Health Data Circulation in France: Between Public Interest and Privacy Enhancing Technologies

Margo Bernelin

CNRS, France

DOI 10.3217/978-3-99161-062-5-016, CC BY 4.0

Abstract. Within the healthcare system, the promises of Privacy-enhancing technologies (PETs) have attracted considerable attention to the point that, in France, personal health data cannot be used for anything other than care if it is not protected by such Technologies. This movement toward exploring more closely 'data circulation-privacy-friendly' solutions emerged about ten years ago in a context where the State was willing to encourage health data circulation for medical research. Indeed, in France, the most important health databases are operated by the State and have the advantage of being comprehensive in terms of population. In a bidding war with other States that were also willing to open their databases for research, the French Government introduced a bill to make the national databases accessible for research. To obtain support for the bill, some Members of Parliament and Senators, but also the French Health Ministry kept putting forward the benefits of Privacy Enhancing-Technologies in protecting health data and focussed on the public interest dimension of sharing health data for research purposes. Analysing the legal landscape and discourse, this paper demonstrates that since 2016, Privacy Enhancing-Technologies have been a key factor in authorising data access alongside the public interest to conduct research. Rather than closing any debate on data privacy, it has actually opened new questions on the efficacy of Privacy Enhancing-Technologies and on what their scope should include.

1 Introduction

With the discussion and recent publication of the Health Data Space Regulation by the European Union (march 2025)[2], the accessibility of health data for research has gained increased attention (Shabani 2022, Aufrechter 2025). Indeed, the new regulation seeks to promote the digitalisation of medical records and other health data by ensuring that all EU citizens have access to an electronic version of these records, or at least part of them (Marelli 2023, Horgan 2022, Mergelin 2024.). Moreover, the regulation requires health data holders, such as hospitals or pharmaceutical companies, to allow others to access their data for research purposes (de Grove Valdeyron 2024). Such a requirement has created significant hope for data sharing for public good but also concerns regarding privacy (Bernelin 2024).

Health data are defined by the 2016 General Data Protection Regulation (GDPR) as 'personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status'[3]. Such data, including medical files, medical screening reports, health insurance data, prescription sheets or event data from health-related IoT devices, are sensitive by nature. Indeed, they reveal very intimate information about one's health issues that he or she might not want to share with anyone. Moreover, the data does not only relates to one's health but also to their family members, as medical history and some hereditary health components or diseases are also included in medical files. In addition, medical records often also include information about employment, family life, hobbies, housing situation or any other element that might have an impact on health.

The variety of data as well as what they can reveal (diseases or their risk of happening, injuries, disabilities, sexual orientation or experiences of abuses) require their strong protection against undue access. In this regard, it is needless to say such data could be used as discriminatory weapons with regards to access to employment or to services such as loans (for instance see Garcia, 2024 p. 2062). Such data can also offer crucial information for malicious individuals that will use it for illegal activities such as phishing, catfishing or even blackmailing (Kleinman 2020, or CybersecurityAsia.net 2024). Their protection is therefore crucial. One way to achieve it is to keep them secret. Physicians, or General Practitioners depending on your location, are already subject to medical secrecy/confidentiality obligations. In France, not only is it a deontological obligation but also a legal one, which requires health care professionals and any other professional

[2] Regulation (EU) 2025/327 of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847.

[3] art. 4 (15) Regulation 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

accessing health information (such as data analysis in hospital or administrative agents) to keep them secret or to face criminal sanctions[4]. To lawfully share patient data with someone else, medical professionals must have compelling reasons such as, in case of an emergency, to save one's life or with the patient's consent. In France, the judges have recently ruled that sharing identifiable health data for teaching purposes with medical students is not legitimate as such and requires, to be lawful, the patient's consent. In this instance, judges held the medical practitioner liable for this breach of confidentiality[5].

Sharing health data is, nevertheless, considered crucial for various reasons. It is, indeed, important to share patient data among a medical team in order to ensure the provision of optimal care for a patient. It is also important, for public health purposes, to have data about a population or specific population groups in order to design and implement health plans and, for public authorities, to make informed decisions. It is also paramount to exploit health data for research purposes, whether the data was collected directly for research or for another purpose, such as patient care. In this regard, various and large datasets are useful, for instance, to predict a patient's trajectory by comparing it with that of another patient. Health data can also be crucial to observe drugs side effects, to train Artificial Intelligence-based Medical Devices used in patient care and more generally for applying *big data* approaches in health to foster the formulation of new hypothesis, detect diseases before they appear or to provide for more personalised treatment (Bernelin and Desmoulin 2020, Mercier 2020). However, processing and sharing health data inherently involve privacy risks.

In this context, ten years ago, the French Parliament authorised the creation of a large and centralised (in Paris) health 'super' database (Information System - IS) in order to make health data available for research purposes. By adapting medical confidentiality rules and data protections laws, the Acts attracted much attention (Bossi Malafosse 2016 and 2016a; Debiès 2016, Desmoulin-canselier 2018, Devillier 2017). Indeed, they have been widely analysed by the legal literature, which has extensively studied the type of data being made available for research (Bernelin 2020, p.26), patients' individual rights (Debiès 2016), the public interest in the circulation of health data (Péchillon 2015, Teller 2022, Pailhès 2018) and the governance mechanism of this Information System (Supiot, 2020). The question of how to protect privacy in this setting was largely left on the back burner, legal scholar envisioning this dimension more as a technical issue that, indeed, required mention (Devillier 2017) but did require deeper analysis and questioning. However, since then and with the entry in force of the General Data Protection Regulation (GDPR), privacy issues have emerged as a focal point in the case law concerning health

[4] Article L1110-4 French Public health Code and article 226-13 French Criminal Code.

