

Data accessibility doesn't mean they're fit for purpose.

On the Importance of Domain Knowledge in the Implementation of a Dutch Health Data Access Body

Open letter to the Dutch Ministry of Health, Welfare and Sport (VWS) and the HDAB-NL program leadership, on behalf of Nivel, Vektis, Dutch Hospital Data (DHD), **Netherlands Comprehensive Cancer Organisation** (IKNL), 16 September 2025.

Contact @Nivel: prof. dr. Robert A. Verheij (r.verheij@nivel.nl); dr. I. Bos (i.bos@nivel.nl).

Summary

The European Health Data Space (EHDS) offers a unique opportunity to use health data for research, policymaking, and innovation. EHDS primarily focuses on improving data availability. However, especially when it comes to healthcare data itself (electronic health records, claims data), data availability does not automatically mean that data are fit for purpose.

To also promote fitness for purpose, it is essential that existing, expert, and trusted data holders — such as Vektis, DHD, Nivel, and IKNL — are given a role in the new governance and data-access model under EHDS.

These organizations possess extensive domain knowledge, proven data quality procedures, and established relationships with stakeholders such as healthcare providers, patients, insurers, and software vendors. Their understanding of the possibilities and limitations of data is necessary to judge whether data are suitable for a specific purpose. Preserving and mobilizing this knowledge within the governance structure yet to be established is a prerequisite for ethical, reliable, accurate, and efficient data use.

This open letter is a joint response by Nivel, Vektis, DHD, and IKNL to an invitation from HDAB-NL to contribute ideas regarding the further development of the Dutch Health Data Access Body.

Context: EHDS and the Health Data Access Body (HDAB)

EHDS presents a unique opportunity for the Netherlands to use health data safely, responsibly, and effectively for research, policymaking, and innovation. By establishing one or more national Health Data Access Bodies (HDABs), the European Union aims to improve the secondary use of data — such as for scientific research, policy development, and evaluation.

EHDS mainly focuses on data availability. Once approved by an HDAB, data holders will be required to provide data within a certain timeframe.

However, using data without sufficient attention to data fitness for purpose undermines the goals of EHDS. For any given question, a proper assessment of data fitness for purpose is essential to prevent errors, misinterpretations, and unintended biases.

Within the EHDS framework, machine-readable metadata standards are being developed. While important and useful, these are often insufficient for properly assessing data fitness for purpose. To determine whether certain data can be used for a research question and to avoid misinterpretation, deeper knowledge is required — for example, about disease progression and treatment practices, or about how the healthcare system is financed.

Such knowledge can only partially be captured in metadata and data quality indicators, yet it is essential for correct interpretation. Healthcare data may reveal much about disease prevalence, but they are also strongly influenced by treatment options and financial incentives to provide certain types of care. Expertise on these matters resides with experienced data custodians such as Vektis, DHD, Nivel, and IKNL.

Assessing data fitness for purpose is not always straightforward. Data combinations, subsets, and selections that initially appear adequate often need revision later in order to obtain valid answers. Domain knowledge is essential in this process, and the involvement of experienced data holders is therefore required — including during the analysis phase.

With a future permit-based system, the HDAB will operate at a distance from the data collection process. There is a real risk that this will lead to procedural assessments without sufficient attention to the context and fitness for purpose of data for the intended purpose. This highlights the need to establish processes in which the knowledge held by data custodians is systematically incorporated.

This does not mean that data custodians must provide intensive guidance for every data request, but it does mean they should actively contribute to shaping the new governance and access model, so that together we ensure both safety and optimal fitness for purpose of data for a given purpose. The exact role may differ between data custodians.

Domain Knowledge for Data Fitness for purpose

Data custodians such as Vektis, DHD, Nivel, and IKNL possess the domain knowledge necessary to assess data fitness for purpose because they:

1. Independently collect and maintain high-quality, continuously updated datasets that play an essential role in monitoring and interpreting policy effects and conducting scientific research. Their data are crucial for policy, research, and monitoring outcomes.
2. Maintain relationships with healthcare professionals, healthcare organizations, and software suppliers (for example, providers of electronic health record systems), which are crucial for interpreting registration data and identifying biases.
3. Maintain relationships with similar institutions in other European countries, enabling them to contribute effectively to better comparability of data across countries.
4. Have many years of experience combining and harmonizing data for policy, supervision, and research purposes.
5. Have operational methods for assessing and improving data fitness for purpose and data validation, including compliance with standards and privacy regulations.
6. Understand the practical and contextual background of the data, which is necessary to use and interpret the data correctly for a given purpose.

Because of all this, these organizations have the expertise needed to translate a research or policy question into a data request. Potential users usually do not come with a request for data itself; they come with a research question, which then has to be translated into a data requirement. Often, users do not know which data they should request and are therefore dependent on the expertise of data custodians.

The importance of this expertise is growing as the combined use of multiple data sources becomes increasingly standard. Scientific and policy questions can rarely still be answered adequately using a single data source.

A Layered Model of Data Governance

We believe it is important that data custodians are involved in shaping a governance framework in which their domain knowledge is safeguarded. We envision a layered model of data governance consisting of:

1. An HDAB responsible for national coordination and enforcement, routing data requests, and overseeing metadata and transparency.

2. Trusted health data holders (such as Vektis, DHD, Nivel, and IKNL) responsible for advising on data fitness for purpose, metadata, contextual knowledge, and interaction with healthcare providers.
3. Source holders (healthcare institutions and suppliers) responsible for supplying data to trusted health data holders, providing contextual knowledge, and handling technical implementation.

For such a model to work, two important conditions must be met:

- Adequate implementation legislation. Clear legal frameworks must define the roles of the HDAB, trusted data source holders, and data custodians. This will create clarity regarding responsibilities, powers, and funding.
- Sufficient financial resources. Establishing and maintaining HDAB registries, quality assurance processes, and metadata standards requires investment. We would like to jointly explore how these resources can best be secured, for example through a national subsidy scheme or a dedicated HDAB budget.

Conclusion

EHDS offers the Netherlands major opportunities for better use of existing health data. However, responsible use of health data requires more than availability: it requires understanding and knowledge of the context in which the data are created. Data do not speak for themselves. Improved machine-readable metadata can only partially solve this issue.

The HDAB should therefore not only provide centralized coordination but also structurally collaborate with existing data custodians, who contribute unique expertise to make data reuse safe, effective, and meaningful.

We invite the Ministry of Health, Welfare and Sport and the HDAB-NL program leadership to engage with us in dialogue and explore how the role of knowledge organizations in ensuring data fitness for purpose can be firmly embedded. Through discussion, collaboration, and practical action, we can jointly create a system in which data availability and data fitness for purpose go hand in hand, and in which cooperation between HDAB and expertise holders is concretely and sustainably organized.