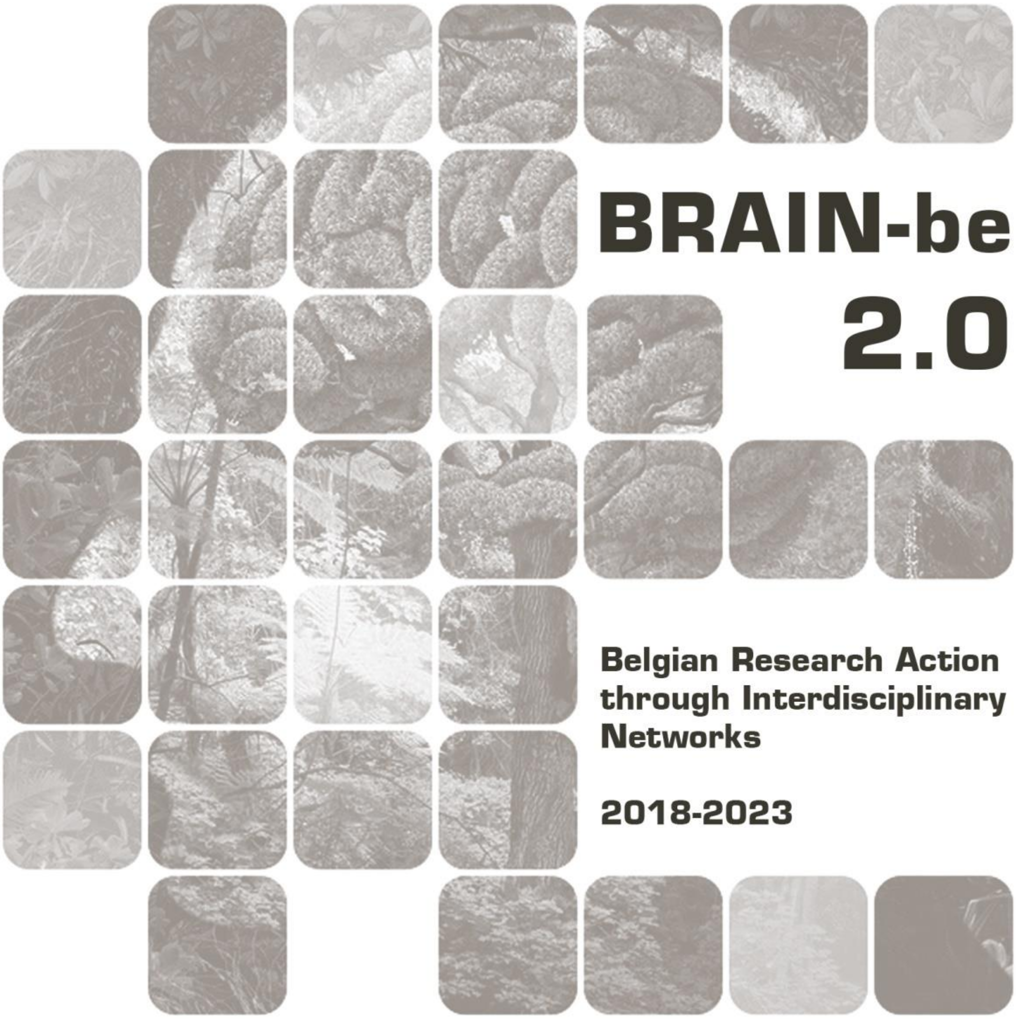

BRAIN-be 2.0

Belgian Research Action through Interdisciplinary Networks

2018-2023

AHEAD

Marlies Saelaert (Sciensano)
Wannes Van Hoof (Sciensano)
Brecht Devleesschauwer (Sciensano)

Pillar 2: Heritage science

NETWORK PROJECT

AHEAD
Towards the Development of a National Health Data Platform

Contract - B2/202/P2/AHEAD

FINAL REPORT

PROMOTORS: Brecht Devleesschauwer (Sciensano)

AUTHORS: Marlies Saelaert (Sciensano)
Wannes Van Hoof (Sciensano)
Brecht Devleesschauwer (Sciensano)

Published in 2023 by the Belgian Science Policy Office
WTCIII
Simon Bolivarlaan 30 bus 7
Boulevard Simon Bolivar 30 bte 7
B-1000 Brussels
Belgium
Tel: +32 (0)2 238 34 11
http://www.belspo.be
http://www.belspo.be/brain-be

Contact person: Georges JAMART
Tel: +32 (0)2 238 36 90

Marlies Saelaert; Wannes Van Hoof; Brecht Devleesschauwer. ***Towards the Development of a National Health Data Platform***. Final Report. Brussels: Belgian Science Policy Office 2023 – 44 p. (BRAIN-be 2.0 - (Belgian Research Action through Interdisciplinary Networks))

Table of Contents

ABSTRACT

CONTEXT

Belgium has a rich health information landscape, including valuable information originating from many different sources, such as administrative sources (e.g. concerning health care use, mortality, causes of death, social security, etc.), population health surveys and health surveillance systems. These different data sources are managed by different academic and governmental actors. To date, however, an integrated health information system is lacking, which hampers the valorisation of these data sources.

OBJECTIVES AND METHODS

The overall aim of the AHEAD project was to offer an active push towards the development of a national health data platform, by taking steps that will allow to move closer to the holy grail of a single integrated health data platform. It achieved this aim through the following specific objectives/work packages:

- WP1: The characterization and documentation of the Belgian health information system, and this by (i) proactive stakeholder engagement and (ii) bringing together all involved parties via the organization of webinars and study days
- WP2: The identification of ethical, legal, and technical bottlenecks in the prospective development of an integrated health data platform, and this by (i) the development of an online citizen consultation on the secondary use of health data and (ii) the development of a business case built around prospective linkages between Sciensano's Health Interview Survey (BHIS) data and administrative data sources
- WP3: The elaboration of recommendations for developing a national health data platform as an accessible, long term monitoring and research infrastructure

During the course of the AHEAD project, the Belgian Health Data Agency (HDA) got established. This provided valuable synergies, as several of the objectives and recommendations of AHEAD fall within the mandate of the HDA. The AHEAD project was therefore able to shift focus to areas that were complementary to the short-term objectives of the HDA.

CONCLUSIONS

A better valorisation of (health) related data within the Belgian health information landscape is of utmost importance to be able to provide answers to policy-relevant research questions. Belgium, being a country with many authorities, services and agencies but short communication lines, might hold the right cards and offer suitable opportunities to build an infrastructure holding longitudinal collections of health and administrative data. The AHEAD project explored these opportunities, but also the challenges, in the development of a national, integrated health data system. The different activities of AHEAD resulted in the following 10 recommendations:

1- Clarify the current Belgian health information landscape and provide and sustain up-to-date metadata
2- Establish a consortium of the willing
3- Standardize linkage procedures
4- Develop an ethico-legal framework
5- Increase digitalization
6- Secure a long-term financial and organizational effort
7- Identify and tackle ethical and legal bottlenecks
8- Respect the data relationship between citizens and health (related) data in the secondary use of health data
9- Maintain a suitable power balance in the secondary use of health data
10- Establish a citizen-powered framework for the secondary use of health-data

KEYWORDS

Belgian health information landscape; Health data reuse; Secondary use of health data; Data linkages; Administrative health data cohorts; Research infrastructures, Citizen consultation

1. INTRODUCTION

Belgium has a rich health information landscape, including valuable information originating from many different sources, such as administrative sources (e.g. concerning health care use, mortality, causes of death, social security, etc.), population health surveys and health surveillance systems. These different data sources are managed by different academic and governmental actors. Up till today, an integrated health information system is lacking, which hampers the valorisation of these data sources.

AHEAD aimed to investigate the possibility of sustainable and longitudinal linkages between the these different (digital) data sources and of integrating this wealth of information they contain.

To do so, AHEAD realized the following specific activities:

- Characterizing and documenting the Belgian health information system.
- Identifying ethical, legal, and technical bottlenecks by developing
    - an online citizen consultation on the secondary use of health data
    - a business case on prospective linkages between information originating from Sciensano's Health Interview Survey and administrative data sources
- Developing recommendations for prospective initiatives on health data linkages and integration.

By realizing these activities, AHEAD aimed to provide an active push to the prospective development of a national health data platform as an accessible and sustainable monitoring and research infrastructure.

2. STATE OF THE ART AND OBJECTIVES

STATE OF THE ART

Public health challenges such as the ageing population, the emergence of chronic diseases and the identification of risk factors for the development of other disabilities call for new ways of using our health information systems. The determinants of a health across the life span (lifestyle, medical interventions and ageing) are all interrelated and result from complex interactions between a wide range of environmental exposures, phenotypic characteristics and genomic factors across the lifespan. Public health research aims to identify the conditions of a healthy development and successful ageing over the life course. Figure 1 depicts the general contributing factors and stresses the need for the identification of correlations between these factors.

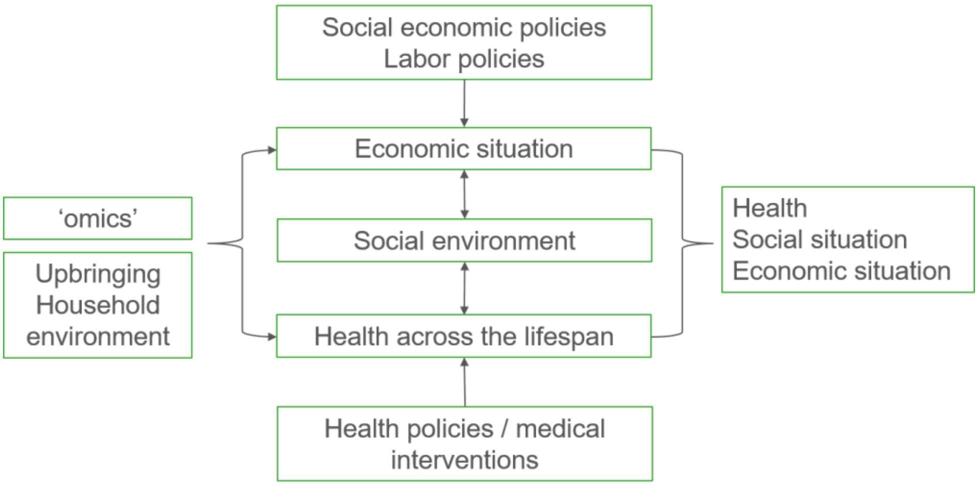

Figure 1. Factors leading to a favourable health status

To identify the correlations and determinants of the factors in Figure 1, linking data from different domains on an individual level is crucial.

