



The exposome concept: how has it changed our understanding of environmental causes of chronic respiratory diseases?

Alicia Guillien ^{1,7}, Manosij Ghosh ^{2,7}, Thomas Gille ^{3,4,5} and Orianne Dumas ⁶

¹University of Grenoble Alpes, French National Institute of Health and Medical Research, French National Center for Scientific Research, Team of Environmental Epidemiology Applied to Reproduction and Respiratory Health, Institute for Advanced Biosciences, Grenoble, France. ²Environment and Health, Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium. ³Inserm UMR 1272 “Hypoxia & the Lung”, UFR Santé, Médecine, Biologie Humaine (SMBH) Léonard de Vinci, Université Sorbonne Paris Nord (USPN), Bobigny, France. ⁴Dept. of Physiology and Functional Explorations, Avicenne University Hospital, Hôpitaux Universitaires de Paris Seine-Saint-Denis (HUPSSD), Assistance Publique - Hôpitaux de Paris (AP-HP), Bobigny, France. ⁵Dept. of Physiology and Functional Explorations, Jean Verdier University Hospital, Hôpitaux Universitaires de Paris Seine-Saint-Denis (HUPSSD), Assistance Publique - Hôpitaux de Paris (AP-HP), Bondy, France. ⁶Université Paris-Saclay, UVSQ, Univ. Paris-Sud, Inserm, Equipe d’Epidémiologie Respiratoire Intégrative, CESP, Villejuif, France. ⁷Both authors contributed equally.

Corresponding author: Alicia Guillien (alicia.guillien@univ-grenoble-alpes.fr)



Shareable abstract (@ERSpublications)

The exposome approach can help us better understand multifactorial respiratory diseases through multidisciplinary collaboration, harmonised resources and use of sophisticated methods addressing combined exposures and longitudinal data. <https://bit.ly/3Ng9MNn>

Cite this article as: Guillien A, Ghosh M, Gille T, *et al.* The exposome concept: how has it changed our understanding of environmental causes of chronic respiratory diseases? *Breathe* 2023; 19: 230044 [DOI: 10.1183/20734735.0044-2023].

Copyright ©ERS 2023

Breathe articles are open access and distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0.

Received: 13 Feb 2023
Accepted: 31 May 2023

Introduction

In the late 20th century, the prevalence and incidence of chronic respiratory diseases increased considerably worldwide [1]. As an example, the prevalence of asthma has doubled (or even tripled) in developed countries in <50 years. It is well known that genetics is involved in the development of chronic respiratory diseases. Nevertheless, this large increase in a short period emphasises the role of environmental factors, which may act alone or together with genetic factors. Moreover, such an increase cannot be explained by a single environmental factor, but is probably due to a large number of environmental factors interacting with each other.

In this context, which is common to many chronic diseases, in 2005 Christopher Wild defined the concept of the exposome to encompass “life-course environmental exposures (including lifestyle factors), from the prenatal period onwards” [2].

The exposome concept: definitions and potential benefits

In his paper, WILD [2] expressed his concern over the lack of measures of environmental exposures that could help identify causes of human disease. He highlighted the need to characterise an individual’s environmental exposure “from conception to death”, complementing information obtained from genome. In a later paper, WILD [3] further described the following three distinct and complementary domains of exposome research (figure 1).

- 1) Internal: internal biological processes such as oxidative stress, inflammation, epigenetic changes, metabolism and the internal microbiome.
- 2) General external: social, economic factors, the urban environment and climate factors.
- 3) Specific external: an individual’s immediate local environment, including exposure to chemicals, diet, physical activity, tobacco and infections.

Building on the definition proposed by Wild, MILLER and JONES [4] expanded the definition: “The cumulative measure of environmental influences and associated biological responses throughout the lifespan, including exposures from the environment, diet, behavior, and endogenous processes”, by including a quantifiable “cumulative measure” of the exposome component and the “differential response” in biological processes.



