Volume 7 | Number 3

EHPL 3 2023

EHPL

EUROPEAN HEALTH & PHARMACEUTICAL LAW REVIEW

Articles

Reports

The Legal Publisher
lexxion

Conducting a Data Protection Impact Assessment in Health Science: A Comprehensive Guide

*Marcelo Corrales Compagnucci, Alan Dahi and Peter Alexander Earls Davis**

This article provides a guide to conducting a data protection impact assessment (DPIA) for data sharing within health science research. Given the sensitivity of data in health sciences, a DPIA is vital to ensure adherence to data protection regulations and safeguard individual rights and privacy. This guide outlines the core components of a DPIA, including defining its purpose and scope, evaluating the necessity of data processing activities, gauging potential risks, and strategizing effective risk mitigation. By demystifying the DPIA process, this article empowers researchers and stakeholders to execute responsible and ethical data practices in line with the General Data Protection Regulation (GDPR) standards. Additionally, it offers practical examples, tools and resources to enhance the efficiency of conducting DPIAs in health science projects.

I. Introduction

Health science research inherently involves innovative exploration using health data. From the perspective of data protection, two crucial aspects demand heightened attention: 1) the recognition of health data as a special category of personal data, and 2) its novel application in research contexts. Given their combined implications, it becomes imperative for professionals in the field to perform what the European General Data Protection Regulation (GDPR) terms as a Data Protection Impact Assessment (DPIA).[1]

However, it is essential to recognize that a DPIA is not a one-size-fits-all procedure. Its specific structure is influenced by the nature of the research project. For instance, a study on genetic diseases will emphasize different facets than research on the epidemiology of infectious diseases. Similarly, a clinical trial-based study will have distinct assessment criteria compared to a meta-analysis research.

Against this backdrop, the article endeavors to present a streamlined and actionable guide to conducting a DPIA, curated especially for those unfamiliar to the intricacies of data protection, while also pointing to more detailed resources for those seeking depth. It encapsulates the pivotal steps requisite for a DPIA, particularly when data sharing in health science research is in play.

As the prevalence of sensitive data in such projects is undeniable, a DPIA becomes crucial in aligning with data protection mandates and safeguarding individual privacy. Our guide delves into aspects such as determining the DPIA's intent, assessing data processing requisites, and strategizing effective risk alleviations.

In essence, this article stands as a foundational resource for initiating DPIAs within health science data sharing collaborations. It advocates ethical data practices and underscores the importance of abiding by data protection legislation. Moreover, readers will find an array of practical tools and resources, enhancing their proficiency in executing DPIAs.

Following this introductory section, Section 2 elaborates on the concept and necessity of a DPIA. Section 3 presents a comprehensive breakdown of the steps involved in conducting a DPIA, while Section 4 offers a concluding overview.

DOI: 10.21552/ehpl/2023/3/5

* Marcelo Corrales Compagnucci is Associate Professor and Associate Director of the Center for Advanced Studies in Biomedical Innovation Law (CeBIL), Faculty of Law, University of Copenhagen (UCPH). For correspondence: <marcelo.corrales13@gmail.com>. Alan Dahi is an independent researcher, German-qualified data protection lawyer, and guest lecturer in data protection at the Leibniz University Hannover. Peter Alexander Earls Davis is a postdoctoral researcher at the University of Copenhagen (UCPH).

1 As explained in more detail in Section 2 below, a DPIA is a structured process prescribed by the GDPR in certain circumstances to identify and minimise risks to the freedoms and rights of individuals that result from the processing of their personal data.

II. Understanding DPIA and its Relevance for Data Sharing

A DPIA serves as a structured approach employed by controllers[2] to assess the potential impact of processing activities on the safeguarding of personal data. In accordance with Article 35(1) of the GDPR, a DPIA becomes obligatory when specific data processing activities are anticipated to pose a substantial risk to the rights and freedoms of individuals. This requirement becomes particularly significant when novel data processing technologies are introduced, or when the characteristics, scope, context, and objectives of the processing intensify such risks.[3]

It is important to note that a DPIA should not be viewed as a one-time compliance exercise but rather as an ongoing process that assists organizations in identifying, assessing, and minimizing the data protection risks associated with a specific processing activity. A successful DPIA integrates data protection considerations into a project from its initial stages through its implementation and eventual completion.[4]

The determination of whether a type of processing is likely to pose a high risk to the rights and freedoms of data subjects can be achieved through a "generic" assessment of the processing activity or because the specific activity has been "blacklisted" by a supervisory authority. This means that certain processing activities may be identified in advance as inherently carrying a high risk, requiring a DPIA to be conducted.[5]

By conducting a DPIA, organizations can proactively identify and address potential data protection risks, implement appropriate safeguards, and ensure compliance with relevant regulations. This process helps organizations prioritize data protection and privacy considerations and fosters a culture of responsible data handling throughout the lifespan of a project.