In Belgium, multiple high-quality data sources exist, containing valuable information on social security, education and health care, as well as clinical and subjective health data, collected using surveys, small cohorts, registers, surveillance, etc. Examples include the Belgian participation in various European and topic-specific cohort-like initiatives, such as the Survey of Health, Ageing and Retirement in Europe, (1)), the Statistics on Income and Living Conditions (SILC), and the European Social Survey (ESS). Examples on the national level are the mortality follow-up of the population census 1991, 2001 and 2011 (Interface Demography Research Group, VUB), as well as initiatives that focus on specific subgroups of the population, such as patients (2), twins (East Flanders Twin Survey, (3)) or children (Limburgs geboortecohort, (4)). Furthermore, IMA/AIM governs a large population-based sample, the Echantillon Permanente Steekproef (EPS), which contains a randomly sampled cohort of health insured individuals in Belgium and consists of three types of data; one regarding the population, one containing a database with reimbursed health care procedures and one with the reimbursed medication. Finally, several administrative datasets, such as the mandatory health insurance data, hospital discharge data, and causes of death data, provide routine information on the health status of the Belgian population.

Most of these rich data sources in Belgium are managed by a variety of academic and governmental actors. This implies that the included information is frequently 'siloed' and that different sets of individual health data remain largely unharmonized and unconnected. This raises a number of issues.

From a research perspective, some crucial health questions are cumbersome, or even impossible, to answer. For example, labor force participation is an important policy objective and substantial resources are made available by the Federal Government, aiming to reduce poverty by increasing the participation rate. There are important barriers and resources that have an impact on labor participation, including health and disability, geographical differences, educational level and household characteristics as well as wellbeing and coping, social support, self-efficacy (e.g. Augner, 2018 (5), Hergenrather et al., 2015 (6) ; Zenger et al., 2013 (7)). To assess their prospective impact on sustained labor force participation and to decrease inequalities, data on these topics should be brought together to explore their complex interactions. Creating a structural administrative data linkage platform that simultaneously foresees solutions towards retrospective or even prospective data collection allows for unraveling the complex determinants of labor force participation. Another example entails the understanding of the barriers to seeking health care and mapping of the unmet needs for health care, which is imperative for developing rational policy to provide services that are accessible, efficient, and (cost-)effective. Addressing these barriers includes socio-demographic aspects, previous health care experiences of patients, physical accessibility, but also family environment and social support (8). Here, data obtained in administrative processes such as the social security system is vital to complement (longitudinal) health survey data and, eventually, to identify direct and indirect effects.

Finally, the exchange of and access to up-to-date health data is important for rapid response to research questions that concern current public health challenges. This has been demonstrated very clearly during the current COVID-19 crisis. In (the aftermath of) this crisis, information is needed on the social, economic, cultural and physical environment (including behavioral data, food, education, work environment, sickness leave/return to work, housing, school) in order to study the effects on health status and population health.

Currently, data linkage is mainly project-based. These projects frequently focus on specific population groups such as patients, specific occupational groups, or specific age groups. Hence, the scope of these studies is limited and many initiatives are cross-sectional, which leaves less opportunities to assess longitudinal exposure-outcome relations (9). However, longitudinal data on (the impact of) major determinants of health and disease and a sound base for targeted policy, policy follow-up and evaluation of the health and social systems require population-based cohort studies.

Moreover, the current process of project-based data linkage implies that researchers spend a lot of time on inventorying available data, requesting the desired data, and linking the data. Sometimes months or even years are spent on reaching consensus with all the data holders, adapting the datasets for researchers and building the pipelines for the data exchange. Routinely linked data or pipelines for fast data exchange, might save valuable time and resources of researchers, and might also aid in supporting individuals working in policy and administration by proving accurate numbers on incidence and prevalence, causes and prognosis with a minimum time lag. In addition, it is significantly cheaper to invest in high-quality (secondary) data analyses on existing data than to initiate new studies.

In conclusion, there is a strong need to facilitate flexible data collection and structural data linkages on many topics regarding population health, which might facilitate research and, subsequently, decision-making processes by citizens, clinicians, public health practitioners and policy makers.

At the European level, there is a strong urge to integrate health information systems, both within and between countries. To respond to the Joint Questionnaire on Non-Monetary Health Care Statistics, and in particular to the upcoming Questionnaire on Diagnosis-Based Morbidity Statistics, countries are required to link multiple data sources to provide accurate and complete estimates. However, the current situation makes it very cumbersome for Belgium to fulfill these data requirements. Furthermore, there is a demand at the European level for the establishment of national nodes, which are organizational entities, often linked to a national institution or governmental unit, that function as a national liaison and bring together relevant national stakeholders in a systematic way (Joint Action on Health Information, 2019).

For the period of 2019-2025, the development of a European Health Data Space (EHDS) is a priority for the European Commission (10). The EHDS would make health data more easily accessible and exchangeable, not only for primary use, i.e. the development of a better health care system, but also for secondary use such as health related research and policy making. The EHDS for health information will pay particular attention to (i) data management and rules for data exchange, (ii) data quality, and (iii) data infrastructures and interoperability.

Within the context of the development of a EHDS, the Joint Action Towards the European Health Data Space helps EU member states and the European Commission to develop and promote concepts for the secondary use of health data, in order to benefit public health as well as health research and innovation in Europe (11). TEHDAS will focus on, for instance, sustainability in the secondary use of health data, reliability and compatibility of health data for secondary use, and an exploration of the role of individuals in the secondary use of health data by actively engaging them in dialogue.

Other countries provide valuable examples on how (information on) health data can be centralized to support conservation and simple access to these collections. Like Belgium, Denmark has many databases and registers containing health data. The Copenhagen Healthtech cluster aims to reduce the administration connected to identification and application for access to Danish health data by providing an overview of the many Danish health data and descriptions of how to apply for access to this data (www.danishhealthdata.com). Norway has a well-developed and well-documented system for register data, *Microdata.no,* which is developed and operated by the Norwegian Centre for Research Data and Statistics Norway, and which provides researchers at approved research institutions with access to register data (www.microdata.no). Researchers can use non-anonymized data through a platform that protects confidentiality. In Finland, a *one-stop shop* for the secondary use of health and social data was established to improve the data security and maximize the use of these data: *Findata* (www.findata.fi). Findata operates in conjunction with Finland's National Institute for Health and Welfare; it permits secondary use of data and provides access to anonymized or pseudonymized data for use in a remote access system for a fixed period of time.

The absence of a unique national system to link the different components of the health and health related data collection tools at the regional and national level has been recognized by several national and international groups as a key weakness of the Belgian health information system (EU Expert Group on Health Systems Performance Assessment, 2019).

Nonetheless, different elements already seem to be in place, including the existence of a unique national register number which can be used to link multiple databases, as well as the existence of platforms for data linkage. Sciensano has already developed a technical platform for data linkage and transfer, Healthdata.be, that aims to facilitate the data exchange between healthcare professionals and researchers to increase public health knowledge in Belgium.

The current situation calls for an active push towards the development of a national health data platform. Figure 2 shows the current and possible scenarios for integration of the Belgian health information landscape. From a research perspective, a single integrated national health data platform would be the "holy grail" (right side of the graph).

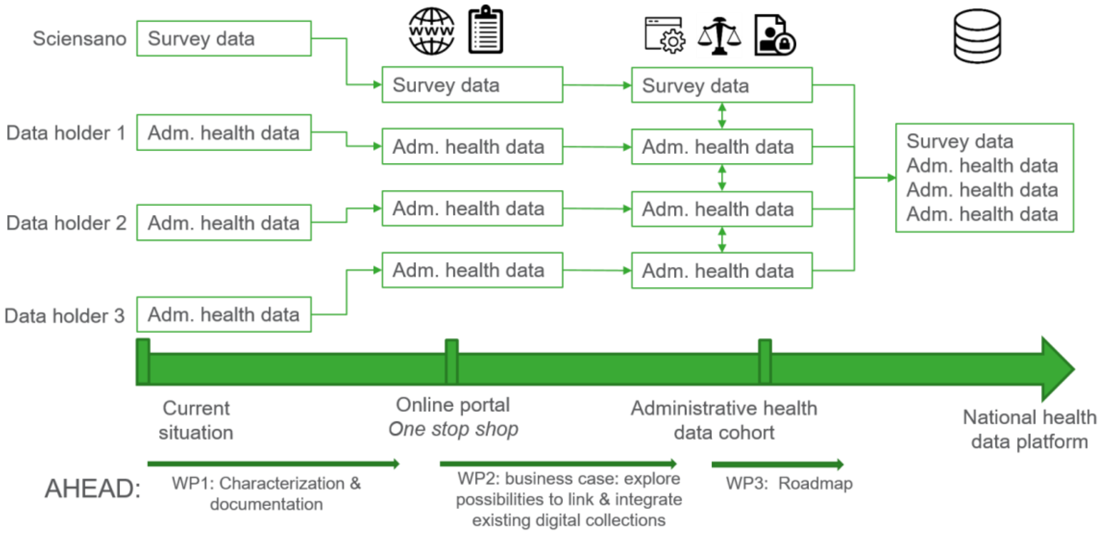

Figure 2. The AHEAD project provides an active push towards the development of a national health data platform

In line with this aim of an integrated national health data platform, the Belgian Ministry of Public Health has asked for the development of a federal Health Data Authority (HDA) in the policy note of 2 November 2020[1]. The HDA should develop and implement a policy strategy regarding health (care) data and, moreover, centralize various health databases in a GDPR-conform way. Consequently, the HDA should become the unique point of contact concerning national health (care) data. This way, the HDA could support scientific research and policy making initiatives that aim for a more qualitative and effective health care system.

On the way towards this "holy grail", there are intermediate steps to take, which by themselves can already provide valuable progress compared to the status quo.

---

[1] https://www.dekamer.be/doc/FLWB/pdf/55/1580/55K1580004.pdf#page=18

OBJECTIVES

The overall aim of the AHEAD project was to offer an active push towards the development of a national health data platform, by taking steps that will allow to move closer to the holy grail of a single integrated health data platform. It achieved this aim through the following specific objectives/work packages:

- WP1: Characterization and documentation of the Belgian health information system
- WP2: Development of a citizen consultation on the secondary use of health data and of a business case built around prospective linkages between Sciensano's Health Interview Survey (BHIS) data and administrative data sources, and this to identify of ethical, legal, and technical bottlenecks
- WP3: Elaboration of recommendations for developing a national health data platform as an accessible, long term monitoring and research infrastructure

3. METHODOLOGY

WP1 THE BELGIAN HEALTH INFORMATION LANDSCAPE

WP1 characterized and documented the Belgian health information landscape and its constituent entities.

In Belgium, multiple high-quality data sources exist, containing valuable information on clinical and subjective health, social security, health insurance data, education, health care, etc.

