Opinion on

the Proposal for a Regulation on the European Health Data Space dated 03.05.2022

The European Commission has introduced the Proposal for a Regulation on the European Health Data Space ("EHDS") dated May 3, 2022 (the "Proposal") with the aim of regulating **primary and secondary use of electronic health data**.

At first sight. the Proposal's guiding principle regarding secondary use is understandable and attractive. With regards to secondary use, the Proposal requires that **existing data silos be centrally cataloged and anonymized data be made widely usable via a defined standardized access process.** Nevertheless, the Proposal falls short of creating a framework that validates the ambitious and comprehensive title "European Health Data Space".

In some European countries, like parts of Scandinavia, the health care system is governed by the state, giving it control over health data. Moreover, in Scandinavia, society's understanding of the role of the state is markedly different from that in, for example, Germany; therefore, control of the secondary use of health data by the state is only logical. It can be argued that this model could be transferred to the whole of Europe.

It also appears that the current draft of the Proposal reflects insufficient coordination between **data protection and GDPR experts** on the one hand **and health experts** on the other hand.

This opinion discusses the problems of the Proposal and introduces **alternative solutions** and **modifications**.

A summary of the secondary use of electronic health data as proposed in Chapter IV of the Proposal, is set forth in the appendix.

Critique 1:
Disempowerment of the Citizen - Loss of Privacy regarding Electronic Health Data

The Proposal aims to **increase the control data subjects have over their electronic health data, while creating** a basis for the secure use of electronic health data by third parties. Throughout the Proposal, the control of data subjects over their data is listed first when discussing the Proposal objectives.[1] The Proposal clearly sets forth the importance of the empowerment of EU citizens to control their electronic health data. In the first Section of the explanatory memorandum, sentence 4 states: *"EHDS will create a common space where natural persons can easily control their electronic health data."*

However, the subsequent text of the Proposal not only fails to achieve this goal, but also **entirely undermines the control of patients over the secondary use of their electronic health data**. For example, Art. 33 (5) of the Proposal allows for the secondary use of electronic health data without consent:
"Where the consent of the natural person is required by national law, health data access bodies shall rely on the obligations laid down in this Chapter to provide access to electronic health data."

The **conference of secondary use without consent** is tantamount to a **complete loss of control by the data subjects**, which cannot be outweighed by the general interest in secondary use of data nor by the merely theoretical possibility that a data subject will benefit from new treatment options.

When balancing societal and individual interests, it is not appropriate to simply disregard the individual position. Most people have societal interest and feel that improving healthcare is such an important goal that they will **voluntarily make their own data available for secondary use**. **The majority of the German population (82%)**, for example, considers it an ethical imperative to use anonymized health data to improve therapy.[2]

The question of what proportion of citizens would allow or deny use of their electronic health data is a hypothetical one, as citizens were never actually provided with the appropriate data infrastructure to enable such a decision. Furthermore, it is incorrect to assume that any and all data from any and all patients are required under all circumstances. First, it cannot be assumed that every scientific question has a data bias based on partial consent, and second, any such bias can often be mitigated by use of aggregate patient reference data. **Obtaining specific consent should be the rule in Europe**, and the **use of data without consent should be the exception and only be utilized** in carefully defined individual cases.

[1] Reasons and Objectives of the Proposal, Chiff. 1., Section 4, Recitals 1, 12, 67, Art. 1 (2a) of the Proposal

[2] Health-Care Barometer PWC 2020, S. 35.

Critique 2:
Nationalization of Data - Data Socialism

The Proposal creates an **obligation for companies** (with the exception of micro-enterprises) and other third parties to transmit, upon request, all electronic health data available to them to "health data access bodies" established by the Member States. This applies to both data obtained through publicly funded measures, and data collected based on private funding.

Recital 40 of the Proposal states:
*The **data holders*** [Note: highlights added] *can be* [...] *health or care providers,* [...] ***private organisations**,* [...] *or* [...] *entities that carry out research with regards to the health sector that **process** the categories of health and **health related data*** [...]. *In order to avoid a disproportionate burden on small entities, **micro-enterprises are excluded from the obligation** to make their data available for secondary use in the framework of EHDS*. [Note: The latter clarifies that, except for microenterprises, all companies fall within the definition of data holder.]

Recital 40 continues:
"[...] ***data**, collected and processed by data holders with the **support of*** [...] ***public funding**, should be made available* [...]*"*

The above recitals describe the collection of data based (partly) on public funding. The Proposal further addresses data resulting from privately funded data collection, without making public funding a prerequisite: *"In some Member States, **private entities**, including private healthcare providers and professional associations, play a pivotal role in the health sector. **The health data held by such providers should also be made available for secondary use."*** Thus, data from multicenter studies, registries of professional associations, or from industrial research or development are also covered by this disclosure requirement.

