

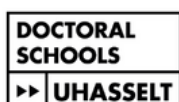
An abstract graphic of a network with nodes and lines, transitioning from purple and blue on the left to yellow and orange on the right, with a black horizontal line at the bottom.

WORKSHOP

HOW TO SET-UP A HEALTH DATA SHARING INITIATIVE?

Monday 22 April 2024 - Embuild Limburg

Report



Research Foundation
Flanders
Opening new horizons



Johnson & Johnson
Innovative Medicine



Content



<u>Introduction</u>	1
<hr/>	
<u>Key takeaways</u>	2-4
<hr/>	
<u>Inspiring #DataSavesLives examples</u>	5-6
<hr/>	
<u>Break-out sessions</u>	7-15
<hr/>	
<u>Break-out 1: Engaging stakeholders and value proposition</u>	8-11
<hr/>	
<u>Break-out 2: Identifying critical questions</u>	12-13
<hr/>	
<u>Break-out 3: Sharing best practices and lessons learned</u>	14-16
<hr/>	
<u>Annex 1 - List of participants</u>	17
<hr/>	



Introduction

The purpose of the workshop "How to set-up a health data sharing initiative?" - organized by UHasselt's [Biomedical Data Sciences](#) research group - was to explore and define the pathways for setting up effective health data sharing initiatives, particularly focusing on the use of real-world data (RWD) and federated learning systems.

Participants from diverse backgrounds, including general practitioners, oncologists, data engineers, and IT professionals participated in the workshop, ensuring a broad range of expertise. The list of participants is provided in the annex of this document. This report captures the key takeaways from the workshop, offering a guide for similar future initiatives.



30 health data enthusiasts participated in this workshop - #DataSavesLives



Key takeaways



1. Stakeholder needs: How to engage different stakeholders and craft value propositions

- **Patients:** Prioritize engaging patient organizations to understand their needs for enhancing health and wellbeing outcomes. Emphasize education about data use to facilitate informed engagement, ensuring patients are aware of how their data will be utilized and the benefits it can bring. Additionally, recognize that some initiatives may involve data from the general public rather than just patients. A data-sharing initiative can provide value beyond learning about treatment outcomes, offering insights into broader health trends and contributing to public health improvements.
- **Healthcare organizations:** Focus on benefits like operational efficiency, adoption of value-based healthcare models, and opportunities for increased funding and exposure to innovative practices.
- **Industry:** Aim for transformative health solutions through partnerships that ensure mutual benefits such as process optimization and strategic alignments.
- **Data holders:** Highlight benefits like increased visibility, improved data quality, opportunities for publishing findings, and access to broader resources.
- **Academia:** Offer incentives such as publication opportunities, financial support, enhanced networking, and access to a larger pool of data for research.

2. Recurrent key themes and challenges

- How can we effectively integrate diverse data sources while maintaining data privacy and quality?
- What are the best practices for ensuring scalability and interoperability in health data initiatives?
- How can AI and machine learning be leveraged to improve health outcomes in the context of these data sharing initiatives?





3. Do's and Don'ts



- **Initiate with clear ideation and planning:** Define the scope, mission, and vision clearly from the beginning.
- **Develop a comprehensive data strategy:** This should include data collection, quality assessment, and enhancement strategies.
- **Maintain flexibility and adaptiveness:** Use adaptive management strategies to respond to changes and feedback throughout the project lifecycle.
- **Stakeholder engagement:** Early and comprehensive engagement of stakeholders is vital. Tailoring value propositions to meet the specific needs and expectations of different groups—ranging from patients and healthcare organizations to industry partners and academia—is essential for ensuring broad-based support and active participation.
- **Governance and leadership:** Strong governance structures and decisive leadership are non-negotiable aspects of managing complex health data initiatives. These elements are critical not only in steering projects through their lifecycle but also in maintaining alignment with ethical standards and regulatory requirements.
- **Infrastructure and integration:** Investing in robust technological infrastructures that support the scalability and integration of health data systems is recommended. Such infrastructures must prioritize data quality and security to foster trust and reliability.
- **Continuous learning and adaptation:** Adopting a mindset of continuous improvement and learning from existing data sharing networks can enhance the effectiveness of new initiatives. It is beneficial to integrate lessons learned from established frameworks to avoid common pitfalls and accelerate progress.
- **Transparent communication and documentation:** Maintaining clear, open lines of communication with all stakeholders and ensuring thorough documentation of processes and outcomes helps in managing expectations and reinforcing the collaborative nature of the initiative.



