levels three orders of magnitude lower (fM– μ M). The ability to identify and quantify low-abundance analytes – most environmental chemicals – is still immature. We will thus often miss chemicals that are present at low levels. Depending on the methodology used, we cannot detect all analytes present (e.g., we can only measure chemicals that are isolated in a given extraction process), and may not be able to define a reference or baseline value. Issues of sample availability and quality, identification of unknown analytes, and capture of non-persistent chemicals remain challenges and must continue to be addressed. Each method also comes with its analytical measurement error. Importantly, biomarkers have not been defined for all environmental exposures (Steckling et al., 2018). Untargeted (feature) measurement promises the discovery of such biomarker (patterns), but also emphasizes the challenges of feature identification, quantification, and reproducibility.

The challenge of method integration

The more exposures are included, the more measurement technologies and types of data have to be combined. This has financial, technical, expertise, and biostatistical consequences. For example, not all technologies can be applied to all subjects (e.g., small children, frail elderly subjects, and permissive/non-permissive work environments). Some measurements are too costly to be applied to large cohort studies. Very few investigators have expertise in all relevant measurement technologies, especially vis-a-vis quality assurance, and thus will require collaborations to increase coverage of the exposome in a given population. Another challenge is the ability to share and work on large omics datasets from different originating laboratories. For this purpose, the Human Toxome Project developed the Human Toxome Collaboratorium (Fasani et al., 2016), a cloud-based platform for data storage with metadata analyses and their documentation. This might be a path toward exposome workflows and reporting. The HEALS project in the EU developed the HEALS geo-database (HEALS GeoDB) as a repository of all geo-referenced measurements and data emanating from the HEALS techniques and methodologies.

The challenge of statistical assessment of increasingly complex data sets

Big data refers not only to the volume of data, but also to their variety and the velocity of data accrual – the so-called 3Vs that are characteristics of big data. This stretches the limits of our existing analytical methodologies. “Big data is transforming the traditional ways of handling data to make sense of the world from which it is collected. Statisticians, for instance, are used to developing methods for analysis of data collected for a specific purpose in a planned way. Sample surveys and design of experiments are typical examples. Big data, in contrast, refers to massive amounts of very high dimensional and even unstructured data which are

continuously produced and stored with much cheaper cost than they are [sic] used to be. High dimensionality combined with large sample size creates unprecedented issues such as heavy computational cost and algorithmic instability” (Pyne et al., 2016). Typically, big data originate from combining data from multiple sources using different technologies over time. Unavoidably, this introduces heterogeneity, experimental variations, and statistical biases that require more adaptive and robust procedures for analysis and curation (Megahed and Jones-Farmer, 2015; Wang et al., 2016; Saidulu and Sasikala, 2017; Reid, 2018¹²; Cox et al., 2018) that increasingly rely on machine learning (also termed artificial intelligence (AI)) (Di Ciaccio et al., 2012; Pyne et al., 2016; Hartung, 2018), also for environmental health (Sarigiannis et al., 2009).

Database challenges

The application of untargeted metabolomics to identify environmental exposures correlated with human health has its own unique challenges (Dennis et al., 2017). The metabolomic aspect of analyses relies heavily upon library searching of spectra for annotation with standard confirmation coming later, which can be quite time-consuming and labor-intensive. The largest reference databases for metabolomics are the Metabolite and Tandem MS Database (METLIN)¹³ and the Human Metabolome Database (HMDB)¹⁴. To date, METLIN and HMDB have largely focused on naturally occurring metabolites. According to Dennis et al. (2017), the number of compounds in METLIN and HMDB that may be potentially relevant to exposure studies has not yet been carefully assessed. The number of databases available for metabolomics continues to expand and has unique utility depending on the research question. A more expansive discussion of metabolomics database resources is available (Go, 2010). To facilitate large-scale exposomic studies, the field may benefit from having database search functionalities specifically dedicated to environmental exposures. As discussed above, discovery experiments are typically most successful when a small subset of relevant features can be targeted for structural identification and follow-up hypotheses. Thus, databases and repositories that curate information on the human exposome (e.g., the HEALS GeoDB) could provide powerful mechanisms for prioritizing features of interest to environmental health scientists.

Bioinformatic challenges

Exposomics requires extensive bioinformatics for data reduction and analysis (Manrai et al., 2017). The NIEHS Exposome Workshop (Dennis et al., 2017) covered bioinformatic needs that are specific to exposomic biomonitoring approaches, seeking to link internal biochemical perturbations to external exposures. “Through pathway analysis and data extraction algorithms, biological pathway perturbations can provide great insight into broad disease processes. Additionally, detection of low-level xenobiotic and unknown chemicals can be greatly enhanced through

¹² Introduction to a series of twelve workshops at the Fields Institute for Research in the Mathematical Sciences

¹³ https://metlin.scripps.edu/landing_page.php?pgcontent=mainPage

¹⁴ <http://www.hmdb.ca>

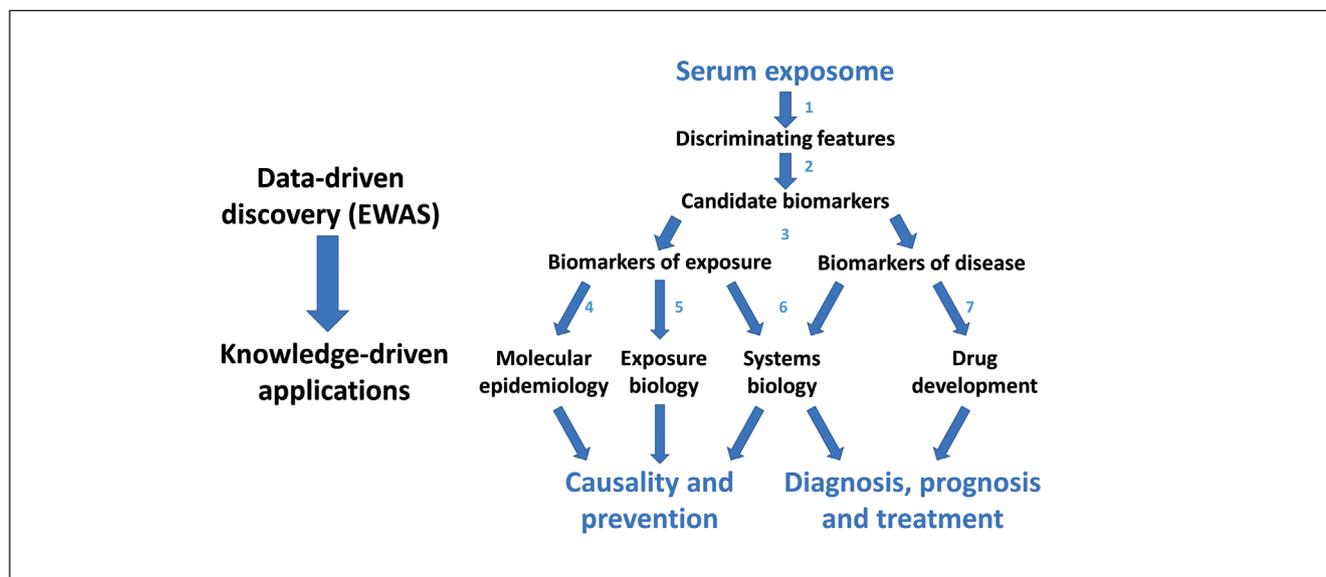


Fig. 2: Scheme for conducting exposome-wide association studies (EWAS) to discover serum biomarkers of exposure and disease and for applying biomarkers to investigate disease causality, prevention, diagnosis and treatment

Redrawn from Rappaport (2012). (1) Diseased versus healthy (case control studies), *untargeted designs*; (2) chemical identification; (3) dose-response; (4) diseased vs. healthy (prospective cohorts), *targeted designs*; (5) identify sources and measure exposures; (6) genomics, epigenetics, transcriptomics & experiments; (7) disease stage and response to therapy.

bioinformatic techniques. Further development of bioinformatic tools and data storage and handling will be key to advancing our understanding of the health impact of complex exposures.” The large numbers of signals in omics measurements compromise the identification of effect signatures and may lead to over-fitted associations with exposure characteristics. This calls for a reduction in data dimensionality through mechanistic reasoning and biomarker identification (Blaauboer et al., 2012) to discriminate relevant signals from epi-phenomena and random responses (Hartung et al., 2012) (see Section 9 below).

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. The key element of analysis will be the pattern recognition in the exposome of people with different exposures or disease manifestations. Pattern recognition is a particular strength of machine learning.

Machine learning comes in different flavors (Jason Brownlee’s Machine Learning Mastery blog lists 14 types¹⁵). The most important forms are supervised learning (we first categorize the data, e.g., exposed/not exposed or healthy/diseased), unsupervised learning (the computer generates clusters based on correlations, etc.), and reinforcement learning (a class of problems where an agent must learn to operate in an environment using feedback without a fixed training dataset, rather a goal to achieve, actions

they may perform, and feedback about performance toward the goal) (Sarigiannis et al., 2018c). A broader discussion goes beyond the scope of this article, but it is obvious that these technologies lend themselves to the mining of exposome datasets.