1. 'Generic' Determination

The GDPR does not provide explicit guidance on how to determine whether a processing activity is likely to result in a high risk to the rights and freedoms of data subjects. However, it does specify certain types of processing activities that can be assumed to entail such a risk. These activities are as follows:

1. **Systematic and extensive evaluation:** This refers to the automated processing, including profiling, that involves the comprehensive assessment of personal aspects concerning individuals. For instance, analyzing and predicting aspects like work performance, economic situation, health, personal preferences, behavior, location, or movements fall under this category.[6]
2. **Large-scale processing of special categories of data:**[7] This includes the processing of sensitive data that reveals racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, data concerning health, data concerning sex life or sexual orientation, criminal convictions, or related security measures. When such data is processed on a significant scale, it is considered to pose a high risk to data subjects' rights and freedoms.[8]
3. **Systematic monitoring of a publicly accessible area:** When there is extensive monitoring of a publicly accessible area, such as through surveillance cameras or other means, on a large scale, it is regarded as a processing activity that can result in a high risk to individuals' rights and freedoms.While the GDPR does not provide an exhaustive list of activities, these examples give an indication of the types of processing that may be considered to carry a high risk. It is essential for research organizations to carefully assess their processing activities in light of these examples and, if necessary, conduct a DPIA to ensure compliance and protect individuals' privacy rights.As these types of processing do not encompass all "high risk" operations, the WP29 Guidelines propose a set of nine criteria

2 A "controller" is the entity who "defines the means and purposes of the processing" (Article 4(7) GDPR). A "processor" is the entity who processes personal data on behalf of the controller, if the controller did not process personal data directly themselves but outsourced the task (Article 4(8) GDPR). For more information regarding the concepts of controllers and processors see Alan Dahi and Marcelo Corrales Compagnucci (2022) Device Manufacturers as Controllers: Expanding the Concept of 'Controllership' in the GDPR. *Computer Law and Security Review* 47: 105762.

3 Article 35(1) GDPR.

4 Article 35(11) GDPR. Article 29 Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679 adopted on 4 April 2017 (as last revised and adopted on 4 October 2017), WP 248 rev.01, pp. 8 et seqq. and 14; CNIL, Privacy Impact Assessment (PIA), Methodology, February 2018 edition, p. 3.

5 Article 35(3) GDPR.

6 See recital 71 GDPR.

7 See Articles 9(1) or 10 GDPR.

8 See recital 75 GDPR.

and assessment rules. If two or more criteria are met, it is likely that a DPIA will be necessary, and the likelihood increases with the presence of more criteria. These criteria are as follows:[9]

4. **Evaluation or scoring:** This includes profiling and predicting performance or behavior.[10] Examples include a biotechnology company providing genetic tests directly to consumers to determine the likelihood of developing a specific disease or health issue, or obesity researchers leveraging machine learning to assign personalized risk assessments based on genetic and epigenetic data.[11]
5. **Automated-decision making with legal or similar significant effect:** This is where processing may result in the exclusion or discrimination of individuals.[12] It is possible, for example, that the use of certain types of artificial intelligence (AI) technology in hospitals may result in bias, both in the algorithm and in the data used to train the algorithm.
6. **Systematic monitoring:** This means the observation, monitoring, and control of data subjects, including data collected via networks and any "systematic monitoring of a publicly accessible area".[13] Examples include mobile applications that track health outcomes, or the use of smart cameras to monitor patients in a hospital canteen.
7. **Sensitive data or data of a highly personal nature:** Data that falls under this category includes special categories of personal data[14] such as health data, biometric, and genetic data. For instance, a hospital processes medical history of treatments and keeps patients' medical records in the hospital information system.
8. **Data processed on a large scale:**[15] To determine whether the processing is of a large scale, the WP29 recommends that the following criteria be considered: a) the number of data subjects, either as a percentage of the relevant population or as a specific number; b) the volume and/or range of data items being processed; c) the duration, or permanence, of the data processing activity; and, d) the geographical extent of the processing activity. For example, the collection of health data from medical wearables for the purpose of generating health profiles, or large virtual biobanks intended to facilitate medical research.
9. **Matching or combining datasets:** This will be the case where two or more data processing operations performed for different purposes or by different controllers are matched or combined in a manner that exceeds the reasonable expectations of the data subject. In a health science context, this may occur where different genetic datasets from the same data subjects but in totally unrelated clinical trials are merged for a new purpose.
10. **Data concerning vulnerable data subjects**: This encompasses anyone who is unable to easily consent to, or oppose, the processing of their data, or exercise their rights, for example due to a power imbalance with the controllers or for other reasons. Among these individuals are children, employees, asylum seekers, mentally ill persons, elderly patients, etc. As such, pseudonymized personal data of children participating in clinical trials or genetic research would be an example where personal data of vulnerable data subjects is processed.
11. **Innovative use or applying new technological or organizational solutions:** Cutting-edge technological or organizational solutions can lead to novel forms of data collection and usage, potentially creating high risks to the rights and freedoms of data subjects. Medical devices as part of the Internet-of-Things would be such an example. Additionally, while not specifically mentioned by the WP29, the use of AI tools will likely trigger the need for a DPIA.[16]

9 Article 29 Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679 adopted on 4 April 2017 (as last revised and adopted on 4 October 2017), WP 248 rev.01, pp. 9-13.

10 Especially from "aspects concerning the data subject's performance at work, economic situation, health, personal preferences or interests, reliability or behavior, location or movements" (recitals 71 and 91 GDPR).

11 The EU Horizon project "Bio-Streams", for instance, aims to carry out the latter example: see further <https://www.bio-streams.eu/objectives/> accessed 5 September 2023.

12 See Article 35(3)(a).

13 See Article 35(3)(c). According to the WP29 Guidelines on Data Protection Officer 16/EN WP 243, "systematic" refers to one or more of the following: a) occurring according to a system; b) pre-arranged, organized or methodical; c) taking place as part of a general plan for data collection; d) carried out as part of a strategy. A "publicly accessible area" is defined by the WP29 as any place open to the general public, for instance a piazza, a shopping center, a street, a marketplace, a railway station or a public library.

14 As defined in Articles 9 and 10 GDPR.

15 It is not outlined in the GDPR what constitutes large-scale, although recital 91 provides some guidance.

16 The EU project "CLASSICA", for instance, looks at exploring ways to integrate AI-driven cancer classification technology into surgical procedures. See further <https://classicaproject.eu/> accessed 5 September 2023.