[5] Strasbourg Tribunal Administratif, 5th Cahmber, July, 4th 2024 (n°2207563).

data sharing for research[6] with the literature questioning the GDPR application to health-related research on data (Marelli *et al*. 2021). Such a situation begs the following questions: What was the role played, at the time, by Privacy Enhancing-Technologies (PETs) for the regulatory creation of a centralised health Data Space in France? How has this role materialised in law? How, post-GDPR and the creation of multiple Health Data Spaces, are these Privacy Enhancing-Technologies referred to by Judges when other socio-ethical issues, such as public interest in research, are in balance with health data Access?

Answering these questions is crucial to understand how acts were enacted, their wording and their current application. Such answers provide material to support a critical analysis of the regulation of health data access, which appears to be caught between technological imaginaries of protection and public interest in research. To answer those questions, one has to return to the very beginning in order to dissect how the Acts were influenced by privacy issues. Indeed, the legal discourse in the parliamentary debates provides privileged insight into how technological objects and the risk associated with them were framed and balanced against socio-ethical values. Their study is paramount for a better understanding of the rules governing health data access prior to the creation of many Health Data Spaces (Hoeyer *et al.* 2024), including the European Health Data Space (EEDS, March 2025), and for analysing their consequences today for health data circulation. Such analysis is missing from the scholarship that has rather primarily focused on the recent EEDS and its complex articulation with the GDPR and the European Union Artificial Intelligence Act 2024 (De Grove-Valdeyron ed. 2025, Quinn *et al.* 2024), on the utilitarian inspiration behind its regulatory model (Lianos 2025), on the users' journey to access data (Forster *et al.* 2025), on Member States' preparedness to implement it (Kessissoglou *et al.* 2024).or on the geopolitical dimension of data access under the EEDS (Donia & Marelli 2025).

Against this backdrop, from an methodological point of view, we examined the parliamentary debates (plenary discussions) and parliamentary work publications (MPs' and Senators' and institutional reports on bills) dedicated to the 2016 and 2019 health data access reforms. We also paid attention to the recent case law on health data access for research in order to get an understanding of and compare how judges refer to Privacy-Enhancing Technologies and balance it the public interest nature of such access. This discourse analysis demonstrates that within the 2016 and 2019 health data access reforms (**part 2**), Privacy Enhancing-Technologies were strongly promoted as the solution for health data circulation for research, to the point that, it was even anticipated in 2016 that health data would be made available in an open access format thanks to

[6] Few recent examples: Conseil d'État (CE), 25th April 2025 n°503163, CE 13th November 2024 n°492895; CE 19th October 2024 n°491644; CE 22nd March 2024 n°492369; CE 9 March 2023 n° 468007; CE 23rd November 2022,n° 456162.

protective Privacy Enhancing-Technologies (**part 3**). However, another key condition was also established and further elaborated in the following years: the requirement of public interest under which health data would be shared for research only if it benefited public interest such as for research purposes. As a result, Privacy Enhancing-Technologies and public interest narratives remain the framework guiding health data circulation for research in France. When privacy risk changed in nature, public interest discourses were used in the case law to authorise research regardless, thereby calling into question PETs as regulatory optimum (**part 4**).

2 The 2016 and 2019 French Health Data Access Reforms

In 2016, the French parliament identified the need to make health data available in open access to better inform the population about health and to make public decision. Such publication required to anonymised health data in order to protect patient's privacy (**2.1**). Anonymisation refers to methods that allow for the complete and irreversible de-identification of personal data, through the removal of direct and indirect identifiable information in data, or by using more sophisticated and often combined technical measures such as data synthesis, differential privacy and other obfuscation tools (for instance Sella *et al.* 2025). Under Privacy laws, anonymised data is no longer considered personal data and is therefore not subject to protection requirements. In comparison, pseudonymisation technics offers less privacy protections, removing identifier but leaving the data potentially open for re-identification. The 2016 Statute, therefore, introduced a dual system where some data would be available in open access as anonymised data while the remaining one would be accessible after pseudonymisation within an Information System (**2.2**). The 2019 Statute, on its part, expanded even more the System and reformed its governance (**2.3**).

2.1 Health Data: The Need for Open Access

In the 2010s in France, the need to access health data for public information and research purposes became paramount. Reports were solicited by the Government from various actors to pave the way for a health data circulation plan, and open access emerged to be the preferred option to do so. The Open data in Health Commission's report, published in 2014, defined open data as the openness and sharing of data published online in open format allowing for unrestricted and free re-use by anyone[7].To justify the need to organise such an open access in health, the report underlined that the former procedure for data access was too complex and lacked clarity for researchers, which was detrimental to

[7] Rapport Commission Open Data (2014), Ministère de la santé, p. 9 (https://drees.social-sante.gouv.fr/IMG/pdf/rapport_final_commission_open_data-2.pdf).

public health. On the other hand, the report indicated that the gathering of anonymised health data and its processing would constitute a source of progress for knowledge and value creation (p.37). Moreover, the publication in an open format would contribute to build a stronger health democracy by enabling patients to access data about the healthcare system and thereby empowering them to make more efficient decisions.