Key takeaways



3. Do's and Don'ts



- **Underestimate the importance of governance:** Strong, decisive leadership and governance structures are crucial.
- **Over-promise on outcomes:** Manage expectations realistically to maintain trust and engagement.
- **Neglect legal and ethical considerations:** Ensure thorough planning around legal and ethical issues, which are often more crucial than technological aspects.
- **Rush the foundational phases:** Take time in the ideation and planning phases to mitigate challenges later on.
- **Ignore stakeholder diversity:** Avoid conflicts by harmonizing value propositions to align with the needs of all stakeholder groups.



Inspiring #DataSavesLives examples



The workshop was kick-started with a [presentation](#) of the mission and strategic activities of the research group Biomedical Data Sciences team at UHasselt. Following this, an introductory roundtable served as a platform for participants to share their motivations and current challenges in the field of health data sharing. Several examples of initiatives demonstrate the potential of collaborative health data sharing to enhance patient care, streamline research, and harness the power of advanced technologies like AI and machine learning.

More information about some of the initiatives that were mentioned during the roundtable discussion:

- **[EHDEN](#) (European Health Data & Evidence Network):** EHDEN is committed to reducing the time needed to provide answers in real-world health research by developing a federated network at scale across Europe. By standardizing data sources to a common data model (OMOP), EHDEN aims to significantly speed up the generation of reliable evidence. The network currently includes 187 Data Partners in 29 countries, harmonizing over 850 million anonymous health records. EHDEN also trains SMEs in data harmonization processes, fostering a strong open science community.
- **[DARWIN EU](#):** The Data Analysis and Real World Interrogation Network (DARWIN EU) provides the European Medicines Agency ([EMA](#)) and national authorities with real-world evidence on the use, safety, and effectiveness of medicines. DARWIN EU supports regulatory decision-making through a catalogue of observational data sources, high-quality real-world data, and the execution of non-interventional studies. This initiative is a key component of the EMA-HMA Big Data Steering Group work plan.
- **[WiNGS](#) (Widely Integrated NGS Platform):** WiNGS is a federated, web-based platform that simplifies the analysis of genome sequencing data while preserving patient privacy. It supports the clinical use of Whole Genome Sequencing (WGS) by optimizing ICT requirements and allowing anonymized analyses across client-hosted hubs. This platform is part of [ELIXIR Belgium](#), a research infrastructure that provides platforms and guidance for research data management and reproducible data analysis, and offers domain specific services for Plant and Biodiversity, Human health and COVID-19 research
- **[Antwerp Health Harbour](#):** Antwerp Health Harbour is a collaborative initiative aimed at strengthening Antwerp as an innovative health cluster. On September 21, 2023, key partners including the City of Antwerp, Netwerk Ziekenhuis aan de Stroom, Ziekenhuisnetwerk Helix, University of Antwerp, and Eerstelijnszone Antwerpen signed an intention declaration to develop a unique ecosystem in digital health. This collaboration focuses on utilizing big data for health care prevention, sharing data and applications, and creating innovative health projects. This partnership is expected to have a significant positive impact on health innovation in the Antwerp region and beyond.



Inspiring #DataSavesLives examples

- **OHDSI (Observational Health Data Sciences and Informatics)**: OHDSI is a multi-stakeholder, interdisciplinary collaborative aimed at deriving value from health data through large-scale analytics. It focuses on generating accurate, reproducible, and well-calibrated evidence to promote better health decisions and care. OHDSI operates as an open-source initiative, with a global network of researchers and observational health databases. The European OHDSI community, led by the Erasmus University Medical Center, actively contributes to the development of the OMOP-CDM and analytical tools. The Belgian OHDSI Node, formed from the Belgian EHDEN community, aims to harmonize health data ecosystems in Belgium and align with European standards, fostering an open-science and open-source approach to health data research.

This round table revealed a keen interest in understanding the challenges of initiating health data sharing projects, with emphasis on managing heterogeneous data sets and integrating them with clinical research. In addition, the need for collaborative efforts and expertise across various disciplines was a recurring theme. The roundtable underscored the importance of overcoming existing challenges such as the complexity of setting up collaborative frameworks, data privacy concerns, and technical hurdles in data integration and analysis to fully realize the benefits of health data sharing.