12. **When the processing in itself prevents data subjects from exercising a right or using a service or a contract**: Insofar as the processing itself leads to allowing, modifying or refusing data subjects' access to a service or entering into a contract, for example where a data subject is rejected by an insurance company because an automatic assessment deems them to have a high risk of having an accident, the processing will require a DPIA.

The determination of whether a processing activity poses a high risk to the rights and freedoms of data subjects necessitates a comprehensive understanding of the activity itself. It is crucial for the controller to have knowledge of the categories of data subjects involved, the types of personal data to be processed, the stakeholders involved, and the nature and purpose of the processing. This awareness enables the controller to assess whether the processing activity will result in significant risks for the data subjects.

Certainly, the process of identifying the specific risks associated with the processing activity is an integral part of the DPIA in its own right. Therefore, the initial decision regarding whether to conduct a DPIA can only be based on a preliminary identification and assessment of the risks, unless one of the specific examples provided by the GDPR, WP29 Guidelines, or Member State blacklists applies.

The inherent uncertainty surrounding the existence of high risks further underscores the importance of conducting a DPIA. In cases where the justification for not conducting a DPIA requires substantial effort, the recommended approach is to conduct a DPIA due to this unavoidable uncertainty.

In summary, a thorough understanding of the processing activity is essential for assessing its potential high risks to data subjects. The DPIA process involves identifying and evaluating these risks, and it is advisable to conduct a DPIA even when the justification for not doing so requires significant effort due to the inherent uncertainty involved.

Given the unique characteristics of data sharing projects in the health science field and the growing prevalence of platforms facilitating data sharing among public and private partners, it is important to recognize that such projects typically involve large-scale data processing.

Furthermore, these projects often utilize automated methods and innovative solutions like AI. Considering the sensitive nature of the data involved and the potential inclusion of vulnerable data subjects, it becomes evident that a default assumption should be established: data sharing projects in the health sciences that involve personal data will most likely require a DPIA.

2. 'Blacklists' Determination by Supervisory Authorities

In addition to the criteria provided by the WP29 Guidelines, many national supervisory authorities have established their own lists of processing activities that are considered high risk and automatically require a DPIA. These lists, commonly referred to as "blacklists," serve as valuable resources for organizations to identify whether their specific processing activities fall within the scope of a DPIA.[17]

To access these blacklists, organizations can visit the respective websites of the supervisory authorities or search the website of the European Data Protection Board (EDPB). The EDPB collects and shares the lists communicated by the supervisory authorities of Member States, in accordance with Article 35(4) of the GDPR.

While the specifics may vary between national authorities, the blacklists generally align with the recommendations provided by the WP29 Guidelines mentioned above. They serve as practical examples of how the criteria outlined in the guidelines are implemented in practice.

III. Detailed Steps for a DPIA

Conducting a DPIA does not come with a universal blueprint.[18] While the GDPR does not prescribe a spe-

17 See <https://edpb.europa.eu/search_en?search=Article+35.4> accessed 3 September 2023.

18 Whilst some templates or software, such as that provided by the French data protection authority (CNIL) can be of assistance to controllers and processors, these cannot be followed blindly and must be tailored to the specifics of the processing operation concerned. See: <https://www.cnil.fr/en/privacy-impact-assessment-pia> accessed 6 September 2023. The Catalan data protection authority provides a specific tool tailored for research and innovation health care projects. See: <https://www.bioeticayderecho.ub.edu/en/presentation-methodology-and-tool-conducting-data-protection-impact-assessment-dpia-health-care> accessed 6 September 2023.

cific method, as described above, it does highlight essential components for the assessment. Factors such as the type of the controller, the type of processing activity, the characteristics of involved data subjects, and the risks they encounter all influence the DPIA process.[19] However, despite these variances, and as reflected in the consistency across guidelines at both Member State and EU levels,[20] we believe there exist fundamental elements that permeate all DPIAs. These foundational principles, which every DPIA should incorporate, are elaborated upon in the subsequent sections.

1. Prepare for the DPIA

If a preliminary review indicates a DPIA is required, preparation becomes essential. This involves a detailed evaluation of the processing activities. To effectively carry out this assessment, several key actions need to be taken:[21]

1. **Formation of a multi-disciplinary team**: Assemble a diverse group of experts from areas like data protection, legal, technical, and organizational operations. Their collective insights ensure a thorough DPIA.
2. **Stakeholder engagement**: Identify and involve parties affected by or involved in the processing activities. This encompasses data subjects and other internal or external entities. Early engagement ensures their insights shape the DPIA.
3. **Scope definition**: Clearly outline the DPIA's boundaries, detailing the specific activities, systems, and data flows under examination. This sharpens the assessment's focus on the most pertinent risks to individuals' rights.

a. Assemble the DPIA Team

The team's composition for conducting the DPIA should directly reflect the nature and specifics of the data processing in question. Ensuring the right expertise is essential for a thorough assessment. For example, the processing of medical data may need medical statisticians on the team while the processing of unemployment data may need social scientists. Certain processing activities may require specialists in ethics, IT security, or data analysis.[22] Below is a breakdown of potential team members based on the type of data processed:

1. **Medical data**: Include medical statisticians or healthcare professionals for insights into handling sensitive health data.
2. **Socioeconomic data**: Social scientists or economists can provide insights into potential societal impacts and biases.
3. **Ethical or privacy concerns**: Ethicists or privacy experts offer guidance on legal requirements and ethical best practices.
4. **Technological systems**: IT security specialists assess security measures, vulnerabilities, and suggest ways to mitigate risks.
5. **Complex data analysis**: Data analysts or scientists provide expertise in data modeling and analytics, ensuring accuracy and fairness in processing.