With those positive arguments, one can almost forget that a crucial factor presiding the adoption of the 2016 Act was also the tragic Mediator scandal in France. In the years prior to the reform, the use of the Mediator drug from Servier Pharmaceuticals for weight loss purposes, which was outside of its marketing authorisation, led to numerous patients' death or health issues that chocked the nation (Roure 2012) and abroad (Chrisafis, 2013). In France, it prompted discussions about more effective whistleblowing procedures, but also on how complicated access to health databases, which would have been crucial to shed light on the issue and provide material evidence to patients and their families, actually was (see Brasselet 2018, p. 339). In response the French Government at the time proposed to introduce provisions before Parliament that would facilitate health data access through open access: the future 2016 Act.

2.2 Health Data: The 2016 Act

The 2016 Health Act is a very large legislative text that introduced crucial rules for health data access for research. The new Act provided that anonymised health data should be accessed in an open format. As a consequence, no licence nor authorisation would be necessary to access it. The new Act also created the National Health Data System (NHDS), a large and centralised information system (a collection of databases) that was considered valuable for research purposes. Indeed, the NHDS encompasses the National Health Insurance database, which compiles all data on care reimbursement whether related to medical acts or prescribed drugs. This very large database contains data on more than 65 million patients (Moulis 2015)! However, the NHDS does not stop here, as it also includes the database dedicated to hospital activities and the national register of deaths for all French citizens. Under the new Act, such data would be made available for research purposes once pseudonymised and under the governance of a new institution: the National Health Data Institute.

The French parliament determined that six purposes could justify access to such a 'national treasure', as it now often referred to (Sénat 2023, Belot 2020):

- to provide information on health, healthcare provision, social care and on their quality;
- for the definition, implementation and evaluation of health and social protection policies;
- in order to understand health expenditures;

- to inform healthcare and social care professional on their activities;
- for health surveillance, monitoring and for safety;
- for research, studies, evaluation and for innovation in the field of health and social care[8] .

From a legal perspective, the evolution of the rules governing health data access - that is personal data (as opposed to anonymised data) - was worth analysing. However, the open access provisions of the Act remained largely unexamined. More generally, legal scholars, at this time, did not investigate the use of non-personal data for research considering, perhaps, that there was nothing to report on the subject. This, perhaps, could be explained by the anticipated application of another text adopted in 2016: the General Data Protection Regulation from the European Union. This much-awaited text was to be articulated with the 2016 Health Act on Health Data Access which drew considerable the attention in the literature. Such a coverage of the 2016 Act likely had the effect of diverting attention from Privacy-Enhancing Technologies, especially anonymisation techniques. Such a disregard for Privacy-Enhancing Technologies by the legal literature took also place in the analysis of the 2019 Health Act that also covered health data access for research.

2.3 Health Data : The 2019 Act

In 2019, the 2016 Act appeared to be underperforming. Indeed, the new NHDS did not function as smoothly as anticipated, and data access procedure remained lengthy for pseudonymised data[9]. To gain efficiency, the 2019 Act made changes on the governance side of the System in order to create the *Health Data Hub*, an agency designed not only to deliver practical access to the data but also to provide support for researchers seeking to use the System. The *Hub* emerged as the centre piece of the reform, its name in English chosen to attract researchers from around the world. However, the new governance was not the only innovation. Indeed, the 2019 Act also expanded the System, adding more databases to the list of available resources. Article L. 1461-1 of the French Public Health Code now provides that merely all data collected for a healthcare acts reimbursed by the French National Health Insurance Systems could be added to the NHDS. What an expansion!

The 2019 Act is more prolific regarding Privacy-Enhancing Technologies and emphasised that health data from the NHDS should be made available once pseudonymised as detailed in a security framework (referential) for personal data

[8] Article L. 1461-1 French public Health Code in its 2016 wording.

[9] Étude d'impact, Projet de loi relatif à l'organisation et à la transformation du système de santé, 13th February 2019, p. 89.

protection[10]. According to the Act, the main feature of this techno-legal framework should include pseudonymisation elements and access traceability measures. In its opinion on draft of the 2019 Act, the French Data Protection Authority, CNIL, cautiously indicated that Privacy-Enhancing Techniques should be sufficiently robust to protect such a large gathering of data (CNIL, 2019). While those elements of the text are central to protect privacy, they were, again, left out of most academic analysis or only briefly mentioned (Supiot 2020, Robin 2021) as a mere technical information. It must be conceded that the term 'Privacy-Enhancing technologies' had not yet emerged in 2016 and 2019 as the crucial notion it is now for describing tools and approaches aimed at preserving privacy. In this regard, the OECD report on Privacy Enhancing-Technologies, published in 2023, marked a turning point in public understanding of privacy solutions, therefore the term was not mentioned by the legal literature at the time, but the technics, mentioned in the Acts and framework, were not dissected even though the privacy discussion was crucial for securing the adoption of those Acts, as privacy concerns were central to the Parliamentary debates.

3 Privacy Enhancing-Technologies for Health Data Sharing: a Crucial Discussion

What role did Privacy Enhancing-Technologies play in the passage of 2016 and 2019 Acts? The analysis shows that anonymisation was central in supporting open access and facilitating the adoption of the 2016 Act (**3.1**), while also raising questions about the role that the Data Protection Authority should have in this scenario (**3.2**). Conversely, pseudonymisation was crucial for the circulation of identifiable data (**3.3**). However, Parliamentary debates revealed confusions regarding the distinction between anonymisation and pseudonymisation, leading to the impression that Privacy Enhancing-Technologies were also used to advance an agenda favouring data circulation rather than being critically assessed for their actual capabilities and limitations (**3.4**).

