2. Characterising the exposome

Exposure is commonly assessed by means of a spectrum of questionnaire data and ecological, environmental, and biological measurements.

2.1. The chemical exposome

Léa Maitre
Cristina Balcells Nadal

A significant subset of the exposome is made up of the chemical exposome. This subset encompasses all chemical species (and their associated transformation products), either of biological or synthetic origin, capable of entering the human body via different pathways, including ingestion, inhalation, and dermal absorption. These molecules can originate from a range of sources that include diet, pharmaceutical drugs, and dietary supplements, personal care and consumer products (PCCPs), as well as water and airborne substances.

The best techniques for measuring these chemicals in human tissues or biofluids involve the use of liquid chromatography (LC) and gas chromatography (GC), coupled to either tandem or high-resolution mass spectrometry (MS/MS or HRMS, respectively), and nuclear magnetic resonance (NMR) spectroscopy (Balcells et al., 2024). Over the years, these analytical techniques have evolved from methods that accurately analyse and quantify a few pre-optimised metabolites or chemicals at a time (*targeted methods*) to wide-scope approaches capable of profiling thousands of chemicals in a biospecimen with little a priori knowledge about them (*untargeted methods*). In what follows, current use cases, examples, and applications of both types of method in exposomics are discussed.

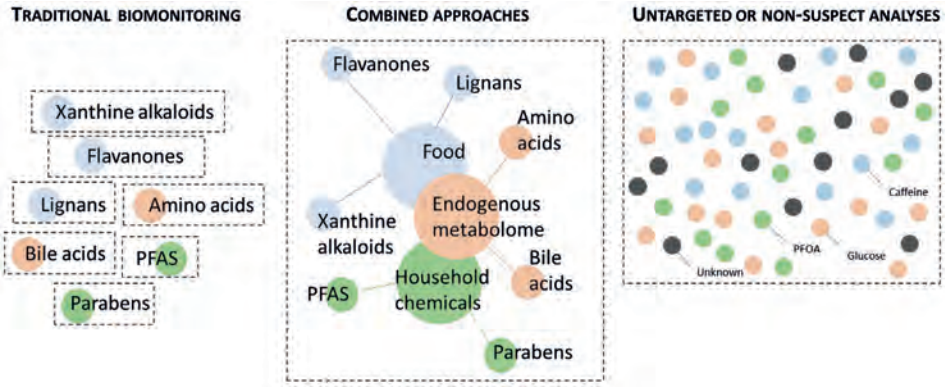

FIGURE 3. Evolution of the characterisation of the chemical exposome.
SOURCE: Created by Cristina Balcells Nadal.

Targeted analyses or traditional biomonitoring

Biological indicators of exposure, which assess internalised doses, are frequently favoured because of their direct relevance to the health outcomes under study. Conventional analytical measurements, commonly known as *targeted analyses*, involve assessing specific chemicals, metabolites, or reaction products in biological mediums such as urine or blood. These established biomonitoring methods have evolved into a fundamental element of exposure assessment in numerous epidemiological studies that aim to establish connections between exposures and health outcomes. Targeted, quantitative methods are still widely used to measure the chemical exposome, given the need for reliable biomonitoring data and the desire to define quantitative exposure-adverse health outcome associations for regulatory authorities and policymakers. Numerous governmental agencies and national laboratories worldwide periodically publish exposure biomonitoring data of their populations (e.g. PARC in the EU and NHANES in the US). Most targeted analytical methods employed to characterise the exposome are tailored to distinct groups of established exposures. Typically, these methods measure either single or up to a few tens of compounds and are validated for a specific biofluid or matrix. Moreover, they can be used to characterise molecules of different origin, from endogenous or dietary metabolites to man-made chemicals.

Examples of the most common families of chemical exposures targeted by current biomonitoring programmes include:

— *Plasticisers*: phthalates.
— *Combustion products*: polycyclic aromatic hydrocarbons (PAHs).

— *Tobacco smoke*: tobacco-specific nitrosamines (TSNAs), cotinine, heterocyclic aromatic amines (HAAs).

— *Pesticides, insecticides, and insect repellents*: pyrethroids, organochlorine pesticides, dialkyl phosphate pesticides and other organophosphate pesticides, carbamate insecticides, insect repellent, and metabolites (DEET).

— *Herbicides*: atrazine and metabolites, 2,4-D, 2,4,5-T and metabolites, sulfonylurea and metabolites.

— *Fungicides.*

— *Industrial and PCCPs*: per- and polyfluorinated substances (PFAS), parabens, phthalates, bisphenols.

— *Pharmaceuticals.*

— *Natural products*: mycotoxins, phytoestrogens.

— *Multiple sources*: volatile organic compounds (VOCs) and metabolites, polychlorinated biphenyls (PCBs), polychlorinated dibenzodioxins (PCDDs) and dibenzofurans.

Another major source of chemical exposure is diet. Some of the most reproducible biomarkers (metabolites) of food intake include:

— *Coffee, tea and cocoa*: caffeine, theobromine, theophylline.

— *Meat*: specific fatty acids and amino acid derivatives.

— *Fish*: specific fatty acids and amino acid derivatives.

— *Fruit and vegetables*: flavones, flavanones, coumarins.

Combined approaches

The realisation that these analytical techniques are suitable for the measurement of most small molecules, regardless of their origin, has paved the way for efforts to cover wider windows of the chemical space in a single analytical assay. In this way, and with the emergence of more sensitive, faster-scanning instrumentation, targeted assays typically aimed at a single chemical family as described above are giving way to more comprehensive targeted strategies, in an effort at simultaneously capturing various combinations of molecules from the endogenous metabolome, the food-related and microbiota-derived metabolomes, pharmaceuticals, environmental contaminants, and household chemicals.

