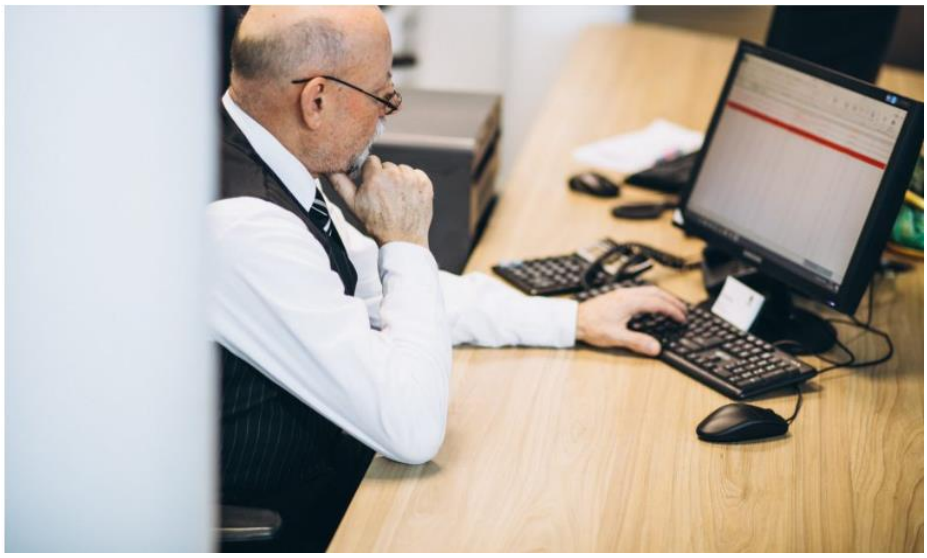
Anonymisation of Electronic Health Record (EHR) data



One of the pillars of the European Health Data Space (EHDS) is the secondary use of health data – for research and policymaking, among other purposes. Enter anonymisation, to make this possible. What is different from EHR data here? And what implications does this hold for the European Electronic Health Record Exchange Format (EEHRxF)?

What is the EEHRxF?

The European Electronic Health Records Exchange Format (EEHRxF) serves as a standardized framework for the seamless cross-border exchange of health information. It encompasses a wide range of essential domains, ensuring comprehensive coverage and interoperability across diverse healthcare systems. These domains include:

- (i) Patient Summaries: Consolidated records detailing a patient's medical history, diagnoses, treatments, and allergies, facilitating comprehensive care coordination.
- (ii) ePrescriptions/eDispensations: Electronic prescriptions and dispensation records enable efficient and secure medication management, promoting patient safety and adherence.
- (iii) Laboratory reports: Standardized formats for laboratory test results ensure consistent interpretation and exchange of vital diagnostic information, enhancing clinical decision-making.
- (iv) Medical images and reports: Standardized exchange formats for medical images and accompanying reports enable seamless sharing of radiological and imaging data, supporting accurate diagnoses and treatment planning.
- (v) Hospital discharge reports: Structured summaries of patient discharge information, including treatment received, medications prescribed, and follow-up instructions, ensure smooth transitions of care and continuity of treatment across healthcare settings.

Together, these components of the EEHRxF empower healthcare professionals with comprehensive, interoperable health information exchange capabilities, ultimately improving patient outcomes and enhancing the quality and efficiency of healthcare delivery.

XpanDH is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.

Electronic Health Records (EHRs):

EHRs store some of our most personal information. XpanDH works to enable interoperable cross-border EHR transfer, allowing for European-wide tracking of a patient's journey. Reaching from an overview of her health status at a given time, the medication she is prescribed, her medical images and related reports, lab test results and hospital discharge report (as outlined in Article 5 of EHDS regulation). This holds enormous potential for improving care quality – and to boost research and policymaking. However, this last use purpose must be only acceptable if individual patients cannot be identified through the data. Otherwise, it would constitute a major violation of their privacy.

The EHDS Regulation approved on 24th April 2024, therefore, foresees that Health Data Access Bodies should make health data available for these purposes only in non-personal or, exceptionally, pseudonymised format (EHDS, Recital 49, 50). Yet interestingly EHDS regulation – and GDPR – do not define anonymisation explicitly. They only define "anonymous information" as 'personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable' (GDPR, Recital 26).

The challenging nature of this task is a circumstance that even the EHDS regulation recognises. It acknowledges that even anonymisation might not eliminate the risk of re-identification of data subjects. Even if re-identification is not possible at a given time and only from a given data set, it might become possible if: 1) the dataset is linked to external data sources – such as social media, or data banks from other areas; 2) technology offers new re-identification methods. (EHDS, Recital 64). For EHRs, this risk is even bigger. The EHDS regulation describes them as being part of a category of health data that offer specifically broad characteristics of re-identification (EHDS, Recital 64). For example, the report of a patient's hospital stays, after she broke her leg, might have been entered into a database. For anonymisation, it is not enough to exclude our patient's name. The Health Data Access Body might have altered her data through so-called k-anonymity. Through this method, information like her exact age (36) might have been transformed to an age span (30-40). In the end, her data set will be indistinguishable from at least some other data subjects in the data set.

Nevertheless, it might still be possible to re-identify our patient – especially when linking her anonymised hospital discharge report data to data from her remaining EHR. There might have been six women aged 30-40 discharged from her hospital with a fractured leg that day. But someone with access to her remaining EHR data – including past imaging of her leg – might be able to re-identify her in the dataset. Given the diverse types of data stored in EHRs, there are many ways to combine this data and single out an individual in another anonymised data set. This risk increases even more when linking to data from external sources – e.g. if our patient posted about her hospital stay on social media.

Since EHR data can potentially facilitate the re-identification of individuals in a supposedly anonymized dataset, some stakeholders perceive EHR interoperability – for instance, via the EEHRxF – as problematic. Although the XpanDH team does not share this perspective, we value this significant stakeholder input in WP5. Therefore, this example of health data anonymization in the EHDS aids us in comprehending resistance to the EEHRxF among certain stakeholder groups.

About the authors:

Carola Schulz (XpanDH WP5 lead, empirica); with support from Marina Grossi (empirica) and Radhika Poojara (ECHAlliance)

XpanDH is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.