3.1 Anonymisation and the Open Access Discussion in Reports

Chapter 5 of the 2016 Health Act was dedicated to health data access and, interestingly, was entitled 'Enabling health data open access'. Very few scholars have noticed the significance of this title (with the exception of Robin 2021 or Péchillon 2015). In the wake – and in the shadow - of the Mediator scandal, 'health data open access' was presented as the Government's new open-data venture following the publication of various other public databases. Under the 2016 Act, article L. 1461-2 of the Public Health Code

[10] Loi n° 2019-774 du 24 juillet 2019 relative à l'organisation et à la transformation du système de santé, article 41.

provided, first of all, that the data from the NHDS would be made available for the public as aggregated data or in a way that prevents the direct or indirect identification of individuals. In other words, only anonymised data (therefore non-personal data) would be made available to the public from the newly created system. Such an option was considered feasible.

In the draft presentation of the 2016 Act before Parliament, the Health Minister (in charge of the text) was clear: ' We are allowing our nation to join the broad open data movement. It is our duty to promote health data for the collective good in the strict respect of privacy'.[11] In response, one Member of Parliament remarked that the draft did not demonstrate a robust technical oversights of health data open access.[12] When the question arose regarding the type of actors that could access the data, the answer was clear: no one would access identifiable data. Therefore, the technical protection of health data became an argument within the other crucial debate: determining who should have access to this data.

The 2013 *Bras Report* dedicated to health data, on which the provisions were based, was, indeed, optimistic that anonymisation techniques could be robust enough to ensure privacy, as the risk of reidentification would be residual (Bras and Loth 2015, p.29). The report indicated, at the time, that it would be very difficult for researchers to extract anonymous data and then compare them with other datasets in order to re-identify individuals (p.28). That particular risk was even judged 'less foreseeable' (p.28). However, the report, as well as the 2016 Act, remained unclear regarding what type of data could be shared openly with a limited re-identification risk. The 2014 *Health Data Report* attempted to provide clarity by listing databases that should be made open. The list included the swimming-pools water quality database and the drug consumption in hospital database (in volume), but it also listed the single parent benefit database (p.45). While the two first quoted databases contain non-personal data, the third one does and therefore would require the use of anonymisation tools. The risk of listing very different databases that are subjected to different privacy risk created much uncertainty on what privacy risks' exposure encompassed.

3.2 Anonymisation and the Data Protection Authority's Role

The parliamentary discussions centred on the techniques that could be used to make health data accessible in an open format. Report number 233 from the Parliamentary Joint Commission (a commission composed of MPs and Senators in cased of disagreement between the two chambers on a bill) evidenced these discussions and

[11] Translation by the author. Marisol Touraine, audition AN, 17 mars 2015 https://www.assemblee-nationale.fr/14/cr-soc/14-15/c1415034.asp

[12] Jean-Pierre Door, AN, 17 mars 2015 https://www.assemblee-nationale.fr/14/cr-soc/14-15/c1415034.asp

stated the necessity to have the French Data Protection Authority (CNIL) to provide clarity by listing robust anonymous technics that could be used to protect health data for open access: «The most sensitive personal data will only be accessible in open data after the application of complete anonymisation procedures declared compliant by the Data Protection Authority (CNIL), and any infringement of the ban on using open data in health for the purposes of identifying an individual may be subject to sanctions» (Translation by the author, CMP 2015).

'The most sensitive personal data' is a concept that contributed to blurring the line regarding the type of data being processed, given that health data are inherently sensitive by nature (without graduation). While the option of a CNIL seal of approval was contemplated by some MPs, it was later rejected as too complex to implement, since no single methodology existed for that purpose (CMP 2015). Despite this difficulty, the Government reassured MPs by suggesting that a later bill would introduce the possibility for the CNIL to list reliable and robust anonymisation techniques. This promising idea of the possibility to establish a list of sound and operational anonymisation idea of mandating the CNIL to list anonymisation technics did not materialise and the Government never brought it back again.

From the reports paving the way to the 2016 Act and its parliamentary discussions, we can conclude that, overall, Open Health Data was pursued at a time where health-related data - whether personal or not - was not easily accessible to the for informational purposes or to researchers. Indeed, in this context it felt paramount to ensure data access for all to facilitate a better understanding of health and healthcare related issues, both for citizens and public decision-making. However, by not sufficiently distinguishing between databases that included personal data and those already composed of statistical or aggregated data, the law created uncertainty regarding on the frontiers of open health data and the feasibility of anonymisation.

3.3 Pseudonymised Data and the Circulation Discussion

The 2016 Health Act provided that when data could not be anonymised, then it was to be made available under another Privacy-Enhancing Technologies tool: pseudonymisation. This was intended to be a residual measure, as the parliamentary reports on the Act demonstrated: 'When health data cannot be completely anonymised, their access for research, studies and public interest evaluation, will be limited and regulated by privacies guaranties' (Commission des Affaires Sociales 2015). The then-adopted article L. 1461-4.-I of the public health code stated that 'the NHDS does not provide individuals' first and family names or their social security number nor their address'. In order to enhance privacy further, the same article indicated that the medical practitioners' own identification numbers would be kept in a separate database. Indeed, Members of Parliament considered it too risky to keep all data together: the larger the dataset, the easier reidentification becomes.

In 2017, the Framework applicable to make data available for research in a pseudonymised form was adopted as a Ministerial Order[13]. The document provided that pseudonymisation should be irreversible and based on robust cryptographic techniques. Two rounds of pseudonymisation should take place, the first pseudonymisation should occur when data was transferred to the NHDS (level 1 pseudonymisation) and another round of pseudonymisation when the data was made available for researchers (level 2 pseudonymisation). Moreover, all accesses to data were to be logged in an access journal in order to identify who access it and when.