Many entities have developed and collected long data series on health. *The Intermutualistic Agency (IMA/AIM)* governs a large collection of health care data, as processed by the health insurance services in Belgium, and aims to support the improvement of the performance, the quality and the accessibility of the Belgian Health Care system and health/invalidity insurance. Furthermore, there is the *National Hospital Stay Database (NHSD)*, a merged database that is created by the TCT (Technische Cel – Cellule Technique) from RIZIV/INAMI, to generate a comprehensive overview on the required care, the financial needs and the reimbursed costs, for each specific pathology and to be able to list these costs for each hospital individually and compare the data between all Belgian hospitals. The NHSD is based on the Hospital Discharge Data (Minimal Hospital Data, MHD), an administrative data source governed by the Ministry of Health in which (anonymized) medical and nursing data of the (non-psychiatric) hospitals is collected to support the governance and health policies of the Ministry of Health and the hospital (networks), and the Hospital Billing Data (HBD) from the National Health Insurance companies, which contains the information regarding the billing data for hospitalized patients.

Regional initiatives involve Intego, an integrated network of general practitioners (GPs) in Flanders, managed by the Academic Centre for General Medicine from KU Leuven. Intego aims to create a large database to centralize health data in primary care. The network incorporates an automated data collection, based on the Electronic Medical Record (EMR) of the patient as registered in their GP's practice.

Other databases contain valuable information in relation to health. *Statbel*, the Belgian statistical office, collects, produces and disseminates reliable and relevant figures on the Belgian economy, society and territory. The collecting is based on administrative data sources such as population structure, movements, births, mortality and migration, as well as surveys, such as data on the Labor Market (Labor Force Survey), Education and training (Survey for the workforce) and poverty and living conditions (form the SILC-survey Statistics on Income and Living Conditions). Another example is the data warehouse of Labor and Social security (governed by the Crossroads Bank for Social Security, CBSS) that aims to link socio-economic data from Belgian social security and other public institutions. At the request of a number of social security institutions, a data warehouse was created in 1999 in which the administrative data is permanently stored. The added value of this data warehouse lies in the linkage of the data of the participating social security institutions. Information on the websites of these institutions already provides an overview of the characteristics of the databases.

However, a clear and accessible overview of these data sources is currently lacking. As a result, when researchers want to set up a new research/linkage project, a lot of time is spent on inventorying which data are available, who is responsible, and which steps should be taken to get access to (parts of) this data collection.

Therefore, WP1 aims to better integrate the Belgian health information system by providing a clear overview of the most important data holders in this field and describing how their data can be accessed and linked. Previous reports (e.g. Inventaris van databanken gezondheidszorg, KCE, 2006) had shown that making an inventory and acquiring data for research is indeed hampered by the fact that the exact characteristics of the databases cannot be obtained in a user-friendly manner. To some extent, this has been addressed by the development of the FAIR Healthdata portal (https://fair.healthdata.be/), which aims to provide an exhaustive overview of existing health databases. While this portal aims to provide a comprehensive technical summary of the databases and their metadata, there is a need for a more focused platform, targeted at researchers, and providing a clear and concise overview of the key health data sources in Belgium, their contact persons, and practical information on how linkages can be set up – which is for instance the aim of the French health data hub (https://www.health-data-hub.fr/).

As the AHEAD project aims to better integrate the Belgian health information system, an inventory will be made of the entities within the Belgian health information system and their databases. WP1 will identify the key data holders in the Belgian health information system and characterize and document their contact and access details, to support valorization of these collections. A key feature

of this inventory is the enabling for the academic community to access an overview of the Belgian health information system, as well as exact characteristics of the databases per institute or source, and links to the FAIR catalogues (12).
To accomplish the central objective of WP1, desk research needs to be combined with active outreach to the different players in the Belgian health information system. To arrive at an up-to-date inventory and documentation, including information about their potential deficiencies, and information on who the contact persons are or where information can be obtained, active stakeholder engagement is indispensable.

Besides, this stakeholder engagement inherently contributes strongly to the secondary objective of WP1, i.e. building, supporting and strengthening the network and interaction between national key players regarding health and administrative data, including both data holders and data users. Realizing a clear overview of available data sources, including information about central contact persons, how to access and obtain the information, and how to analyze and link data for secondary uses such as research and policy making and all of this in an efficient and more longitudinal way, requires mutual trust between data holders and data users and a climate of dialogue and open discussion. Hence, WP1 also aims to bring key players into contact with each other and this through online and offline events such as webinars, seminars, study days, etc.

Throughout the development of the AHEAD project, several projects and initiatives have been launched, both on a national and European level, of which the central aims strongly align with those of the AHEAD project. At the Belgian level, this includes the establishment and organisation of a Health Data Agency (HDA), of which the legislative proposal has been approved on march 9th 2023. The Belgian HDA pursues the central objectives of facilitating the availability of health (related) data, developing and implementing a policy strategy concerning health (related) data, and stimulating innovation as well as scientific and policy-supporting research. Particular tasks include, but are not limited to, providing a data catalogue, facilitating health data access requests, and communicating about the secondary use of health (related) data (13). This will happen through a clear website and dashboards that, for instance, clearly explain for which purposes and with which impact health (care) (related) data are used (14).

Considering these ambitions of the HDA, their strong affinity with the AHEAD WP1 objective of making an (online) inventory of the key entities and their databases within the Belgian health information landscape (including access details), and other parallel initiatives such as the FAIR Healthdata portal (https://fair.healthdata.be/) and the Health Information Portal (https://www.healthinformationportal.eu/health-information-portal), WP1 of the AHEAD project slightly shifted its focus away from the development of an online inventory and, instead, focused more on initiatives to strengthen the network and interactions between key stakeholders regarding the accessibility, linkage, and secondary use of health data. Hence, not only a webinar on data needs was organized yet also a study day on health information and policy to bring together data holders and data users and to mutually clarify objectives, resources, and needs and encountered bottlenecks (cf. infra, Result-section WP1).

WP2.1 CITIZEN CONSULTATION

Generally, WP2 explored the possible ethical, legal, and technical bottlenecks for an integrated health information system.

To do so, firstly, a public citizen consultation has been set up to better understand citizens' perceptions and expectations regarding the secondary use of health data (for instance for research and policy making) and their role in its governance.

Panels with an active involvement of citizens are tried and tested participatory methods in which citizens investigate and discuss a socially complex and often controversial theme. Citizens' perspectives provide important new information, both to policymakers and other stakeholders (researchers, medical practitioners, etc.), as they can be regarded experts in their own experiences and views. Online platforms want to stimulate engagement, reflection and exchange with and between citizens. The specific purpose of the citizen consultation within the AHEAD project is to ensure a bottom-up approach where citizens’ voices are considered from the start and are included in the co-construction of the ethical, legal and societal framework for the secondary use of health data.

WP2.2 BHIS+ BUSINESS CASE ON PROSPECTIVE LINKAGES

As an additional way to explore and document ethical, legal, and technical bottlenecks regarding a more integrated health data landscape, WP2 of the AHEAD project develops a business case built around prospective linkages between Sciensano's Health Interview Survey (BHIS) data and administrative data sources. This business case builds on expertise and know-how established during previous data linkage projects at Sciensano, as well as during the BELCOHORT project, which aimed to establish the feasibility a prospective population cohort as research infrastructure in Belgium.

Background and study aim

The Belgian situation lacks a conceptual data model in which objectives are set and health and health care data are organized. Based on this conceptual model the role of the responsible organizations for providing the datasets and the procedures and methods to exploit data can be drawn up. Sciensano is responsible for managing some of the major national health surveys such as the Belgian Health Interview Survey (BHIS). Within AHEAD, a business case is developed that is built around the use of the BHIS data as a starting point for the creation of an administrative health data cohort, referred to as the “BHIS+”.

The BHIS is the health survey in Belgium carried out every 4 to 5 years. Each wave comprises of about 10,000 surveyed individuals in some 6,000 households. The survey is commissioned and co-financed by the “Commission of Commissioners”, consisting of the cabinets and administrations of the Ministries and organized by Sciensano. To date, six waves have been conducted in 1997, 2001, 2004, 2008, 2013, and 2018. The purpose of the BHIS is to monitor the health status of the general population as well as health determinants including health behaviors, medical care consumption and social and demographic characteristics. The repeated cross-sectional design of the BHIS enables the assessment of health trends and provides evidence for the evaluation of health policy. The target population of the BHIS consists of all persons with residence in Belgium, including the institutionalized elderly, with no restrictions on age or nationality. The National Population Registry (NPR) is used as the sampling frame. This registry contains information on gender, age, address, citizenship, marital status, etc. of each individual. Data collection is undertaken using face-to-face interviews at the participant’s home. The interviews are supplemented with a self-administered questionnaire (for the participants aged 15+) covering more sensitive topics like mental health, use of illicit drugs and sexual behavior.

Based on the data collected within the BHIS2018, WP2 of the AHEAD project explores and documents the possibilities of routinely linking the BHIS data with (administrative) health data from other sources within the Belgian health information system. Hence, the BHIS+ business case will document the possibilities and challenges for prospectively setting up an administrative health data cohort in Belgium that could function as a research infrastructure.

Figure 3 proposes a model in which the participants of the BHIS are followed up over time by refreshing their data every year through linking their survey data to the data from the registers. Other large population based data collection in Europe, such as the UK Biobank, also conduct follow-up for health-related outcomes through linkages to routinely available national datasets such as hospital admissions, mortality, cancer registries, primary care, cancer screening data, and disease-specific registers (15). Another example is the Estonian Biobank, where the follow-up is also conducted by enrichment of phenotypic data by updating the health data from medical records available in national health databases and hospital registries (16).

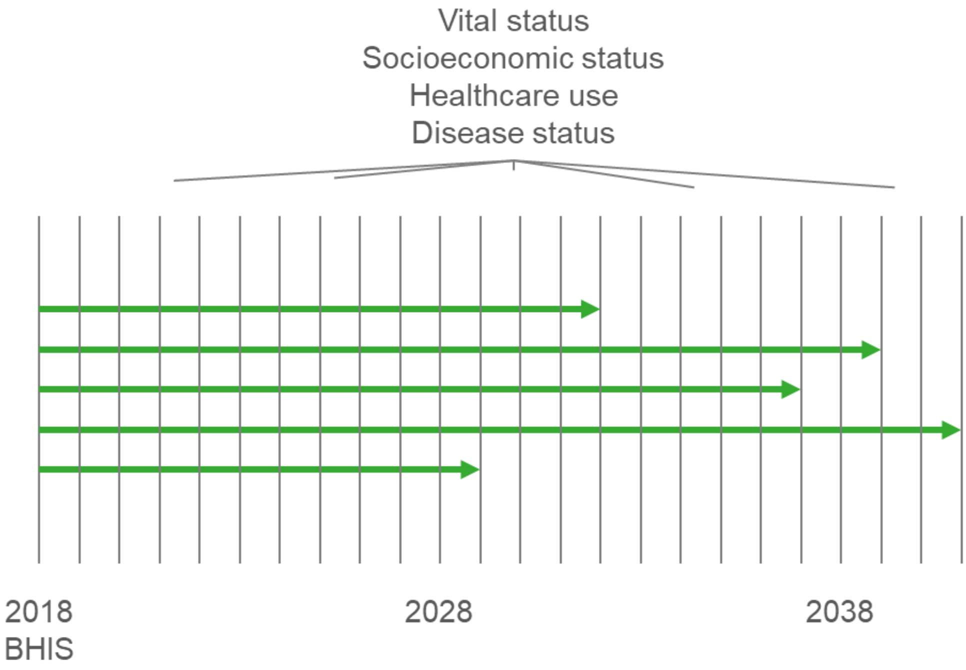

Figure 3. Schematic model of the BHIS+ administrative health data cohort. Each arrow represents a participant, from the date of survey until the date of death, while each vertical line represents a refreshment of the data of this subject.

