



CrossMark

# The exposome concept: a challenge and a potential driver for environmental health research

Valérie Siroux<sup>1,2,3</sup>, Lydiane Agier<sup>1,2,3</sup> and Rémy Slama<sup>1,2,3</sup>

**Affiliations:** <sup>1</sup>INSERM, IAB, Team of Environmental Epidemiology applied to Reproduction and Respiratory Health, Grenoble, France. <sup>2</sup>Université Grenoble Alpes, IAB, Team of Environmental Epidemiology applied to Reproduction and Respiratory Health, Grenoble, France. <sup>3</sup>CHU de Grenoble, IAB, Team of Environmental Epidemiology applied to Reproduction and Respiratory Health, Grenoble, France.

**Correspondence:** Valérie Siroux, Institut Albert Boniot, Inserm/Université Grenoble Alpes U1209, Equipe d'épidémiologie environnementale appliquée à la reproduction et à la santé respiratoire, Rond point de la chantourne, 38706 La Tronche, France. E-mail: valerie.siroux@univ-grenoble-alpes.fr

**ABSTRACT** The exposome concept was defined in 2005 as encompassing all environmental exposures from conception onwards, as a new strategy to evidence environmental disease risk factors. Although very appealing, the exposome concept is challenging in many respects. In terms of assessment, several hundreds of time-varying exposures need to be considered, but increasing the number of exposures assessed should not be done at the cost of increased exposure misclassification. Accurately assessing the exposome currently requires numerous measurements, which rely on different technologies; resulting in an expensive set of protocols. In the future, high-throughput 'omics technologies may be a promising technique to integrate a wide range of exposures from a small numbers of biological matrices. Assessing the association between many exposures and health raises statistical challenges. Due to the correlation structure of the exposome, existing statistical methods cannot fully and efficiently untangle the exposures truly affecting the health outcome from correlated exposures. Other statistical challenges relate to accounting for exposure misclassification or identifying synergistic effects between exposures. On-going exposome projects are trying to overcome technical and statistical challenges. From a public health perspective, a better understanding of the environmental risk factors should open the way to improved prevention strategies.



@ERSpublications

The exposome concept presents challenges in terms of measurement and assessment of exposome–health relationships <http://ow.ly/4nuw4Q>

## Introduction

The exposome concept was initiated within the field of epidemiology, and the term was coined by WILD [1] in 2005 to encompass “the totality of human environmental exposures from conception onwards, complementing the genome” (figure 1) [1–3]. Therefore, the exposome concept calls for providing a description of lifelong (from the prenatal period) exposure history. It was developed to highlight the need for more comprehensive environmental exposure data. In 2012, WILD [2] defined three overlapping domains within the exposome that refer to different factors to be measured: 1) a general external environment including factors such as the urban–rural environment, climate factors, social capital or education; 2) a

---

Editorial comments in *Eur Respir Rev* 2016; 25: 104–107.

Received: April 14 2016 | Accepted after revision: April 30 2016

Support statement: This work was supported by the European Community Seventh Framework Programme FP7/2007–2013 (grants 308333; HELIX project). Funding information for this article has been deposited with Open Funder Registry.

Conflict of interest: Disclosures can be found alongside the online version of this article at [err.ersjournals.com](http://err.ersjournals.com)

Provenance: Submitted article, peer reviewed.

Copyright ©ERS 2016. ERR articles are open access and distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0.