Break-out sessions

Three parallel break-out sessions were introduced to tackle the following challenges:

- **Break-out 1: Engaging stakeholders and value propositions:** This break-out focused on identifying key stakeholders for health data initiatives and creating compelling value propositions to engage them effectively.
- **Break-out 2: Identifying critical questions that need to be addressed when initiating health data sharing initiatives:** This break-out session aimed to pinpoint essential questions that should be addressed as early as possible to ensure smooth execution of large-scale collaborative efforts using health data.
- **Break-out 3: Sharing best practices and lessons learned:** This session discussed best practices and lessons learned ('do's and don'ts') from existing initiatives.





Break-out 1: Engaging stakeholders and value proposition

This break-out session focused on meticulously identifying and prioritizing key stakeholders essential to health data initiatives, highlighting the unique value each contributes. The comprehensive list of stakeholders included patients, healthcare organizations, data holders, industry, academia, and the general public, with particular emphasis on prioritizing based on potential impact and level of involvement.

Patient organizations: Emphasized as crucial for directly voicing patient needs and priorities, this group's value lies in enhancing health and wellbeing outcomes, advocating for patient-focused initiatives, creating transparency, and fostering trust across healthcare ecosystems. Additionally, educating patients about data use was identified as a vital component for informed engagement.

- **Enhanced health outcomes:** By integrating patient data from diverse sources, healthcare providers can tailor treatments more precisely to individual needs, leading to better health outcomes. For example, patient registries for rare diseases can use shared data to identify effective treatment protocols faster.
- **Advocacy and prioritization:** Patient organizations gain a stronger platform to influence which patient groups should be prioritized for research and funding. For instance, cancer patient organizations can lobby for more focused research on less common but deadly forms of cancer, ensuring these groups are not overlooked.
- **Transparency and trust:** Initiatives that include patient organizations in the decision-making process foster greater transparency. For example, involving patients in the design and governance of health data platforms can increase their trust in how their data is used.
- **Educational outreach:** Educating patients about the benefits and use of their data can enhance engagement and informed consent. Campaigns and workshops can be conducted to explain how data sharing leads to better healthcare services.





Break-out 1: Engaging stakeholders and value proposition

Healthcare organizations: Discussions highlighted the benefits of helping these organizations streamline operations, adopt value-based healthcare models, and utilize benchmarking tools to improve service delivery. Also noted were advantages such as increased funding opportunities and exposure to innovative practices.

- **Streamlined operations:** Sharing data can help healthcare organizations streamline their processes. For example, interoperable electronic health records (EHRs) reduce redundant tests and procedures, saving time and resources.
- **Innovative practices:** Access to a wider pool of data allows these organizations to adopt innovative practices more quickly. For example, data from different hospitals can be analyzed to identify best practices in patient care.
- **Value-based healthcare:** Implementing value-based healthcare models becomes more feasible with comprehensive data. For instance, hospitals can track patient outcomes and costs more accurately, helping them transition from fee-for-service to value-based care.
- **Increased funding opportunities:** Collaborative data initiatives often attract funding from various sources, including government grants and private investors. Healthcare organizations can benefit from these additional funds to enhance their services.
- **Benchmarking tools:** Data sharing enables healthcare organizations to benchmark their performance against peers, identifying areas for improvement and setting realistic targets.

Industry: The session clarified that while industry partners seek access to data, the ultimate goal should be transformative health solutions, not merely data access. Partnerships, like those with the EHDEN foundation, underscore the necessity for mutual benefits such as process optimization and strategic alignments in healthcare innovations.

- **Access to high-quality data:** Industries can access extensive and high-quality datasets, which are essential for research and development. For example, pharmaceutical companies can use shared clinical data to expedite drug discovery and development.
- **Market expansion:** By leveraging shared data, companies can identify unmet needs and develop new products, thereby increasing their market share. For example, a company might discover a need for a new type of insulin pump through patient data.
- **Business process optimization:** Shared data can streamline business processes, such as supply chain management and product development cycles, making operations more efficient.
- **Strategic alliances:** Partnerships within data sharing networks, like those facilitated by EHDEN, can lead to strategic alliances that promote innovation and market growth.



Break-out 1: Engaging stakeholders and value proposition

Data holders: Benefits for data holders in participating in health data networks include increased visibility, enhanced data quality through collaborative learning, opportunities for publishing findings, and access to broader resources like funding and advanced analytical tools.