The Proposal also creates the possibility of **state supervisory bodies directly monitoring** medical devices and medicinal products and thus, their manufacturers by pooling all data, since all **data generated by medical devices** must be **disclosed** to the health data access bodies (Art. 33 (1) (k)).

Furthermore, the Proposal specifies that the health data access bodies can **give priority to government use over private use requests in the event there is limited capacity** (Recital 51).

Experience has shown that the development of new drugs, treatments, and medical devices is more successful when private parties are involved equally. Private participation is essential to providing the best possible medical care for society at any given time (for example, as was shown with the vaccine development process during the COVID-19 pandemic).

The health data access bodies' control function does not apply to certain requests from public entities. For example, a data permit is not required when a public body requests secondary use of health data to support the tasks assigned to it (very broadly defined):

Pursuant to Art. 48, a data permit is not required if the health data access body performs its tasks pursuant to Art. 37(1)(b), (c):
"By derogation from Article 46 of this Regulation, a data permit shall not be required to access the electronic health data under this Article. When carrying out those tasks under Article 37 (1), points (b) and (c), the health data access body shall inform public sector bodies and the Union institutions, offices, agencies and bodies, about the availability of data [...]*"*

Art. 37 (1) reads:
"Health data access bodies shall carry out the following tasks:
[...]
b) ***support public sector bodies in carrying out the tasks*** *enshrined in their mandate,* ***based on national or Union law****;*
c) support Union institutions, bodies, offices and agencies in carrying out tasks enshrined in the mandate of Union institutions, bodies, offices and agencies, based on national or Union law;
[...]*"*

Ultimately, it is not market-based mechanisms but **exclusively state interests**, which are inevitably **subject to political factors and influence**, that decide who gets access to which data. However, equal access to data for private-sector entities makes more sense, as this is the only way to harness their innovative power for the common good.

Critique3:
Privacy vs. the Dystopia of the "Transparent Citizen"

The exclusive state control of health data access bodies is regulated in Article 36 (1)*: "****Member States shall designate one or more health data access bodies*** *responsible for granting access to electronic health data for secondary use. Member States may either establish one or more new public sector bodies or rely on existing public sector bodies or on internal services of public sector bodies that fulfil the conditions set out in this Article."*

In some countries, hardly anyone considers a **state monopoly on access to data** to be inappropriate; however, many citizens in Europe fortunately have a more critical view. By **linking data with other personal data** available to each state, as well as having **access to "trustees"** entrusted with the pseudonymization of data, such as the RKI in the case of Germany, it is possible, in principle, to **break down anonymization**. **Health data in plain text** would then be added to all data that is already available.

The EHDS should generally protect the **privacy** of data subjects and discourage the creation of a **"transparent citizen"** (through multinational institutions or networked approaches) rather than encourage it (through politically controlled authorities and centralized approaches).

In addition to the transparency of health data vis-à-vis the state, whether through access to the pseudonymization sites or through linkage with other profile data, the possible **transparency vis-à-vis third parties** must also be considered. There is a risk that the **privacy of data subjects** will be compromised, despite a ban on re-identification, due to **large platform companies with access to other personal data** and corresponding profiles.

Therefore, **the provision of anonymized and pseudonymized individual records envisioned by the Proposal is problematic**.

Art. 44 (3) reads:
"Where the purpose of the data user's processing cannot be achieved with anonymised data, taking into account the information provided by the data user, the health data access bodies shall provide access to electronic health data in pseudonymised format. [...]*"*

In political discussions, we often talk about **anonymized and personal data**. However, this binary distinction is a central issue. Allegedly anonymized individual data sets can often be **re-identified to the original persons by linking them to other data sets**.[3] The personal data for said linkage may not currently exist or be readily available, but this could change as a result of a **data leak or future technical development**.[4] For this reason, anonymized individual data sets should be **protected in a similar way as personal data**. However, Effective and complete anonymization of aggregated data from a sufficient number of data sets is possible. A distinction must also be made between whether one only has access to individual data sets or whether they are actually transmitted. Finally, **genetic data from biomaterial** can never truly be anonymized. So instead of only distinguishing between anonymized and personal data, the security measures and access mechanisms for secondary use of the data should be graded according to the level of identifiability.

The present Proposal inadequately addresses this general dilemma. The problem is the **combination of disempowerment of data subjects** by consent-free data access and the **insufficient security mechanisms**. If patients themselves could actively **decide** and **revoke consent** for data processing, the situation would be completely different.

[3] e.g. Re-Identification of the Netflix Price Dataset, 2006.

