

## Can patient data from the national eArchive be used for research?

The electronic archive of patient records (eArchive) is a data repository where health care organizations store patient data from electronic health record systems. Although the primary purpose of using data of the electronic health records is direct health care delivery, the data could also be used for secondary purposes such as research. Using patient data from the eArchive for research purposes would enable the exploitation of large data sets, for example in evaluation studies on the effectiveness of treatment methods or in studies on population health and welfare. If in addition to patient data administrative data were also stored in the eArchive, we could also study the efficiency of the health service, or compare the services provided by different health care organizations against each other.

The International Medical Informatics Association (IMIA) encourages European Union Member States to publish a framework to enable large-scale sharing of data, and methods in the meaningful use of data from electronic health records for research and for healthcare service development. Such a framework can guide and facilitate widespread collection, storage, aggregation, linkage and transmission of health data. The framework would provide appropriate protections for legitimate secondary use.

The importance of healthcare data to clinical research is irrefutable. However, across Europe, there is a diversity of cultures, policies, regulations and operational arrangements for access to healthcare data. At the same time, the heterogeneity of health care environments and services across Europe is an opportunity for generating richer clinical pharmacoepidemiology insights. There is a long history of enhanced data collection into population based disease registries in individual EU Member States. In Finland healthcare data is collected and stored in several registries. This is done afterwards using net forms, files taken from information systems and paper forms. In addition to patient data, administrative data is collected in health care organizations. This data should also be able to combine with patient data and this kind of registry materials would enable a broader range of health service research. The collection of data in a structured format allows data to be used for purposes of monitoring the quality of care and population health.

Reusing the data for research entails data standardization, the use of unified terminologies, semantic interoperability of data (exploitation of the data of all different systems), as well as the use of appropriate methods and data supplemented by administrative data, for example, a variety of registration data and patient satisfaction or patients' quality of life data.

The comparability of data from different electronic health record systems affects data content, data structures and data format. The comparability and completeness of the data influence the methods used in data collection and data extraction. Unified documentation methods improve the quality of data.

Increasing public awareness and the transparency of data use is needed to promote the secondary use of health data in research. In discussions the focus must be on data access, use and control, not on ownership. Privacy policies, for example, the anonymisation of data and security issues must be considered.

Further development of the eArchive must take account of reusing data for research. All data which is documented in health care organizations should be stored in the eArchive. Finally, reusing data from the eArchive in research entails that such data be of good quality.

**Kristiina Häyrinen**  
**Editor-in-Charge**