- **Increased visibility:** By contributing to large data networks, data holders can increase their visibility in the research community and beyond. For example, a hospital that shares its data with a national health database can attract more research collaborations.
- **Enhanced data quality:** Collaborative learning and shared standards improve the quality of data. For instance, consistent data entry protocols across institutions can reduce errors and improve the reliability of research findings.
- **Opportunities for publication:** Participation in data networks often leads to new research findings and increased opportunities for publication in prestigious journals.
- **Access to resources:** Data holders can gain access to broader resources, including advanced analytical tools and additional funding, facilitating more comprehensive research projects.



Break-out 1: Engaging stakeholders and value proposition



Academia: For academic institutions, the primary incentives include publication opportunities, financial support, enhanced networking, and access to a larger pool of data for research purposes. Educational opportunities both in teaching and learning were also emphasized as key motivators.

- **Publication opportunities:** Access to a larger pool of data enhances the scope and impact of academic research, leading to more publications in high-impact journals. This is particularly significant when utilizing naturally occurring, real-life data, which is often not feasible to collect and finance within the typical timeframe of a research project.
- **Funding support:** Collaborative projects often attract funding from a variety of sources, including governmental and private sector grants. The translational nature of the research questions addressed through health data initiatives aligns well with the priorities of many funding bodies, increasing the likelihood of securing grants.
- **Networking:** Participation in health data networks fosters collaboration with other institutions and researchers, broadening academic and professional networks. These initiatives promote open science on two levels: facilitating the reuse of data and promoting team science while ensuring compliance with regulations such as GDPR.
- **Access to data:** Researchers gain access to diverse datasets, enabling more robust and comprehensive studies. The ability to work with and reuse data collected in natural settings without the need to invest in data collection themselves is a significant advantage, making the research process more sustainable and cost-effective.
- **Educational opportunities:** Health data sharing initiatives provide valuable teaching and learning opportunities. For example, students can engage in real-world data analysis projects, preparing them for future careers in data science and healthcare research. Additionally, lessons learned and peer-to-peer knowledge leveraging initiatives can be launched, identifying recurrent frameworks and the opportunity for setting up training programs within the collaborative networks, adding significant value for academia.
- **Citizen science:** Health data initiatives offer the opportunity to commit to citizen science by including citizens throughout various steps of the research project. This approach is gaining importance in competitive EU funding channels, promoting science for and by the people, and enhancing the societal impact of academic research.

Critical discussion points included the challenge of engaging stakeholders without compromising trust among the group. Achieving a harmonized value proposition that aligns with all stakeholder groups is paramount to avoid conflicts and ensure broad-based buy-in. The session also highlighted the dynamic nature of selling these value propositions, underscoring the importance of adaptable engagement strategies and the pitfalls of over-promising on outcomes.

