HEALTH DATA TALK SERIES:

Data linkages in Belgian (administrative) RWD: the missing link

UNLOCKING THE POTENTIAL OF REAL WORD DATA IN BELGIUM: THE VALUE AND POSSIBILITIES OF DATA LINKAGES IN HEALTHCARE

Thanks to a long tradition of systematically collecting administrative healthcare data, Belgium possesses a remarkable set of databases that positions us to play a leading role in the secondary use of data (for legitimate purposes). For example, Belgium's Minimal Hospital Dataset (Minimale Ziekenhuisgegevens/Résumé Clinique Minimum), maintained by the Federal Public Service of Health, collects detailed information on every hospital admission in Belgium. This includes demographic details (like age and gender), information on the diagnoses (in ICD coding[1]), the length of stay, whether the patient was admitted to an intensive care unit or entered the hospital via the emergency room, and much more….

Another rich source is the IMA database. Every time someone visits a doctor, gets a prescription filled, or undergoes a treatment that is reimbursed by the Belgian social security system, a record of that action is created (a claim in a fee for service system). IMA collects this information from the seven Belgian health insurance funds, so it has data on almost all medical care that is reimbursed in Belgium. In addition, these databases also contain demographic information on patients (such as age, gender, socio-economic status, educational level), but also data on, for example, sickness leave. Almost the entire Belgian population is covered by the social security system, meaning these datasets are highly representative. Its completeness (due to fee for service system) and representativeness make it a highly valuable source for the secondary use of health data, as it enables researchers to track patients anonymously or pseudonymously over time and gain detailed insights into their healthcare use. Re-using these anonymous or pseudonymised health data insights have a high value for health care and health decisions.[2]

Belgium has many other valuable administrative health/health related data sources, such as:

- Statbel databases with cause-of-death statistics,
- The Cancer Registry, a national database that collects all cancer diagnoses
- Registers of medical implants in which information on implants placed in patients is collected,
- Health surveys from Sciensano,
- The Intego database with primary care data,
- Belgian Resident Assessment Instrument (BelRAI)
- …

Most of them are already included in the Belgian dataset catalogue of the Health Data Agency (HDA) describing the meta-data of the datasets. What makes these datasets even more powerful is that, **in theory, they can all be linked** using a unique identifier such as the national INSS[3] number (Identification Number Social Security). This opens up incredible opportunities for scientific and policy research. Linking and analysing health data, of course, in a pseudonymised format, from different sources greatly increases its value, helping to improve diagnosis, personalize treatments, enhance patient outcomes, identify unsafe practices, reward quality care, detect fraud, assess long-term treatment effects, and ultimately also support innovation in healthcare[4].

The depth and richness of these data makes Belgium uniquely positioned to conduct Real-World Data (RWD) and Real-World Evidence (RWE) studies for medical innovations by both academic and private sector researchers.

THE CHALLENGES IN DATA LINKAGE FOR RESEARCH IN BELGIUM AND HOW TO MOVE FORWARD

However, Belgium is not yet reaching its full potential. In a survey conducted by OECD[4] in 2022, Belgium scored high when it came to **data governance.** Data governance is defined by the rules and systems to ensure data is

---

[1] International Classification of Diseases (ICD); Belgium currently uses ICD-10, for more information, click here

[2] For more information on the value of secondary us of data, see Health Data Talk Series, HDA Academy: Value of secondary use of data for health care users and Value of secondary use for innovation

[3] Identificatie Nummer Sociale Zekerheid (INSZ), or Identification Number Social Security (INSS) in English. It is a unique identifier used in Belgium for individuals within the social security system

[4] OECD (2022), Health Data Governance for the Digital Age: Implementing the OECD Recommendation on Health Data Governance, OECD Publishing, Paris.

handled responsibly, ethically, with transparency and securely. But based on the survey, there was room to improve when it came to **availability, maturity, and actual use** of this data. Regarding the same OECD survey, Belgium scored 57% on data sharing with the public sector, 0% on data sharing with the for-profit sector, and only 33% on linking datasets for research, statistics, and monitoring.

Based on the first-hand experience of a data request made by the author (Degefaye) as part of his PhD project, the DepEnD (Deprescribing at the End of Life) project, which involved accessing data from various agencies, the obstacles he faced were numerous when requesting data linkages for his research. For example, multiple approvals from different bodies with different forms causes a long and administrative procedure for researchers. The timeline to receive response and approval of the data request was cumbersome. The quality of the data was not optimal: errors in the linkage, implausible values, missing variables, etc.

Since 2023, the Belgian government took a proactive role in addressing these barriers. The creation of the Belgian **Health Data Agency (HDA)** and their full support for the implementation of the **European Health Data Space (EHDS)** are expected to improve the secondary use of health data. The current efforts initiated by the HDA and other stakeholders to improve transparency, streamline procedures, increase staffing, standardize data-sharing practices, and ensure better inter-agency coordination will be critical to succeed and to position Belgium as a frontrunner in RWD research, further building on its strong e-health and data infrastructure.

MOVING FORWARD: RECOMMENDATIONS TO IMPROVE DATA LINKAGES

The value of health data linkages requires each partner in the ecosystem to contribute.

- **Governments** play a role in enabling structural conditions such as centralised access mechanisms, harmonised procedures, and long-term investment in infrastructure and research.
- **Data agencies** contribute by ensuring quality assurance, transparent communication, and streamlined access processes, supported by adequate technical capacity and a more optimal use of the national registration number as a unique identifier that enables data linkage.
- **Researchers** can contribute through early engagement, thoughtful variable selection, and knowledge sharing, while anticipating practical constraints such as delays and funding implications.

Also the EHDS aims to enable secure and interoperable health data exchange across Member States. Pilot projects such as HealthData@EU and joint actions like TEHDAS2 are already testing the practical implementation of such linkages. These initiatives illustrate how shared infrastructure, harmonised standards, and trust frameworks can help overcome longstanding barriers to secondary data use.

Together, these complementary efforts help to solve the missing link in data linkages benefiting researchers, policymakers, and ultimately, patients.

Signed by Degefaye Zelalem Anlay (Vrije Universiteit Brussel) and the Health Data Talk Series group

Hendrik De Rocker
(APB)

Inge Van de Velde
Stefan Gijssels
(Patient Expert Center)

Isabelle Huys
(KULeuven)

Katleen Janssens
(Health Data Agency)

Karen Crabbé
(pharma.be)

Nick Marly
(Cabinet public health)

Stefanie Devos
(beMedTech)

Sofie Bekaert
(King Baudouin Foundation)

Sofie De Broe
Wannes Van Hoof
(Sciensano)

Teodora Lalova-Spinks
(Ghent University, KULeuven)