Untargeted or non-suspect analyses

With only a few hundred chemicals routinely measurable using targeted methods, exposomic approaches are critical for understanding the thousands of chemicals people are exposed to on a daily basis through direct chemical exposures, as well as the consequences of such exposures (e.g. oxidative stress markers).

Moreover, all the information obtainable from targeted methods, even wide-scope approaches, necessarily concerns previously characterised molecules, which means the potential for the discovery of unknown exposures or exposure biotransformations is limited. By resorting to untargeted biomonitoring approaches, such as high-resolution metabolomics (HRM), thousands of chemical species can be monitored using just a relatively small amount of biological specimen (≤ 100 μL) and for the cost of a single traditional biomonitoring analysis of 8-10 target chemicals.

In principle, HRM provides the most comprehensive description of small molecular composition possible, including biomarkers of exogenous exposure as well as endogenous metabolites, which together make up a major component of the internal exposome. This field is currently in a stage of rapid development, capable of measuring and annotating hundreds and thousands of small molecules in each analytical run. This is gradually allowing light to be shed on the "dark exposome" or "unknown" chemical risk factors of disease (i.e. not yet identified as suspected risk factors and for which no high-accuracy measurement tools are available). One of the main challenges in measuring the chemical exposome is covering the range of compound abundance in the human body. The concentration of endogenous metabolites, food biomarkers, and drugs present in the blood can span roughly eight orders of magnitude; however, when combined with environmental pollutants, the required range for detecting all exposome compounds present in the body increases to over ten orders of magnitude from femtomoles to millimoles (Rappaport et al., 2014). This exceeds the linear dynamic range of modern mass spectrometers by 10,000-100,000 fold. However, recent developments in separation science, by increasing the resolution of existing separation methods (ultra-performance liquid chromatography or UPLC), and augmenting the complexity of data to detect more compounds (ion mobility spectrometry or IMS), are addressing this problem. Another strategy involves the use of (ultra) high-resolution mass spectrometry ((U)HRMS), which allows a radical increase in the number of detected features and an enhancement of the mass resolution, by exploiting such instruments as Orbitraps and Fourier-transform ion cyclotron resonance (FT-ICR). Other options to boost the number of features obtained include combining complementary stationary phases (e.g. reversed-phase chromatography and hydrophilic interaction liquid chromatography, or GC and UPLC) or removing high abundant analytes and concentrating low abundant exposome compounds, as used in targeted or semi-targeted analyses with the integration of standard reference compounds to allow for quantification (Gil-Solsona et al., 2021).

Finally, the development of mass spectral libraries promises improved annotation of metabolic features. Often referred to as "dark matter", the majority of features measured by untargeted MS cannot be annotated with high confidence

to a given compound, where the level of confidence can range from unknown (level 5, simply knowing the mass of the molecule); elucidating its molecular formula (level 4, thanks to the isotope pattern detected); assigning functional groups or a compound class (level 3, by detecting diagnostic fragments); assigning a probable structure (level 2, by matching its fragmentation pattern to a spectral library) to finally validating the annotation with a chemical standard (level 1) (Schymanski et al., 2014).

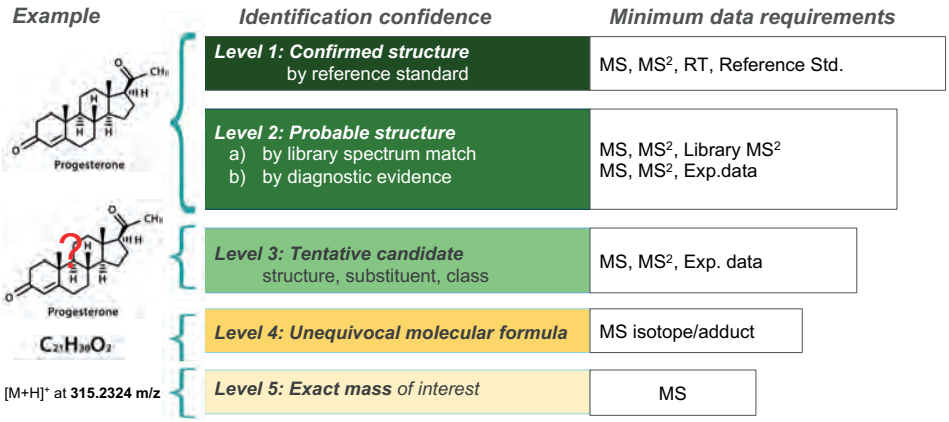

FIGURE 4. Identification confidence levels in high resolution mass spectrometric analysis.
SOURCE: Created by Cristina Balcells Nadal.

Indeed, more than 40 million compounds are listed in PubChem and ChemSpider but spectra are available for only one hundredth of these. Metabolome-wide databases are building on endogenous compound libraries and incorporating environmental toxicants, food contaminants and supplements, as well as drugs and their biotransformation products (METLIN; The Human Metabolome Database (HMDB); Warth et al., 2017) and metabolic databases dedicated to biomarkers of exposure to environmental risk factors are also being developed (e.g. http://exposome-explorer.iarc.fr/). Likewise, other tools are becoming available to help identify compounds uncovered in untargeted analyses and which rely on additional compound characteristics such as fragmentation patterns (tandem MS), exact mass (HRMS), retention time modelling, or ion mobility (collision cross section or CSS), aided by advanced computation and machine learning (Dührkop et al., 2015).