Despite privacy protection being adopted and enforced, the 2019 Act' s discussions demonstrated that privacy issues remained of importance. Indeed, while health data sharing for research was no longer questioned as such, the ways to protect privacy remained central to the new Act. Agnès Buzin, the then Health Minister, indicated during the debate that the most rigorous protection for Health Data was considered paramount by MPs.[14] In front of the Senate, she further indicated that access to data should be made through privacy-preserving approaches.[15] By emphasising on technical requirements for data sharing, the Minister did not question the necessity of data sharing, nor did she challenge the rights that individuals should retain over their data.

Even more privacy protections were added to the 2019 Act in order to prevent re-identification. Indeed, while in 2016, the law authorised to reidentify individual from the database if one could fear a significant health risk, to enrol the person in a trial or because it was necessary for the study (former article L. 1461-4 CSP), the 2019 Act suppressed this possibility to ensure that the pseudonymisation process remained irreversible.[16] The discussion, nevertheless, lacked clarity, the Health Minister indicating that the data from the HDH was now anonymised (Senate hearing 6, June 2019) even though it was actually pseudonymised, with a stronger prohibition on reidentifying patients.

However, the 2019 debate shifted focus away from open data to concentrate on access to identifiable data. Indeed, in practice, health research requires the most complete datasets possible in order to obtain sufficient elements to draw valid and useful conclusions. The trade-off between privacy protection and utility is very real and while open access health data is useful, the 4P approach to medicine (personalised, preventive, predictive, participatory) requires more than aggregated or statistical data. As a result, the way to ensure a smoother access to health data was addressed in 2019

[13] Translation by the author. Arrêté du 22 mars 2017 relatif au référentiel de sécurité applicable au Système national des données de santé

[14] Agnès Buzin, Assemblée Nationale, 18th March 2019

[15] Agnès Buzin, Sénat, 6th June 2019, https://www.senat.fr/seances/s201906/s20190606/s20190606_mono.html .

[16] Étude d'impact, Projet de loi relatif à l'organisation et à la transformation du système de santé, 13th February 2019, p.92.

without challenging the principles provided by the 2016 Act that – namely, data secrecy, confidentiality, traceability of access, and public interest reason to access data, the latter often being weighed against Privacy Enhancing-Technologies protections.

4 Privacy Enhancing-Technologies and Public Interest: a Dynamic Duo

The literature on the 2016 and 2019 Health Acts also emphasised that research projects will need to tick the 'public interest' box in order to gain access to the data. However, the notion is hardly defined by law (**4.1**) and is therefore left to interpretation by agencies and courts (**4.2**). Courts have, indeed, played a crucial interpretative role by establishing a balance that research project shall achieve between privacy protection and public interest, with the presence of the latter being enough to authorise research project when privacy protection were in doubt. Such a practical application of the statutory provisions quite differs from what Parliament foreseen in 2016 and 2019, namely, having two strong criteria that would not be watered-down by the other one (**4.3**).

4.1 The Public Interest Criterion in the 2016 and 2019 Acts

The 2016 Act introduced that when pseudonymised data is to be accessed, the access proposal should be justified by a 'public interest' dimension. In the parliamentary discussion in front of the Lower House (*Assemblée Nationale*), this 'public interest' criterion was deemed crucial when access was to be granted for identifiable data (*nominative* in French).[17] Therefore, the 2016 Act introduced a balance between privacy risk and protections on one hand, and the public interest of researches, on the other. Such a safeguard was intended to protect against undue access to health data, for instance by insurance companies seeking to increase premium or by other companies aiming to target advertising in the healthcare domain. However, public interest and what it entails is not defined or elaborated in the law, and, as others have noted public interest is a hard concept to grasp (Morlet-Haïdara 2022). One way to recompose its content is to examined the details in the French Public Health Code on what type of research could be made using the System. In this regard, Article L. 1461-1 III states that research project that will be allowed to access the NHDS must contribute to:

'1° The Information on health, healthcare provision, medical and social care and their quality;

2° The definition, implementation and evaluation of health and social protection policies;

[17] Parliamentary debates 19th march 2015 https://www.assemblee-nationale.fr/14/cr-soc/14-15/c1415040.asp .

3° Keeping track of healthcare expenditure, health insurance expenditure and medico-social expenditure;

4° The Information of health and medico-social professionals, structures and establishments about their activities;

5° Health surveillance, monitoring and safety;

6° To research, studies, evaluation and innovation for health and medico-social care'.

We can argue that these six borad purposes together give substance to the notion of public interest and clarify what is expected even though, in the law, the public interest criterion is an additional requirement to this list.

Article L. 1461-1V of the Public Health Code guides us further in defining the frontier of the 'public interest' criterion as it specifies the instances where access to the NHDS is prohibited. Indeed, access will not be granted to the System when the intended research seeks to:

'1° Promote medical products to health professionals or health establishments;

2° Exclude cover under insurance contracts or changes to insurance premiums for an individual or a group of individuals presenting the same risk.'

From the Parliamentary debate and the provision adopted in the Public Health Code, we can conclude that the public interest criterion does not, *de facto,* exclude private actors from the System, such as pharmaceutical or insurance companies, but rather require them not to use the System in a way that would negatively affect individuals or groups of individuals. The notion of public interest therefore aligns with a form of common interest that does not conflict with individual interest.

However, this criterion still lacks content and required both the HDH in charge of assessing proposals to access the NHDS, but also the Data Protection Authority (CNIL) and Courts to interpret it.

