BMJ Health & Care Informatics

Use of informatics to characterise the exposome of COVID-19

Fernando Martin-Sanchez,[1] Miguel Atienza-Maderuelo,[2] Guillermo Lopez-Campos,[3] Paloma Collado[4]

To cite: Martin-Sanchez F, Atienza-Maderuelo M, Lopez-Campos G, *et al*. Use of informatics to characterise the exposome of COVID-19. *BMJ Health Care Inform* 2021;**28**:e100371. doi:10.1136/bmjhci-2021-100371

FM-S and MA-M are joint first authors.

Received 08 June 2021
Accepted 26 July 2021

[1]National School of Public Health, Carlos III Health Institute, Madrid, Spain
[2]IMIENS, UNED, Madrid, Spain
[3]Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast, Belfast, UK
[4]Psychobiology, UNED, Madrid, Spain

Correspondence to
Dr Fernando Martin-Sanchez; fmartin@isciii.es

THE IMPORTANCE OF EXPOSOME IN PRECISION MEDICINE

Exposome refers to the set of external exposures that affect an individual's health from conception to death.[1] This term includes all non-genetic risk factors which, interacting with the genetic endowment (genome), produce a state of health or disease (phenome). Considering these non-genetic data in precision medicine contributes to improving clinical decision making and biomedical research. It is still too frequent to find in the literature the description of research projects on certain diseases that only refer to the study of the genotype and the phenotype (https://www.icpermed.eu/en/icpermed-medicine.php). The same can be observed in database resources (https://www.ncbi.nlm.nih.gov/gap/). This indicates that either the exposome is neglected or the phenome is mistakenly considered to include the exposome (eg, smoking).

INTRODUCING THE CONCEPT OF EXPOTYPE

There is a scarcity of accepted standards to represent human exposure data. Therefore, representing individuals' exposome data constitutes a challenge for biomedical informatics, considering that it represents a source of big data that changes throughout time and space.[2] To overcome this, the concept of expotype emerges as a fundamental term to facilitate a holistic approach to health.

Introducing the concept of *expotype*[3] is a non-trivial issue, and so it is presented here by analogy with the terms *genotype* and *phenotype*. As the genome represents the complete DNA of a human organism, including all its genes, the term genotype refers to the genetic information of one specific individual. Likewise, the concept of phenome represents all those features presented by a human organism, and the phenotype is defined as the structural and/or functional characteristics that can be observed about an individual, produced by the interaction between its genotype and its context. Finally, by analogy, the *expotype would represent a specific set of exposures accumulated by an individual during a certain time/space window.* In this sense, to understand how diseases are developed, it is mandatory to assess how the environmental risk factors (expotypes) interact with the genomic information of an individual (genotype), generating specific phenotypes.[4] Collecting standardised data about individual genotypes, expotypes and phenotypes in research repositories offers the possibility of conducting integrative studies about the pathophysiology of complex diseases.

DESCRIPTION OF A COVID-19 EXPOTYPE

Multiple knowledge domains need to be integrated to efficiently address current health challenges, such as the COVID-19 pandemic. SARS-CoV-2 has put great pressure on health systems, and it is an appropriate example of how external exposure plays a key role in affecting human health.

Building a COVID-19 expotype requires, besides the virus biology, considering all those factors related to its transmission and development, gathering individual data for health and research purposes. In general, environmental factors such as temperature and humidity play a key role in viral transmission. Recent studies show that ecological factors (policy, health behaviours, physical environment and clinical care) are associated with COVID-19 case fatality rate.[5] Furthermore, exposure to ultraviolet radiation favours the synthesis of vitamin D in the body, which plays an important role in the immune defence against viral and bacterial infections.[6] Additionally, a relevant correlation has been found between exposure to particulate matter (pollution) and COVID-19 death rates.[7] On the one hand, pollution actively contributes to comorbidity, worsening the health of those exposed and therefore their prognosis

fighting the virus. On the other hand, particulate matter may also be a possible virus carrier. Other environmental factors such as noise pollution and tobacco use have also been related to COVID-19.

SARS-CoV-2 effects are also related to social, cultural and behavioural factors. Measuring time spent by individuals outdoors and indoors is essential, given that most of the infections are related to extended human contact periods, mostly occurring in locations such as restaurants or gyms. Safety measures such as using masks or maintaining social distance are also behavioural risk factors to be analysed in conjunction with the expotype data.

Individuals with obesity show greater likelihood of requiring special care and increased risk of death.[8] Moreover, research has shown how food consumption influences brain neurotransmitter systems, affecting mood. Bidirectionally, mood can also influence food choices, leading to repetitive cycles of overeating, weight gain and depression. Hence, data related to eating habits also form part of COVID-19's expotype structure,[9] whereas psychological status would be part of the phenotype.

REPRESENTING EXPOTYPES WITH THE ISO-EN13606 MODEL (ARCHETYPES)

Today, digital health methods allow us to track individual exposure data. However, a standard model to process these data is still lacking. We present here a mechanism for data representation based on the concept of expotype, meaning specific exposures of an individual during a certain time/space window, and its representation using *archetypes*, based on the CEN/ISO-EN13606 data model international standard.[10] Such standard data models will allow formalising and managing knowledge as it evolves, easily adapting it to new research and development needs.

Twitter Fernando Martin-Sanchez @fermarsan

Contributors FM-S and GL-C conceived of the presented idea. FM-S and MA-M developed the data models. PC provided the biomedical framework. MA-M wrote the manuscript with support from FM-S, PC and GL-C. All authors provided critical feedback and contributed to the final manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Commissioned; externally peer reviewed.

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;14:1847–50.
2 Cui Y, Balshaw DM, Kwok RK, *et al*. The Exposome: embracing the complexity for discovery in environmental health. *Environ Health Perspect* 2016;124:A137–40.
3 Martin-Sanchez FJ, Lopez-Campos GH. The new role of biomedical informatics in the age of digital medicine. *Methods Inf Med* 2016;55:392–402.
4 Martin-Sanchez F, Bellazzi R, Casella V, *et al*. Progress in characterizing the human Exposome: a key step for precision medicine. *Yearb Med Inform* 2020;29:115–20.
5 Kathe NJ, Wani RJ. Determinants of COVID-19 case fatality rate in the United States: spatial analysis over one year of the pandemic. *J Health Econ Outcomes Res* 2021;8:51–62.
6 Borella E, Nesher G, Israeli E, *et al*. Vitamin D: a new anti-infective agent? *Ann N Y Acad Sci* 2014;1317:76–83.
7 Sciomer S, Moscucci F, Magrì D, *et al*. SARS-CoV-2 spread in northern Italy: what about the pollution role? *Environ Monit Assess* 2020;192:325.
8 Dietz W, Santos-Burgoa C. Obesity and its implications for COVID-19 mortality. *Obesity* 2020;28:1005.
9 Colomina M-T, Sánchez-Santed F, Conejo NM, *et al*. The Psychoexposome: a holistic perspective beyond health and disease. *Psicothema* 2018;30:5–7.
10 Tapuria A, Kalra D, Curcin V. Feasibility of using en 13606 clinical archetypes for defining Computable phenotypes. *Stud Health Technol Inform* 2020;270:228–32.