Break-out 2: Identifying critical questions



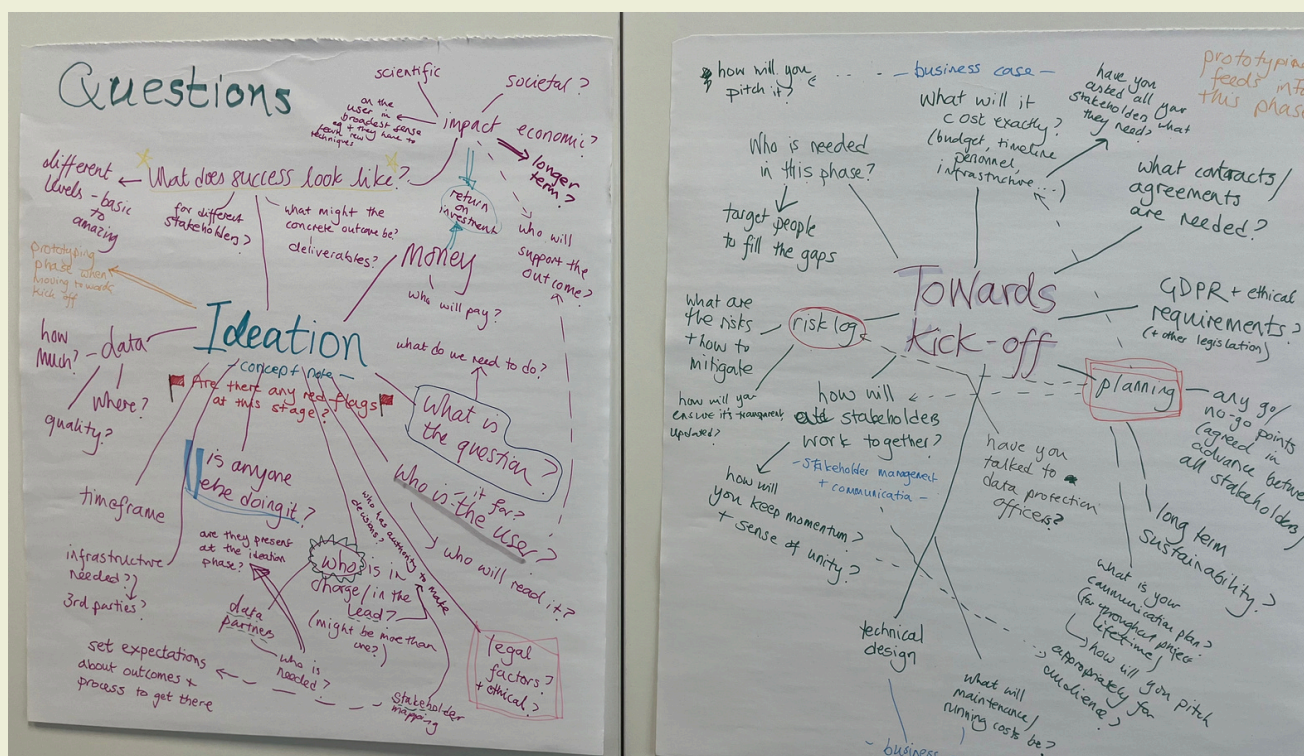
This break-out session aimed to pinpoint essential questions that should be addressed as early as possible to ensure smooth execution of large-scale collaborative efforts using health data. Key questions to reflect upon include:

- **Defining the “why”:** Clearly define why you want to set up a health data sharing initiative. What is the question you want to answer through data-driven insights? What will be the scientific, societal, and economic impacts?
- **Timing:** What is the timeframe that is acceptable or desired to set up this initiative?
- **Defining measures of success:** What does success look like at different levels, from basic to amazing, and what might the concrete outcomes be?
- **Identifying overlap:** Is anyone else doing a similar initiative? It might be better to join forces rather than reinventing the wheel.
- **Data requirements:** What data is needed, how much is required for statistical significance, and where can it be found? Is there a metadata catalogue available, or is this part of the project? What is the data quality, and what level of quality is required?
- **Financial and in-kind support:** Identifying who will support the initiative financially and in-kind, who owns the idea, and the types of infrastructure required. Who will pay, and what is the expected return on investment?
- **Infrastructure needs:** What infrastructure is needed? Do we need to involve third parties, or is the infrastructure present at the data partners' respective sites?
- **End-user identification:** Who is the end-user, who will read or use the results, and who will support the outcome? Who will pay, and what is the expected return on investment?
- **Legal and ethical considerations:** Emphasizing the need for robust planning around legal and ethical elements, often deemed more crucial than technological aspects. Are there any legal/ethical factors that should be considered? Who has the authority to make decisions?
- **Leadership and participation:** Who is in charge and who is needed in the ideation phase? Are all necessary parties present?
- **Expectation management:** Set clear expectations about the outcomes and define high-level processes required to get there.
- **Risk management:** Maintain a risk log to identify and mitigate potential risks, e.g. how to ensure transparency and keep the initiative up-to-date? Have you consulted with data protection officers?



Break-out 2: Identifying critical questions

- **Stakeholder collaboration:** Define how stakeholders will work together, establish a stakeholder management and communication strategy, and try to keep the momentum and sense of unity within the multi-stakeholder consortium.
- **Technical design and business case:** Develop a detailed technical design and business case, considering maintenance and running costs, budget, timeline, personnel resources, and infrastructure needs. Engage all stakeholders to reflect on their specific requirements.
- **Contracts and agreements:** Determine the necessary contracts and agreements, ensuring compliance with GDPR, ethical requirements, and other relevant legislation.
- **Planning and sustainability:** Set go/no-go points agreed upon by all stakeholders, ensure long-term sustainability, and develop a comprehensive communication plan throughout the project's lifetime. How will you pitch to the broader community and other interested parties?