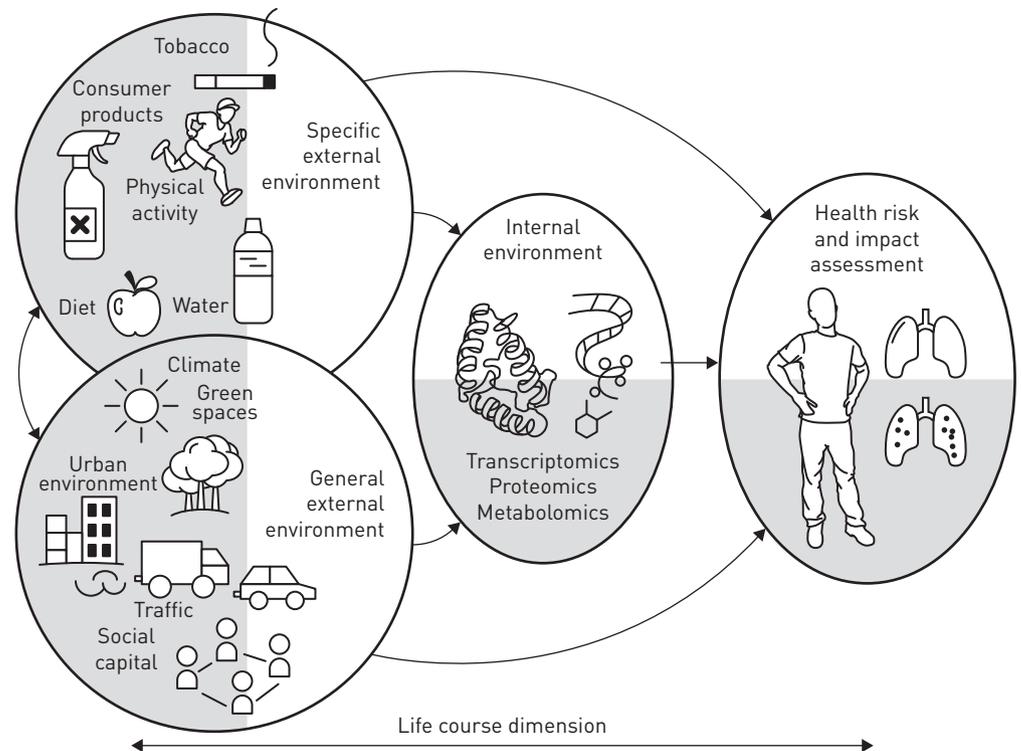


FIGURE 1 The effects and interactions between the different domains constituting the exposome (specific and general external environments and internal environment [2]) and health risk. Reproduced from [3] with permission.

specific external environment including diet, physical activity, tobacco, infection, occupation, *etc.*; and 3) an internal environment to include internal biological factors, such as metabolic factors, gut microflora, inflammation, oxidative stress or ageing. Therefore, this internal environment includes the biological/toxicological manifestations of exposures in the body. More recently, MILLER and JONES [4] proposed an alternative definition of the exposome, which explicitly incorporates behavioural risk factors, the body's response to environmental influences, and the endogenous metabolic processes that can alter or process the chemicals to which humans are exposed.

The number of citations of "exposome" in PubMed has increased over recent years; from six in 2005 to 150 in 2010 and >1600 in 2015. The justification for the development of this concept is at least two-fold. First, assessing many exposures simultaneously is expected to provide a more accurate assessment of the impact of the environment on human health; secondly, now that genetic polymorphisms can be accurately assessed at a relatively low cost, assessing environmental exposures with a similar degree of accuracy would enable efficient study of the interplay between the environment and the genome. A corresponding financial investment is needed to complement the huge investment in the genetic field in the early 2000s (and allow well-balanced analyses of gene–environment interaction effects on health).

### Motivation and promises

There is increasing evidence that genetic variants explain a limited fraction of the variability in the risk of chronic diseases, leaving a potentially large role for environmental exposures and interaction between environmental and genetic factors. In the context of allergic diseases, although genetic research provided insight into the pathways and mechanisms involved in disease occurrence, the identified genetic variants only explain a small proportion of the variability in disease risk [5]. The same applies for other chronic diseases such as type II diabetes or obesity [6]. However, a wide range of environmental and lifestyle risk factors have been associated with asthma and allergic diseases (*e.g.* viral infections, allergens, diet, urbanisation, farm animals, pets, smoking, air pollution, occupational exposure, *etc.*) showing the relevance of considering the exposome in the aetiology of these diseases [7]. Nevertheless, current knowledge on the environmental factors that contribute to allergic diseases remains limited. This might be due to the fact that the effect of the exposures considered to date is underestimated (*i.e.* due to measurement errors, inappropriate risk model, *etc.*), the main environmental risk factors remain unknown (*e.g.* due to our current limitation in investigating tools), or that exposures mostly impact health in synergy with other exposures or biological or behavioural factors.

Simultaneously considering a large set of exposures in human environmental health studies is a major evolution. In daily life populations are simultaneously exposed to a wide range of factors. For several of these factors, exposures are close to existing guidance values, as shown by biomonitoring data [8]. To date, most epidemiological studies focused on a single (or a limited number of) environmental factor(s) and, therefore, were unable to totally understand co-exposures and account for the effect of combined exposure to several environmental compounds. This might be an important limitation. Indeed, synergy between different environmental factors may exist, as illustrated by a study showing that an organochlorine pesticide and oestrogen can form a supramolecular ligand of the pregnane X receptor (a key regulator of the body's defence against foreign substances) and that their combination produces a cellular effect much higher than that predicted from the individual effects of each substance considered separately [9]. In addition, there may be confounding by co-exposures (*i.e.* the estimated effect of an exposure could be masked or overestimated because of a correlation of this exposure with other environmental factors), which can be tackled when simultaneously considering the exposures.