By integrating a range of professionals suited to the specific needs of the DPIA, the team can approach the processing activities comprehensively, ensuring all risks and angles are considered.

It is important to note that a Data Protection Officer (DPO),[23] where appointed, will not by default be part of the DPIA team. According to Article 39(1)(c) of the GDPR, the DPO should provide the controller advice on the DPIA only "where requested". In contrast, the DPO is statutorily tasked with advising the controller and its employees on their obligations under the GDPR (and as such on when a DPIA is obligatory), as well as to monitor the performance of any DPIA conducted. In this respect, the DPO is a key stakeholder rather than a DPIA team member.

19 Illustrating this for big data analytics, Georgios Georgiadis and Geert Poels 'Towards a privacy impact assessment methodology to support the requirements of the general data protection regulation in a big data analytics context: A systematic literature review' *Computer Law & Security Review,* 44(2022), 105640, ISSN 0267-3649.

20 Specifically for this article, CNIL, Privacy Impact Assessment (PIA), Methodology, February 2018 edition; German Datenschutzkonferenz (DSK), Kurzpapier Nr. 5, Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO, 17.12.2018; Article 29 Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679 adopted on 4 April 2017 (as last revised and adopted on 4 October 2017), WP 248 rev.01.

21 See, e.g., German Datenschutzkonferenz (DSK), Kurzpapier Nr. 5, Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO, 17.12.2018, p. 2.

22 ibid.

23 The DPO is an independent advisory and monitoring role within the controller and processor. The controller or processor must involve the DPO in all issues that relate to the processing of personal data. See Articles 36-39 of the GDPR.

b. Identify the Stakeholders

Identifying stakeholders is a critical step in conducting a DPIA as it ensures that the assessment takes into account the perspectives and concerns of all relevant parties. In addition to the DPO, controllers, processors, and other organizations involved in the processing activities, it is important to recognize the role of data subjects as stakeholders in the DPIA process.[24]

While directly contacting all affected data subjects may not always be feasible, it is essential to explore alternative avenues to involve them in the assessment. One approach is to identify and engage with representatives of data subjects, such as consumer rights agencies, patient societies, union representatives, or workers' councils. These groups can provide valuable insights into the potential impact of the processing activities on the rights and interests of data subjects.[25]

Additionally, considering the option of conducting a poll or survey of representative data subjects can further enhance the inclusiveness of the DPIA. This approach allows for a broader understanding of data subjects' perspectives, concerns, and expectations regarding the processing activities. It provides an opportunity to gather valuable feedback and insights directly from those individuals whose data is being processed, enabling a more comprehensive and informed assessment.

By involving stakeholders, including representatives of data subjects, the DPIA can benefit from diverse viewpoints and expertise, ensuring that the assessment considers a wide range of potential risks and impacts. It also promotes transparency and accountability by demonstrating a commitment to involving all relevant parties in the decision-making process related to data processing.

Furthermore, stakeholder engagement can contribute to building trust and maintaining positive relationships with data subjects, as it demonstrates a genuine commitment to protecting their rights and interests. It allows for effective communication channels to address any concerns or questions they may have, fostering a sense of empowerment and involvement in the data protection process.

Overall, the identification and involvement of stakeholders, including data subjects and their representatives, are essential for a robust and comprehensive DPIA. It ensures that the assessment captures the full spectrum of interests and perspectives, leading to more informed decision-making and the implementation of appropriate measures to safeguard the rights and freedoms of data subjects.

c. Determine the Scope of the DPIA

Defining the scope of a DPIA is a crucial step that lays the foundation for conducting a comprehensive and effective assessment. The scope determines the boundaries and focus of the DPIA, guiding the assessment process and ensuring that all relevant aspects of the processing activity are properly evaluated. To precisely determine the scope, it is essential to have a deep understanding of the processing activity itself. This involves identifying the technical components, systems, and processes involved, as well as the specific types of personal data that will be processed. By gaining clarity on these aspects, the scope can be defined in a way that captures all the pertinent elements that may pose risks to data subjects' rights and freedoms.[26]

When determining the scope, it is important to consider the potential threat sources that may affect the processing activity. Threat sources most intuitively include security risks, such as external factors such as unauthorized access, data breaches, or malicious attacks, as well as internal security threats like system vulnerabilities or human errors.[27] However, it may also include other risks to the data subject, such as the potential for unfair or bias decisions from an AI system, the risk that data thought to be anonymous data becomes re-identified, the negative consequences flowing from inaccurate personal data, or the threat that data will be re-used (for another pur-

---

24 German Datenschutzkonferenz (DSK), Kurzpapier Nr. 5, Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO, 17.12.2018, p. 3.

25 See also e.g., CNIL, Privacy Impact Assessment (PIA), Methodology, February 2018 edition, p. 8; German Datenschutzkonferenz (DSK), Kurzpapier Nr. 5, Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO, 17.12.2018, p. 3.

26 See CNIL, Privacy Impact Assessment (PIA), Methodology, February 2018 edition, p. 4; German Datenschutzkonferenz (DSK), Kurzpapier Nr. 5, Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO, 17.12.2018, p. 2.

27 See, e.g., Kirkham, T, Armstrong, D, Djemame, K, Corrales Compagnucci, M, Kiran, M, Nwankwo, I, Jiang, M & Forgó, N 2012, Assuring Data Privacy in Cloud Transformations. in *Proceedings of the 11th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (IEEE TrustCom-12)*. IEEE, pp. 1063.

pose) within the organization. By identifying and considering these threat sources within the scope, the DPIA can thoroughly assess the potential risks and develop appropriate safeguards to mitigate them effectively.