In light of the exposome’s data-driven approach, analytic plans will need to accommodate the complex and high-dimensional nature of environmental mixtures that reflect large to small biomolecules or biomarkers (e.g., proteome and metabolome), while recognizing their interrelatedness in model specification and statistical analysis. As exposomic analyses generate specific hypotheses about cause and effect, analytic techniques will ultimately need to include multiple testing, false discovery thresholds, joint modeling of exposures and health outcomes across the lifespan (Li et al., 2019), paucity of data on multiple time scales due to the varied nature of exposures, and the need for replication to validate initial associations (Sarigiannis et al., 2018c).

The quality assurance challenge

Exposomics requires carefully collected and well-maintained biospecimens. This includes the methods of specimen collection and storage, as well as study design and available metadata. Quality assurance (QA) continues with the data-generating technologies employed and, in the case of metabolomics, this is in its infancy (Bouhifd et al., 2015b; Beger et al., 2019). An example from our own work is Myint et al. (2017), showing joint bounding of peaks across samples improving the differential analy-

¹⁵ <https://machinelearningmastery.com/types-of-learning-in-machine-learning/>



sis in MS-based metabolomics. The QA of data processing has been mentioned above. We will later argue that mechanistic *in vitro* models might enhance the mechanistic understanding and thereby corroborate identified patterns and clusters of findings. This means that the QA aspects of *in vitro* work (Hartung, 2007; Pamies and Hartung, 2017; Pamies et al., 2017, 2018) and their reporting (Hartung et al., 2019; Krebs et al., 2019) apply also to exposomics. Ultimately, Good Exposome Practices (GEP) will have to be defined to allow assessment of study quality.

The chemical mixture challenge

Environmental factors act in concert, changing phenotype and disease risks. The sheer number of possible combinations of exposures makes the identification of mixture-phenotype associations an analytic challenge. According to Patel (2017), “*The total number of sets of exposures that can exist is immense... For example, if our study had measured three (e.g., serum lead, cadmium, and arsenic) exposures in total, the number of possible combinations ... is 3, including (lead, cadmium), (lead, arsenic), and (arsenic, cadmium). Suppose we measure 100 exposures (N = 100): the total number of combinations of size two that could co-occur is 4950. Of size 5, the total combinations are on the order of millions (exactly 75,287,520). One can imagine the scale of the number of combinations when scaling up to 100s of exposures of the exposome. Therefore, a primary challenge of identification of mixtures in phenotypes is the expansive number of exposure combinations possible.*” Identifying mixture effects is resource-intensive and requires big data (i.e., sample sizes and power). Afterwards, the verification of synergistic relationships is not as straightforward as for single exposure-phenotype associations, including (Vrijheid, 2014): (1) the evaluation of risk estimates for many single exposures in an agnostic, hypotheses-generating manner (EWAS, i.e., exposome-wide association study) (Fig. 2); (2) the evaluation of risk estimates for combined exposures through data-driven dimension reduction methods; and (3) the evaluation of risk estimates for groups of subjects sharing a similar exposome, such as in a Bayesian profile regression analysis.

9 The exposome meets the toxome – mechanistic toxicology to help understand exposome patterns

In the end, the patterns associated with exposures and effects need to make biological sense. They must reflect the bodies’ response. Therefore, the idea that mechanistic understanding can support or even validate (Hartung et al., 2013) findings is very attractive. Our mechanistic knowledge was previously poorly organized in the scientific literature. Independently, Hartung and McBride (2011) and Ankley et al. (2010, 2016) suggested formalized frameworks, termed pathways of toxicity (PoT) (Kleensang et al., 2014) and adverse outcome pathways (AOP), respectively. The former gave rise to the Human Toxome Project (Bouhifd et al., 2015a), while the latter led to much broader activities, including at the Organisation for Economic Co-operation and Development (OECD) level, which marginalized the human toxome approach.

Some conceptual differences are important for the exposome approach: While AOP essentially organize existing knowledge to reduce biases and to organize complex chains of events from exposure to adversity, the PoT employ multi-omics approaches from *in vitro* models to postulate quantitative networks of events. Although based on different criteria (Hartung et al., 2012), these two approaches establish a framework and serve as a gateway for systems toxicology (Sturla et al., 2014; Sauer et al., 2015; Hartung et al., 2017). More generally, while AOP are ideally structured for regulatory toxicology and integrated testing strategies, PoT – with their emphasis on a network-level description of molecular events – are more tailored for understanding the complexity of exposomics, which involves multiple chemicals with multiple targets over a broad timespan. Figure 3 shows how an untargeted exposome can lead to a PoT hypothesis, which then allows its targeted evaluation.

The topic of combining AOP/PoT with the exposome was addressed in a workshop organized by the Integrated Project EXPOSOME at the UFZ Helmholtz Centre for Environmental Research in Leipzig, Germany, in 2015 (Escher et al., 2017). The workshop identified the consideration of endogenous and non-chemical stressors as well as mixture effects through which the exposome enriches the AOP concept. AOPs could further help to anchor system-wide responses to dominant modes or mechanisms of action, and thus increase the confidence in a hypothesis by providing mechanistic plausibility. Consideration of individual susceptibility to exposures is another element of the exposome that enriches the toxome. In this context, susceptibility is affected by genetics, prior exposure pattern, and age of exposure. Aggregate exposure pathways (AEPs) (Teeguarden et al., 2016) have been put forward as a concept to relate external exposures to internal actions on the receptor (i.e., sites of interaction with the body) and triggering the molecular initiating event (MIE) of the AOP. To further explore the utility and benefit of future application of AOPs in exposome assessment, the Leipzig workshop recommended (Escher et al., 2017) strengthening the following research topics:

- 1) “*More AOPs should be developed and deposited in central databases (e.g., <https://aopwiki.org/>). With every new AOP developed the capacity of exposome assessment based on biological responses will increase significantly.*”
- 2) “*Evolutionarily conserved cellular toxicity pathways may serve as common denominators for integrated effect assessments. We advocate the use of [key events] KEs across species for a systems biology-assisted approach to exposome assessment.*”
- 3) “*Chronic exposure (both in terms of exposure duration and delay of effects) represents the typical environmental situation and is most relevant for human health. AOPs for chronic toxicity are at present not well-described. ... an AOP-based exposome assessment should target chronic endpoints and disease outcomes.*”
- 4) “*The research question is an important driver of the type of exposome/AOP research, such as the identification of the main chemical risk drivers in relation to mode of action, the complex multifactorial influences of mixtures or identification of a threshold level that can explain adverse effects.*”

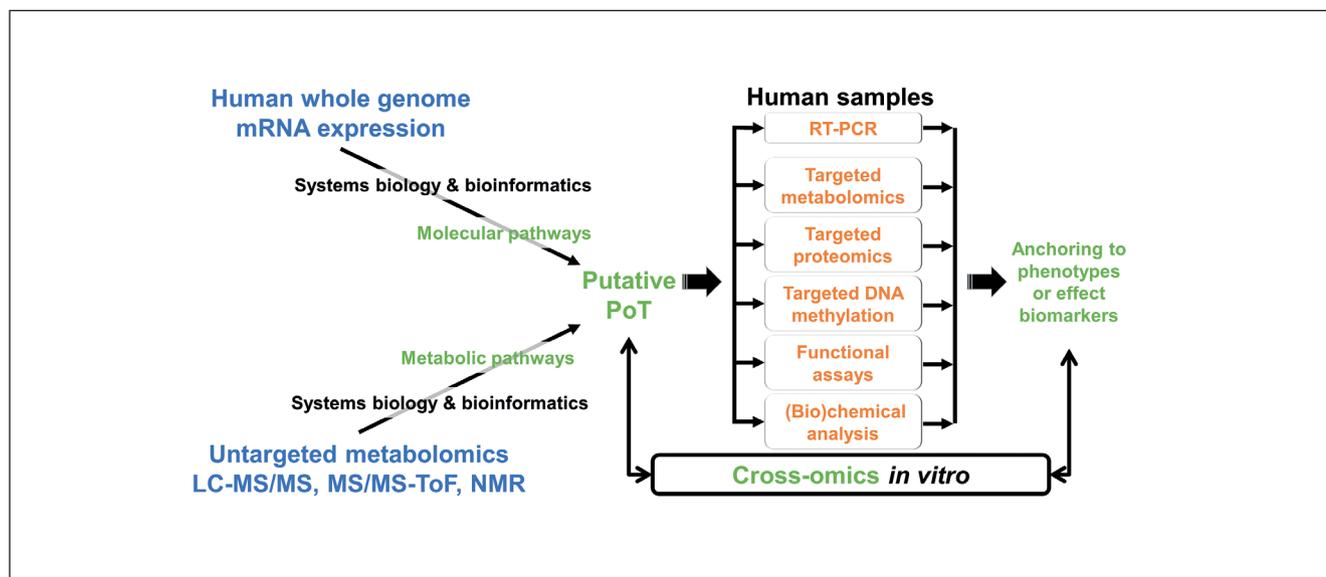


Fig.3: The interlink of mechanistic toxicology and the exposome

Mostly untargeted technologies (blue) are used to broadly characterize the exposome and then deduce a putative pathway of toxicity (PoT). Using targeted approaches (orange), this is validated to anchor to adversity. The elements typical for a PoT/AOP approach (green) interlink these.