4.2 The Interpretation of the Public Interest Criterion.

The practice of assessing data access proposals led different actors to give life to the public interest criterion set out in the law. For the Health Data Hub (HDH), the notion of public interest should be understood as requiring research proposals submitted to tick the following boxes:

- « The aim of the project should be clear, intelligible and truthful
- The benefit of the project should be direct/indirect for groups of individuals, the society or the scientific community (bettering the healthcare system, research, increase of knowledge, etc.)

- Effort should be made toward transparency and result's publication of results, but also documentation, software used and link to public repositories
- Step should be taken to ensure scientific integrity measures, the quality of studies and prevent bias in results, to ensure the implication of research professionals, to put in place a proper scientific governance, to open [research] outcomes or methods in order to foster discussion and results' verifiability ».[18]

For the HDH, the public interest is therefore more about scientific integrity, transparency and openness, than about a precise field of research such as those described in the Public Health Code. The HDH goes further and specifies that the public interest criterion is not incompatible with commercial interests, underlining the fact that private actors can also access the System. However, the question of private actors' access to such data was not fully resolved and led the Data Protection Authority (CNIL) and the courts to,provide their own interpretation of the public interest notion.

Indeed, the CNIL and the *Conseil d'État* (the French highest administrative court) rejected access to the NHDS by journalists who wanted to publish an article on the list of 'best hospitals in France'. While in 2015, before the passing of the Act, the French Health Minister reassured that journalists would be able to access the NHDS for such studies as their work serves public interest[19], the CNIL did not felt as positive. For the CNIL and the *Conseil d'État* (that was later solicited with a demand to withdraw the CNIL's opinion), the methodology behind the data access proposal as well as the anticipated findings were questionable and lacking public interest.[20] According to the CNIL, the results would have negatively affected public information. In this light, we can conclude that the public interest criterion is assessed based on the scientific rigor of the methodology proposed to access the data. Accordingly, the CNIL and *Conseil d'État*'s decisions underline the need for a methodology strongly grounded within the scientific literature.

4.3 Public Interest *vs.* Privacy Enhancing-Technologies: a Question of Balance?

Since 2020, when it comes to public research, the CNIL and especially the *Conseil d'État*, both seem to weigh two different and, as demonstrated, equally important criteria: public interest and the use of Privacy Enhancing-Technologies, to the point where Daniel Kadar and his colleagues raised the following question: 'What if, in the Covid-19 era, health took precedence over the protection of health data?' (translation by the author, 2021) In other words, what if public interest in research were to trump privacy risk in the authorisation

[18] Translation by the author. HDH website : https://www.health-data-hub.fr/interet-public .

[19] Parliamentary hearing, 3th March 2015, https://www.assemblee-nationale.fr/14/cri/2014-2015/20150194.asp.

[20] CE, 30/06/2023, n° 469964, Sté d'exploitation de l'hebdomadaire Le Point; CNIL n° 2022-103, 20th October 2022.

of health data access? Such a question strongly emerged as the *Conseil d'État* in the above-mentioned case law[21] always authorises health data access when privacy doubts exist on the ground that the risk might only be residual while public interest into research commands to authorise access to data. The latest case law is an example of such findings and related to the creation of a 'super health database' for research purposes.

Indeed, in February 2025 the CNIL authorised the creation of a large database derived from the NHDS for the DARWIN EU project that is led by the European Medicines Agency for the EU. The project aims, in France, to create a database concerning 10 million patients in order to make the data available for the purpose of 'routinely estimate the prevalence and incidence of drugs and vaccines use in France using standardised indicators' and the prevalence or incidence of certain pathologies.[22] For the CNIL, such purposes satisfy the public interest criterion and the Privacy Enhancing-Technologies solutions applied appear robust: two level of pseudonymisation for the data and the erasure of the initial dataset within 3 months from its creation. However, the robustness of such protection was raised in front of the judges as data storage will be handled by an American Company: Microsoft.

For the HDH, in charge of the DARWIN EU project in France, Microsoft was the only available choice, no other French or European contractor being capable of storing such a large quantity of data. However, this choice raised concerns for charities dedicated to health as the company operates under Unites States of America (USA) Law. Indeed, since 2001, the USA laws authorised its federal agencies to access data held by their national companies. As a result, even if the data is stored in Europe, the US security agencies can still gain access to it, which would constitute an undue access for French citizens' personal data. For the plaintiffs (charities), the volume of data is so important that its double pseudonymisation would not be sufficient to protect the individuals' identities if federal US agencies were to access it.

For the judges and the CNIL, such a risk is deemed to be residual. The judges do not provide further details, but emphasised the public interest nature of the DARWIN EU project. Point 7 of the ruling states that there is a 'public interest in not delaying the carrying out of the studies relating to the estimation of the incidence and prevalence of pathologies in the general population planned as part of the DARWIN EU project' (translation by the author). Thus, the public interest prevails even when a risk of access exists, provided it is assessed as residual. In this regard, a previous decision of the *Conseil* d*'État* in 2024, on equivalent fact stated that 'it cannot be entirely ruled out that the authorised data processing, which is particularly sensitive in view of its nature as

---

[21] Conseil d'État (CE), 25th April 2025 n°503163, CE 13th November 2024 n°492895; CE 19th October 2024 n°491644; CE 22nd March 2024 n°492369; CE 9 March 2023 n° 468007; CE 23rd November 2022,n° 456162.

