Occupational Health: Accessing and coding biomonitoring data acquired as part of the authorities' monitoring activities

Case study 3: Extraction of Cd and Cr occupational exposure data for comparison with exposure of the general population.

on behalf of the Federal Office of Public Health (FOPH)

contract ID:14200640

Hasnaa Chettou[1], Murielle Bochud[2] & David Vernez[1]

[1]Department of Occupational and Environmental Health, Unisanté, University Center for Primary Care and Public Health, Lausanne, Switzerland

[2]Department of Epidemiology and Health Systems, Unisanté, University Center for Primary Care and Public Health, Lausanne, Switzerland

DOI : https://doi.org/10.16908/pub.2025.005

This report was produced by Unisanté in the context of a mandate (Proof-of-concept-DigiSanté) including 3 case studies as part of the DigiSanté program. This case study illustrates the secondary use of health-related data, as supported in the DigiSanté project "Datenraum für die gesundheitsbezogene Forschung" (DigiSanté Programm zur Förderung der digitalen Transformation im Gesundheitswesen, package 4). It relates to data use categories (1) Academic research and (4) Scientifically developed information provision for stakeholders in healthcare system and health policy (authorities, etc.).

Context

The Horizon Europe PARC project, in which Switzerland is participating, requires the collection of exposure data for several living environments. This need stems from the modeling efforts carried out in PARC workpackage 6.2.1. - *Exposure aggregation through different sources and routes related to the general and occupational environments*. This project will contribute to the development of decision-making tools for aggregating exposure to chemical pollutants and can also guide decision-makers or

regulators on priority sources of exposure requiring action or special attention. Considering aggregate exposure is a key element in understanding exposure from a One Health perspective. **At present, only fragmented information, distributed across disjointed data sources, is available, and efforts to harmonize and link data are needed to gain an overall view. In this sense, the objectives of the European PARC WP6.2 are aligned with the intentions of the Swiss DigiSanté program**.

The aim of WP 6.2.1 is to develop methods for decision-making on the relevance of aggregating human exposure to chemical pollutants, considering all relevant exposure sources (occupational, general environment, domestic environment) and routes (inhalation, food ingestion, skin exposure). A typical example is the concomitant exposure to heavy metals found in certain professional activities (metallurgy, mechanics, watchmaking, etc.), in our general environment and in our everyday food consumption. Considering these sources separately does not allow us to assess the body burden of the population, or to correctly evaluate the risk posed by these substances.

In parallel with the conceptual development of these methods, case studies applicable to specific chemicals are planned to test their feasibility and practical challenges. The development of these case studies requires the use of exposure scenarios for selected substances, including Cadmium (Cd) and Chromium (CrVI).

Data sources

Three sources of exposure data of particular interest for Switzerland were identified.

- **The Swiss Health Study** (SHeS-pilot): this cross-sectional population-based study provides data on blood cadmium (Cd) and chromium (Cr) concentrations as well as gender, age, smoking status, diet, and profession.
- **The SUVA biomonitoring data**: in its role as an enforcement body, SUVA carries out biomonitoring in companies including workers who are particularly exposed to certain risks. The collected data include, in addition to age and sex, data on occupational exposure, in particular, workplaces descriptions and activity sector, and also Cr(VI) and Cd biomonitoring data.
- **The Swiss Kidney Project on Genes in Hypertension (SKIPOGH)**: this multi-centric population-based and family-based cohort collected data on the determinants of blood pressure and renal function in the canton of Vaud, Geneva and Bern. It is a SNFS-funded research project. The collected data include blood and 24-hour urinary excretion of Cd and Cr as well as the usual socio-demographic variables, including profession.

To exploit this information in an aggregate exposure perspective, it is necessary to be able to **combine the databases through common exposure determinants** (e.g., the same occupational activity, the same domestic environment, the same city, etc.). That is, **no classical record-linkage via a unique person-identifier** is needed, but databases of relevance need to be combined via **common unifying parameters/variables**. This type of linkage/connection would allow to enrich the available information in a single database. For instance, the Suva dataset could be used to define exposure class or exposure estimates to heavy metal according to the profession of the exposed person. The use of this information in general population databases could provide indications of the contribution of different sources of exposure (dietary, occupational) to total body burden. At the same time, it could provide a basis for preventive measures in both the domestic and occupational contexts.