10 Combining the exposome and predictive (*in silico*) toxicology tools

Computational toxicology (Hartung and Hoffmann, 2009) is fueled by new approaches, growing computational power, and curated databases (Patlewicz and Fitzpatrick, 2016). The description of an application of big data and AI in toxicology by Luechtefeld et al. (2018b) is an example from our own research which illustrates the enormous computational effort. The approach, developed in collaboration with Underwriters Laboratories (UL), a safety standard and testing organization, and by creating the spin-off, ToxTrack LLC¹⁶ (Hartung, 2016; Luechtefeld and Hartung, 2017; Luechtefeld et al., 2018a), was to organize similar chemicals to derive a prediction, similar to a read-across – a RASAR (read-across-based structure/activity relationship). Briefly, the approach combined several reliable data sources (PubChem¹⁷, ECHA (Luechtefeld et al., 2016), Integrated Chemical Environment (ICE)¹⁸...) and used the most common Chemical Similarity (PubChem2D) and Tanimoto (Jaccard) metrics. Network features using proximity to positive and negative neighbors and data fusion making use of other toxicity, biological and physicochemical endpoints (not only the property to be predicted) were employed. In total, 74 properties were included, which means that each chemical was characterized by a 222-dimensional vector (74 for the chemical itself and 74 each for the closest negative and positive neighbor). Machine learn-

ing (logistic regression, random forest) results in probabilistic hazard estimates using computing clusters (Apache Spark pipeline), allowing massive scale computing. It took a 180-core Amazon cloud server two days to calculate the similarity map for the 10 million structures in our database at about \$5,000 in computing costs. More than 10 trillion one-to-one comparisons were made. The resulting map has chemicals that are similar to each other close together, and those less similar more distant from each other. Now, any chemical, whether included in the original 10 million structures or not, can be placed into the map with “only” half a billion operations performed in less than one second using a standard computer.

To evaluate the quality of predictions, all chemicals with a known classification were predicted in a five-fold cross-validation, which means that 190,000 predictions were made, pretending there was no information on the substance and then comparing the prediction to the known result. This was done for all chemicals, even where no close or contradictory information on a chemical’s neighbors was available. A remarkable 87% correct results were obtained for nine common toxicological hazards.

In parallel, the original ECHA database was scrutinized for chemicals for which multiple animal test results were available. 350-750 chemicals with repeat tests were found for each of the six most commonly used toxicity tests. They were, on average, 81% reproducible. However, this number is inflated by the non-toxic substances, which typically remain negative in re-

¹⁶ <https://toxtrack.com>

¹⁷ <https://pubchem.ncbi.nlm.nih.gov>

¹⁸ <https://bit.ly/2QrAg15>



tests. Only 69% reproducibility was found for toxic chemicals. The UL Cheminformatics Suite¹⁹ has been available since December 2018. The expanded database now contains 30 million structures, 300,000+ with biological and 60,000+ with animal data. The FDA has announced the formal evaluation of the tool for cosmetic ingredients and food additives.

It is now a most promising exercise to combine this approach with exposure information such as toxicants identified in exposomics studies. At this stage, this is limited by uncertainty in feature identification and the fact that the most relevant chronic and systemic toxicities have not yet been implemented.

11 What is in the box for alternatives to animal testing?

The current risk paradigm is essentially based on a hazard and an exposure assessment. Owing to the complexity of exposure assessment, with the notable exception of drugs with defined doses, the emphasis is typically placed on hazard. Some legislation is even *entirely* hazard-based, such as European biocides and plant-protection-product legislation or the 1958 US Delaney clause (which banned carcinogenic substances from being added to food²⁰).

Hazard is traditionally assessed by animal testing. Placing stronger emphasis on exposure, or anticipating it before or concomitant to hazard assessment, gives opportunities to tailor and shortcut animal testing. An example is the concept of the threshold of toxicological concern (TTC) (Hartung, 2017), applicable where exposure assessment shows that expected doses are far below levels likely to produce a health effect. Hartung and Leist (2008), to the best of our knowledge, first suggested an internal TTC, where only the bioavailable fraction (e.g., blood levels) is used to establish whether a critical concentration is reached, “*The idea might be carried further by identifying internal TTCs, i.e. threshold peak plasma levels of toxicological concern, by measuring actual bio-availability. Increasingly, it is recognised how valuable toxicokinetic information for chemicals would be in order to carry out risk assessments and especially to integrate in vitro and in vivo data.*”

Partosch et al. (2015), Coecke et al. (2013), as well as Hartung (2017) have developed the argument for an internal TTC further. The concept gained some traction with a push from the cosmetics industry (Ellison et al., 2019). Here, we clearly overlap with the blood level determinations of toxicants and their metabolites in the exposome. Substances not measured at relevant levels in blood might be weeded out or “put on the back-burner” for hazard assessment, saving unnecessary animal testing. Notably, this is, in the context of the exposome, only possible for substances that are in broad use – typically existing high-production volume chemicals or where data from human trials can be produced. The leverage of an exposure-driven testing approach (Rowbotham and Gibson, 2011; Sewell et al., 2017), i.e., tailoring testing needs to actual exposure, is obvious.

A noteworthy example is the work carried out in HEALS, where environmentally relevant doses were used for the assessment of neurodevelopmental perturbations caused by co-exposure to phthalates and metals based on *in vitro* assays (Sarianni et al., 2018a) and human cohort data. HepaRG cells were exposed to mixtures of DEHP, DiNP, and BBzP phthalates, lead, and mercury. These were the most abundant pollutants in the REPRO PL (Polańska et al., 2011) and PHIME (Trdin et al., 2019) cohorts that studied environmental causation of neurodevelopmental disorders in neonates and children across Europe. The effective concentrations of the chemicals *in vitro* were estimated by extrapolation from human biomonitoring data through internal dosimetry modeling using the INTEGRA computational platform (Sarianni et al., 2016). Cross-omics analyses were performed on the treated cell models, including transcriptomics, proteomics, and metabolomics. Cross-omics data were analyzed using bioinformatics algorithms involving normalization, quality control, advanced statistics, and cross-omics pathway analysis with GeneSpring GX V.14.9. Integrated pathway-level analysis of transcriptomics and proteomics data revealed that co-exposure to phthalates and heavy metals leads to perturbation of the urea cycle due to alterations in the expression levels of arginase-1 and -2, argininosuccinate synthase, carbamoyl-phosphate synthase, ornithine carbamoyltransferase, and argininosuccinate lyase. Co-mapping of proteomics and metabolomics data revealed that their common drivers are responsible for the homeostasis of metabolic pathways related to choline, phosphatidylcholine, phospholipases, and triacylglycerol metabolism. The identification of the urea, phosphatidylcholine biosynthesis I and phospholipase metabolic pathways is of particular interest since these pathways also have been identified in human samples from the REPRO PL and PHIME cohorts using untargeted metabolomics analysis and have been associated with impaired psychomotor development in children from the age of three to six. The importance of the study becomes evident in the capacity of the *in vitro* system to identify the key pathway perturbations associated with these adverse outcomes in humans. This example paves the way for successful prediction of human adverse outcomes using *in vitro* models treated with doses that represent real-life exposure levels, highlighting the importance of the exposome concept in predictive toxicology. Thus, the combination with mechanistic *in vitro* models might enhance the mechanistic understanding and thereby corroborate identified patterns and clusters of findings.

12 Outlook and conclusions

Albert Einstein (1879 -1955) once said, “*We cannot solve our problems with the same thinking we used when we created them.*”²¹ The exposome is such an out-of-the-box approach, as it gives up the single chemical/single hazard paradigm, which has explained some public health problems but has perhaps also misdirected

¹⁹ <https://msc.ul.com/en/products/cheminformatics/>

²⁰ https://en.wikipedia.org/wiki/Food_Additives_Amendment_of_1958

²¹ https://www.brainyquote.com/quotes/albert_einstein_121993

us by focusing on a few but not necessarily the most important ones. Juarez (2019) summarized the enormous potential of the exposome in public health: “*The exposome provides a systems science approach to bringing together and organizing data needed to model the relationships, mechanisms, and pathways among and between external exposures, endogenous exposures, health outcomes, and population-level health disparities. It holds promise for identifying completed exposure pathways from source of exposure in the natural, built, social, and policy environments to route of entry into the body, biomarkers of exposure, biomarkers of disease, disease phenotype, clinical outcomes, and population level disparities, across the lifespan, and between generations.*”

As laid out above, the exposome paradigm brings together exposure science advances, mechanistic toxicology, molecular epidemiology, and bioinformatics/big data analytics. It promises to improve approaches in some key challenges for toxicology and public health, namely:

- Co-exposures at low doses
- Chronic/long-term exposures at low doses
- The role of (epi)genetics and age-related susceptibility in toxicological responses
- Understanding intersecting pathways of toxicity
- Improved public health protection via precision prevention

Dennis et al. (2017) emphasized the rigor and expense for traditional biomonitoring method development, which translates into a slow and costly process for new chemicals of interest. They also state that such analyses often require relatively high volumes of sample, “*Typically 0.5-1 mL for a single method (~ 10 mL urine and > 20 mL serum to measure the 250-300 currently biomonitored chemicals), which can be limiting for certain biospecimen types and age groups under study. Yet targeted analyses are valuable given the accuracy and depth at which a chemical of interest can be assessed. By coupling traditional biomonitoring methods with broader exposomic approaches, the benefits of both strategies can be fully realized... Through untargeted biomonitoring approaches such as high-resolution metabolomics (HRM), > 1,500 metabolites can be monitored with a relatively small amount of biological specimen ($\leq 100 \mu\text{L}$) and for the cost of a single traditional biomonitoring analysis of 8-10 target chemicals... To understand the complexity of exposures faced throughout the lifespan, both traditional and nontraditional biomonitoring methods should be used. Through hybrid approaches and the integration of emerging techniques, biomonitoring strategies can be maximized in research to define the exposome.*”

Another interesting opportunity lies in the utility of exposome studies in vulnerable populations: Smith et al. (2015) have proposed using exposomics tools to quantify cumulative risk and suggested engaging affected communities in participatory exposome research. Potential populations of interest include migrant populations, newborn infants, and communities located near multiple sources of pollution.

