EMERTONDATA

# HEALTHCARE DATA SHARING: FASTEN YOUR SEATBELT

By Aimé Lachapelle & Victor Pinault

## How major healthcare systems are aligning to unleash the power of data sharing to EMEA.

It is commonly acknowledged that **healthcare data has tremendous potential as a tool to improve overall public health** and unlock significant innovation in both the public and private healthcare sectors. Use cases based on the sharing and processing of healthcare data (e.g., machine learning, AI, etc.) are numerous.

Recent notable examples include leveraging data-driven algorithms to improve cancer diagnosis, reduce dangerous drug interactions, and predict the best treatment option based on a patient's specific symptoms, history, and demographics, etc.

As far as the healthcare data is disseminated across all actors, **a key enabler is healthcare data sharing,** on which this paper focuses.

This requires overcoming a range of **distinct challenges, including careful consideration of patient privacy concerns, anonymization of data** without loss of relevance, and the **harmonization of data across disparate sources and IT solutions.**

**These barriers to sharing healthcare data are gradually being broken down, driven by a few major national initiatives across Europe** (Health Data Hub in France, NHS in the UK, MII in Germany) and Israel, which we review in this paper.

If the regulation regarding healthcare data sharing remains at its early stage, **should start-ups and corporates start launching data sharing projects?**

**When should they enter the game? And what is missing to seize the opportunity to increase data sharing, alongside the national initiatives?**

## THE EUROPEAN DIRECTIVE ON DATA SHARING AVOIDS INCLUDING HEALTHCARE WITHIN ITS SCOPE

2019 has seen the adoption of a new directive on open data and the re-use of public sector information (PSI) by the European commission. It introduces the concept of **high value data sets** which are held by public sector bodies and public undertakings and must be made available to support their re-use[1] by a wider audience. To date, these priority datasets include geospatial, earth
observation and environment, meteorological, companies and company ownership and mobility data.

**The current directive notably excludes both healthcare data (avoiding the complexities of personal data) and privately owned data.**

**All of these high value data sets have an important** common characteristic: they do not require extensive anonymization work before being released to the public and they are predominantly owned by public entities; two good reasons for the European Commission not to consider healthcare data within the scope of the first directive.

Sharing the precious public healthcare data of millions of citizens at the European level takes time.

Furthermore, the data is fragmented, redundant and incomplete. It is also owned by private parties who might not be so keen on sharing it for free. Nevertheless, major recent national initiatives are opening the way for a more expansive European level sharing of data.

The private sector has already seen its fair share of data management constraints since the introduction of the GDPR (General Data Protection Regulation), but this will not be the status quo going forward.

The forced sharing of data for private actors such as PSD2 (Payment Services Directive 2) is a vivid example of how sharing can be imposed to corporates (banks in this case), and the next iteration of the GDPR is also in development.

**The public sharing of healthcare data is gaining momentum and European corporates (pharmaceutical companies, medical facilities, insurers, etc.) are not immune from regulated sharing. Nevertheless, healthcare actors can have
a range of business reasons to proactively begin sharing their data.**

## THE ASYMMETRY OF ACCESS TO DATA IN THE HEALTHCARE SECTOR SUPPORTS THE CASE FOR INCREASED SHARING

Health data can be segmented into two main categories: **individual data** (related to the patients' health status, the medical expert's diagnoses payments and reimbursements) and **aggregated data** (related to facilities or practitioners, including overall medical payments & reimbursements).

These data sets are generated by different entities, each with varying levels of access to the same data. Exhibit 1 shows a synthesis of data ownership, even though each country has its specificity.

**Although the use of healthcare data is often of a sensitive nature, use cases demonstrate that data sharing can lead to various benefits, including improvements in public health outcomes, cost efficiency enhancements, and an improved ability to simulate and model healthcare economics.**

Selected prospective use cases are explored in this direction:

| | Type of data | Medical practitioners and facilities | Public and private insurers | Pharmaceutical companies |
|---|---|---|---|---|
| Granular data | **Health status** (e.g. disability, death) | ✓ | ✓ | ✗ |
| | **Diagnoses, prescription, test results** | ✓ | ✗ *Partial knowledge of prescriptions* | ✗ *Partial knowledge of prescriptions via health information providers* |
| | **Overall medical payments & reimbursements** | ✗ | ✓ *No sharing between public and private insurers of each other's reimbursements* | ✗ |
| Aggregated data | **Operational performance** (e.g. quality of Service, procedures, costs) | ✓ | ✗ *A priori no access to medical facilities/practitioners' KPIs* | ✗ |

Access to the data
✓ Full ✗ Limited ✗ None

Source: Emerton Data analysis