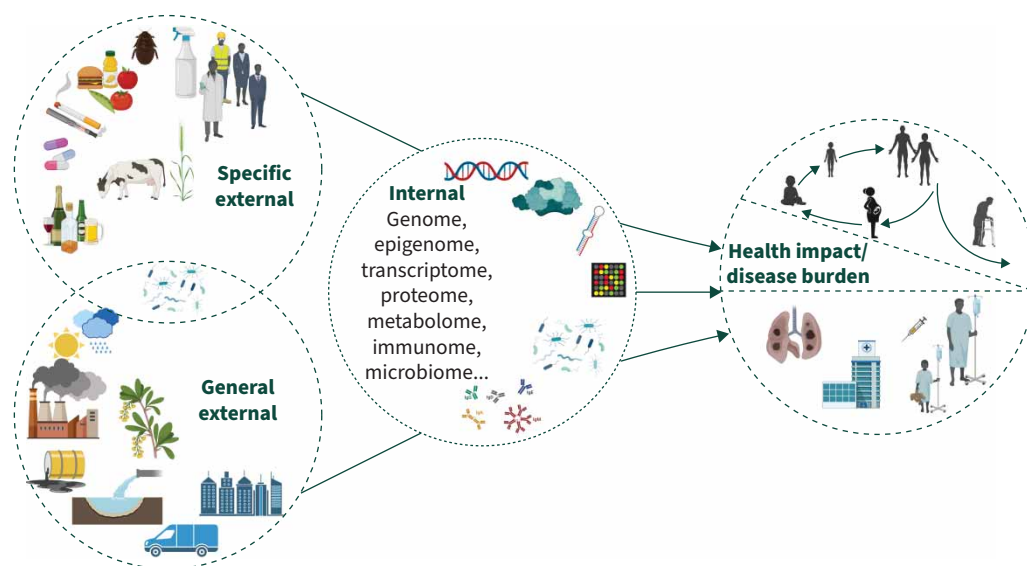


FIGURE 1 Description of the exposome.

Until recently, epidemiological studies focused on one single exposure or family of exposures. This first step, along with biological studies, identified the major risk/protective factors for chronic respiratory diseases, such as tobacco smoking, a history of viral infections, occupational exposures or living on a farm. However, this approach was limited by the issue of publication bias and also by the difficulty in considering “mixture effects”. Wild’s paper highlighted the need to collect a wide range of environmental exposures that must be jointly considered to better match reality, where humans experience many exposures simultaneously on a daily basis.

Given that exposome research is in its nascent stage, it requires multidisciplinary collaboration (exposure sciences, toxicology, biology, epidemiology, statistics, *etc.*) to define the scope and design of such studies, overcome analytical challenges and develop computational tools to meaningfully associate the exposome to health outcomes. The following sections present recent developments in the field and the remaining challenges of exposome studies.

How to assess the exposome: attempts and challenges

The exposome concept hinges on integration of information from multiple disciplines and in the assessment of the totality of all exposures, including environmental, chemical and lifestyle factors.

Several large scale exposome initiatives have been launched in recent years, including the EXPOsOMICS project (<https://exposomics-project.eu/live-exposome.pantheonsite.io/index.html>), the HELIX project (www.projecthelix.eu), the European Human Exposome Network (EHEN, a cluster of nine exposome projects; www.humanexposome.eu), and the Human Exposome Project (<https://humanexposomeproject.com>). Over the past decade, there have been considerable improvements in biochemical/analytical techniques such as using high-resolution mass spectrometers. Advances have been made for characterisation of the internal exposome (*e.g.* epigenetic changes, RNA expression, proteins, metabolites, microbiome) in a wide range of biological matrices, using both targeted and untargeted/agnostic approaches and high-throughput platforms. In addition to biological markers, exposome projects also rely on the development and refinement of wearable sensors and other passive sampling devices, and other smart technologies for high-resolution assessment of an individual’s exposome.