Dennis et al. (2017) also emphasized the “fit-for-purpose” concept, which addresses the balance between overall cost of anal-

ysis and the degree of analytical rigor required to use the internal exposure measure results for a given purpose. “*In instances where legal implications exist or regulatory decisions are to be made, maximum analytical rigor is required. But for exploratory studies and for many epidemiologic studies, statistical power derived from a larger number of samples, but with sufficient precision to detect differences, is often preferred. In these cases, relaxation of analytical rigor may translate into lower costs that, in turn, could enable the number of samples analyzed to increase. Furthermore, in untargeted approaches, authentic standards are not always necessary to evaluate a chemical’s relationship to disease or alterations in biomolecular concentrations. In addition, many ‘add-on’ studies use samples collected for different analyses for which the sample collection/storage may represent more imprecision, thus not warranting the increased cost of strict analytical rigor.*” Box 1 shows further recommendations from the NIEHS workshop:

Box 1: Recommendations NIEHS Workshop 2015
(Dennis et al., 2017)

- 1: Encourage secondary analyses of samples collected for traditional targeted chemical studies.
- 2: Evaluate and use standardized measurement platforms with measurement harmonization.
- 3: Use existing resources and databases to obtain information on current exposures that may be important.
- 4: Provide guidance for the use of existing databases and develop tools to allow searches across multiple databases.
- 5: Foster and facilitate discussion with people from different disciplines to discuss the reality of targeted and untargeted analytic capabilities.
- 6: Develop chemistry methods to enable the detection of low-abundance chemicals and to enable differentiation of endogenous molecules from exogenous molecules.
- 7: Develop bioinformatics techniques to enhance detection of unknown chemicals using untargeted methods.
- 8: Encourage development of pharmacokinetic models.

As Buck Louis and Sundaram (2012) summarized, “*Moving forward requires greater trans-disciplinary awareness of the exposome paradigm, a willingness for novel thinking and exploration including recognition of the need to measure both the external and internal environments... and identifying suitable resources for both proof-of-concept or other types of observational research.*” The agnostic approach of the exposome, however, can be problematic for grant funding, which still favors hypothesis-driven research.

A number of international activities are driving the development of the exposome concept and its implementation. On the US side, the NIEHS has been a driving force of exposomics²², making it part of both the 2012-2017 and 2018-2023 NIEHS

²² <https://www.niehs.nih.gov/research/supported/exposure/bio/index.cfm>



Strategic Plans. In 2013 and 2018, NIEHS funded the Health and Exposome Research Center: Understanding Lifetime Exposures (HERCULES) at Emory University²³, which is conducting exposome-focused research and also developing new tools and technologies for assessing the exposome. In 2015, NIEHS launched the Children's Health Exposure Analysis Resource (CHEAR), a major new infrastructure to provide researchers access to laboratory analyses of biological samples from NIH-funded studies on children's health, which has supported more than 30 studies. In 2019, NIEHS started an expanded program, called the Human Health Exposure Analysis Resource (HHEAR)²⁴, which provides a larger community of researchers access to centralized, high-quality exposure assessment services. The ultimate goal of understanding the qualitative and quantitative relationships between exposures and their socio-demographic determinants and the profiles determined by the omics techniques is pursued by the establishment of the Johns Hopkins Exposome Collaborative (Box 2). The challenge lies in the quality of the data, and integrated analysis and statistical modeling of all the big data sets generated, which is why we propose to combine the exposome with AI (Hartung et al., 2019). We propose to establish a functional framework, allowing for full integration of genomics with high-quality exposomics data (GxE), by leading the field towards cutting-edge standardized methodologies and approaches, best practices (Good Exposome Practices, GEP) and ethical standards, community participation, and AI-driven big data analyses to better address public health disparities.

Box 2: The Exposome Collaborative @ Johns Hopkins University



The Exposome Collaborative at Johns Hopkins University was funded in January 2019. Its purpose is to congregate the intellectual and material resources housed under the various disciplines within environmental health sciences and engineering and to evaluate the exposome in a holistic manner.

Goals of The JHU Exposome Collaborative:

- Bring together expertise in all the different disciplines needed to better characterize the exposome in human health studies.
- Develop the needed tools and analyses for exposome assessment.

The ultimate goal is not only to provide a functional exposome framework, but also to lead the field in, e.g., establishing best-practices, ethics standards, standardized methodologies, and multi-modal data analyses using artificial intelligence. Public health studies implementing the established exposome framework will have a competitive advantage over other studies that, e.g., solely focus on genomics or traditional exposure assessments, as the exposome framework will agnostically be able to capture yet unexplained causes of health disparities and identify desperately needed targets for intervention.

The JHU Exposome Collaborative was kicked-off with a launch event on November 8, 2019, to which the authors of this article contributed. Seed funding was provided through the generous donation of Krieger School of Arts & Sciences alumnus Yu Wu (MS '89, PhD '96) and Whiting School of Engineering alumna Chaomei Chen (MS '88) to the Department of Environmental Health and Engineering.

Steering Committee & Principal Investigators:

Founders: Gurumurthy Ramachandran, PhD and Thomas Hartung, MD PhD

Director: Fenna Sillé, PhD

Associate Director: Alexandra Maertens, PhD

Members: Carsten Prasse, PhD, Kirsten Koehler, PhD, Ana Rule, PhD, Andre Kleensang, PhD, Hugh Ellis, PhD, Keeve Nachman, PhD, Scot Miller, PhD, Lesliam Quiros-Alcala, PhD, Meredith McCormack, MD

Contact: exposome@jhu.edu

Across the Atlantic, the European Commission established the European Exposome Cluster. The aim of this cluster, which ran from 2012 until 2019, was to create a framework for the exposome. Three projects were part of this cluster, i.e., EXPOsOMICS²⁵, HELIX²⁶ and HEALS²⁷. The CROME-LIFE project²⁸ ran from 2013 to 2018, focusing on applications of the exposome in the environment and health problems in the Mediterranean region. In addition, the NEUROSOME²⁹ integrated training network, running from 2017 to 2021 with European and American participation, promises to unravel the neurological exposome, focusing on exposome-based investigation of environmental fac-

²³ <http://emoryhercules.com/>

²⁴ <https://hhearprogram.org>

²⁵ <http://www.exposomicsproject.eu/>

²⁶ <http://www.projecthelix.eu/>

²⁷ <http://www.heals-eu.eu/>

²⁸ <https://www.crome-life.eu/>

²⁹ <https://www.neurosoma.eu/>

tors causing or contributing to neurodevelopmental and neurodegenerative disease. An upcoming EU Human Exposome Project will put €100 million into the exposome; its launch event has been announced³⁰. In Japan, an Exposome Conference was held (Zhang et al., 2019) and the Japan Environment and Children's Study (JECS) identifies harmful factors in the environment that affect children's growth and health from birth to 13 years of age (2011-2032) with some exposomics aspects.

The Global Exposome Harmonization Project was recently initiated between US universities including Columbia, Mt. Sinai, Emory and Mayo Clinic with INSERM (France), Masaryk University (Czech Republic), Utrecht University (The Netherlands), University of Antwerp (Belgium), the Helmholtz Society (Germany), University of Vienna (Austria), and Imperial College London (UK). It aims at validation (inter-/intra-laboratory), harmonization of exposome measures and standardization of operating procedures, radical transparency, shared pooled standards and shared bioinformatic platforms. This promises to be a major advance for the field, but little information is currently publicly available.

The biomarker concept offers great promise in refining exposure assessment, establishing the biological plausibility of an exposure-disease association, and evaluating intervention studies. Biomarker techniques, such as the omics approaches, have shown promise and can provide information on mode-of-action and dose-response relationships (Thomas et al., 2011). As these techniques evolve, estimation of internal dose and response markers will be a critical test of the new technologies for application in risk assessment strategies (DeBord et al., 2015).