Additionally, the scope should clearly outline which specific elements or components of the processing activity will be included in the assessment and which ones will be excluded. This ensures that the DPIA focuses on the areas that are most relevant and critical in terms of potential risks to data subjects' privacy and rights. By defining the scope in this manner, the controller gains a precise understanding of what aspects will be covered by the assessment, allowing for a more targeted and effective evaluation.

It is important to note that defining the scope is not a one-time task but rather an iterative process that may require ongoing adjustments as the assessment progresses. As the DPIA unfolds and more information is gathered, the scope may need to be refined and expanded to ensure a comprehensive analysis of all relevant aspects.

In summary, defining the scope of a DPIA is a critical step that requires a deep understanding of the processing activity. It allows for the identification of relevant threat sources, determines the boundaries of the assessment, and ensures a targeted evaluation of the potential risks to data subjects' rights and freedoms. By establishing a precise scope, the controller can effectively plan and execute the DPIA, leading to enhanced data protection measures and compliance with applicable regulations.

2. Conduct the DPIA

Having established the foundation for the DPIA, the next step is to proceed with the assessment itself. The DPIA assessment consists of two distinct parts, each addressing different aspects of data protection and security:[28]

1. **Assessment of GDPR's Processing Principles:** In this part, the focus is on evaluating how the processing activities align with and adhere to the fundamental principles outlined in the GDPR. This involves a thorough examination of the specific data protection controls that are in place or planned to ensure compliance. The assessment aims to determine whether the processing activities meet the requirements of lawfulness, fairness, and transparency in the collection, use, and handling of personal data. It also assesses aspects such as purpose limitation, data minimization, accuracy, storage limitation, and accountability. By scrutinizing these principles, the DPIA ensures that appropriate measures are implemented to protect individuals' privacy rights and uphold the legal obligations under the GDPR.[29]
2. **Assessment of Security Risks:** The second part of the DPIA focuses on evaluating the security measures and safeguards in place to protect the personal data from unauthorized access, loss, alteration, or disclosure. This assessment aims to identify and assess specific security risks associated with the processing activities. It involves analyzing the technical and organizational measures implemented to ensure the confidentiality, integrity, and availability of the data. The assessment considers factors such as the storage and transmission of data, access controls, encryption, backups, incident response procedures, and staff training. By conducting this security assessment, the DPIA aims to identify vulnerabilities or weaknesses in the data processing infrastructure and recommend measures to mitigate or address those risks effectively.[30]

It is important to note that the term "assessment" within the context of DPIA can be somewhat misleading. The DPIA can be conducted either prior to or after the processing activities,[31] and the assessment itself may involve evaluating hypothetical or planned measures as well as auditing existing measures already implemented by the controller. The risk

28 See CNIL, Privacy Impact Assessment (PIA), Methodology, February 2018 edition, p. 3.

29 ibid, p. 5.

30 ibid, p. 6; See also, e.g., Kiran, M, Khan, AU, Jiang, M, Djemame, K & Corrales Compagnucci, M 2013, 'Managing Security Threats in Clouds', Digital Research 2012. The 8th International Conference for Internet Technology and Secured Transactions , London, United Kingdom, 09/12/2013 - 12/12/2013; Djemame, K, Barnitzke, B, Corrales Compagnucci, M, Kiram, M, Jiang, M, Armstrong, D, Forgó, N & Nwankwo, I 2013, 'Legal Issues in Clouds: Towards a Risk Inventory', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 371, no. 1983, pp. 1-17.

31 Typically, an *a priori* DPIA is conducted for new processing activities, whereas an *a posteriori* DPIA is necessary for processing activities that have undergone significant changes or instances where the controller realizes, after the processing has already commenced, the requirement for a DPIA. This is reflected in the law itself, recognizing that a DPIA is not a one-off exercise but an ongoing process (Article 35(11) GDPR).

assessment criteria should be regularly updated to ensure emerging risks are promptly identified and addressed.[32]

For example, when assessing the transparency principle, which is closely linked to the data subject's right to information, the evaluation can take the form of auditing existing privacy notices to ensure they provide clear and comprehensive information to individuals. In cases where no privacy notice is yet in place, the assessment may involve making recommendations and providing guidance on developing appropriate notices.

a. Assessing the Data Protection Controls

A "control" is a measure put in place to ensure that the data protection requirements stemming from the processing principles are met. Controls can be of a technical or organizational nature.[33]

The processing principles are:

1. Lawfulness, fairness, and transparency (Article 5(1)(a) of the GDPR);
2. Purpose limitation (Article 5(1)(b) of the GDPR);
3. Data minimization (Article 5(1)(c) of the GDPR);
4. Accuracy (Article 5(1)(d) of the GDPR);
5. Storage limitation (Article 5(1)(e) of the GDPR);
6. Integrity and confidentiality (Article 5(1)(f) of the GDPR); and
7. Accountability (Article 5(2) of the GDPR).

Additionally, the DPIA should also assess how the controller will meet the data subject rights in Articles 12 et seqq of the GDPR, i.e., the:

1. Right to information (Articles 12-14 of the GDPR).;
2. Right to access (Article 15 of the GDPR);
3. Right to rectification (Article 16 of the GDPR);
4. Right to erasure (Article 17 of the GDPR);
5. Right to restriction of processing (Article 18 of the GDPR);
6. Right to data portability (Article 20 of the GDPR);
7. Right to object (Article 21 of the GDPR); and
8. Rights concerning automated individual decision-making (Article 22 of the GDPR).

The data protection principles form the foundation for all data protection obligations, including rights. Though these rights can be evaluated under the principles, it is typically more straightforward to review them in a distinct section.

Certain processing activities may require further evaluations, such as those involving international data transfers.[34] Additionally, where health data is processed on a large scale, controllers are required to designate a DPO.[35] However, detailing every potential obligation a controller might face is outside this article's purview.