[4] e.g. MyHeritage Dataleak, 2018.

Critique 4:
Weakening of Europe as a Center of Business and Science

So far, difficult access to health data for secondary use in science has been a disadvantage for Europe. The present concept could make the situation even worse. The **requirement to disclose** all data collected in primary use **destroys the motivation** to gain a **head start, for example, in publishing research results** by the expensive collection of particularly relevant and structured data as part of a research project. It also creates the incentive to **collect data not in Europe**, but in markets where there is no risk that third parties can use one's own data to publish better findings earlier through more skillful evaluation or linkage with other data.

Accordingly, the draft regulation does not incentivize or protect **"added value – generators,"** such as companies and **scientific workgroups** that have generated added value by collecting and structuring data. Without such incentives or protection, such companies and workgroups would otherwise run the risk of being put on equal footing with other companies or scientific workgroups which merely want to use data.

Furthermore, the protection of intellectual property is not addressed sufficiently: *"Such data* [Note: "[...] *data benefiting from specific legal protection such as intellectual property from medical device companies or pharmaceutical* [...]"] *should be made available for public and regulatory activities, supporting public bodies to carry out their legal mandate, while complying with, where relevant* ***and possible****, the protection enjoyed by commercial data." (Recital 40)*

Combining multiple elements of a data space concept into centralized application centers steers toward **state economy in healthcare data processing**, rather than enabling **innovation and competition**. A Centralized approach requires the state to create the necessary infrastructure. As with many other issues, the state is acting as a player, **distorting competition** for the best data infrastructure rather than facilitating it. Innovation and competition among industrial providers should be encouraged instead of tying the dynamics of development to a specific government institution for data access. This is the only way to create a structure that will sustainably develop **resilience of the healthcare system** and **technical sovereignty of Europe**. For this to happen, however, the complex structures of a health data space must be broken down into individual components, with standardized interfaces and broad interoperability.

PROPOSALS

Therefore, the following modifications to the current concept are proposed:

Privilege for "Added Value - Generators".

There should be a privilege for "added value generators." This could protect companies that have previously generated added value by collecting, structuring, and processing data (e.g. with AI); these companies should not run the risk of being put on an equal footing with companies that merely want to use data. Individuals or institutions that collect data for publication, such as multicenter studies or clinical registries, should be allowed to use this data exclusively for 24 months, especially to protect their interest in issuing research publications. Accordingly, it should be possible to restrict the disclosure of new data in whole or in part during this period.

A disclosure requirement should apply to private companies only if the data collection was predominantly financed with public funds. Data collected exclusively as part of multicenter studies, by professional societies as part of nongovernmental registries, or by companies in connection with private research or product development should be explicitly excluded from the disclosure requirement.

Limitations on Data available through Access Bodies.

- Data use without patient consent should be possible only for data that is collected from all applicable patients, for example, because of national laws requiring such data collection.
- Data that is only collected on a consent-based basis must be excluded from consent-free processing.
- Transmission of individual records by the data holder should be excluded, especially for pseudonymized data. Permitted data access options:
    - Aggregated data from more than 100 individual records (only these are truly anonymous).
    - Federated machine-learning directly on the datasets of (multiple) data holders for AI applications.
- Retention should be possible only for basic data and with specific legal basis.
- Access bodies should only act as an agent for data to be exchanged only between the data collection point (data holder) and the data user after issuance of a data permit, without receiving the data themselves.

Patient-Centered Standardized Consent

The European Regulation on European Data Governance (Data - Governance - Act), which came into force on June 23, 2022, already stipulates in Article 25 the requirements for a European consent form for the use of personal data for altruistic purposes. This concept is ignored in the current Proposal. However, it should be taken up and implemented.

- A consent structure must be defined in such a way that each use can be clearly assigned to a consent-free or a consent-based access.
- It must be transparent to the patient, regardless of the type of digital data collected, which data type can be used by which data user group for which clinical application area in which data access mode and to which data storage criteria, either consent-free or consent-based.
- This transparency must also map out when a consent must be specifically granted (opt-in), or is deemed granted, and when consent must be specifically revoked (opt-out). The consent must regulate processing exclusively, with extensions or restrictions not permitted; otherwise meaningful management of consent parameters is not possible.

Accreditation Process for independent Application Bodies

Similar to the accreditation of a notified body for CE certification of medical devices, independent application bodies could emerge. Separating access permit from physical access using an appropriate digital certificate also improves the level of data protection.

European Pseudonymization Body

An independent multinational body for pseudonymization would significantly improve the level of data protection, as this body (similar to the EMA) could not be influenced by political interests of individual states.