Improved exposure assessments can feed into the systems biology approach to evaluate how exposures disrupt normal biological processes. Systems toxicology is the integration of classical toxicology with quantitative analysis of large networks of molecular and functional changes occurring across multiple levels of biological organization (Hartung et al., 2012, 2017; Sturla et al., 2014). The connectivity theory (Schmutz et al., 2004) proposes a new environmental health impact assessment paradigm, addressing the complexity related to the interplay of genetic, epigenetic, environmental, dietary, and sociodemographic parameters. This requires coupling high-dimensional biological analysis and system science using big data analytics and bioinformatics. Connectivity applied on the exposome elucidates toxicity pathways (Rappaport, 2018; Vrijheid et al., 2014; Sarigiannis, 2019a) and unravels causal associations between environmental stressors and health, paving the way towards precise prevention and targeted policy interventions (Turner et al., 2018). Determining prevalent co-exposures and related main effects (Sarigiannis and Karakitsios, 2018) using machine-learning methods and data analytic approaches ("exposome-wide association studies", i.e., EWAS) (Fig. 2) on large cohorts is expected to enhance our capacity to identify the most harmful environmental chemical mixtures (Sarigiannis, 2019b). The feasibility of using mechanistic analysis of toxicity pathways using multi-omics data (Turner et al., 2018) has been demonstrated, for example, in the NIH Human Toxome Project (Bouhifd et al.,

2015a; Kleensang et al., 2014). Analysis of downstream metabolic products or altered gene expression could provide the entry points for precision medicine (Hartung and McBride, 2011), while the integration of genome and exposome data can support precise prevention, both on population and individual levels (Sarigiannis et al., 2018b).

At the Johns Hopkins Exposome Collaborative launch, Denis Sarigiannis ended his presentation with the following remarks: "*Exposome science can overhaul the current environmental health risk assessment paradigm. This requires the combination of:*

- *Refined external and internal exposure assessment*
- *High-dimensional biology and system science aiming at integration using big data analytics (multi-omics) and bioinformatics*
- *Deeper integration is needed regarding human and in vitro data; there is higher convergence as we move to pathways that regulate homeostasis.*
- *Cross-omics analysis and interpretation of metabolic pathways allow us to target our interventions against the mechanism that is highly perturbed – 'pathway activation' based interventions.*

This has implications on both community and personalized levels: At community level, the most cost-efficient policies should be prioritized (exposome informed policy-making) and at individual level, targeted dietary and lifestyle recommendations should be derived, for minimizing exposure, as well as the propagation of the adverse outcome pathways."

On the same occasion, Gary W. Miller concluded by saying, "*For years the exposome has been a non-descript, fuzzy, and, at times, rather unscientific idea. Today, we can say that the exposome represents a new science – a new way of approaching how the environment influences health. Truly a new discipline. As such, the field must establish and defend its identity, develop its core principles (i.e. the constellation), and forge a path forward (journal, society, regular meetings, websites, etc.)."*

We are witnessing a scientific revolution in the making in toxicology (Hartung, 2008).

References

- Ankley, G. T., Bennett, R. S., Erickson, R. J. et al. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem* 29, 730-741. doi:10.1002/etc.34
- Ankley, G., Escher, B., Hartung, T. et al. (2016). Pathway-based approaches for environmental monitoring and risk assessment. *Chem Res Toxicol* 29, 1789-1790. doi:10.1021/acs.chemrestox.6b00321
- Baccarelli, A. A. (2019). The human exposome: A new "omic" ready for prime time. *J Am Coll Cardiol* 74, 1329-1331. doi:10.1016/j.jacc.2019.07.029
- Baker, M. (2013). The 'omes puzzle. *Nature* 494, 416-419. doi:10.1038/494416a
- Beger, R. D., Dunn, W. B., Bandukwala, A. et al. (2019). Towards quality assurance and quality control in untarget-

³⁰ https://europa.eu/newsroom/events/launch-event-european-human-exposome-network_en



- ed metabolomics studies. *Metabolomics* 15, 4. doi:10.1007/s11306-018-1460-7
- Ben-Shlomo, Y. and Kuh, D. (2002). A life course approach to chronic disease epidemiology: Conceptual models, empirical challenges and interdisciplinary perspectives. *Int J Epidemiol* 31, 285-293. doi:10.1093/ije/31.2.285
- Blaauboer, B. J., Boekelheide, K., Clewell, H. J. et al. (2012). The use of biomarkers of toxicity for integrating in vitro hazard estimates into risk assessment for humans. *ALTEX* 29, 411-425. doi:10.14573/altex.2012.4.411
- Bouhifd, M., Hartung, T., Hogberg, H. T. et al. (2013). Review: Toxicometabolomics. *J Appl Toxicol* 33, 1365-1383. doi:10.1002/jat.2874
- Bouhifd, M., Andersen, M. E., Baghdikian, C. et al. (2015a). The human toxome project. *ALTEX* 32, 112-124 doi:10.14573/altex.1502091
- Bouhifd, M., Beger, R., Flynn, T. et al. (2015b). Quality assurance of metabolomics. *ALTEX* 32, 319-326. doi:10.14573/altex.1509161
- Buck Louis, G. M. and Sundaram, R. (2012). Exposome: Time for transformative research. *Stat Med* 31, 2569-2575. doi:10.1002/sim.5496
- Buekers, J., David, M., Koppen, G. et al. (2018). Development of policy relevant human biomonitoring indicators for chemical exposure in the European population. *Int J Environ Res Public Health* 15, 2085-3003. doi:10.3390/ijerph15102085
- Cantone, I. and Fisher, A. G. (2013). Epigenetic programming and reprogramming during development. *Nat Struct Mol Biol* 20, 282-289. doi:10.1038/nsmb.2489
- Coecke, S., Pelkonen, O., Leite, S. B. et al. (2013). Toxicokinetics as a key to the integrated toxicity risk assessment based primarily on non-animal approaches. *Toxicol In Vitro* 27, 1570-1577. doi:10.1016/j.tiv.2012.06.012
- Coulter, J. B., O'Driscoll, C. M. and Bressler, J. M. (2013). Hydroquinone increases 5-hydroxymethylcytosine formation through ten eleven translocation 1 (TET1) 5-methylcytosine dioxygenase. *J Biol Chem* 288, 28792-28800. doi:10.1074/jbc.M113.491365
- Cox, D. R., Kartsonaki, C. and Keogh, R. H. (2018). Big data: Some statistical issues. *Stat Probab Lett* 136, 111-115. doi:10.1016/j.spl.2018.02.015
- Dagnino, S. and Macherone, A. (2019). *Unraveling the Exposome: A Practical View*. Cham, Switzerland: Springer Nature. doi:10.1007/978-3-319-89321-1
- DeBord, D. G., Burgoon, L., Edwards, S. W. et al. (2015). Systems biology and biomarkers of early effects for occupational exposure limit setting. *J Occup Environ Hyg* 12, Suppl 1, S41-S54 doi:10.1080/15459624.2015.1060324
- DeBord, D. G., Carreón, T., Lentz, T. J. et al. (2016). Use of the "exposome" in the practice of epidemiology: A primer on -omic technologies. *Am J Epidemiol* 184, 302-314. doi:10.1093/aje/kwv325
- Dennis, K. K., Marder, E., Balshaw, D. M. et al. (2017). Biomonitoring in the era of the exposome. *Environ Health Perspect* 125, 502-510. doi:10.1289/EHP474
- Di Ciaccio, A., Coli, M. and Angulo Ibanez, J. M. (2012). *Advanced Statistical Methods for the Analysis of Large Data-Sets*. Heidelberg, Germany: Springer. doi:10.1007/978-3-642-21037-2
- Duncan, T. M., Reed, M. C. and Nijhout, H. F. (2013). A population model of folate-mediated one-carbon metabolism. *Nutrients* 5, 2457-2474. doi:10.3390/nu5072457
- Ellison, C. A., Blackburn, K. L., Carmichael, P. L. et al. (2019). Challenges in working towards an internal threshold of toxicological concern (iTTC) for use in the safety assessment of cosmetics: Discussions from the Cosmetics Europe iTTC Working Group workshop. *Regul Toxicol Pharmacol* 103, 63-72. doi:10.1016/j.yrtph.2019.01.016
- Escher, B. I., Hackermüller, J., Polte, T. et al. (2017). From the exposome to mechanistic understanding of chemical-induced adverse effects. *Environ Int* 99, 97-106. doi:10.1016/j.envint.2016.11.029
- Eskes, C., Boström, A.-C., Bowe, G. et al. (2017). Good cell culture practices & in vitro toxicology. *Toxicol In Vitro* 45, 272-277. doi:10.1016/j.tiv.2017.04.022
- Fasani, R. A., Livi, C. B., Choudhury, D. R. et al. (2016). The human toxome collaboratorium: A shared environment for multi-omic computational collaboration within a consortium. *Front Pharmacol* 6, 322. doi:10.3389/fphar.2015.00322
- Feil, R. and Fraga, M. F. (2011). Epigenetics and the environment: Emerging patterns and implications. *Nat Rev Genet* 13, 97-109. doi:10.1038/nrg3142
- Fraga, M. F., Ballestar, E., Paz, M. F. et al. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 102, 10604-10609. doi:10.1073/pnas.0500398102
- Gluckman, P. D. and Hanson, M. A. (2006). *Developmental Origins of Health and Disease*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511544699
- Go, E. P. (2010). Database resources in metabolomics: An overview. *J Neuroimmune Pharmacol* 5, 18-30. doi:10.1007/s11481-009-9157-3
- Goodman, J. I. (2017). Incorporation of an epigenetic evaluation into safety assessment: What we first need to know. *Curr Opin Toxicol* 3, 20-24. doi:10.1016/j.cotox.2017.05.001
- Goodsaid, F. and Mattes, W. B. (eds.) (2013). *The Path from Biomarker Discovery to Regulatory Qualification*. Academic Press. doi:10.1016/C2011-0-07034-5
- Hartung, T. (2007). Food for thought ... on cell culture. *ALTEX* 24, 143-147. doi:10.14573/altex.2007.3.143
- Hartung, T. (2008). Towards a new toxicology – Evolution or revolution? *Altern Lab Anim* 36, 635-639. doi:10.1177/026119290803600607
- Hartung, T. and Leist, M. (2008). Food for thought ... on the evolution of toxicology and phasing out of animal testing. *ALTEX* 25, 91-96. doi:10.14573/altex.2008.2.91
- Hartung, T. (2009). Toxicology for the twenty-first century. *Nature* 460, 208-212. doi:10.1038/460208a
- Hartung, T. and Hoffmann, S. (2009). Food for thought ... on in silico methods in toxicology. *ALTEX* 26, 155-166.