Study type

The study design of the BHIS+ study is a qualitative prospective, observational study. Semi-structured interviews are conducted with individual health data experts from the Belgian health information landscape.

A qualitative methodology enables an in-depth exploration and analysis of the topic and the provided information and, hence, is considered most suitable for the purpose of this study, which had an exploratory goal.

Study population

Interviews are conducted with national health data experts, including ethico-legal experts, involved in the exchange, linkage, reuse, and management of health data sources within the Belgian health information landscape.

Inclusion and exclusion criteria

Purposeful sampling is used to recruit potential study participants. Health data experts, including ethico-legal experts, working in Belgian institutes are selected according to their knowledge, role and/or experience with the exchange, linkage, reuse, and management of health data. We aimed for a study population to be as representative as possible of the different types of stakeholders involved in the linkage and reuse of health data.

Based on their professional role and previous experiences with health data linkages and research infrastructures, we considered three subgroups of participants:

- Group A : Health data experts that have been involved in the setting up of previous linkages with data from the BHIS
- Group B: Health data experts that have previous experiences with (routine) data linkages and/or research infrastructures (e.g. health data experts working at Technische Cel/Cellule Technique (TCT), Intego, IMA/AIM, …)
- Group C: Health data experts working at (administrative) health data institutions or research institutions that might be involved in the BHIS+ business case (e.g. data protection officers, academics, the Belgian Health Data Authority (HDA), the Information Security Committee (ISC), the Data Protection Authority (DPA), Statbel, IMA/AIM, the Belgian Cancer Registry (BCR), the Crossroads Bank for Social Security (CBSS), the Database Minimal Hospital Data (DMHD), …)

The possibility was retained to include additional stakeholders that would be identified by other participants as especially knowledgeable or experienced stakeholders.

Sample size

Maximum 30 health data experts would be included in this study.

Recruitment

Sciensano has good relations with other stakeholders and administrative health data sources within the Belgian health information landscape, hence we were able to organize the recruitment of suitable participants mostly directly.

Potential participants were contacted directly through professional contact details or suitable institutions were contacted through general contact information and asked for their willingness and availability to participate. In case of a positive answer, the contact details were asked of those persons that might be most suitable to represent the institution. In case of large institutions, we could contact more than one representative to cover as many practices, experiences and perspectives as possible.
Additional participants could be recruited by snowball sampling.

No standardized recruitment letters have been used. Instead, potential participants received a personalized recruitment letter (email), tailored to their professional role and/or experience with (routine) health data linkages and research infrastructures.

Interview topics

A semi-structured interview guide has been used for all interviews. This interview guide aimed to specifically consider the following topics:

Ethical bottlenecks

When developing systemic health data collections, linkages and analyses, the fundamental ethical principle of autonomy demands that the wishes of citizens are respected by asking their opinion or willingness to get involved or not – as stipulated by current legal frameworks. However, epidemiological research might use citizens' data without their consent and, hence, result in a potential loss of autonomy. Nevertheless, this might as well be leveled by an ethically sound data management plan and quality assurance.

The balance between fundamental ethical values, such as respect for autonomy and confidentiality, and the benefits that big data and linked data collections can bring to public health research, is carefully researched.

Legal bottlenecks

The current legal framework allows for project-based data linkage between different data sources. Potential possibilities and barriers regarding the linkage of health and administrative data on a more structural basis within the limits of the current legal framework are identified. Relevant issues might include a longitudinal commitment of data holders, intellectual property issues, clear rules for engagement for partnerships and solid procedures to draw samples from this data and share the data, all according to the FAIR (Findable, Accessible, Interoperable, Reusable) data principles.

Belgium already hosts some initiatives that link data on a more structural basis, including the Technische Cel/ Cellule Technique (TCT) initiative and the Permanent Sample (Echantillon Permanente Steekproef, EPS) facilitated by the Intermutualistic Agency IMA/AIM. The legal frameworks for these initiatives are also explored.

Technical bottlenecks

Both Sciensano and other national data sources have built up a large expertise on managing large scale health data and have a longstanding history in (the technicalities of) data linkage, albeit on ad hoc basis. However, to scale these experiences up into a more structural, longitudinal cohort, several technical questions remain unanswered and more expertise might be needed. To facilitate the prospective development of an appropriate data management strategy, we thoroughly identify potential technical bottlenecks.

In order to map technical bottlenecks, some previous linkages of Sciensano's health data collection with other entities (e.g. the HISLink projects, where BHIS data are linked to IMA/AIM healthcare use data) are explored and experienced difficulties as well as proposed solutions are determined.
Furthermore, Sciensano has developed a technical platform for data linkage and transfer, Healthdata.be, that aims to facilitate the data exchange between healthcare professionals and researchers to increase public health knowledge in Belgium. Healthdata.be's expertise and know-how are indispensable in order to map the difficulties in standardizing and homogenizing the preservation of health data for research purposes.

Ethics approval

The qualitative interview study In the context of the BHIS+ business case has been granted ethical approval by the Commission of Medical Ethics at Ghent University Hospital. All participants signed an informed consent form including a statement on the pseudonymised publication of study results.

WP3 RECOMMENDATIONS

Based on the results of WP1 and WP2, recommendations for developing a national health data platform as an accessible, long term monitoring and research infrastructure are developed.

4. SCIENTIFIC RESULTS AND RECOMMENDATIONS

WP1 THE BELGIAN HEALTH INFORMATION LANDSCAPE

Interviews with main stakeholders

As a preparatory work to fine-tune our insight in the Belgian health information system and to reach out to national key players (both data holders and data users), we attended as an observer at more than 15 interviews that were organized in the context of the TEHDAS project. TEHDAS helps EU member states and the European Commission to develop and promote concepts for the secondary use of health data to benefit public health and health research and innovation in Europe. More information can be found on the TEHDAS project website.
Throughout the TEHDAS interviews, key stakeholders regarding national health and administrative data were interviewed about the Belgian health information landscape, more specifically about topics including data sources, data quality, data infrastructure, data governance, need for (human, technical, financial) resources, and capacity building.
The Belgian factsheet on its health information system can be found through this link.

Webinar "Data needs during the COVID-19 crisis"

In collaboration with the HELICON project, we organised a webinar on data needs in times of COVID-19. (HELICON is a Belspo BRAIN-be project (2021-2025) that aims to unravel the social inequalities and the long-term and indirect health effects of the COVID-19 crisis in Belgium. More information can be found on the HELICON website.)

The joint webinar took place on April 27th, 2022.

The COVID-19 pandemic has created unprecedented challenges for the scientific community. Belgian researchers have increasingly gained insight in procedures and tools needed to navigate this crisis and in this process, data has been vital to understand, manage and prevent the spread of the virus and its pervasive consequences. With creative approaches and unique collaborations, several data problems have been solved successfully. Other data needs, however, have remained unfulfilled.
The webinar focused on potential obstacles and gaps in the available research infrastructure but it also provided examples of innovative data use and research opportunities during the COVID-19 crisis. Also the challenging balancing act between data accessibility and data security was discussed. This knowledge may be instrumental for the COVID-19 research, as well as for managing potential future health crises.

The program of the webinar looked as follows:

Prof. Dr. Niel Hens *University Hasselt and University Antwerp*	Pandemic preparedness: the need for better data from a modelling perspective
Dr. Nienke Schutte *Sciensano-PHIRI*	European perspective on COVID-19 data and comparability
Dr. Laura Van den Borre *VUB & Sciensano*	Data protection and research in (post-)COVID times
Patrick Lusyne *Statbel*	Closing remarks

The recording and presentations of the webinar are available through this link.

Study day "Health information and policy"

On October 6th 2022, we organized a study day in Brussels on health information an policy making in Belgium. The study day brought together key players of the Belgian health information landscape. Both in keynote lectures and in a panel discussion, common developments and challenges were discussed in depth. Additionally, this day provided ample time for networking.

The program of the study day looked as follows:

Brecht Devleesschauwer *Head of Health information service, Sciensano*	Welcome
Stefaan Demarest *Project Leader Health Interview Survey, Sciensano*	The Belgian Health Interview Survey: past, present and future
Vanessa Vanrillaer *Federal matters Expert, Mutualités Libres/Onafhankelijke Ziekenfondsen*	Valorisation of health information to support policy making
Hans Constandt *Program Manager, Belgian Health Data Authority*	Towards the establishment of the Belgian Health Data Authority
Network Lunch	
Tina Van Havere *Advisor, Cabinet of the Minister of Health* Caroline Costongs *Director, EuroHealthNet*	Perspectives on knowledge translation
Panel discussion: from health information to policy	
Closure	

About 80 people, including researchers, healthcare workers, and policy makers, attended the study day. Afterwards, the presentations of the study day were sent by e-mail to the attendees.

After positive feedback, it has been decided to turn the Health information and policy study day into an annual initiative, initiated by the AHEAD project and continued by the Health information service of Sciensano. Currently, preparations for the second editions in November 2023 are fully ongoing (cf. infra).

Scientific publications

- Schutte, N., Saelaert, M., Bogaert, P. et al. Opportunities for a population-based cohort in Belgium. Arch Public Health **80**, 188 (2022). https://doi.org/10.1186/s13690-022-00949-5

Another output of WP1 considers a scientific paper on the development of a Belgian population-based cohort. This paper specifically focusses on the challenge of a better integration of the Belgian health information system, also from a longitudinal perspective, and the opportunities of integrating (some of) these data sources into a sustainable population-based cohort. Population-based cohort studies are key in providing longitudinal data on the impact of major determinants of health, disease and disabilities and they are a powerful design to capture exposure-outcome relations and explore causality (17). They provide a sound base for targeted policy, policy follow-up and evaluation of the health and social systems.

When reviewing the existing Belgian cohort initiatives focusing on health and/or behaviour, such as the Echantillon Permanente Steekproef (a large population-based sample governed by the Intermutualistic Agency), the Belgian Health Interview Survey and the Food Consumption Survey (both managed by Sciensano) or more specific initiatives such as the HIV cohort (18) or East Flanders Twin Survey (3), some limitations may be identified. Some of these projects focus on specific population groups, occupational groups, or age groups. Accordingly, the scope of these studies is limited, based on specific research questions with a specific exposure and a limited number of outcome measures. Secondly, many initiatives are cross-sectional and, hence, they leave less opportunities to assess longitudinal exposure-outcome relations. Furthermore, some studies miss a systematic link between their research outputs and health. However, accessing and exchanging timely health data is important for rapid response to current public health challenges. Belgium has a rich collection of administrative data sources, yet they are mainly 'siloed' and the current linkage processes are project-based, complex, while they also lack transparency and require a lot of time and resources. The Belgian health information system needs a more flexible data collection tool to facilitate data collection and structural data exchange on many topics regarding population health and social systems, facilitating research and, subsequently, decision-making processes by citizens, clinicians, public health practitioners and policy makers.