However, key challenges remain in setting up exposome studies. One of the most obvious difficulties is the assessment of the “totality” of exposure over an individual’s life course. Given that an exposure does not have the same effect at various different stages of life, with some periods especially critical to health and disease, measurement must consider the temporal variability. Towards reconstruction of the exposome in its entirety, and in the absence of life-course data, several projects have used different workable approaches. The ATHLETE project [5] focuses on the effects of a wide range of general and specific external exposome determinants on mental, cardiometabolic, and respiratory health outcomes in the first

two decades of life. The EXIMIOUS project [6] relies on a “meet in the middle” approach, combining several study populations covering the entire lifespan, including both general population and birth cohorts, as well as disease cohorts. The EPHOR project [7] relies on large-scale pooling and harmonisation of existing European cohort data into a “mega cohort”, systematically looking at multiple exposures and diseases, including obstructive lung disease. Pooling of existing data is often supplemented with the collection of new high-resolution external and internal exposure data, and multi-omics data. Estimation of occupational exposures can be improved using harmonised job exposure matrices. In other studies, data and text mining approaches, adverse outcome pathway analysis and systems biology are integrated to understand of the impact of the exposome on health. Experimental approaches are also used in addition to cohorts, such as in REMEDIA study [8], where naïve COPD or cystic fibrosis preclinical rodent models are exposed in an atmospheric simulation chamber where the complexity of the real atmosphere can be mimicked over a relative short period of time.

While large scale data collection and pooling are expected, at the European (and global) level, another major obstacle in exposome research is implementation of the FAIR (findable, accessible, interoperable, reusable) guiding principles for data sharing and management, and related ethical and legal issues. In this context another EHEN project, the Human Exposome Assessment Platform (HEAP) [9], aims to design and implement an ethical governance structure that can serve as an example of exposome data management. Furthermore, a harmonised meta(data) catalogue is expected as an outcome of the EHEN projects and will be made available through the EHEN Molgenis data platform [10].

Linking the exposome to respiratory health: new statistical challenges

Once data have been collected and checked for quality, challenges arise related to the analysis of large amounts of data.

With the advent of exposome studies, gathering hundreds of exposure variables, the first association studies consisted of fitting one separate regression model for each exposure variable based on the methods employed in genome-wide association studies. In a simulation study comparing six regression-based methods to assess exposome–health associations [11], this method, known as an exposome-wide association study (ExWAS), reached the highest sensitivity to detect true predictors, but this came with the price of the highest rate of false discovery proportion. Nonetheless, this method has the advantage of giving an estimate and a p-value for each exposure, thus providing a list of exposures associated or not associated with the health outcome, along with an effect size and p-value.

To account for multiple exposures in a unique model, the ExWAS is often complemented by a multivariable model, including a variable selection step such as the deletion/substitution/addition algorithm or the LASSO method [12]. For chemical exposures, some methods are also able to deal with their “mixture” effects, usually known as “multipollutant models”. For example, weighted quantile sum (WQS) regression is based on the assumption that a low level of exposure to a single exposure factor could have no or a non-detectable effect, but a combined exposure to multiple factors at low doses could have an effect. WQS regression estimates an exposure index through a weighted sum of each exposure categorised in quantiles and thereafter assesses the association of this index with a health parameter and can highlight the top contributors of the WQS index [13]. This method makes the strong assumption that all exposures have an effect on the outcome in the same direction, which does not always correspond to reality. Therefore, risk factors and protective factors have to be considered in two separate models. Another multipollutant model is the Bayesian kernel machine regression (BKMR), which aims to analyse several biomarkers jointly, looking for non-linear effects and interactions [14]. This method can also perform variable selection.

To get a step further by considering the complex structure of the exposome, where correlations are often stronger between two exposures from the same family rather than two exposures from different families [15], recent studies have performed clustering analyses to identify groups of individuals sharing the same exposure patterns and then assess their associations with health outcomes [16]. Three main classes of clustering analysis exist: hierarchical clustering (*e.g.* ascending hierarchical clustering), partitioning (*e.g.* k-means), and model-based clustering based on mixture models. These methods can be unsupervised (considering only exposures to identify clusters) or supervised (considering both the exposure and the outcome to identify clusters), and are strongly relevant in the context of multifactorial diseases where several exposures can interact and have a synergistic effect on health. Partitioning and model-based clustering rely on random sampling, which might lead to instability issues, especially with increased numbers of exposures. To improve the stability of the model, it is preferable to first proceed to a variable selection step. Hierarchical clustering produces stable results but is not suitable for mixed exposures (both continuous and categorical exposures). Finally, some statistical methods are able to deal with the causal structure existing within the exposome or

with intermediary data. The internal exposome and biological layers are considered as intermediary layers between the external exposome and health. The incorporation of these intermediary layers could help to reduce the false discovery proportion in exposome–health association studies [17], but this requires large amounts of data (both in terms of exposure and participants).