The exposome concept considers the lifelong exposure history, and therefore calls for changes in exposure over time to be taken into account. To date, most environmental health studies assessed the environmental exposure at a single time-point and, therefore, might not cover the highest risk exposure window.

Finally, a key potential benefit of considering the exposome concept is to limit the risk of selective reporting (*i.e.* testing several exposures and only reporting the most significant associations) and the number of false positives in the epidemiological literature. Indeed, in many epidemiological studies a large number of exposures have been assessed, but their association with health is usually reported on an exposure-by-exposure basis without correcting for multiple testing or showing the number of statistical tests performed [10].

In addition to providing insight into the relationship between environment and health, a more descriptive benefit of the exposome concept would be to provide insights on the correlation between environmental exposures, and also on the existence of subgroups of the population carrying a large burden in terms of environmental exposures. This issue pertains to the concept of environmental justice [11].

### Large projects related to the exposome

Several exposome-related research initiatives have been launched, in particular in Europe. The HELIX project ([www.projecthelix.eu/fr](http://www.projecthelix.eu/fr)) focuses on the early-life exposome and aims to combine all environmental hazards that mothers and children are exposed to and link this to the health, growth and development of children (including respiratory health) [12]. The EXPOSOMICS project ([www.exposomicsproject.eu/](http://www.exposomicsproject.eu/)) aims to develop a new approach to assess environmental exposures, with a specific focus on air pollution and water contaminants, and to link these exposures to biochemical and molecular changes in the body. The Heals project proposes “the functional integration of ‘omics derived data and biochemical biomonitoring to create the internal exposome at the individual level” ([www.heals-eu.eu/](http://www.heals-eu.eu/)). In the USA, the Hercules project focuses on providing key infrastructure and expertise to develop and refine new tools and technologies for the discovery, evaluation and application of the exposome (<http://emoryhercules.com/>).

### Challenges related to exposome measurement

Although very appealing, the exposome approach is complex, both in terms of measuring it (several hundred exposures are to be considered over the life course) and analysing its relationship to health [10].

One obvious challenge in the implementation of the exposome approach in humans relates to the assessment of “the totality” of environmental exposures. RAPPAPORT [13] has proposed two complementary strategies to characterise the exposome. In the bottom-up approach, the sources of important exogenous exposures are identified and the level of exposure is assessed through environmental measurements (*i.e.* air, water or food). In the top-down approach, the circulating levels of parent compounds and/or its metabolites and by-products are assessed individually through biological assays in blood, urine or other tissues. This top-down approach does not directly provide information on the sources and environmental levels of exposure (although pharmacologically based pharmacokinetic models could be used for the latter purpose).

The assessment of the exposome can rely on many different tools and technologies. These include geographic information system-based environmental models (in particular for air pollutants and noise), environmental sensors, personal dosimeters (*i.e.* for air pollutants, radiations, *etc.*), photos (to document diet, cosmetic or drug use), biomarkers, *etc.* (figure 2). New sensors relying on information technology are being constantly developed. For example, silicone wristbands, which were able to identify 49 compounds (including polycyclic aromatic hydrocarbons, consumer products, personal care products, pesticides and phthalates) among 30 volunteers, were suggested as an easy-to-use personal passive sampler [14]. In the longer term, high-throughput ‘omics technologies (transcriptomics, epigenomics, lipidomics, metabolomics) could help identify exposure biomarkers, and could even allow integration of a wide range of individual