- doi:10.14573/altex.2009.3.155
- Hartung, T. and McBride, M. (2011). Food for thought ... on mapping the human toxome. *ALTEX* 28, 83-93. doi:10.14573/altex.2011.2.083
- Hartung, T., van Vliet, E., Jaworska, J. et al. (2012). Systems toxicology. *ALTEX* 29, 119-128. doi:10.14573/altex.2012.2.119
- Hartung, T., Hoffmann, S. and Stephens, M. (2013). Mechanistic validation. *ALTEX* 30, 119-130. doi:10.14573/altex.2013.2.119
- Hartung, T. (2016). Making big sense from big data in toxicology by read-across. *ALTEX* 33, 83-93. doi:10.14573/altex.1603091
- Hartung, T. (2017). Thresholds of toxicological concern – Setting a threshold for testing below which there is little concern. *ALTEX* 34, 331-351. doi:10.14573/altex.1707011
- Hartung, T., FitzGerald, R., Jennings, P. et al. (2017). Systems toxicology - Real world applications and opportunities. *Chem Res Toxicol* 30, 870-882. doi:10.1021/acs.chemrestox.7b00003
- Hartung, T. (2018). Making big sense from big data. *Front Big Data* 1, 5. doi:10.3389/fdata.2018.00005
- Hartung, T., de Vries, R., Hoffmann, S. et al. (2019). Toward good in vitro reporting standards. *ALTEX* 36, 3-17. doi:10.14573/altex.1812191
- Jirtle, R. L. and Skinner, M. K. (2007). Environmental epigenomics and disease susceptibility. *Nat Rev Genet* 8, 253-262. doi:10.1038/nrg2045
- Jones, D. P. (2016). Sequencing the exposome: A call to action. *Toxicol Rep* 3, 29-45. doi:10.1016/j.toxrep.2015.11.009
- Juarez, P. D. (2019). The public health exposome. In S. Dagnino and A. Macherone (eds.), *Unraveling the Exposome: A Practical View* (23-62). Cham, Switzerland: Springer Nature. doi:10.1007/978-3-319-89321-1
- Juberg, D. R., Bus, J. and Katz, D. S. (2008). The opportunities and limitations of biomonitoring. *Mackinac Center, Policy Brief February 2008*. <https://www.mackinac.org/archives/2008/s2008-01.pdf>
- Judson, R. S., Kavlock, R. J., Setzer, R. W. et al. (2011). Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. *Chem Res Toxicol* 24, 451-462. doi:10.1021/tx100428e
- Kleensang, A., Maertens, A., Rosenberg, M. et al. (2014). Pathways of toxicity. *ALTEX* 31, 53-61. doi:10.14573/altex.1309261
- Krebs, A., Waldmann, T., Wilks, M. F. et al. (2019). Template for the description of cell-based toxicological test methods to allow evaluation and regulatory use of the data. *ALTEX* 36, 682-699. doi:10.14573/altex.1909271
- Lee, Y. H., Ma, H., Tan, T. Z. et al. (2012). Protein arginine methyltransferase 6 regulates embryonic stem cell identity. *Stem Cells Dev* 21, 2613-2622. doi:10.1089/scd.2011.0330
- Li, L., Lu, X. and Dean, J. (2013). The maternal to zygotic transition in mammals. *Mol Aspects Med* 34, 919-938. doi:10.1016/j.mam.2013.01.003
- Li, N., Friedrich, R., Maesano, C. N. et al. (2019). Lifelong exposure to multiple stressors through different environmental pathways for European populations. *Environ Res* 179A, 108744. doi:10.1016/j.envres.2019.108744
- Luechtefeld, T., Maertens, A., Russo, D. P. et al. (2016). Global analysis of publicly available safety data for 9,801 substances registered under REACH from 2008-2014. *ALTEX* 33, 95-109. doi:10.14573/altex.1510052
- Luechtefeld, T. and Hartung, T. (2017). Computational approaches to chemical hazard assessment. *ALTEX* 34, 459-478. doi:10.14573/altex.1710141
- Luechtefeld, T., Marsh, D., Rowlands, C. et al. (2018a). Machine learning of toxicological big data enables read-across structure activity relationships (RASAR) outperforming animal test reproducibility. *Toxicol Sci* 165, 198-212. doi:10.1093/toxsci/kfy152
- Luechtefeld, T., Rowlands, C. and Hartung, T. (2018b). Big-data and machine learning to revamp computational toxicology and its use in risk assessment. *Toxicol Res* 7, 732-744. doi:10.1039/C8TX00051D
- Maertens, A., Bouhifd, M., Zhao, L. et al. (2017). Metabolomic network analysis of estrogen-stimulated MCF-7 cells: A comparison of over-representation analysis, quantitative enrichment analysis and pathway analysis versus metabolite network analysis. *Arch Toxicol* 91, 217-230. doi:10.1007/s00204-016-1695-x
- Manrai, A. K., Cui, Y., Bushel, P. et al. (2017). Informatics and data analytics to support exposome-based discovery for public health. *Annu Rev Public Health* 38, 279-294. doi:10.1146/annurev-publhealth-082516-012737
- McKie, R. (2001). Revealed: The secret of human behaviour. *The Observer* (London). (Cited from Miller and Jones, 2014).
- Megahed, F. M. and Jones-Farmer, L. A. (2015). Statistical perspectives on “big data”. In S. Knoth and W. Schmid (eds.), *Frontiers in Statistical Quality Control 11* (29-47). 2nd edition. Cham, Switzerland: Springer. doi:10.1007/978-3-319-12355-4_3
- Mikkelsen, T. S., Ku, M., Jaffe, D. B. et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553-560. doi:10.1038/nature06008
- Miller, G. W. (2013). *The Exposome: A Primer*. Waltham, MA: Academic Press, Elsevier, Inc.
- Miller, G. W. and Jones, D. P. (2014). The nature of nurture: Refining the definition of the exposome. *Toxicol Sci* 137, 1-2. doi:10.1093/toxsci/kft251
- Myint, L., Kleensang, A., Zhao, L. et al. (2017). Joint bounding of peaks across samples improves differential analysis in mass spectrometry-based metabolomics. *Anal Chem* 89, 3517-3523. doi:10.1021/acs.analchem.6b04719
- Niedzwiecki, M. M., Walker, D. I., Vermeulen, R. et al. (2019). The exposome: Molecules to populations. *Annu Rev Pharmacol Toxicol* 59, 107-127. doi:10.1146/annurev-pharmtox-010818-021315
- NRC – National Research Council (2006). *Human Biomonitoring for Environmental Chemicals*. Washington, DC: The National Academies Press. doi:10.17226/11700
- NRC (2010). *The Exposome: A Powerful Approach for Evaluat-*