Great advances have been made to use methods that are more relevant to these large and complex datasets, and these efforts must be continued; in particular, by including more sophisticated methods addressing combined exposures and handling of longitudinal data.

The exposome and respiratory health: what have we learnt from pioneer studies?

Among exposome studies focusing on respiratory health and assessing a wide range of exposures belonging to at least two different exposure families, several have been based on cohorts of children. An early study, based on the Kingston Allergy Birth Cohort in Canada [18], reported that prenatal smoke exposure, mould or dampness in the home, and use of air fresheners were associated with increased respiratory symptoms, while breastfeeding, older siblings and increased gestational age were associated with a decreased risk of respiratory symptoms. Another study, based on the HELIX project involving six European cohorts, identified three prenatal exposures (perfluorononanoate, perfluorooctanoate and distance to nearest road) and nine postnatal exposures (copper, ethyl-paraben, five phthalate metabolites, house crowding and facility density around school) that were associated with decreased forced expiratory volume in 1 s % predicted [12]. In the same population, a lower distance between the residence and the nearest road, higher di-isononyl phthalate and a lower level of particulate matter absorbance were associated with increased risk of rhinitis [19]. All these studies applied the ExWAS method and the associations did not remain significant after correction of the p-value for multiple tests. Among studies based on adult populations, two French studies performed clustering analyses to identify profiles of exposure associated with respiratory health. In a population of adults with ever-asthma (EGEA cohort), a specific profile of combined lifestyle and environmental factors, characterised by heavy smoking, poor diet, higher outdoor humidity and proximity to traffic, was associated with reduced lung function, while none of these factors showed an association individually in the ExWAS [20]. Finally, in a large web-based cohort (the NutriNet-Santé cohort) three profiles of early-life combined exposure (“high passive smoking–own dogs”, “poor birth parameters–day-care attendance–city centre”, or “>2 siblings–breastfed”) and one profile of combined lifestyle factors (“unhealthy diet–high smoking–overweight”) were associated with greater asthma symptoms and poorer asthma control than reference profiles (“farm–pet owner–moulds–low passive smoking” for early-life exposures and “healthy diet–nonsmoker–thin” for lifestyle factors) [16].

The literature on exposome–respiratory health associations remains scarce and existing studies highlight the need to use larger sample sizes in order to improve statistical performance of exposome studies, but also to apply comprehensive approaches jointly considering a wide range of exposures and including multiple periods of exposure.

Conclusion

The papers by Christopher Wild defined the concept of the exposome and emphasised the importance of considering environmental factors (in a broad sense, including all non-genetic factors) to complement the effect of the genome in understanding of the aetiology of diseases. In this context, several large-scale projects have been launched in recent years, resulting in an increasing number of epidemiological studies focusing on the effect of the exposome on health, including respiratory health. The increased use of the exposome approach is a step forward for a better understanding of the development of complex diseases, through the integration of biological processes over a lifetime. This relies on multidisciplinary expertise and collaboration, harmonisation and accessible resources for implementation. To help with these aspects, an accessible toolbox containing data models, guidelines and protocols has been designed (the EHEN toolbox [21]). In terms of public health, these methods should help develop preventive strategies based on multifactorial rather than individual factors, for example, by improving land-use in urban areas or the promotion of healthy lifestyles.

Conflict of interest: T. Gille reports personal fees from ROCHE S.A.S., other from OXYVIE (oxygen provider), other from VIVISOL France (oxygen provider), other from MENANIRI France, outside the submitted work. The remaining authors have nothing to disclose.

Support statement: M. Ghosh is supported by funding from EXIMIOUS (grant agreement No 874707) and EPHOR (grant agreement No 874703) projects funded by the European Union’s Horizon 2020 research and innovation programme. A. Guillien is funded by the ATHLETE project that has received funding from the European Union’s

Horizon 2020 research and innovation programme under grant agreement No 874583 and by the EPIMEX project, funded by Fondation pour la Recherche Médicale (FRM), grant number ENV202004011870. This publication reflects only the authors' view and the European Commission is not responsible for any use that may be made of the information it contains.