Each assessment should clarify how specific requirements are addressed and justify the chosen measures for the given processing. Measures are deemed appropriate when they cannot be *reasonably* improved, taking all specifics of the processing into account. The greater the risks for the data subject, the higher the bar for "appropriateness".[36]

i. Purpose Limitation

The principle of purpose limitation dictates that data should be collected for specific, explicit, and legitimate reasons and must not be processed or used in ways that deviate from these original purposes. In essence, the data should remain true to its initial objective and should not be repurposed without further lawful basis, unless the new purpose aligns closely with the original one. This is crucial in ensuring that data usage remains responsible and honors the individual's privacy.[37]

Specifying the purpose is not always an easy task. In its *Opinion 03/2013 on purpose limitation*, the Article 29 WP explained that "each separate purpose should be specified in enough detail to be able to assess whether collection of personal data for this purpose complies with the law, and to establish what da-

32 A good comparison of this issue is seen in general audits. Initially, a company's accountants might spend significant funds and resources on a risk assessment approach. However, they often reuse the same risk analysis annually, leading to a growing disparity between the actual situation and the yearly report. Comparable patterns can be observed in audit firms that consistently use the same risk evaluation process within the social capital network for many years. See Marcelo Corrales Compagnucci, *Big Data, Databases and "Ownership" Rights in the Cloud* (Springer, 2019), p. 152.

33 See e.g., CNIL, Privacy Impact Assessment (PIA), Methodology, February 2018 edition, pp. 7, 11.

34 Ibid. p. 5.

35 Article 37(1)(c) of the GDPR.

36 See e.g., German Datenschutzkonferenz (DSK), Kurzpapier Nr. 5, Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO, 17.12.2018, p. 3.

37 Article 5(1)(b); 6 and 26 of the GDPR.

ta protection safeguards to apply".[38] The purpose can neither be too broadly specified, nor should it be artificially broken down into micro-purposes. Either can negatively affect both the data subject and the controller. From the perspective of the individual providing the data, the information shared by the doctors might be partial or unclear. They might be unaware that a particular "purpose" involves various data processing steps.

As a way of example, consider that a patient is admitted to a hospital for a major surgery. During the admission process, he provides comprehensive personal and health details, including his medical history, genetic information, lifestyle habits, and contact information. In this scenario, the patient's data is collected solely for his medical treatment, monitoring, and ensuring his well-being during his stay. A few months later, the hospital opens a new wellness center focusing on holistic health, nutrition, and exercise. Even though the hospital has the patient's contact and health information on file, the purpose limitation principle restricts the hospital's ability to lawfully use this data to send him promotional material or tailored nutrition and exercise plans related to the new wellness center. The patient's information was initially gathered for medical treatment purposes, not for marketing or additional health services.

For this reason, it is important to be clear from the outset and inform the patients about the purposes involved in the data processing. In this way, individuals are protected from having their data processed in ways they are not aware of or did not consent to. Individuals are given an increased degree of control over their personal data and organizations can be held accountable for how they manage their data.[39]

It should be noted that further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes is not incompatible with the original purpose provided the processing is in accordance with Article 89(1) GDPR. Article 89(1) GDPR requires that such processing is subject to technical and organizational safeguards to protect the rights and freedoms of the data subject.[40] Moreover, relying on the research exemption to the purpose limitation has been described as "precarious"[41] if other entities than the original controller(s) are involved in the further processing. The DPIA should carefully consider these intricate matters if Article 89(1) is invoked by the controller.

ii. Lawfulness, Fairness, Transparency

According to the GDPR, data shall be processed lawfully, fairly, and transparently, in relation to the data subject.[42]

iii. Lawfulness

For data processing to be deemed lawful, it must rest upon established legal grounds, as outlined in Articles 6 and 9 of the GDPR:

- Art. 6 pertains to the "lawfulness of processing," which encompasses grounds such as obtaining the data subject's consent or if the processing is integral to the execution of a contract.
- Art. 9 addresses the "processing of special categories of data," which includes sensitive data. Detailed exploration of these articles will be undertaken in subsequent discussions.

Each legal basis has its specific requirements. Taking consent as an example, the assessment will need to explain both why consent is the appropriate legal basis and how the solicited consent will meet the requirements for consent to be valid, namely of being freely given, specific, informed, and unambiguous.[43]

While it is relatively straightforward to apply the legal bases when data is used for its specific intended purpose (as referenced in "purpose limitation" above), challenges arise when the data is repurposed. For instance, in health sciences research where data

38 Article 29 Data Protection Working Party, 'Opinion 03/2013 on purpose limitation, Working Paper 203', adopted on 2 April 2013, p. 16 <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf> accessed 3 September 2023.

39 Article 5(1)(b) GDPR.

40 For an elaboration of requirements in the context of research biobanks, see further Ciara Staunton, Santa Slokenberga, and Deborah Mascalzoni (2019), *The GDPR and the research exemption: considerations on the necessary safeguards for research biobanks*. *European Journal of Human Genetics* 27, 1159-1167.

41 Becker, R., Thorogood, A., Ordish, J., & Beauvais, M. J. (2020). COVID-19 research: navigating the European general data protection regulation. Journal of Medical Internet Research, 22(8), e19799.

42 See Articles 5, 6 and 9 of the GDPR.

43 See Article 4(11) of the GDPR.

reuse is common, if the new purpose deviates from the original intent, Article 6(4) of the GDPR stipulates that such "further processing" must be compatible with the purpose for which the data was originally collected.