In the context of the Belcohort project (2018-2020), stakeholder meetings and an international workshop have been organized to explore the opportunities for establishing a population-based cohort. Emerging from these meetings and workshop, some key recommendations have been formulated that can be found below (Result-section, under "WP3 Recommendations").

Side results

- At the Department of Epidemiology and public health of Sciensano, we established a Health Information Taskforce. The Taskforce includes colleagues from different units, including the Health indicators unit, EU Health information system service, and Health surveys unit, as well as colleagues from the Cancer centre. Monthly meetings are organised to update colleagues on ongoing or planned projects and initiatives concerning health data reuse, data linkage, etc. The Taskforce is intended to serve as a long-term initiative supporting synergies between projects involving the (re)use and implementation of health information.

- In the 2020-2026 management plan of Sciensano, the aim of achieving a better Knowledge Translation (KT) has explicitly been integrated. The central objective of KT is to strengthen the uptake of scientific outputs in practice and to strive for evidence-based policy making. This is a transversal task for which all directorates at Sciensano must work together: scientists producing the evidence, COM (e.g. by ensuring a targeted and tailored dissemination of results), P&O (e.g. by providing trainings). To help integrating the concept of KT into Sciensano's working practices, a KT working group has been established. The working group will define, promote and implement the KT vision of the institute. The KT group also includes a group of KT ambassadors of all different scientific directorates. They represent the interests of their scientific directorate in the implementation of KT, pilot KT training modules, co-design future KT activities at the institute, and encourage and inspire colleagues to participate in KT activities organized within Sciensano. Marlies Saelaert is a KT ambassador since the establishment of the KT working group.

Follow-up activities

- We plan on publishing a scientific paper on the main (administrative) health data sources in Belgium, including access procedures. We are currently drafting the paper.

- We are currently set to organize the next edition of the Health information and policy study day. The study day will take place Thursday November 23rd 2023 in the Royal Library of Brussels.
  In general, the study day will address various topics regarding the production, analysis and valorisation of health data and health information. More specifically, this edition will focus on the ways health data is collected and used at the local level. Before noon, several interactive workshops will be organised on the production, analysis and availability of health data with relevance for the local level. In the afternoon, various presentations on the use of health information and its implementation in local policies will be scheduled.

  The program looks as follows:
    - 09h30 – 10h00: Registration + coffee

    - 10h00 – 10h15: Welcome + General intro (Brecht Devleesschauwer, Michael Creek)
    - 10h15 – 10h40: Workshop 1 - The Local Health Interview Survey (Lize Hermans, Sciensano)
    - 10h45 – 11h10: Workshop 2 - Food environments and their impact on population health (Stefanie Vandevijvere, Sciensano)
    - 11h15 – 11h40: Workshop 3 - The IMA Atlas (Andreas De Bleser, Intermutualistic Agency)
    - 11h45 – 12h00: Plenary feedback from the workshops

    - 12h00 – 13h00: Lunch

    - 13h00 – 13h30: Presentation by Elke Van den Brandt (Brussels Minister of Mobility, Public Works and Road Safety)
    - 13h30 – 13h45: Presentation population health management dashboards - FOD Volksgezondheid / SPF Santé Publique – Ann Marie Borg
    - 13h45 – 14h05: Christian Massot (Hainaut Health Observatory) in interaction with Andreas De Bleser (Intermutualistic Agency)
    - 14h05 – 14h25: Liesbet Meyvis (City of Antwerp's health department) in interaction with Lize Hermans (Sciensano)
    - 14h25 – 15h10: Panel discussion
    - 15h10 – 15h25: Wrap-up and closing (Brecht Devleesschauwer, Michael Creek)

    - 15h30 – 16h30: Reception

- Together with academics, we will further explore the possibility of organising a course on the Belgian health information landscape. In this course, topics such as major data sources, data access, data linkage, FAIR principles, legal grounds, ethical reflections, citizen engagement, … could be covered.

- In line with the further establishment and implementation of the Belgian Health Data Agency and their planned activities, we will consider the added value of further developing the website www.healthinformation.be, as a centralized online portal on key data holders in the Belgian health information portal.

WP2.1 CITIZEN CONSULTATION

Main results

The online public consultation called Healthy Data wanted to stimulate engagement, reflection and exchange with and between citizens on the secondary use of health data. Citizens' contributions aimed to feed future recommendations concerning the awareness and involvement of citizens in the reuse of data. The e-consultation also contributed to the identification of ethical, legal, and technical bottlenecks in health data reuse and in the development of a more integrated health information system in Belgium.

On the 13th of December 2021, the citizen e-consultation Healthy Data was launched. The consultation was open for citizens' submissions until the 8th of May 2022. The Healthy Data platform was available in English, French and Dutch and was open to all European citizens.

Information materials (pedagogical tools, an informative website with real life cases and current issues, an interactive comic book, …) were developed in layman's terms to allow citizens to get familiar with the concept of health data and the current Belgian health information system. Subsequently, citizens were encouraged to share their values, opinions and questions and formulate advice on the secondary use of health data and ethical issues of advanced and longitudinal data linkages. Therefore, the platform consisted of an interactive test and open discussion boards.

An international communication strategy has been set up to broadly disseminate the initiative and circulate invitations to participate in the online consultation. The e-consultation has been disseminated among policy makers, patient organisations, youth organisations, schools, academics and young researchers, the general public, health associations, … An educational toolkit for teachers at secondary schools is available on the platform and a monthly newsletters was delivered to subscribers.

Between December 2021 and May 2022, the Healthy Data platform attracted 25.000 visitors and collected almost 6,000 contributions on the purposes, conditions and citizen involvement regarding the secondary use of health data.

Once the consultation closed, a thematic analysis of the contributions was conducted to produce an interim report. The aim of this report was to provide an overview of key patterns and trends as well as a descriptive qualitative analysis of the contributors' opinions. This information was used as a starting point for discussion with national and European stakeholders. The interim report was published in June 2022 and can be consulted here.

Following the end of the consultation, three virtual workshops were organised in September 2022 and attended by national stakeholders from the three pilot countries. These national workshops were followed by a hybrid European workshop held in Brussels on the 11th of October 2022. 150 stakeholders participated in the workshops. During the national workshops, the results of the public consultation were presented and feedback was gathered from their national contexts. At the European workshop, the results of the interim report and the outline of the citizen recommendations were presented. The European stakeholders were then asked to discuss this outline and provide feedback. The overall aim of the four workshops was to seek stakeholders' input on how best to translate citizens' views into citizen recommendations concerning the awareness and involvement of citizens in the secondary use of data.

The resulting recommendations are embedded in three major themes that had been identified throughout the analysis of citizens' contributions: the data relationship and the power balance.

- Firstly, citizens feel that anyone who is using their data, is using a piece of them, of their identity and history, and, hence, is engaged in a relationship with them. Citizens think they need to be respected as a partner in this relationship. Specific recommendations that result from this theme include the ability of accessing understandable information about the secondary use of health data, the possibility of having access to their own data and of being informed about the ways their data are used for secondary purposes, the inclusion of their values when concepts such as individual benefits and the common good are being defined and delineated, the need for a plurality of views and actors in decision-making processes, the possibility of continuous engagement in the life-cycle of their data.
- Secondly, citizens consider health data as very powerful, both for the good and for the bad. Therefore, they want a good balance between risk mitigation and benefit maximisation to ensure a proportionate secondary use of health data. Specific recommendations that result from this theme include the possibility of exercising control over secondary use, the protection of people's identity, the importance of transparent intentions of data users and of aligning these intentions with purposes citizens support, the need for strong accountability mechanisms, and the need for fostering both a legal framework and IT solutions for data protection.
- Finally, citizens call for the development of a regulatory framework for secondary use that does not only include appropriate governance structures but also respect for central ethical values as they are interpreted by citizens.

The final 12 recommendations on how citizens want their health data to be used for secondary purposes and on their desired role in the governance framework, are discussed in the result-section under "WP3 Recommendations" and in the dissemination and valorisation-section under "Support to decision making".
An in-depth discussion on the recommendation can be consulted in the final report of the Healthy Data e-consultation that was published in March 2023.

Healthy Data was a joint initiative of the AHEAD project and the EU Joint Action TEHDAS ("Towards a European Health Data Space"). The e-consultation was organized by Sciensano (Belgium), the Health Data Hub (France) and the NHS Confederation (United Kingdom).

Scientific publications

- Saelaert, M., Mathieu, L., Van Hoof, W., & Devleesschauwer, B. (2023). Expanding citizen engagement in the secondary use of health data: an opportunity for national health data access bodies to realise the intentions of the European health data space. *Archives of Public Health*, *81*(1), 168. https://doi.org/10.1186/s13690-023-01182-4

Side results

Throughout the AHEAD project, we got involved as a partner in the BECODIGITAL project.

The overall objective of the BECODIGITAL project is to develop a practical and scientifically grounded roadmap to support citizen co-creation through digital technologies while ensuring that this co-creation is performed in an inclusive manner, and to validate it in a federal government context. This roadmap and associated toolkit will be developed following design science research's best practices. The BECODIGITAL project focuses on three key innovative aspects related to co-creation: (1) stakeholders' pre-conditions and engagement issues related to co-creation, (2) citizen co-creation via rapidly evolving digital technologies providing new ways to engage citizens and (3) the evaluation of the outcomes of digital co-creation.

Sciensano is involved in the project as a partner providing use-cases regarding citizen co-creation.

Follow-up activities

- Further dissemination of the final report and recommendations
- Further in-depth analysis of citizens' contributions and publication of a scientific paper on societal and ethical issues concerning health data reuse

WP2.2 BHIS+ BUSINESS CASE ON PROSPECTIVE LINKAGES

Data analysis

The BHIS+ business case on the possibilities and challenges regarding the development of prospective linkages and a research infrastructure was designed as a qualitative prospective, observational study with semi-structured interviews (cf. *Infra*). Between September 2022 and March 2023, 26 semi-structured interviews have been conducted. Participants included experts on national health or administrative data and/or on (structural) data linkages within the Belgian health information landscape, as well as academics and legal, technical and ethical experts.

The interviews lasted between 45 minutes and 110 minutes. All interviews have been recorded, transcribed and thematically analysed with an inductive approach.

As described by Braun and Clarke (19), the analysis consisted of data immersion, code generation, theme identification, theme revision, theme definitions and production of the final report.

The software program NVivo20 was used to support data analysis.