Break-out 3: Sharing best practices and lessons learned



This session discussed best practices and lessons learned ('do's and don'ts') from existing initiatives.

- **Phase 1: Ideation Phase**

- During the ideation phase, there was a strong consensus that the focus should be on expectation management. This included checking the availability of data and defining the research questions the health data sharing initiative aims to answer. Building trust among stakeholders was of utmost importance, along with engaging all necessary parties from the beginning. This could be facilitated through organizing demos and sending newsletters to communicate transparently. It was also emphasized that involving a data architect or a representative of the execution team is crucial.
- Key considerations also included ensuring the formation of a consortium with all key stakeholders represented from the start, ensuring clear role definitions, and defining the added value of the initiative in the ecosystem. Involving the legal department early on was highlighted to manage collaboration agreements and governance documentation. Developing use cases to demonstrate the initiative's value was also deemed important.

- **Phase 2: Towards Kick-off**

- To maintain momentum, organizing regular meetings and extending communication strategies through widespread newsletters and social media campaigns were recommended. Investing in change management trajectories was identified as essential for successful implementation. Additionally, it was advised to start thinking about future steps and decide whether to set up the initiative independently or collaborate with an existing one based on current knowledge.
- Leadership should be provided by a "friendly dictator" or ultimate orchestrator, ensuring a clear project plan with defined deliverables and intermediate milestones. Meetings should always conclude with results or a call-to-action, avoiding "waterhead-structures." It was also emphasized not to strive for perfection but to view the process as a learning journey. Clear responsibilities and feasibility assessments were crucial before initiating the project.