Further detailed and constructive proposals can be provided as an additional attachment to this document in the short term

Annex

Summary of the regulations regarding secondary use of electronic health data as proposed in Chapter IV of the Proposal for a Regulation on the European Health Data Space dated May 3, 2022 (the "Proposal").

Chapter IV of the Proposal aims to facilitate the secondary use of electronic health data, e.g., for research, innovation, policy making, patient safety or regulatory activities. It defines several types of data that can be used for certain purposes and specifies prohibited purposes (e.g., use of data against individuals, commercial advertising, increasing insurance, development of dangerous products) (Explanatory Memorandum para. 5 p.23).[5]

Secondary use involves sharing health data of data holders with data users through a public data access body and a secure processing environment.

Member States must establish a health data access point for secondary use of electronic health data and ensure that data holders make the electronic data available to data users (see Explanatory Memorandum Chiff. 5). Public data access bodies have the following responsibilities, among others: Collect and compile health data from the various data holders, process requests for data access and issue data permits, electronically publish data permits as well as applications, and provide metadata on available data sets, assist public bodies in the performance of their duties and take all necessary measures to maintain the confidentiality of intellectual property rights and trade secrets (cf. Art. 37(1)). Any natural or legal person may submit a data access application to an access body (cf. Art. 45(1)). Public bodies do not require data permits if data is used for the purpose of fulfilling the tasks assigned to them (cf. Art. 48). Health data may only be shared in anonymized form or, in justified cases, in pseudonymized form (cf. Art. 44 (3)). Pseudonymization encryption keys for data assignment are held only by the health data access bodies (cf. Recital 49).

Data holder

According to Art. 2 (2), "*data holder*" means "*any natural or legal person, which is an entity or a body in the health or care sector, or performing research in relation to these sectors, as well as Union institutions, bodies, offices and agencies who has the right or obligation, in accordance with this Regulation, applicable Union law or national legislation implementing Union law, or in the case of non-personal data, through control of the technical design of a product and related services, the ability to make available, including to register, provide, restrict access or exchange certain data.*"

Private companies are also data holders and are therefore obligated to disclose the electronic health data available to them.

[5] References refer to the Proposal

Data categories

This applies, inter alia, to the following data categories:

- *EHRs (Art. 33 (1) lit. a).*
- *data impacting on health, including social, environmental behavioural determinants of health; (Art. 33 (1) lit. b).*
- *relevant pathogen genomic data, impacting on human health (Art. 33 (1) lit. c).*
- *human genetic, genomic and proteomic data (Art. 33 (1) lit. e).*
- *population wide health data registries (public health registries) (Art. 33 (1) lit. h).*
- *electronic health data from medical registries for specific diseases (Art. 33 (1) lit. i).*
- *electronic health data from medical devices and from registries for medicinal*
- *products and medical devices (Art. 33 (1) lit. k).*
- *electronic data related to insurance status, professional status, education,*
- *lifestyle, wellness and behaviour data relevant to health (Art. 33 (1) lit. n).*

Health data containing protected intellectual property and trade secrets of private companies are also made available for secondary use (cf. Art. 33 (4)). The data holder shall put the electronic health data at the disposal of the health data access body within 2 months from receiving the request from the health data access body (cf. Art. 41 (4)).

Data user

Data user means a natural or legal person who has lawful access to personal or non-personal electronic health data for secondary use (cf. Art. (2) lit. z).

Permissible purposes

Secondary use will be permitted for the following purposes (Art. 34 (1)):

- *activities for reasons of public interest in the area of public and occupational health, such as protection against serious cross-border threats to health, public health surveillance or ensuring high levels of quality and safety of healthcare and of medicinal products or medical devices.*
- *to support public sector bodies or Union institutions, agencies and bodies including regulatory authorities, in the health or care sector to carry out their tasks defined in their mandates.*
- *to produce national, multi-national and Union level official statistics related to health or care sectors.*
- *education or teaching activities in health or care sectors.*
- *scientific research related to health or care sectors.*
- *development and innovation activities for products or services contributing to public health or social security, or ensuring high levels of quality and safety of health care, of medicinal products or of medical devices.*
- *training, testing and evaluating of algorithms, including in medical devices, AI systems and digital health applications, contributing to the public health or social security, or ensuring high levels of quality and safety of health care, of medicinal products or of medical devices.*

- *providing personalised healthcare consisting in assessing, maintaining or restoring the state of health of natural persons, based on the health data of other natural persons.*

Art. 33 (5) states the following: *"Where the consent of the natural person is required by national law, health data access bodies shall rely on the obligations laid down in this Chapter to provide access to electronic health data."* This lays the foundation for secondary use of electronic health data without the consent of the data subject.