Main results

After a first coding process of all interviews, resulting in 591 codes, the analysis of interviews led to the identification of themes concerning: Experiences with (routine) data linkages; The purpose of linkages and research infrastructures; Legal grounds for data reuse; The legal framework for data reuse; Data governance and management and infrastructures; Professional collaboration.

Experiences with (routine) data linkages

Unfortunately, health (related) data were often described as performing poorly on FAIRness (Findability, Accessibility, Interoperability, Reusability). The difficulty of identifying where specific data can be found as well as the overall unavailability of some data were mentioned often. Data were said to lack accessibility, since access procedures are complex and opaque, and waiting times for data request are long. Regarding data's interoperability, participants mentioned how Belgian health data are not sufficiently structured for secondary use. Their lack of structure and standardization makes them difficult to exploit. As a consequence, many participants called for a "FAIRification" of health and administrative data.

Regarding data linkages between health and administrative data, several participants confirmed that, currently, the Belgian health information landscape is characterized by a system of ad hoc linkages. Linkages are always set up in the context of a specific study with a delineated purpose. Hence, setting up linkages on record-level is always a long and challenging process and requires going through the entire application procedure every time again.

Because of the lack of FAIR data and the difficulties regarding linkages, many participants expressed the need for a change in the Belgian health information landscape. As a positive touch, several participants perceived a general benevolence among administrations and data providers and they believed that a change in the Belgian health data system can actually be realised by "a coalition of the willing".

The purpose of linkages and research infrastructures

Many participants responded positively to the added value of an administrative health data cohort that would link several datasets and would be accessible as a research infrastructure. It would take away many of the experienced problems mentioned above, especially the burden of setting up the same (similar) linkages over and over again. The most commonly cited incentive for integrating one's data into a research infrastructure concerned the more intensive use and therefore upgrading of one's data. Data providers hold an enormous amount of data which, however, they frequently cannot exhaustively exploit themselves. Many of them would like their data to be used more actively, so the data's value increases and gets more materialized.

Several interview participants had some experiences with broad data linkages. One example was the HISLink (in which HIS data are linked with IMA data). This linkage has been set up to answer all kind of questions that might be suggested by the RIZIV/INAMI. This way, the requested set of data was considered proportional to the objective, and very broad linkages, including a lot of data, could be set up. This linked dataset can be used to answer a variety of questions without requiring additional authorisations.

However, setting up this linkage requires permission every time, which causes it to remain an ad hoc linkage. Moreover, while the HISLink is aligned with a broad a range of objectives, it cannot be considered a real research infrastructure, since the linked dataset is only accessible for researchers of Sciensano.

However, many participants stated that we are currently not used to a data governance model in which large data sets are broadly and structurally linked without predefined research question. Hence, there were many questions regarding the added value and efficiency of setting up a research infrastructure. These questions often reflected the underlying concern that is would not be justifiable to set up a cohort/infrastructure for interesting questions that might potentially emerge in the future. Cohorts should always align with the objective one wants to realise and need to serve a specific purpose. Moreover, not everyone's purposes are the same and, hence, not all research questions can be answered by the same cohort. Specific research questions will still require ad hoc linkages.

These reflections seemed to be grounded on previous experiences with linkages that always had been set up in the context of a specific purpose/research question.

The specificity of current data linkages was linked to issues concerning data protection and confidentiality. To avoid the re-identification of data subjects, the principles of data minimization and proportionality need to be respected, which means that the need for certain data and variables needs to be thoroughly justified. Moreover, this specificity was linked to the concept and practice of an informed consent. If a research project and consent form are not sufficiently specific, this consent might be considered not truly informed.

Aligning a research infrastructure with a more specific (research) purpose, was also considered necessary to get approval for the development of this infrastructure, for instance by the Information Security Committee (ISC). It was doubted whether a general linkage between large data sets would be approved of by the ISC. A suggested mid-term solution was to not leave research questions completely open and unknown, yet to define them as generic and long-term oriented as possible. Researchers could try "stretching" the objectives they want to realize with a particular linkage.

Legal grounds for data reuse

The specific purpose of current ad hoc linkages is strongly related to the required legal ground for data use and reuse. For current data linkages, participants often referred to institution-specific royal decrees, missions or mandates. A lot of involved stakeholders have a specific mission and mandate that provides them with the legal ground for collecting, storing, (re)using, and linking data. However, this mandate might be rather specific and delineated. It includes clear operational purposes and modalities of data (re)use, as well as a description of the data that are necessary and proportional to realize these objectives. Hence, it might impede data reuse and linkage for broad and unspecified purposes.

Some participants suggested that, in the context of developing a research infrastructure, it might be useful to try to enlarge the purposes that are included in an institution's legal mandate and to include the (structural) set-up of linkages more explicitly.

Besides institution-specific mandates, also more general grounds, such as public interest or informed consent, were referred to as legal grounds for data reuse and linkage.

Public interest might be the most commonly invoked legal ground for secondary use, yet it lacks conceptual clarity. There are neither official criteria for it (e.g. when is a research project beneficial for different populations or communities?), nor a general consensus about its meaning (e.g. can the public interest be served by commercial parties and can a certain degree of reciprocity be imposed on commercial parties for having access to public health data?).

Regarding informed consent, interview participants mentioned that there often is the problem that at the time of data collection no consent has been obtained for further research or linkages, since, at that time, it was not possible to predict scientifically valuable applications of these data in the future. Also for the development of research infrastructures, informed consent was rarely considered a suitable legal basis. It was often perceived as complicating scientific research, for instance during recruitment procedures or when a participant later wishes to withdraw from the research population. There were concerns among participants that working with an informed consent holds the risk of biased opt-in or opt-out choices by certain subpopulations and of research samples that are too small to be representative. Moreover, it was discussed whether informed consent actually realizes its central objective. Informed consent forms are often very complex and difficult to understand for participants. Hence, they might sign a form without actually understanding its consequences. It should also be avoided that informed consent procedures turn into constant and endless requests that overburden citizens and, eventually, leave them indifferent. This holds the risk of achieving the opposite of a key objective of informed consent, namely giving citizens some control over the use of their personal data.

Finally, some participants questioned whether informed consent as a legal ground for data reuse and linkage is actually in line with other potential grounds and values. For instance, opting out for data reuse could be considered as contradicting public interest. Citizens should realize that many of their data are collected anyway, for instance for purposes of reimbursement, and that reusing these data in a context of public health (for instance for a better planning of hospital capacity) might easily expand their value to other citizens without any risk of harm. For these types of reuse, informed consent could be considered unnecessary. Some participants also considered reuse in this context as a way of reciprocity for benefiting from public health care and health insurance.

Legal framework for data reuse

In light of the current practice of ad hoc linkages for specific purposes, most participants believed that structural linkages for broad purposes and longitudinal cohorts require an adjusted legal framework. They considered the current legal framework as no longer adapted to fast moving developments in data linkage procedures. Regulations like the GDPR and Belgian health laws are no longer keeping pace with current linkage practices and structures. When developing the GDPR, the current scope of secondary use of health data could not be foreseen and hence, it offers too little guidance for current approaches in scientific research. This implies that, especially when interpreted quite strictly, these regulations and legislations can easily block linkage procedures or, as suggested by some participants,

An adjusted legal framework should allow for broad linkages and the sharing and reuse of large datasets. It also should flexible, since a framework that is too detailed and specific will quickly collide with future technologies or scientific challenges and, hence, become unsuitable for setting up a longitudinal cohort or research infrastructure.

On a Belgian level, several participants referred to the development of the Health Data Agency (HDA) as a potentially facilitating element in the adjustment and/or development of a suitable legal framework. In March 2023, the legislative proposal concerning the establishment of a federal HDA has been approved. This HDA aims to facilitate the availability of health (related) data, develop and implement a policy strategy concerning health (related) data, and stimulate innovation and research (13). In these ambitions, the HDA would like to focus on five central stakeholders, namely researchers, the industry, healthcare professionals and providers, policymakers, and the general population. In our qualitative study, interviewed participants expressed expectations towards the HDA such as a clarification of the current health data landscape in Belgium, the establishment of a central point of contact regarding health data reuse, a facilitation and standardization of health data reuse, and the establishment of a legal and official governance framework for reuse.

For some, the risk of a rigid legal infrastructure that, over time, would rather impede than facilitate data linkages, was a reason to not focus on legal adjustments but more on professional collaboration and trust while using the opportunities that are already available within existing legislation. This idea was linked to the belief that existing legal frameworks, including for instance the GDPR, do not fundamentally impede or block scientific research. Several participants explained how the GDPR has set the framework for the secondary and scientific use of data, as long as one respects the conditions under which it should happen. The main aim of the GDPR is to respect the data subjects' rights, which results in principles such as privacy by design, proportionality, and transparency towards data subjects. As long as these principles are respected, the GDPR will not block health data reuse for purposes such as scientific research, so it was stated by several participants.

Data governance and management

Both within or beyond the limits of the current legal framework, several participants emphasized the need for a formal and official governance structure regarding the secondary use of health data and the development of structural linkages and cohorts. A solid governance structure, as well as information and transparency, were considered the basis and precondition for more fluent data sharing and linkage procedures with which all stakeholders must concur.

Also, a solid governance structure might be more important than complex legal discussions on, for instance, the ownership of data (linkages) or consent procedures for data subjects. A governance model that guarantees a correct reuse of healthdata and that is trusted by citizens might be more valuable and feasible than consent procedures that overburden data subjects.

Many participants indicated that before a suitable governance structure for a research infrastructure could be built, some key questions and concerns should first be answered.

A first question concerns whether different research infrastructures could exist along each other or whether we should evolve towards one overarching infrastructure. Over the next years, various research infrastructures will probably be developed. They might have too much overlap and also strongly differ in efficiency and performance, which might affect policy decisions about which platforms to support and maintain. Hence, one overarching infrastructure might be more efficient, economical and sustainable. On the other hand, Belgium's political structure and the dispersion of data over different governance levels might profoundly challenge the development of one overarching infrastructure.

A second concern regarded the services that could still be offered in the context of a research infrastructure. Many study participants emphasized how (re)using and linking data in a correct way, often requires support and guidance by the data holders who are very familiar with the specificities, interpretation, possibilities and analysis of the data. Setting up a central research infrastructure holds the risks of losing professional expertise regarding the data and the incorrect reuse of these data. Providing adequate data trainings, rich data documentation sets and smooth communication lines between all players could partially help solve this issue, but might not be sufficient.

Another governance issue that was repeatedly mentioned regarded the question who should host the intended research infrastructure and be responsible for it. Several participants suggested that institutions might be eager to do so, for instance because of practical reasons such as being familiar with one’s own infrastructure and tools for analysis. However, also the idea of “data is power” and the associated reluctance to externalise one’s data, might underlie this desire to host a research infrastructure. Several participants mentioned, however, that hosting (structural) data linkages and a research infrastructure implies a great responsibility. Financial and technical resources are needed to enable both the safe hosting of data, including, for instance, solid access procedures, as well as suitable tools for analysis. Both requirements were considered related, since a lack of performant tools for analysis can result in attempts to copy or download data and analyse them in insecure environments.