[22] CNIL (2025) : Deliberations n°2025-013 et 2025-014.

health data and of the scientific and economic potential of its use, may be the subject to access requests by US authorities, on the basis of the laws of that country, *via* the intermediary of the host's parent company'.[23] Therefore even when privacy risk remains, the public interest nature of proposed research positively influences the outcome of the ruling, raising questions about the original intentions of Parliament. Some might argue that privacy issues and public interest considerations are equally taken into consideration, but the rulings consistently appear to downplay privacy concerns. The reason might be perhaps because in 2016 and 2019, Parliament mostly anticipated private criminal ill-intent toward data access but not foreign State access to it and therefore courts are less able to tackle this issue or do not perceive it as a pressing one.

5 Conclusion

In 2016 and 2019 the French Parliament adopted new Acts that organised a System to make health data available for research. The study of the role played by Privacy-Enhancing Technologies in the parliamentary debate demonstrates that privacy concerns were very important for MPs. Indeed, health data were only to be shared under strong Privacy Enhancing-Technologies approaches such as robust anonymisation and double pseudonymisation schemes. Such protections were designed to prevent undue access to health data by private actors such as companies and illegal data re-identification of patients. The 2016 and 2019 Acts also introduced another criterion for data access: public interest. While this public interest remains evanescent by nature, it seems to refer both to the scientific quality of the research proposal and to its objective of benefiting the society.

In practice, the case law in front of courts demonstrates that the judges give considerable weight to the second element, public interest, over the first. Indeed, the limits of the Privacy Enhancing-Technologies are not thoroughly discussed even though they could be, as the cases often concern thousands and even millions of patients. Such a vast quantity of data could make re-identification technics and attacks easier. However, parliamentary debates had evidenced that privacy issues were imitatively linked to researchers and private companies' access. The case law shows that privacy risk is more complex, intertwined, and not restricted to private interests. Here foreign platforms such as Microsoft create new privacy risks, as their national government might have an interest in the data. Faced those new risks, courts are reluctant to engage in diplomatic rows and prefer focussing on the anticipated research's public interest in order to authorise their conduct especially since there is no national digital platform in France that has the material capacities to store such data volumes.

[23] Translation by the author. CE, n°491644, 19/10/ 2024

In order to address this new privacy challenge, the French Parliament adopted in 2024 a new Act that requires the NHDS and the HDH to use a French or and EU-based cloud for data storage within 18 months (art. 31 I)[24]. However, that deadline was postponed by the fact that some delegated legislation, that has not yet been published, are required to introduce more detailed provisions about the storage modalities. As a result, application of this new requirement has unfortunately been delayed leaving for a now an uncomfortable *status quo*.

[24] loi n°2024-449 visant à sécuriser et à réguler l'espace numérique.

References

Aufrechter, Cyril (2025): Le règlement européen des données de santé est publié ! In Dalloz actualité 20/05/25.

Belot, Laure (2020): Les données de santé, un trésor mondialement convoité. In Le Monde, March, 3rd 2020.

Bernelin, Margo (2024): Anonymisation des données et cybersécurité en santé : un droit hésitant ? in Droit, Santé et Société 67 (2), pp.20.

Bernelin, Margo (2020): Données massives et santé publique: entre redefinitions et ruptures normatives. In ADSP n°112, p.25-27.

Bernelin, Margo, Desmoulin, Sonia (2020): Données massives et santé publique. In ADSP n°112, p.2.

Bossi Malafosse, Jeanne (2016): Les nouvelles règles d'accès aux bases médico-administratives. In Dalloz IP/IT p.205.

Bossi Malafosse, Jeanne (2016a): La donneé de santé dans les systèmes d'information: du soin à la santé publique. In Communication Commerce électronique n° 10 étude 18.

Bras, Pierre-Louis, Loth, André (2014): Rapport sur la gouvernance et l'utilisation des données de santé, Drees.

Brasselet, Renatto (2018): La circulation de la donnée à caractère personnel relative à la santé: disponibilité de l'information et protection des droits de la personne, PhD Thesis, https://hal.science/tel-02188518v1.

Chazard, Emmanuel (2020): Big Data, data reuse en santé: un chemin semé d'embpuches nécessitant une approche pluridisciplinaire. In ADSP n°112, p.51-53.

Chrisafis, Angelique (2013): France shaken by fresh scandal over weight-loss drug linked to deaths. In The Guardian, January, 6th 2013.

CNIL (2019): Délibération n° 2019-008 portant avis sur un projet de loi relatif à l'organisation et à la transformation du système de santé (demande d'avis) n° 19001144.

Combes, Stéphanie (2022): Le Health Data Hub, levier pour la valorisation des données de santé. In Annales des Mines - Réalités industrielles, Août (3), pp. 59-62.

Commission Mixte Pariataire (CMP 2015), Rapport Projet de loi de modernisation de notre système de santé, n°233.

Commission des affaires sociales (2015), rapport ,°3215, 10th December 2015.

CybersecurityAsia.net (2024): Cyber Predators Target Vulnerable Victims: Hackers Blackmail Hospitals, Trade Patient Data and Find Partners Through Darknet Ads', sept. 25th 2024 (https://cybersecurityasia.net/hackers-target-healthcare-darknet-ads/).

Debiès, Elise (2016): L'ouverture et la réutilisation des données de santé: panorama et enjeux.In RDSS p.697.

Desmoulin-Canselier, Sonia (2018): L'évaluation des médicaments à l'ère de la médecine des données. In RDSS p.1043.

Devillier Nathalie (2017): Chapitre 6. Les dispositions de la loi de modernisation de notre système de santé relatives aux données de santé. In Journal international de bioéthique et d'éthique des sciences 28(3), pp. 57-61.