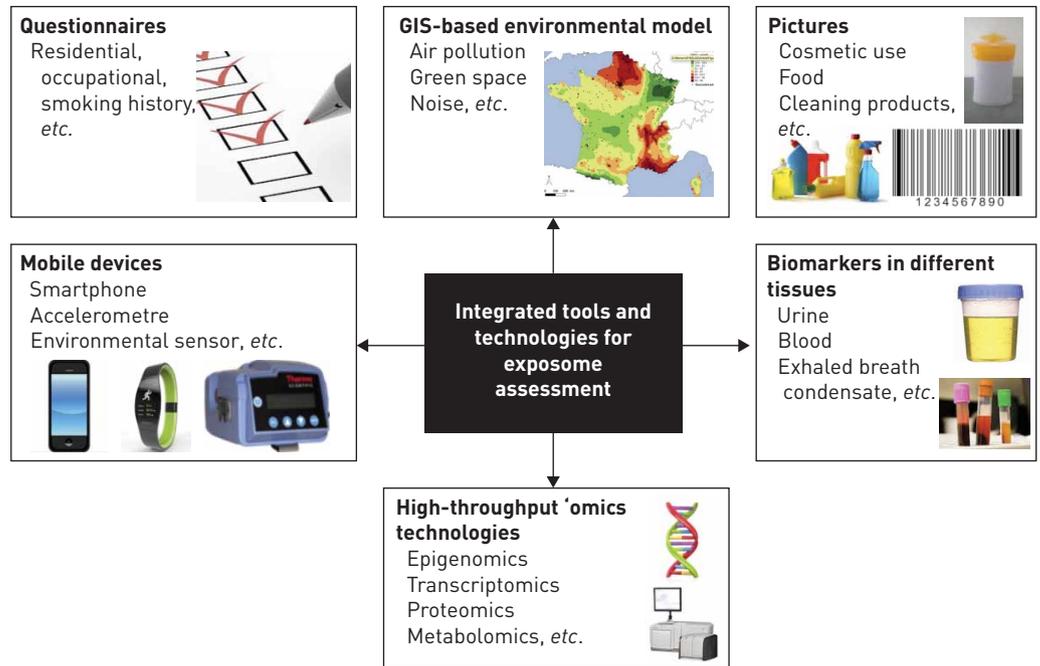


FIGURE 2 Tools for the assessment of the exposome.

exposures in a single measurement. Until this becomes a reality, comprehensively measuring the exposome remains expensive. The characterisation of a family of environmental exposures by biomarkers in a population of 500 subjects typically costs €50 000–€100 000, excluding the costs of recruiting volunteers and collecting and storing biospecimens. Though the costs of 'omics techniques are decreasing, relying on them in exposome studies remains expensive when repeated measurements (to cover the lifespan) on different "omics" (to cover the different layers) are performed (figure 3). This is compounded by the fact that efficiently considering many exposures in relation to health is expected to require larger sample sizes than epidemiological studies focused on a single exposure.

The improvements of biochemical assays over the past decades should not let us forget the challenges related to the use of exposure biomarkers when assessing the health effects of chemicals with high within-subject temporal variability. A simulation study illustrated that the use of a single biospecimen to assess the level of exposure biomarkers with high variability over time (intra-class coefficient of correlation of 0.2, corresponding approximately to what is observed with Bisphenol A) might lead to an attenuation bias in the dose–response function by ~80% (the attenuation factor in the case of classical-type error corresponds to one minus the intra-class coefficient of correlation) [15]. In the case of a compound that is more stable over time, the attenuation bias will be lower (typically ~40% for specific phthalates metabolites). Increasing the number of exposures considered in a study without making effort to limit this type of bias in the exposure measurements would be irrelevant. Similar bias might exist when 'omics technologies are used as exposure biomarkers, since the level of DNA methylation of some CpG sites or the concentration of specific proteins or metabolites vary over time. Bias can be limited by collecting several biospecimens per subject, and thereafter pooling them or using measurement error models [15]. This means that exposome studies in which exposures are assessed from a single biospecimen in each subject is likely to suffer from differential exposure misclassification between exposures [10]. Therefore, the consequence is that one may erroneously attribute the effect of a compound varying strongly over time to another exposure, varying less strongly over time but correlated to the average level of the first compound. Finally, the statistical power of exposome studies is likely to be inhomogeneous across components of the exposome. All in all, managing to increase the number of exposures considered without increasing exposure misclassification in the assessment of each exposure is probably one of the main general challenges of exposome research.

### Assessing the exposome–health relationships

Analysing the association between the exposome and health requires many exposure families to be simultaneously considered in a single statistical approach, and not separately as it is classically done. This allows selective reporting of the associations with the lowest p-values to be avoided, but brings additional statistical challenges compared to former studies.