Such an assessment requires a comparison of the new purpose with the original purpose. This assessment should take into account any links between the purposes, the context in which the personal data was collected (including the data subject's expectations), the types of personal data being processed, any possible consequences of the new processing on the affected data subjects, and the existence of technical and organizational measures to safeguard the rights and freedoms of the data subjects.[44]

In addition to ensuring that the planned processing of personal data has an appropriate legal basis, the lawfulness principle further implies that processing of personal data should be lawful in a broader sense.[45] That means adherence to the GDPR's other rules and principles (including those discussed herein), but also other legal obligations that the controller may be under. For example, certain medicinal research may fall within scope of the Clinical Trials Regulation (CTR)[46] which, as put by the EDPB, "constitutes a sectoral law containing specific provisions relevant from a data protection".[47] The DPIA should consider the impact of these specific provisions on the legality of the processing of personal data. The CTR includes specific considerations for vulnerable populations,[48] more detailed rules relating to the procurement of informed consent,[49] and provisions relating to good clinical practice[50] directed towards ensuring the trial is conducted ethically.[51]

iv. Fairness

Fair processing means that data must not be acquired or processed through deceptive methods, unfairly, or without the knowledge of the data subject.[52] Furthermore, the principle of fairness requires the use of clear and comprehensible language, especially when conveying information to children. Recital 39 of the GDPR also highlights the importance of providing transparent information content.[53]

The concept of fair processing is somewhat vague. Nevertheless, it can be captured with the questions: "*Can the data subject have expected such processing to occur?*" and "*Am I facilitating the exercise of the data subject rights?*".[54] The principle of fairness is especially relevant when the use of data is innovative, where new technologies are used for the processing, and where "further processing" takes place. This is often seen in scenarios where data is repurposed, a practice frequently observed in health science research projects that leverage retrospective data.

Consider a hypothetical university research scenario: The Metabolic Research Center embarks on a study observing the glucose levels of 70 students. Every two hours, these students record their sugar intake and subsequent energy levels. All 70 participants have explicitly given their consent for this study and the specific use of their data by the university. The purpose of this research study is to investigate the correlation between sugar intake and its immediate impact on energy levels in university students, aiming to understand metabolic responses and potential implications for dietary recommendations. Subsequently, the university sees an opportunity to

44 Article 29 Data Protection Working Party, 'Opinion 03/2013 on purpose limitation, Working Paper 203', adopted on 2 April 2013, pp. 20 et seqq, <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf> accessed 3 September 2023.

45 Cécile de Terwangne, Article 5 Principles relating to processing of personal data, Chapter II, p. 314. In: Christopher Kuner, Lee A. Bygrave and Christopher Docksey (eds) *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press, 2020).

46 Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC (Text with EEA relevance)s OJ L 158, 27.5.2014, p. 1–76.

47 European Data Protection Board, Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical Trials Regulation (CTR) and the General Data Protection regulation (GDPR) (art. 70.1.b))
Adopted on 23 January 2019 <https://edpb.europa.eu/sites/default/files/files/file1/edpb_opinionctrq_a_final_en.pdf> accessed 3 September 2023.

48 See Article 10 of the CTR.

49 See Chapter V of the CTR.

50 See Article 47 of the CTR.

51 Recital 80 of the CTR, for instance, refers to the Declaration of Helsinki ethical principles.

52 Cécile de Terwangne, Article 5 Principles relating to processing of personal data, Chapter II, p. 314. In: Christopher Kuner, Lee A. Bygrave and Christopher Docksey (eds) *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press, 2020).

53 Recital 39 GDPR.

54 See EDPB, 'Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects', Version 2.0, 8 October 2019, para. 12.

use these data for a new project, focusing on diabetes risk factors, managed by a different research center and a distinct research team. While the university (acting as the controller) could in principle use the collected data for this new project, given the related objectives, a DPIA may find that it is preferable the university to notify the participants and request additional consent, if feasible. Adopting such a strategy may, further, abide by the university's research ethics code, and ensure that the principle of fair data processing is maintained.

v. Transparency

A fundamental concept of the GDPR is transparency in the use and processing of personal data. According to the GDPR, the principle requires that "any information addressed to the public or to the data subject be concise, easily accessible and easy to understand, and that clear and plain language and, additionally, where appropriate, visualisation be used".[55] Transparency is therefore about the "how" of all communication, whereas the "what" is addressed elsewhere, for example in Article 13 of the GDPR.

Transparency and fairness are intrinsically linked,[56] such that opaque data processing may also breach the fairness principle. As such, particular care with transparency is required where the use of data is innovative or where new technologies are used for the processing. This may include, for example, the use of data-driven AI techniques whose logic is difficult to ascertain or explain.[57]

vi. Data Minimization

The principle of data minimization mandates that the collection of personal data should be restricted to what is necessary for the intended processing objectives. Essentially, this tenet ensures that only the necessary data is utilized for a given purpose, safeguarding individuals' data protection rights.[58] This is especially pertinent to sensitive data categories, notably genomics and health data. The prominence of data minimization in a health science context is further enhanced by Article 89, the so-called "research exemption" provision, that allows for the derogation of data protection rights in certain instances relating to, *inter alia*, scientific research. These derogations are counterbalanced by specific mention of "appropriate safeguards... [to] that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation."[59]

Through adherence to the principle of data minimization, individuals gain increased control over their data, while organizations are required to demonstrate transparency and accountability in their data management practices. Two elements underpin the data minimization principle:[60]

- **Necessity**: It emphasizes collecting data solely when essential for the processing objectives.
- **Proportionality**: It assesses if the desired outcome could be realized with a lesser quantum of personal data.