Regarding data security, there was a general consensus that in reuse practices, data should be pseudonymized or anonymized as much as possible. The more data are linked, the higher the risk of re-identification and, hence, the higher the requirements for privacy and data protection. However, this clear stance is complicated in practice by the equally necessary balance between data protection and data usability. Masking data to the fullest can decrease their research value. Another complicating factor is that the required degree of data protection depends on the sensitivity of data. Health (related) data are usually considered very sensitive (much more than, for instance, certain categories of socio-economic data) and therefore, setting up structural linkages and research infrastructures that include health (related) data requires a broad range of data protection mechanisms, as well as a high level of trust, both among professional stakeholders and citizens.

Many participants indicated that, over and above the importance of technical safeguards for data protection, such as small cell risk analyses or respecting the principle of data proportionality, researchers’ and other professional stakeholders’ attitude towards the reuse of health (related) data is even more important. Eventually, sensitizing data users about the value and sensitivity of health data and developing a deontological framework and responsible attitude among data experts will outweigh the efficiency of technical safeguards that will always remain fragile.

Professional collaboration

In addition to the shortcomings regarding the FAIRness of health (related) data and the legal and governmental framework for their reuse, several participants wondered whether there really is sufficient motivation and willpower among all stakeholders to truly take the leap forward and profoundly change the Belgian health data ecosystem. At some levels of decision-making, there seems to be a lack of awareness regarding the true need and urgency of this issue, which results in postponing actual restructuring. All stakeholders should be convinced by the added value of more open data sharing and linkage procedures and they should trust that contributing to this infrastructure is a safe and efficient investment. However, and unfortunately, this element of trust might not be fully realized between all involved stakeholders and decision levels, a problem that might be enhanced by the complex political structure of Belgium. A powerful element in this potential lack of trust, is the persistent idea of “data is power”. This idea can partially undermine the willingness of data holders to contribute to structural linkages or research infrastructures, since they might fear losing their power or position in the health data landscape. This way, conceptions and feelings of (individual) data ownership might hinder (collective) scientific and societal progress.

However, not everyone agreed on this lack of trust between stakeholders. Several participants emphasized their good relationships and cooperation with other players in the field. Even though there is a lack of formal governance regarding data reuse and even though setting up linkages might require a lot of time, the health data landscape is rather well-organised and efficient on an informal level. Different institutions are supportive towards each other and have collaborated efficiently for many years.

Whether inter-institution trust has already been realised or requires some further refinement, most of the interviewees agreed and explicitly mentioned that this trust is required for the development of structural linkages and a research infrastructure. All stakeholders should trust the health data ecosystem, the accuracy and value of the research that is conducted with it, and the purposes that are realised by it. This trust should be realised step-by-step, by successful collaborations and by leaving behind ideas of data ownership and competition. When a research infrastructure would be established, this should be done from a collaborative point of view, with win-win situations and distinct added value for all parties involved.

Along with this establishment of trust, a better communication and collaboration between all players in the health data landscape could be realised. Currently, stakeholders were said to not always be sufficiently aware of each other’s available datasets, procedures for data sharing, realised data linkages, etc. More structural, open and efficient meeting opportunities between all stakeholders were suggested as a potential solution. By installing a consortium of data holders, data users, authorising bodies, etc., institutions could support each other in the identification of current data flows, feasible data linkages, request procedures, common and valuable research objectives, and opportunities for collaboration. Making this consortium multidisciplinary will assure that different perspectives and concerns, such as data security, data quality, technical issues, legal structures, etc., will simultaneously be considered. Even though developing and sustaining such a consortium will take time, it was considered to be beneficial over time through mutual win-wins.

Follow-up activities

- Publication of a scientific paper on the main facilitators and bottlenecks concerning structural data linkages and the development of a longitudinal cohort that is available as a research infrastructure.

WP3 RECOMMENDATIONS

Based on the results of WP1 and WP2, we provide 10 recommendations on the possibility of sustainable and longitudinal linkages with health (related) data and on the prospective development of a national health data platform as an accessible and sustainable research infrastructure.

The recommendations below are based on:

- Schutte, N., Saelaert, M., Bogaert, P. et al. Opportunities for a population-based cohort in Belgium. Arch Public Health **80**, 188 (2022). https://doi.org/10.1186/s13690-022-00949-5
- TEHDAS. Qualitative study to assess citizens' perception of sharing health data for secondary use and recommendations on how to engage citizens in the EHDS. 2023 Mar. Available from: https://tehdas.eu/results/tehdas-consultation-citizens-support-the-secondary-use-of-health-data-when-it-matches-their-ethical-values/
- The thematic analysis of the BHIS+ Business case

Clarify the current Belgian health information landscape and provide and sustain up-to-date metadata
Belgian health related data are very rich yet very dispersed among different institutions. Moreover, Belgian health (related) and administrative data are in serious need of FAIRification. Even for health data experts, it is not always easy to know where and which data are available. Making the data findable, by developing good and up-to-date metadata, would be a very valuable first step. A clear and up-to-date inventory of health (related) data would save data users a lot of time. Subsequently, harmonizing and standardizing datasets would significantly facilitate the preparation of foreseen data linkages. It should also be clear what variables actually mean and what their possibilities and limits are.
Establish a consortium of the willing
Another preparatory step in the development, governance, and support of an administrative health data cohort and research infrastructure could be the establishment of a consortium of the willing. This would be an active network of key partners in academia, administrations, and the national health information system in general (including both data holders and data users) that combines relevant research questions and methodologies, know-how and technical advances, and that keeps communication lines short and builds trust (20). Different stakeholders should better inform each other about their working procedures, data flows, access and linkage procedures, governance structure, etc. Also installing a central point of contact regarding the secondary use of health data and setting up linkages was recommended. Various means should be found to maintain this

consortium as a long-term, sustainable infrastructure to efficiently pool expertise and resources.
Standardize linkage procedures
Within the Belgian health information landscape, stakeholders should aim for a standardization of linkage procedures. A stronger standardization of linkage procedures could facilitate ad-hoc linkages and ease the current burden of, for instance, frequently changing requirements and procedures.
Develop an ethico-legal framework
Sustainable and longitudinal linkages between Belgian data sources are missing, which hampers the valorisation of these data sources (21). It is crucial that data from different data sources can be linked with each other, and this on an individual level and on a more routine basis. Future surveys therefore should always foresee the necessary legal and ethical provisions to allow prospective routine linkages. Only this way, a complete picture of a person's health and the environment they are part of, can be obtained and quick responses can be provided to public health challenges.
Increase digitalization
Online panels, consulted by means of surveys, can collect information on lifestyle, health behaviour, diseases, and expectations and experiences of health care users. Hence, they can produce high-quality data at low marginal costs (22). In time, these panels can mature by scaling up the data collection (e.g. by including data collected in a clinical or experimental setting and/or by repeating the data collection) and by enriching the data through data linkages. This way, online panels can be a valuable tool for public health research.
Secure a long-term financial and organizational effort
The actual implementation of a population-based cohort requires long-term financial and organizational effort. Currently, temporary research grants mostly lead to ad hoc data linkages. Instead, stable and long-term funding should allow for long-term research goals, consistent (meta) data accessibility, and structural data management (20). Moreover, organizational efforts are needed to increase adherence and avoid drop-out of cohort participants.
Identify and tackle ethical and legal bottlenecks
Establishing routine linkages between health and administrative data in a sustainable population-based cohort and research infrastructure might face fundamental ethical and legal bottlenecks. Currently, a legal framework is missing regarding a more structural health data collection, linkage and dissemination. From an ethical perspective, concerns regarding confidentiality, protecting the data subject's rights and freedoms, and the necessity of technical and/or organizational measures to ensure data protection and ethical compliance, might arise (23). Also citizens' involvement in health data reuse, for instance by obtaining informed consent or by transparent communication on research purposes or results, is an important issue that needs to be further addressed.
Respect the data relationship between citizens and health (related) data in the secondary use of health data
Every time health data is used, citizen perceive that a piece of them, of their identity and history, is being used, including in the case of secondary use. Therefore, anyone using their data is engaged in a relationship with them. They need to be respected as partners in this relationship. Every element of this relationship needs to be given proper attention and requires continuous work.

- Citizens would recommend being able to access information about the secondary use of health data, in an understandable way, allowing them to be more engaged.
- Citizens would recommend having access to their data and know how they are used for secondary purposes. However, they want to choose how and when they are informed about the uses of their data.
- Citizens would recommend that their values should inform what is beneficial to individuals and what constitutes the common good.
- Citizens would recommend that decision-making processes rely on a plurality of views and actors to increase their trustworthiness, as for them the latter depends on who is involved in these instances.
- Citizens would recommend being given the opportunity to be involved in the lifecycle of health data, as they need to be engaged on a continuous basis. Otherwise, their relationship with data custodians and users can deteriorate.

Maintain a suitable power balance in the secondary use of health data

Citizens consider data as power: it can be used to greatly benefit everyone, but this power could also be used against citizens. They need a good balance between risk mitigation and benefit maximization to ensure proportional use of health data for purposes we support.

- Citizens would recommend being provided with the opportunity for meaningful and active decision-making in the secondary use of health data, as they value the ability to exercise control.
- Citizens would recommend to ensure the protection of individuals’ identity, which they perceive as one of the best ways to balance the harms and benefits of the secondary use of health data.
- Citizens would recommend that data users’ intentions should be transparent and in line with purposes citizens support, as they think some users might share their values more than others
- Citizens would recommend that accountability could be enhanced through transparent and stronger mechanisms
- Citizens would recommend to foster good IT solutions to protect their data, beyond having a strong legal framework in place.

Establish a citizen-powered framework for the secondary use of health-data

All the above ideas and concerns should be assembled in a regulatory framework that governs secondary use in actual practice and in a way that is truly supported by citizens. The development of such a citizen powered framework does not only include appropriate governance structures but also respect for central ethical values as they are interpreted by citizens.

- Citizen would recommend that stakeholders respect principles that align with citizens’ ethical values.
- Citizen would recommend having a framework which facilitates the secondary use of health data for purposes and benefits that they support, while minimising the potential risks they identify.

5. DISSEMINATION AND VALORISATION

The AHEAD project developed in parallel with and collaborates with a number of major activities that have been launched. At the European level, the Joint Action Towards a European Health Data Space (TEHDAS) provides new synergies and dissemination opportunities. At the Belgian level, the development of a Health Data Agency provides a strong impetus to reach the objectives of the AHEAD project.

Valorization is at the heart of the AHEAD project. Throughout the project, specific activities are set up to captivate different target audiences, i.e., researchers, stakeholders, data holders, policy makers, and general public. Different communications are applied, tailored to the target audience, to maximize the good transfer and utilization of our results.