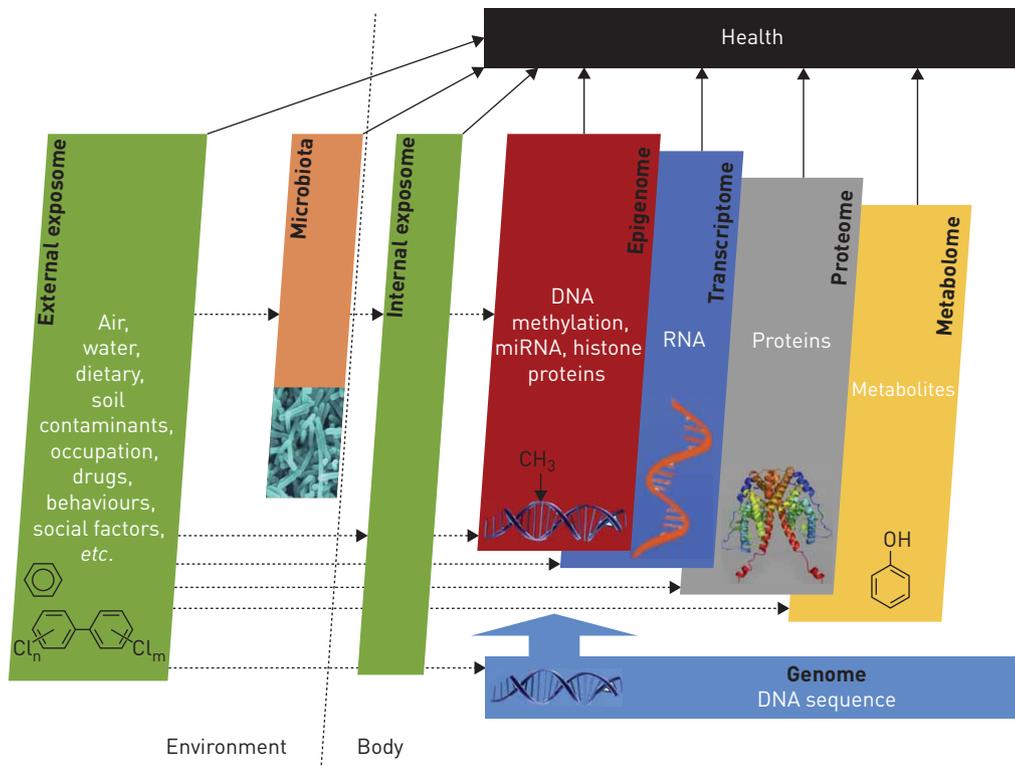


FIGURE 3 Schematic diagram of some of the 'omics layers and pathways of influence of the exposome on health. Not all interactions between the various 'omics layers are represented.

The first study using an exposome approach has been conducted in the context of type 2 diabetes and used the environment-wide association study (EWAS) approach; an agnostic method in which a large number of exposures (in this study 266) were successively and independently tested for their association with the outcome [16]. The statistical approach is analogous to that of genome-wide association studies. However, the correlation structure of the exposome differs from the one observed in genetics. Indeed, in genetics, single nucleotide polymorphisms (SNPs) are often only correlated with their close neighbours, *i.e.* with SNPs located nearby on the gene. When a SNP is found to be significantly associated with a health outcome, it is usually claimed that the genetic region of the SNP rather than the specific SNP identified is involved. In contrast, the correlation structure of the exposome, as recently described in the frame of the HELIX project [17], reveals strong levels of correlation within families of exposure (median within exposure family,  $r=0.45$ ) but also sometimes between families of exposure. Indeed, 11% of the exposures were correlated at a level  $>0.5$  with at least one exposure outside their family. A similar situation has been described in the context of the National Health and Nutrition Examination Survey [18]. In this context, the EWAS approach is not able to untangle the associations that are significant because the exposure is a causal exposure truly effecting the health outcome from the associations that are significant because the exposure is correlated to one of the causal exposures; this happens even if a multiple linear regression step including all potential "hits" is added after the EWAS step [19]. This can lead to a high rate of false positives. Several statistical methods have been proposed to cope with this issue [20] and some of them have been applied [18, 21]. However, systematic simulation studies are needed to characterise the efficiency of these approaches in the context of the exposome. A simulation study conducted in the frame of the HELIX and EXPOsOMICS projects showed that, due to the correlation within the exposome, the linear regression-based statistical methods that were investigated (one for each of the large families of variable selection approaches applied to regression-based methods) were only moderately efficient to differentiate true predictors from correlated covariates [19]. Further issues relate to measurement error and the identification of synergistic effects between exposures, for which it remains unclear which statistical model would perform best and which sample size would be required in exposome studies to provide sufficient statistical power.