- ing *Environmental Exposures and Their Influences on Human Disease*. Washington, DC: The National Academies.
- NRC (2012). *Exposure Science in the 21st Century: A Vision and a Strategy*. Washington, DC: National Academic Press. doi:10.17226/13507
- Ogden, C. L., Carroll, M. D., Kit, B. K. et al. (2012). Prevalence of obesity in the United States, 2009-2010. *National Center for Health Statistics (NCHS) Data Brief, No 82*, 1-8. Hyattsville, MD.
- Pamies, D. and Hartung, T. (2017). 21st century cell culture for 21st century toxicology. *Chem Res Toxicol* 30, 43-52. doi:10.1021/acs.chemrestox.6b00269
- Pamies, D., Bal-Price, A., Simeonov, A. et al. (2017). Good cell culture practice for stem cells and stem-cell-derived models. *ALTEX* 34, 95-132. doi:10.14573/altex.1607121
- Pamies, D., Bal-Price, A., Chesne, C. et al. (2018). Advanced good cell culture practice for human primary, stem cell-derived and organoid models as well as microphysiological systems. *ALTEX* 35, 353-378. doi:10.14573/altex.1710081
- Partosch, F., Mielke, H. and Stahlmann, R. et al. (2015). Internal threshold of toxicological concern values: Enabling route-to-route extrapolation. *Arch Toxicol* 89, 941-948. doi:10.1007/s00204-014-1287-6
- Patel, C. J. (2017). Analytic complexity and challenges in identifying mixtures of exposures associated with phenotypes in the exposome era. *Curr Epidemiol Rep* 4, 22-30. doi:10.1007/s40471-017-0100-5
- Patlewicz, G. and Fitzpatrick, J. M. (2016). Current and future perspectives on the development, evaluation, and application of in silico approaches for predicting toxicity. *Chem Res Toxicol* 29, 438-451. doi:10.1021/acs.chemrestox.5b00388
- Pino, A., Chiarotti, F., Calamandrei, G. et al. (2017). Human biomonitoring data analysis for metals in an Italian adolescents cohort: An exposome approach. *Environ Res* 159, 344-354. doi:10.1016/j.envres.2017.08.012
- Plasschaert, R. N. and Bartolomei, M. S. (2014). Genomic imprinting in development, growth, behavior and stem cells. *Development* 141, 1805-1813. doi:10.1242/dev.101428
- Plusquin, M., Saenen, N. D. and Nawrot, T. S. (2019). Epigenetics and the exposome. In S. Dagnino and A. Macherone (eds.), *Unraveling the Exposome: A Practical View* (127-146). Cham, Switzerland: Springer Nature. doi:10.1007/978-3-319-89321-1
- Polańska, K., Hanke, W., Jurewicz, J. et al. (2011). Polish mother and child cohort study (REPRO-PL) – Methodology of follow-up of the children. *Int J Occup Med Environ Health* 24, 391-398. doi:10.2478/v10001-009-0037-0
- Puntmann, V. O. (2009). How-to guide on biomarkers: Biomarker definitions, validation and applications with examples from cardiovascular disease. *Postgrad Med J* 85, 538-545. doi:10.1136/pgmj.2008.073759
- Pyne, S., Rao, B. and Rao, S. B. (eds.) (2016). *Big Data Analytics: Methods and Applications*. Heidelberg, Germany: Springer. doi:10.1007/978-81-322-3628-3
- Rakyan, V. K., Down, T. A., Balding, D. J. et al. (2011). Epigenome-wide association studies for common human diseases. *Nat Rev Genet* 12, 529-541. doi:10.1038/nrg3000
- Ramirez, T., Daneshian, M., Kamp, H. et al. (2013). Metabolomics in toxicology and preclinical research. *ALTEX* 30, 209-225. doi:10.14573/altex.2013.2.209
- Ramirez, T., Strigun, A., Verlohner, A. et al. (2018). Prediction of liver toxicity and mode of action using metabolomics in vitro. *Arch Toxicol* 92, 893-906. doi:10.1007/s00204-017-2079-6
- Rappaport, S. M. and Smith, M. T. (2010). Epidemiology. Environment and disease risks. *Science* 330, 460-461. doi:10.1126/science.1192603
- Rappaport, S. M. (2012). Biomarkers intersect with the exposome. *Biomarkers* 17, 483-489. doi:10.3109/1354750X.2012.691553
- Rappaport, S. M., Barupal, D. K., Wishart, D. et al. (2014). The blood exposome and its role in discovering causes of disease. *Environ Health Perspect* 122, 769-774. doi:10.1289/ehp.1308015
- Rappaport, S. M. (2016). Genetic factors are not the major causes of chronic diseases. *PLoS One* 11, e0154387. doi:10.1371/journal.pone.0154387
- Rappaport, S. M. (2018). Redefining environmental exposure for disease etiology. *NPJ Syst Biol Appl* 4, 30. doi:10.1038/s41540-018-0065-0
- Reid, N. (2018). Statistical science in the world of big data. *Statist Probabil Lett* 136, 42-45. doi:10.1016/j.spl.2018.02.049
- Rowbotham, A. L. and Gibson, R. M. (2011). Exposure-driven risk assessment: Applying exposure-based waiving of toxicity tests under REACH. *Food Chem Toxicol* 49, 1661-1673. doi:10.1016/j.fct.2011.03.050
- Saidulu, D. and Sasikala, R. (2017). Machine learning and statistical approaches for big data: Issues, challenges and research directions. *Int J Appl Eng Res* 12, 11691-11699.
- Sarigiannis, D. A., Karakitsios, S. P., Gotti, A. et al. (2009). Bayesian algorithm implementation in a real time exposure assessment model on benzene with calculation of associated cancer risks. *Sensors* 9, 731-755. doi:10.3390/s90200731
- Sarigiannis, D. A., Karakitsios, S. P., Handakas, E. et al. (2016). Integrated exposure and risk characterization of bisphenol-A in Europe. *Food Chem Toxicol* 98, 134-147. doi:10.1016/j.fct.2016.10.017
- Sarigiannis, D. A. and Karakitsios, S. P. (2018). Addressing complexity of health impact assessment in industrially contaminated sites via the exposome paradigm. *Epidemiol Prev* 42, 37-48. doi:10.19191/EP18.5-6.S1.P037.086
- Sarigiannis, D., Papaioannou, N., Kapretsos, N. et al. (2018a). Multi-omics analysis reveals that co-exposure to phthalates and metals disturbs urea cycle and choline metabolism. *Toxicol Lett* 295, S136. doi:10.1016/j.toxlet.2018.06.710
- Sarigiannis, D. A., Karakitsios, S. P., Handakas, E. et al. (2018b). Informatics and data analytics to support exposome-based discovery: Part 1 – Assessment of external and internal exposure. In M. D. Lytras and P. Papadopoulou (eds.), *Applying Big Data Analytics in Bioinformatics and Medicine* (115-144). Hershey, PA: IGI Global. doi:10.4018/978-1-5225-2607-0

- Sarigiannis D. A., Gotti, A., Handakas, E. et al. (2018c). Informatics and data analytics to support exposome-based discovery Part 2: Computational exposure biology. In M. D. Lytras and P. Papadopoulou (eds.), *Applying Big Data Analytics in Bioinformatics and Medicine*. Hershey, PA: IGI Global. doi:10.4018/978-1-5225-2607-0
- Sarigiannis, D. A. (2019a). The exposome: A new tool for improved health risk assessment. In B. Roig, K. Weiss and V. Thireau (eds.), *Management of Emerging Public Health Issues and Risks*, xxiii-xiv. Cambridge, MA: Academic Press. doi:10.1016/B978-0-12-813290-6.02002-3
- Sarigiannis, D. A. (2019b). The exposome paradigm in environmental health. In P. Papadopoulou, C. Marouli and A. Misseyanni (eds.), *Environmental Exposures and Human Health Challenges* (1-24). Hershey, PA: IGI Global. doi:10.4018/978-1-5225-7635-8
- Sauer, J. M., Hartung, T., Leist, M. et al. (2015). Systems toxicology: The future of risk assessment. *Int J Toxicol* 34, 346-348. doi:10.1177/1091581815576551
- Schaevitz, L., Berger-Sweeney, J. and Ricceri, L. (2014). One-carbon metabolism in neurodevelopmental disorders: Using broad-based nutraceuticals to treat cognitive deficits in complex spectrum disorders. *Neurosci Biobehav Rev* 46, 270-284. doi:10.1016/j.neubiorev.2014.04.007
- Schmutz, J., Wheeler, J., Grimwood, J. et al. (2004). Quality assessment of the human genome sequence. *Nature* 429, 365-368. doi:10.1038/nature02390
- Sewell, F., Aggarwal, M., Bachler, G. et al. (2017). The current status of exposure-driven approaches for chemical safety assessment: A cross-sector perspective. *Toxicology* 389, 109-117. doi:10.1016/j.tox.2017.07.018
- Smirnova, L., Sittka, A. and Luch, A. (2012). On the role of low-dose effects and epigenetics in toxicology. *Exp Suppl* 101, 499-550. doi:10.1007/978-3-7643-8340-4_18
- Smirnova, L., Harris, G., Leist, M. et al. (2015). Cellular resilience. *ALTEX* 32, 247-260. doi:10.14573/altex.1509271
- Smirnova, L. and Maertens, A. (2019). miRNA as a marker for in vitro neurotoxicity testing and related neurological disorders. In M. Aschner and L. Costa (eds.), *Cell Culture Techniques. Neuromethods* (Chapter 14, 255-281). 2nd edition. Vol. 145. New York, NY: Humana Press, Springer Protocols. doi:10.1007/978-1-4939-9228-7_14
- Smith, M. T., de la Rosa, R. and Daniels, S. I. (2015). Using exposomics to assess cumulative risks and promote health. *Environ Mol Mutagen* 56, 715-723. doi:10.1002/em.21985
- Smith, Z. D., Chan, M. M., Mikkelsen, T. S. et al. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* 484, 339-344. doi:10.1038/nature10960
- Steckling, N., Gotti, A., Bose-O'Reilly, S. et al. (2018). Biomarkers of exposure in environment-wide association studies – Opportunities to decode the exposome using human biomonitoring data. *Environ Res* 164, 597-624. doi:10.1016/j.envres.2018.02.041
- Sturla, S. J., Boobis, A. R., FitzGerald, R. E. et al. (2014). Systems toxicology: From basic research to risk assessment. *Chem Res Toxicol* 27, 314-329. doi:10.1021/tx400410s
- Thomas, R., Clewell, H. J., Allen I. et al. (2011). Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. *Toxicol Sci* 120, 194-205. doi:10.1093/toxsci/kfq355
- Teeguarden, J. G., Tan, C., Edwards, S. et al. (2016). Completing the link between exposure science and toxicology for improved environmental health decision making: The aggregate exposure pathway framework. *Environ Sci Technol* 50, 4579-4586. doi:10.1021/acs.est.5b05311
- Torres-Padilla, M.-E., Parfitt, D.-E., Kouzarides, T. et al. (2007). Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* 445, 214-218. doi:10.1038/nature05458
- Trdin, A., Snoj Tratnik, J., Mazej, D. et al. (2019). Mercury speciation in prenatal exposure in Slovenian and Croatian population – PHIME study. *Environ Res* 177, 108627. doi:10.1016/j.envres.2019.108627
- Tsaioun, K., Blaauboer, B. J. and Hartung, T. (2016). Evidence-based absorption, distribution, metabolism, excretion and toxicity (ADMET) and the role of alternative methods. *ALTEX* 33, 343-358. doi:10.14573/altex.1610101
- Turner, M. C., Vineis, P., Seleiro, E. et al. (2018). EXPOsOMICS: Final policy workshop and stakeholder consultation. *BMC Public Health* 18, 260. doi:10.1186/s12889-018-5160-z
- van der Heijden, G. W., Ramos, L., Baart, E. B. et al. (2008). Sperm-derived histones contribute to zygotic chromatin in humans. *BMC Dev Biol* 8, 34. doi:10.1186/1471-213X-8-34
- Vineis, P., Khan, A. E., Vlaanderen, J. et al. (2009). The impact of new research technologies on our understanding of environmental causes of disease: The concept of clinical vulnerability. *Environ Health* 30, 54. doi:10.1186/1476-069X-8-54
- Visscher, P. M., Brown, M. A., McCarthy, M. I. et al. (2012). Five years of GWAS discovery. *Am J Hum Genet* 90, 7-24. doi:10.1016/j.ajhg.2011.11.029
- Vrijens, K., Bollati, V. and Nawrot, T. S. (2015). MicroRNAs as potential signatures of environmental exposure or effect: A systematic review. *Environ Health Perspect* 123, 399-411. doi:10.1289/ehp.1408459
- Vrijheid, M. (2014). The exposome: A new paradigm to study the impact of environment on health. *Thorax* 69, 876-878. doi:10.1136/thoraxjnl-2013-204949
- Vrijheid, M., Slama, R., Robinson, O. et al. (2014). The human early-life exposome (HELIX): Project rationale and design. *Environ Health Perspect* 122, 535-544. doi:10.1289/ehp.1307204
- Wacholder, S., Chanock, S., Garcia-Closas, M. et al. (2004). Assessing the probability that a positive report is false: An approach for molecular epidemiology studies. *J Natl Cancer Inst* 96, 434-442. doi:10.1093/jnci/djh075
- Wang, C., Chen, M.-H., Schifano, E. et al. (2016). Statistical methods and computing for big data. *Stat Interface* 9, 399-414. doi:10.4310/SII.2016.v9.n4.a1
- Waterland, R. A. and Michels, K. B. (2007). Epigenetic epidemiology of the developmental origins hypothesis. *Annu Rev Nutr*