References

- 1 Eder W, Ege MJ, von Mutius E. The asthma epidemic. *N Engl J Med* 2006; 355: 2226–2235.
- 2 Wild CP. Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev* 2005; 14: 1847–1850.
- 3 Wild CP. The exposome: from concept to utility. *Int J Epidemiol* 2012; 41: 24–32.
- 4 Miller GW, Jones DP. The nature of nurture: refining the definition of the exposome. *Toxicol Sci* 2014; 137: 1–2.
- 5 Vrijheid M, Basagaña X, Gonzalez JR, et al. Advancing tools for human early lifecourse exposome research and translation (ATHLETE): project overview. *Environ Epidemiol* 2021; 5: e166.
- 6 Ronsmans S, Hougaard KS, Nawrot TS, et al. The EXIMIOUS project—Mapping exposure-induced immune effects: connecting the exposome and the immunome. *Environ Epidemiol* 2022; 6: e193.
- 7 Pronk A, Loh M, Kuijpers E, et al. Applying the exposome concept to working life health: the EU EPHOR project. *Environ Epidemiol* 2022; 6: e185.
- 8 Benjdir M, Audureau É, Beresniak A, et al. Assessing the impact of exposome on the course of chronic obstructive pulmonary disease and cystic fibrosis: the REMEDIA European project approach. *Environ Epidemiol* 2021; 5: e165.
- 9 Merino Martinez R, Müller H, Negru S, et al. Human exposome assessment platform. *Environ Epidemiol* 2021; 5: e182.
- 10 MOLGENIS. European Networks Health Data and Cohort Catalogue. <https://data-catalogue.molgenisccloud.org/catalogue/catalogue/#/networks-catalogue>
- 11 Agier L, Portengen L, Chadeau-Hyam M, et al. A systematic comparison of linear regression-based statistical methods to assess exposome–health associations. *Environ Health Perspect* 2016; 124: 1848–1856.
- 12 Agier L, Basagaña X, Maitre L, et al. Early-life exposome and lung function in children in Europe: an analysis of data from the longitudinal, population-based HELIX cohort. *Lancet Planet Health* 2019; 3: e81–e92.
- 13 Wu Y, Song J, Li Y, et al. Association between exposure to a mixture of metals, parabens, and phthalates and fractional exhaled nitric oxide: a population-based study in US adults. *Environ Res* 2022; 214: 113962.
- 14 Benka-Coker W, Hoskovec L, Severson R, et al. The joint effect of ambient air pollution and agricultural pesticide exposures on lung function among children with asthma. *Environ Res* 2020; 190: 109903.
- 15 Tamayo-Uria I, Maitre L, Thomsen C, et al. The early-life exposome: description and patterns in six European countries. *Environ Int* 2019; 123: 189–200.
- 16 Guillien A, Bédard A, Dumas O, et al. Exposome profiles and asthma among French adults. *Am J Respir Crit Care Med* 2022; 206: 1208–1219.
- 17 Cadiou S, Basagaña X, Gonzalez JR, et al. Performance of approaches relying on multidimensional intermediary data to decipher causal relationships between the exposome and health: a simulation study under various causal structures. *Environ Int* 2021; 153: 106509.
- 18 North ML, Brook JR, Lee EY, et al. The Kingston Allergy Birth Cohort: exploring parentally reported respiratory outcomes through the lens of the exposome. *Ann Allergy Asthma Immunol* 2017; 118: 465–473.
- 19 Granum B, Oftedal B, Agier L, et al. Multiple environmental exposures in early-life and allergy-related outcomes in childhood. *Environ Int* 2020; 144: 106038.
- 20 Guillien A, Lepeule J, Seyve E, et al. Profile of exposures and lung function in adults with asthma: an exposome approach in the EGEA study. *Environ Res* 2020; 196: 110422.
- 21 European Human Exposome Network (EHEN). Exposome toolbox. www.humanexposome.eu/toolbox/