Practical application and obstacles

Data description and data acquisition

For SHeS-pilot data

- a description of the request process and data request application form are available on Unisanté website (Unisanté Data Repository1), under a mandate of FOPH and clear governance structure.
- study description and codebooks are also available online, enabling to understand the content of the database and to decide on its relevance for the study before requesting it. Contact persons are available on the same website, making it easy to reach the right persons.
- a description of the study, meta-data and data request process are available on the Maelström catalogue SHeS-pp - Maelstrom Research (maelstrom-research.org)
- profession is recorded in the database

For SKIPOGH data

- a description of the study, meta-data and data request process are available on the Maelström catalogue (https://www.maelstrom-research.org/study/skipogh) and a description of the study is also available on the Unisanté website (Swiss Kidney project on genes in hypertension (SKIPOGH) | Unisanté (unisante.ch))
- profession is recorded in the database

Both databases, SHeS-pilot and SKIPOGH, record the profession, thus creating a basis for interoperability between data sources. However, experience shows that registering the profession requires a harmonized approach to enable systematic coding (e.g., with an ISCO code).

For SUVA data

- To the best of our knowledge, SUVA data are not findable through a metadata catalogue. Also, the data description, and the procedure for requesting data are not available online.
- We were only aware of Suva's data and the possibility of calling on it through our regular exchanges with them.
- Multiple interactions were necessary to find out what data are available at SUVA, whether they are relevant to the project and how to request them. Data cannot be requested without ethical approval.

  Suva needs to justify that the transfer of data is part of their mission (no research mandate).

  Their data is nominative (companies and individuals), and the data is not coded.

  In order to transfer the data, Suva must therefore anonymize (companies) or de-identify/pseudonymize the data (individuals).

[1] https://data.unisante.ch/index.php/catalog/37

It is necessary to go through the ethics committee to ensure that data can be re-used without asking individuals for their consent (that is the data is sufficiently de-identified).
Suva needs the agreement of the ethics commission prior to providing the data.

- If data request from Suva is done in the context of a research project funded by SECO, SECO can ask directly Suva to get access to the dataset in its role as authority, which facilitates considerably the process.

- Suva is the legal authority for occupational disease prevention and control. As such, they may impose a biological monitoring of the workers in high-risk occupational activities (Art. 50 VUV). However, Suva does not have the legal means to share the collected person-related data with research.

- SUVA data are not available in a standardized format and need to be de-identified / anonymized and structured before they can be provided to researchers.
  To our best of knowledge, Suva does subcontract the biological sample analysis, presumably for technical reasons. The biomonitoring results are structured (e.g. excel files), but the descriptors of the exposure situations are not coded and must be extracted from the Suva field reports,

Data processing and combining

The use of SHeS data requires the harmonization of certain variables so that they can be used and analyzed:

- **Answers requiring harmonization**:
  This concerns data from questionnaires in free-field form, where participants may provide similar information in different ways. For instance, in the SHeS data, participants described their exposure to smoke or dust in free-form text, resulting in varied responses regarding the type, source, or name of the chemical emission (e.g., wood dust, welding fumes, steel, etc.). Consequently, the level of detail and utility of the information differs among participants.
  The questions could provide more informative answers if they target specific pollutants rather than mixtures or emissions and are presented as multiple-choice question.

- **Questions requiring harmonization**:
  Questions should enable resulting data to be combined with other databases, and facilitate the processing of responses (e.g., multiple-choice questions):

  - Data from the food frequency questionnaire (FFQ) can be more effectively used and combined with other data if foods or food groups are classified according to a standardized system, such as FoodEx.

  - NOGA code (Nomenclature générale des activités économiques) was requested for the participants' occupational activity, but more detailed classification systems with additional hierarchies, like the International Standard Classification of Occupations (ISCO), could be used. These systems facilitate identifying the worker's tasks and linking this information to exposure data from sources like the SUVA database. However, currently, occupation and activities are described in reports by SUVA, so it has to be recoded first.

  - Questions that can be understood/interpreted in different ways can lead to difficulties in interpreting the answers as they may be understood differently by different participants. This can lead to uncertainty about the participant's intent and, in some

cases, the complete loss of information if responses vary significantly or are unclear. For example, the question "How often do you smoke on average per week?" could be interpreted as asking either for the number of cigarettes smoked per week or the number of days per week the participant smokes. Standard questions and coding schemes from reliable expert sources should be used or – if feasible – record linkage with existing high-quality databases being considered.

Improvement suggestions

- Enhance findability of existing databases in Switzerland, e.g., via the I14Y metadata catalogue
- Put in place the legal prerequisites and infrastructures to facilitate access to and re-use (including record-linkage) of the databases of public bodies and enforcement agencies in the public interest.
- Use of international coding systems (e.g. ISCO for professions, FoodEx for foods, etc.) to enable data to be combined with other national or international data sources.
- Identify key data likely to be used as a link between databases in a OneHealth perspective. Key data like sex/gender, age, education, occupation and geolocation should be
  - coded in a mandatory manner across the national databases,
  - harmonized across all databases.
- Use of targeted questions to ensure consistent and informative responses among the survey participants. Similar questions should be harmonized between databases.
- Support non-personal data re-use to provide evidence to help industry understand health-related issues in their working environment and take appropriate health protection measures.