In WP1, we actively reached out to professional stakeholders in the Belgian health information landscape, by organizing a webinar and study days on health information and policy with the concerned stakeholders, data holders, and policy makers.

In WP2, we disseminated information and results regarding the secondary use of health data among the general public by organizing a public, online consultation. Additionally, we involved the relevant stakeholders in Belgium regarding health data use, linkages and governance, by inviting them to participate in the qualitative interview study of the BHIS+ business case.

Generally, we will continue our monthly meetings with the Health Information Task Force at Sciensano and we will keep on strengthening our relations and collaborations with important stakeholders in the Belgian health information landscape, such as policy makers, academics, research groups, citizen involvement organisations, …

PARTICIPATION/ORGANISATION OF SEMINARS

Organization

- Webinar "Data needs during the COVID-19 crisis"; 27/04/2022 (https://www.brain-helicon.be/webinars/47-data-needs-during-the-covid-19-crisis)
- Debate evening "Gezonde Data", De Maakbare Mens; 25/05/2022 (https://www.demaakbaremens.org/mee-met-het-idee-gezonde-data/)
- Study day "Health information and policy"; 06/10/2022, Brussels
- Stakeholder workshops "Healthy Data", 15/09/2022 (Belgian workshop, online), 16/09/2022 (UK workshop, online), 19/09/2022 (French workshop, online), 11/10/2022 (European workshop, Brussels)
- Study day "Health information and policy: How can health data and local policy interact successfully?; 23/11/2023, Brussels

Oral presentations

- Sciensano EpiTuesday Seminars, 07/09/21, Brussels, Belgium (videoconference)
  Pitch presentation on the AHEAD project
- "Health data reuse for the common good – The need for a more defined ethical framework"; ESPMH conference, August 2022, Warsaw
- "Healthy Data – On online citizen consultation on the secondary use of health data"; December 2022, NIC Europe (online)

Participation

Participation in several international online seminars, such as on/by the European Health Forum (28/09/21), WHO Data Governance (30/09/21), Statistics Netherlands Data Strategy (24/11/21), Smals Privacy Enhancing technologies (21/12/21), Unica Citizen Engagement (15/02/22), GOVTRUST (17/02/22), …

Participation in several (inter)national (online) seminars, such as on/by the Health Data Hub, PHIRI, TEHDAS, VITO, Foresight informed policy making, Deliberative democracy, AI and data sharing, Synthetic data, Citizen science, Belgian Association of Public Health, etc.

SUPPORT TO DECISION MAKING

In the context of the Healthy Data citizen e-consultation, 12 recommendations have been formulated on how citizens would like to be engaged in the secondary use of health data and the future European Health Data Space. The recommendations have been communicated to the European Commission.

The 12 recommendations included the following:

1. Citizens would recommend being able to access information about the secondary use of health data, in an understandable way, allowing them to be more engaged.

2. Citizens would recommend having access to their data and know how they are used for secondary purposes. However, they want to choose how and when they are informed about the uses of their data.

3. Citizens would recommend that their values should inform what is beneficial to individuals and what constitutes the common good.

4. Citizens would recommend that decision-making processes rely on a plurality of views and actors to increase their trustworthiness, as for them the latter depends on who is involved in these instances.

5. Citizens would recommend being given the opportunity to be involved in the lifecycle of health data, as they need to be engaged on a continuous basis. Otherwise, their relationship with data custodians and users can deteriorate.

6. Citizens would recommend being provided with the opportunity for meaningful and active decision-making in the secondary use of health data, as they value the ability to exercise control

7. Citizens would recommend to ensure the protection of individuals' identity, which they perceive as one of the best ways to balance the harms and benefits of the secondary use of health data.

8. Citizens would recommend that data users' intentions should be transparent and in line with purposes citizens support, as they think some users might share their values more than others

9. Citizens would recommend that accountability could be enhanced through transparent and stronger mechanisms

10. Citizens would recommend to foster good IT solutions to protect their data, beyond having a strong legal framework in place.

11. Citizen would recommend that stakeholders respect principles that align with citizens' ethical values.

12. Citizen would recommend having a framework which facilitates the secondary use of health data for purposes and benefits that they support, while minimising the potential risks they identify.

6. PUBLICATIONS

Scientific publications

- Schutte, N., Saelaert, M., Bogaert, P. *et al.* Opportunities for a population-based cohort in Belgium. *Arch Public Health* **80**, 188 (2022). https://doi.org/10.1186/s13690-022-00949-5

- Saelaert, M., Mathieu, L., Van Hoof, W., & Devleesschauwer, B. (2023). Expanding citizen engagement in the secondary use of health data: an opportunity for national health data access bodies to realise the intentions of the European health data space. *Archives of Public Health*, *81*(1), 168. https://doi.org/10.1186/s13690-023-01182-4

Reports

- TEHDAS. Healthy Data, an online citizen consultation about health data reuse – intermediate report. 2022. Available from: https://tehdas.eu/app/uploads/2022/10/tehdas-an-online-citizen-consultation-about-health-data-reuse-intermediate-report.pdf
- TEHDAS. Qualitative study to assess citizens' perception of sharing health data for secondary use and recommendations on how to engage citizens in the EHDS. 2023. Available from: https://tehdas.eu/results/tehdas-consultation-citizens-support-the-secondary-use-of-health-data-when-it-matches-their-ethical-values/

7. ACKNOWLEDGEMENTS

Composition of the AHEAD follow-up committee

Name	Institution
Petronille Bogaert	Sciensano
Joachim Cohen	VUB
Pascal Coorevits	Ugent
John Crombez	Ugent
Koen Decancq	Uantwerpen
Carl Devos	KCE
Thibaut Duvillier	Kruispuntbank voor Sociale Zekerheid
Guillaume d'Ansembourg	RIZIV-INAMI
Sofie De Broe	Sciensano
Emmanuel de Hemricourt de Grunne	eHealth
Karin De Ridder	Sciensano
Sylvie Gadeyne	VUB
Tom Goffin	Ugent, Metamedica
Laetitia Lempereur	FOD VVVL
Oliver Lenaerts	Rijksarchief
Patrick Lusyne	Statbel
Nick Marly	Kabinet Vandenbroucke
Heidi Mertes	Ugent, Metamedica
Sergio Perelman	ULiège
Koen Putman	VUB
Bart Scholiers	KULeuven
Nienke Schutte	Sciensano
Mahsa Shabani	Ugent, Metamedica

Karin Smets	FOD VVVL
Bert Vaes	Intego
Johan Van Bussel	healthdata.be
Johan Van der Eycken	Rijksarchief
Tom Van Renterghem	FOD VVVL
Johan Vanoverloop	IMA/AIM
Wolf Wauters	eHealth
Steering Committee	
Brecht Devleesschauwer	Sciensano
Wannes Van Hoof	Sciensano
Marlies Saelaert	Sciensano

8. REFERENCES

1. Börsch-Supan A. Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 7 [Internet]. SHARE–ERIC; 2019. Available from: 2019. https://doi.org/10.6103/SHARE.W7.700

2. Van Beckhoven D, Buvé A, Ruelle J, Seyler A, Sasse A, Belgian HIV Cohort Study Group. A national cohort of HIV-infected patients in Belgium: design and main characteristics. Acta Clin Belg. 2012;67(5):333–7.

3. Derom C, Vlietinck R, Thiery E, Leroy F, Fryns JP, Derom R. The East Flanders Prospective Twin Survey (EFPTS). twin res. 2002 Oct 1;5(5):337–41.

4. Janssen BG, Madlhoum N, Gyselaers W, Bijnens E, Clemente DB, Cox B, et al. Cohort Profile: The ENVIRonmental influence *ON* early AGEing (ENVIR *ON* AGE): a birth cohort study. Int J Epidemiol. 2017 Jan 15;dyw269.

5. Augner C. Health, education and employment status of Europeans aged 60 to 69 years: results from SHARE Survey. Industrial Health. 2018;56(5):436–40.

6. Hergenrather KC, Zeglin RJ, McGuire-Kuletz M, Rhodes SD. Employment as a Social Determinant of Health: A Systematic Review of Longitudinal Studies Exploring the Relationship Between Employment Status and Physical Health. Rehabilitation Research, Policy, and Education. 2015 Jan 1;29(1):2–26.

7. Zenger M, Berth H, Brähler E, Stöbel-Richter Y. Health Complaints and Unemployment: The Role of Self-Efficacy in a Prospective Cohort Study. Journal of Social and Clinical Psychology. 2013 Jan;32(1):97–115.

8. EU Expert Panel on effective ways of investing in health, 2016. Access to health services in the EU.

9. EU Expert Group on Health Systems Performance Assessment (HSPA). Tools and methodologies to assess the efficiency of health care services in Europe. 2019.

10. European Health Data Space; [Internet]. Available from: https://ec.europa.eu/health/ehealth/dataspace_en

11. Joint Action Towards the European Health Data Space – TEHDAS [Internet]. Available from: https://tehdas.eu/

12. Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016 Mar 15;3(1):160018.

13. Legislative proposal on the establishment and organisation of the Health data agency DOC55 3065/008 [Internet]. 2023 [cited 2023 Apr 25]. Available from: https://www.dekamer.be/FLWB/PDF/55/3065/55K3065008.pdf

14. Explanatory Memorandum to the Draft legislation on the establishment and organisation of the Health data agency DOC55 3065/001 [Internet]. 2022 [cited 2023 Apr 25]. Available from: https://www.dekamer.be/FLWB/PDF/55/3065/55K3065001.pdf

15. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLOS Medicine. 2015 mrt;12(3):e1001779.

16. Leitsalu L, Haller T, Esko T, Tammesoo ML, Alavere H, Snieder H, et al. Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. Int J Epidemiol. 2015 Aug;44(4):1137–47.

17. Szklo M. Population-based cohort studies. Epidemiologic reviews. 1998;20(1):81–90.

18. Van Beckhoven D, Buvé A, Ruelle J, Seyler L, Sasse A. A National Cohort of Hiv-Infected Patients in Belgium: Design and Main Characteristics. Acta Clinica Belgica. 2012 Oct 1;67(5):333–7.

19. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3(3):77–101.

20. Bogaert P, Verschuuren M, Van Oyen H, van Oers H. Identifying common enablers and barriers in European health information systems. Health Policy. 2021 Dec 1;125(12):1517–26.

21. European Commission, Directorate-General for Health and Food Safety. Tools and methodologies to assess the efficiency of health care services in Europe : an overview of current approaches and opportunities for improvement : report by the Expert Group on Health System Performance Assessment [Internet]. Publications Office; 2019. Available from: https://data.europa.eu/doi/10.2875/77000

22. Blom AG, Gathmann C, Krieger U. Setting Up an Online Panel Representative of the General Population: The German Internet Panel. Field methods. 2015;27(4):391–408.

23. Amram D. Building up the “Accountable Ulysses” model. The impact of GDPR and national implementations, ethics, and health-data research: Comparative remarks. Computer Law & Security Review. 2020 Jul 1;37:105413.