## Conclusion

The exposome is a broad concept requiring the involvement from a wide range of disciplines (exposure sciences, toxicology, biology, epidemiology, statistics, *etc.*). The exposome is in its infancy, with many technical and statistical challenges that need to be overcome. In terms of assessment, several hundreds of time-varying

exposures need to be accurately assessed using various technologies. Nevertheless, increasing the number of exposures assessed should not be done at the cost of increased exposure misclassification. Estimating the association between a large set of exposures and health leads to statistical challenges, including the difficulty to efficiently untangle the exposures truly affecting the health outcome from correlated exposures and the difficulty in identifying synergistic effects between exposures. These challenges should not discourage research relying on the exposome approach and the on-going projects will hopefully allow steps forward to be made. If the parallel between genomic and exposomic research holds, huge efforts in terms of consortium building, development and validation of statistical tools and measurement devices will be necessary, requiring massive financial resources. From a public health perspective, a better understanding of the environmental risk factors will open the way to improved prevention strategies.

## References

- 1 Wild CP. Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev* 2005; 14: 1847–1850.
- 2 Wild CP. The exposome: from concept to utility. *Int J Epidemiol* 2012; 41: 24–32.
- 3 Vrijheid M. The exposome: a new paradigm to study the impact of environment on health. *Thorax* 2014; 69: 876–878.
- 4 Miller GW, Jones DP. The nature of nurture: refining the definition of the exposome. *Toxicol Sci* 2014; 137: 1–2.
- 5 Portelli MA, Hodge E, Sayers I. Genetic risk factors for the development of allergic disease identified by genome-wide association. *Clin Exp Allergy* 2015; 45: 21–31.
- 6 Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; 518: 197–206.
- 7 Beasley R, Semprini A, Mitchell EA. Risk factors for asthma: is prevention possible? *Lancet* 2015; 386: 1075–1085.
- 8 Aylward LL, Kirman CR, Schoeny R, et al. Evaluation of biomonitoring data from the CDC National Exposure Report in a risk assessment context: perspectives across chemicals. *Environ Health Perspect* 2013; 121: 287–294.
- 9 Delfosse V, Dendele B, Huet T, et al. Synergistic activation of human pregnane X receptor by binary cocktails of pharmaceutical and environmental compounds. *Nat Commun* 2015; 6: 8089.
- 10 Slama R, Vrijheid M. Some challenges of studies aiming to relate the Exposome to human health. *Occup Environ Med* 2015; 72: 383–384.
- 11 Tyrrell J, Melzer D, Henley W, et al. Associations between socioeconomic status and environmental toxicant concentrations in adults in the USA: NHANES 2001–2010. *Environ Int* 2013; 59: 328–335.
- 12 Vrijheid M, Slama R, Robinson O, et al. The human early-life exposome (HELIX): project rationale and design. *Environ Health Perspect* 2014; 122: 535–544.
- 13 Rappaport SM. Implications of the exposome for exposure science. *J Expo Sci Environ Epidemiol* 2011; 21: 5–9.
- 14 O’Connell SG, Kincl LD, Anderson KA. Silicone wristbands as personal passive samplers. *Environ Sci Technol* 2014; 48: 3327–3335.
- 15 Perrier F, Giorgis Allemand L, Slama R, et al. Within-subject pooling of biological samples as a way to reduce exposure misclassification in biomarker-based studies of chemicals with high temporal variability. *Epidemiology* 2016; 27: 378–388.
- 16 Patel CJ, Bhattacharya J, Butte AJ. An Environment-Wide Association Study (EWAS) on type 2 diabetes mellitus. *PLoS One* 2010; 5: e10746.
- 17 Robinson O, Basagana X, Agier L, et al. The pregnancy Exposome: multiple environmental exposures in the INMA-Sabadell Birth Cohort. *Environ Sci Technol* 2015; 49: 10632–10641.
- 18 Patel CJ, Ioannidis JP. Placing epidemiological results in the context of multiplicity and typical correlations of exposures. *J Epidemiol Community Health* 2014; 68: 1096–1100.
- 19 Agier L, Portengen L, Chadeau-Hyam M, et al. A systematic comparison of linear regression-based statistical methods to assess exposome-health associations. 27th Conference of the International Society for Environmental Epidemiology. São Paulo, 2015.
- 20 Chadeau-Hyam M, Campanella G, Jombart T, et al. Deciphering the complex: methodological overview of statistical models to derive OMICS-based biomarkers. *Environ Mol Mutagen* 2013; 54: 542–557.
- 21 Lenters V, Portengen L, Smit LA, et al. Phthalates, perfluoroalkyl acids, metals and organochlorines and reproductive function: a multipollutant assessment in Greenlandic, Polish and Ukrainian men. *Occup Environ Med* 2015; 72: 385–393.