- 27, 363-388. doi:10.1146/annurev.nutr.27.061406.093705
- Welter, D., MacArthur, J., Morales, J. et al. (2014). The NHGRI GWAS catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 42, D1001-D1006. doi:10.1093/nar/gkt1229
- Wild, C. P. (2005). Complementing the genome with an “exposome”: The outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev* 14, 1847-1850. doi:10.1158/1055-9965.EPI-05-0456
- Wild, C. P. (2009). Environmental exposure measurement in cancer epidemiology. *Mutagenesis* 24, 117-125. doi:10.1093/mutage/gen061
- Willett, W. C. (2002). Balancing life-style and genomics research for disease prevention. *Science* 296, 695-698. doi:10.1126/science.1071055
- Zhang, P., Arora, M., Chaleckis, R. et al. (2019). Tackling the complexity of the exposome: Considerations from the Gunma University initiative for advanced research (GIAR) exposome symposium. *Metabolites* 9, E106. doi:10.3390/metabo9060106
- Ziller, M. J., Müller, F., Liao, J. et al. (2011). Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet* 7, e1002389. doi:10.1371/journal.pgen.1002389

Conflict of interest

TH consults Underwriters Laboratories (UL) on computational toxicology and receives shares from their respective sales. He also holds stock options and consults ToxTrack.

Acknowledgements

Stephen Rappaport was supported by grant P42ES004705 (NIEHS). The mentioned work on computational toxicology was mainly executed by Thomas Luechtefeld during his PhD and within ToxTrack. Thomas Luechtefeld was supported by an NIEHS training grant (T32 ES007141). This work was supported by the EU-ToxRisk project (An Integrated European “Flagship” Program Driving Mechanism-Based Toxicity Testing and Risk Assessment for the 21st Century) funded by the European Commission under the Horizon 2020 program (Grant Agreement No. 681002). This work has also received funding from the European Union’s Seventh Programme for research, technological development and demonstration under grant agreement No 603946 (Health and Environment-wide Associations via Large population Surveys-HEALS). The fruitful discussions with further members of the Johns Hopkins Exposome Collaborative, namely Drs Hugh Ellis, Keeve Nachman, Scot Miller and Meredith McCormack, is gratefully appreciated. The most valuable input of many colleagues inside and outside our team is gratefully appreciated.

Glossary

(Dennis et al., 2017; DeBord et al., 2016; and other sources)

Aggregate exposure pathways (AEPs) – The AEP concept has been developed complementary to the AOP describing the key events from source via external exposure (including environmental and dietary exposure) to exposure in the organism, termed target site exposure.

Biomarker – A characteristic that can be objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention

Biomonitoring – The measurement of the body burden of toxic chemical compounds, elements, or their metabolites, in biological substances

Children’s Health Exposure Analysis Resource (CHEAR) – NIEHS project to advance understanding about how the environment impacts children’s health and development (2015-)

Eco-exposome – The extension of exposure science from the point of contact between a stressor and receptor inward into the organism and outward to the general environment, including the ecosphere

Epigenetics – The study of changes in organisms caused by modification of gene expression rather than alteration of the genetic code itself

Epigenetic scar – Long-term changes in epigenetics induced by stressors, which affect regulation of gene expression

Exposome – Concept describing the totality of exposure experienced by an individual during their life and the health impact of those exposures (Wild, 2005), redefined (Miller and Jones, 2014): The cumulative measure of environmental influences and associated biological responses throughout the lifespan, including exposures from the environment, diet, behavior, and endogenous processes.

EXPOsOMICS – European-funded project under the FP7 Exposome Program that aims to develop a new approach to assessing environmental exposures in adults, particularly through the use of omics techniques

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

Feature – Raw data output from mass spectrometry analysis, which includes an accurate mass-to-charge ratio with associated retention time (RT) and ion intensity. A feature can represent one or more chemicals/metabolites, so data extraction methods are critical to interpretation

Geographic Information Systems (GIS) – Systems that manage geographic data



HELIX (The Human Early-Life Exposome) – European-funded project under the FP7 Exposome Program focused on understanding the early-life exposome through novel exposure measurement and data-driven methods

HERCULES (Health and Exposome Research Center: Understanding Lifetime Exposures) – US NIEHS – funded center at Emory University focused on providing infrastructure and expertise to develop and refine new tools and technologies to advance exposome research as well as on promoting environmental health sciences research overall

HEALS (Health and Environment-wide Associations based on Large population Surveys) – A European Commission-funded project focused on integrating omics data and traditional biomonitoring measurements with alterations in outcomes such as gene expression and metabolic regulation to assess environmental exposures and human health associations

HHEAR (Human Health Exposure Analysis Resource) – US NIEHS program that provides a larger community of researchers access to centralized, high-quality exposure assessment services

High-resolution metabolomics – Mass spectrometry technique that can detect > 10,000 features through instrumentation such as time-of-flight, Fourier transform ion cyclotron resonance, and orbitrap mass analyzers

International Exposome Center, the I3 Care Centre – International Exposome Center between the Netherlands, Hong Kong and Canada aiming at expanding the exposome concept through the development of new investigative tools to discover and quantify modifiable risk factors

Japan Environment and Children's Study (JECS) – Study to identify harmful factors in the environment affecting children's growth and health, and to investigate the relationship between such factors and children's health condition involving 100,000 parent-child pairs followed from birth to 13 years of age (2011-2032)

Multiplexing – Fractionation of samples to remove higher-level chemicals, enabling detection of lower-abundance chemicals

Metabolomics – The characterization, identification, and quantitation of metabolites resulting from a wide range

of biochemical processes in living systems, typically by MS and NMR technologies.

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

Reality mining – Analysis of behavioral and self-reported data extracted from social networks and other portable devices' applications

Reverse causality – A situation in which the outcome precedes and causes the exposure instead of the other way around. This is of particular concern in retrospective and cross-sectional studies.

Semi-targeted/hybrid approaches – Exploits the advantages of both targeted and untargeted approaches: for example, using metabolomics for discovery of potential exposures followed by targeted analysis for a more fully quantitative measure

Systems toxicology – Systems biology applied to toxicology, i.e., the integration of classical toxicology with quantitative analysis of large networks of molecular and functional changes occurring across multiple levels of biological organization

Testing multiple variables for associations – Statistical inference problem especially with the use of omics methodologies, which involve the simultaneous testing 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.

Traditional biomonitoring/targeted analyses – Analyses of biological samples for specific chemicals: either exposures or markers of exposures

Untargeted analyses – Agnostic analyses that can measure a broad set of endogenous and exogenous metabolites in one sample run