Break-out 3: Sharing best practices and lessons learned

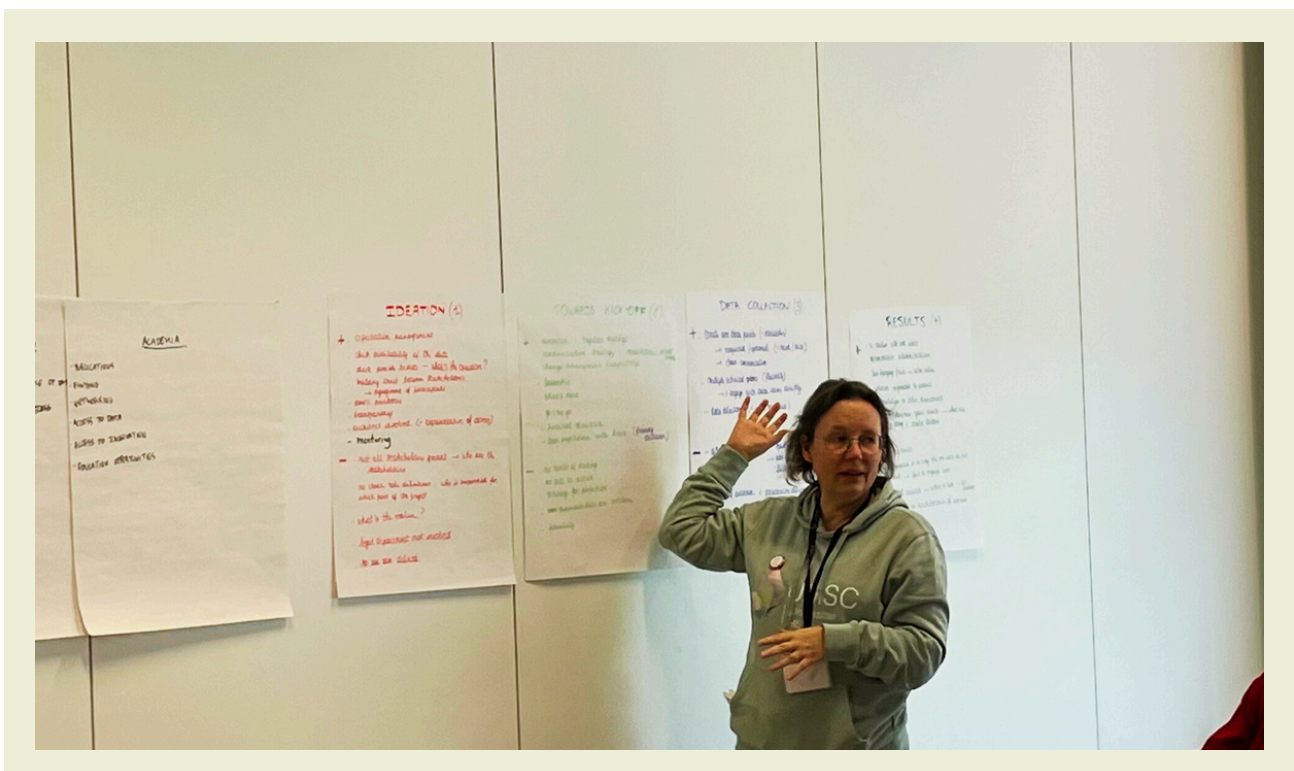


• Phase 3: Data Collection

- In the data collection phase, it was crucial to iterate on which data points or variables to collect, distinguishing between required and optional ones. Clear communication of these requirements to all partners was necessary. Flexibility in technical solutions and engagement with different data teams were emphasized. Clear data definitions and timelines for measurement were also critical.
- Efforts should be made to avoid merging datasets with varying data quality. Benchmarking and installing pass or fail metrics, along with providing feedback and improvement recommendations to data sources, were important. Additionally, designing the database and implementing pseudonymization strategies before starting data collection were necessary steps.

• Phase 4: Results Phase

- In the results phase, it was important to co-create the aspired results with end-users in mind. Documenting inclusion and exclusion criteria and starting with "low-hanging fruit" to demonstrate value was advised. Leveraging lessons learned and applying the insights to other departments could create added value. Results should be showcased and presented in a way that end-users can understand, ensuring successful adoption.
- Strategies to avoid overwhelming users with too many results were also discussed. The architecture of tools and outputs should be adapted to meet end-user needs effectively.





Break-out 3: Sharing best practices and lessons learned

Recurrent themes across all phases included:

- **Expectation Management:** Starting with manageable goals, ensuring data availability from the outset, and establishing a foundational level of trust among all stakeholders.
- **Communication Strategies:** Keeping stakeholders engaged through regular updates, newsletters, and demonstrations to ensure ongoing commitment and awareness.
- **Leadership and Governance:** The need for strong, decisive leadership to drive the project, sometimes requiring a firm hand to keep the project on track and focused.
- **Change management** was identified as crucial for embedding the results within operational settings, such as hospitals, and adapting to feedback and evolving project needs. The flexibility in choosing the right technological environment (e.g., Azure, on-premises, Google) was also discussed, emphasizing the need for adaptability in data management and usage. The session further highlighted the importance of engaging end-users early in the design process, especially in tools development like dashboards, to ensure the outputs are practical and meet the users' needs effectively.
- **Engaging stakeholders without compromising trust among the group.** Achieving a harmonized value proposition that aligns with all stakeholder groups is paramount to avoid conflicts and ensure broad-based buy-in. The session also highlighted the dynamic nature of selling these value propositions, underscoring the importance of adaptable engagement strategies and the pitfalls of over-promising on outcomes.



Annex 1 - List of participants

Name	Affiliation
Maryna Borshchivska	DNAlytics
Mona Bové	OLV Aalst
Dilza Campos	VIB Data Core
David Clijsters	Cegeka
Nicky Daniels	UHasselt
Bram De Caluwé	AZ Klina
Freija Descamps	edenceHealth
Nick Devreker	AZ Delta
Stijn Dreesen	Jessa Ziekenhuis
Axel Faes	UHasselt
Anouk Geenen	UHasselt
Lotte Geys	UHasselt
Thibault Helleputte	DNAlytics
Anne Helme	MS International Federation
Andreas Kremer	ITTM S.A.
Kim Luyckx	imec
Gianny Mestdach	OLV Aalst / AZORG
Rod Middleton	Swansea University
Liesbet M. Peeters	UHasselt
Hilde Philips	UAntwerpen
Noëlla Pierlet	ZOL / UHasselt
Peter Rijnbeek	Erasmus MC Rotterdam
Nishkala Sattanathan	UAntwerpen
David Smeets	J&J Innovative Medicine
Vasilis-Spyridon Tseriotis	Aristotle University of Thessaloniki
Emma Van de Velde	Möbius
Nicky Van Der Vekens	AZ Maria Middelaes Hospital
Annelies Verbiest	UZA
Ilse Vermeulen	UHasselt
Stéphanie Warnon	GHDC
Mandy Melissa Jane Wittens	SMIT-IMEC-VUB