Donia, Joseph & Luca Marelli (2025): Anticipating ethical and social dimensions of the European Health Data Space: A rapid systematic review. In Health Policy, Volume 162,105443.

Forster Rachel *et al.* (2025): User journeys in cross-European secondary use of health data: insights ahead of the European Health Data Space, In *European Journal of Public Health*, Volume 35, Issue Supplement_3, pages iii18–iii24.

Garcia, anna Cristina Bicharra., Garcia, Marcio Gomes Pinto. & Rigobon, Roberto (2024): Algorithmic discrimination in the credit domain: what do we know about it?. In *AI & Soc*, n° 39, 2059–2098.

de Grove-Valdeyron, Nathalie (2024): Espace européen des données de santé: enjeux et défis pour l'utilisation secondaire des données de santé. Entre gouvernance des données et interlligence artificielle: quelle place pour la poursuite de l'intérêt général ?In Obavia, pp.18-23.

de Grove-Valdeyron, Nathalie ed. (2025): *Espace Européen des Données de Santé et IA. Toulouse*: Presses de l'Université Toulouse Capitol.

Horgan, Denis, *et al.* (2022): European Health Data Space—An Opportunity Now to Grasp the Future of Data-Driven Healthcare. In Healthcare 10, no. 9: 1629.

Hoeyer, Klaus, *et al.* (2024): Health in data space: Formative and experiential dimensions of cross-border health data sharing. In Big Data & Society, 11(1).

Kadar, Daniel, Abdesselam, Stéphanie, Gaillard, Laetitia (2021): Données de santé : un vecteur d'innovation sous trop haute surveillance ?. In La Revue des juristes de Sciences PO n° 21, p.10.

Kessissoglou, Irini *et al.* (2024): Are EU member states ready for the European Health Data Space? Lessons learnt on the secondary use of health data from the TEHDAS Joint Action. In European Journal of Public Health, Volume 34, Issue 6, pages 1102–1108.

Kleinman, Zoe (2020):, 'Therapy patients blackmailed for cash after clinic data breach', BBC.com, oct. 26th 2020 (https://www.bbc.com/news/technology-54692120).

Lianos, Ioannis (2025): Access to Health Data, Competition, and Regulatory Alternatives: Three Dimensions of Fairness. In Journal of Competition Law & Economics, nhaf016.

Marcus, J. Scott; Martens, Bertin; Carugati, Christophe; Bucher, Anne; Godlovitch, Ilsa (2022): The European Health Data Space. IPOL- Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament Policy Department studies.

Marelli, Luca *et al.* (2021): Big Tech platforms in health research: Re-purposing big data governance in light of the General Data Protection Regulation's research exemption', In Big Data & Society, 8(1).

Marelli, Luca *et al.* (2023): The European health data space: Too big to succeed? In Health Policy V.135, 104861.

Megerlin, Francis (2024): Espace européen des données de santé : portée de la proposition de règlement. In Reccueil Dalloz 119.

Mercier, Sandra (2020): Médecine Génomique; vers une médecine predictive. In ADSP n°112, pp.30-32.

Morlet-Haïdara, Lydia (2018): Le système national des données de santé et le nouveau régime d'accès aux données. In RDSS p.91.

Morlet-Haïdara, Lydia (2022): Problématiques juridiques posées par le Big Data et les outils institutionnels de la recherche en santé. In Santé Publique, . 34(3), 335-344.

Moulis Guillaume *et al.* (2015): French health insurance databases: What interest for medical research?. In La Revue de Médecine Interne, Volume 36, Issue 6, 2015, pp. 411-417.

OECD (2023): Emerging Privacy-enhancing technologies Current regulatory and policy approaches.

Pailhès, Bertrand (2018): Comment définir et réguler les « données d'intérêt général » ?. In Annales des Mines - Enjeux numériques, 2(2) pp. 39-43.

Pechillon, Éric (2015): L'accès ouvert aux données de santé : la loi peut-elle garantir tous les risques de dérives dans l'utilisation de l'information ?. In L'information psychiatrique, 91(8), pp.645-649.

Quinn, Paul and Erika Ellyne, Cong Yao (2024): Will the GDPR Restrain Health Data Access Bodies Under the European Health Data Space (EHDS)?. In Computer Law & Security Review, Volume 54,105993.

Robin, Agnès (2021): Chapitre 3. Open data et santé : quelles modalités pour la diffusion et l'exploitation des données de santé ?. In Journal international de bioéthique et d'éthique des sciences 32(2) pp. 33-44.

Roure, Thomas (2012): L'affaire Mediator : retour sur 18 mois de scandale. In *Le Monde*, 14th May 2012.

Shabani, Mahsa (2022): Will the European Health Data Space change data sharing rules? In Science 375,1357-1359.

Sella, Nadir *et al.* (2025): Preserving information while respecting privacy through an information theoretic framework for synthetic health data generation. In npj Digit. Med. 8, 49.

Sénat (2023) : Données de santé, rapport d'information n° 873 (2022-2023).

Supiot, Elsa (2020): Du secret médical à la mise à disposition des données de santé - le Health data hub. In Revue des contrats, 112, p.94.

Tamba, Julie (2025): Interopérabilité des dossiers médicaux : ce qui change avec l'espace européen des données de santé. In Journal de droit de la santé et de l'assurance maladie, 43.

Teller, Marina (2022): La régulation des données de santé : entre intérêt général et intérêts particuliers Introduction au cahier spécial. In Revue internationale de droit économique, t.XXXVI(3) pp. 5-11.