However, the expansive nature of big data processing frequently conflicts with the principles of data minimization. To address this challenge, organizations should adopt a prudent strategy for data collection. Employing pseudonymized data adds a protective measure, ensuring that any resulting analysis remains detached from identifiable individuals. Additionally, the utilization of synthetic or anonymized data can serve as an alternative when feasible.[61]

vii. Accuracy

All personal data must be accurate and, where necessary, kept up-to-date; all reasonable steps must be

55 Recital 58 of the GDPR. According to the Article 29 Data Protection Working Party, 2018, 5 "the concept of transparency in the GDPR is user-centric rather than legalistic." Hence, the comprehensibility and presentation of information play a crucial role.

56 Article 29 Working Party, 'Guidelines on transparency under Regulation 2016/29', WP260 rev.01, para. 2.

57 See further Tim Hulsen, Explainable Artificial Intelligence (XAI): Concepts and Challenges in Healthcare. *AI* 2023, *4*, 652-666.

58 Article 5(c) GDPR. See also, Kevin Mc Gillivray (2021), *Government Cloud Procurement: Contracts, Data Protection, and the Quest for Compliance* (Cambridge University Press, 2021), p. 108.

59 See further Christian Wiese Svanberg. Article 89 Safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, pp. 1240-1251. In: Christopher Kuner, Lee A. Bygrave and Christopher Docksey (eds) *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press, 2020).

60 Aurelia Tamò-Larrieux, *Designing for Privacy and its Legal Framework: Data Protection by Design and Default for the Internet of Things* (Springer, 2018), p. 91.

61 Giulia Schneider, *Health Data Pools Under European Data Protection and Competition Law, Health as a Digital Business* (Springer, 2022), p. 306.

taken to ensure that inaccurate information is erased or rectified without delay, taking into consideration the purpose for which the information is processed.[62] While the "principle of accuracy" predominantly applies to verifiable facts, it is worth noting that subjective value judgments, even if they have potential personal implications, are not encompassed by this principle.

To ensure data accuracy, controllers are not only tasked with offering data subjects the means to modify their details but also to foster an environment where subjects can proactively manage and update their data. Consequently, it is recommended that controllers integrate both technical and organizational strategies to promptly detect and correct data discrepancies.[63]

Data accuracy is crucial for data sharing projects in the health science that use data for research purposes or to make decisions, as inaccurate or out-of-date data can lead to wrong conclusions. In addition to the effects on the quality of the research, inaccurate data can have substantial negative effects on the data subject.

viii. Storage Limitation

Generally, personal data cannot be stored indefinitely. Storage should be commensurate to the duration of the processing and the purpose for which the data was collected. However, it is possible to store personal data for longer periods of time if they are processed solely for the purpose of archiving in the public interest, scientific or historical research, or statistical analysis.[64]

While research and statistical analysis can permit longer retention periods than would otherwise be appropriate, as mentioned in the context of data minimization, controllers should be aware that archival consistent with this principle will require special measures to ensure that the data is no longer used for any other purpose.[65]

In any case, the decided upon storage periods must be defined, either by concreate measures of time, such as three years, or by criteria that allow calculation, such as three years after the research project was concluded. A controller must also be able to explain how the storage periods were determined and decided upon. In this vein, a DPIA is an apt instrument for considering, and illustrating consideration of, the legally appropriate length of time that personal data is to be stored. Moreover, a controller must identify and implement technical and organizational mechanisms that ensure proper deletion or anonymization of the data when the storage period expires.

ix. Integrity, Confidentiality and Availability

To maintain the integrity and security of personal data, it is vital to safeguard it against unauthorized or unlawful processing, as well as accidental loss, destruction, or damage. The DPIA should consider how this is to be achieved through various technical and organizational strategies.[66]

Measures like data encryption, physical security features such as locks and access controls, and the use of secure networks are essential to prevent unauthorized data access. Further, controllers must ensure that encryption and pseudonymization keys remain within the EU/EEA, under the jurisdiction and technical supervision of an EU entity.[67] It is also advisable to leverage advanced encryption and processing techniques, such as homomorphic encryption.[68] In addition, implementing a comprehensive Information Security Management System (ISMS) like ISO 27001, complemented by a Privacy Information Management System (ISO 27701), can enhance data protection efforts. Such systems empower individuals with greater control over their personal data, facilitating secure storage and granting them discretion over data sharing.[69]

---

62 Article 5(1)(d) and 16 GDPR.

63 Cécile de Terwange, Article 5. Principles relating to processing of personal data, Chapter II, p. 317. In: Christopher Kuner, Lee A. Bygrave and Christopher Docksey (eds) *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press, 2020).

64 Article 5(1)(e) GDPR.

65 Articles 5(1)(e), 89(1) GDPR.

66 Articles 5, 24 and 32 GDPR.

67 Paulius Jurcys, Marcelo Corrales Compagnucci and Mark Fenwick, 2022, The Future of International Data Transfers: Managing New Legal Risk with a 'User-Held' Data Model, *Computer Law & Security Review* vol. *46,* [105691].

68 Corrales Compagnucci, M, Meszaros, J, Minssen, T, Arasilango , A, Ous , T & Rajarajan, M 2019, 'Homomorphic Encryption: The 'Holy Grail' for Big Data Analytics & Legal Compliance in the Pharmaceutical and Healthcare Sector?', *European Pharmaceutical Law Review*, vol. 3 , no. 4, pp. 144-155.

69 Corrales Compagnucci, M, Aboy, M & Minssen, T 2021, 'Cross-Border Transfers of Personal Data after Schrems II: Supplementary Measures and New Standard Contractual Clauses (SCCs) ', *Nordic Journal of European Law*, vol. 4, no. 2, pp. 37-47.

When researchers and clinicians, for instance, utilize cloud-accessible database services to manage patient data, the DPIA should emphasise that confidentiality and integrity of information is a key priority. Both are crucial pillars of the security framework, joined by the equally vital aspect of data availability during emergencies. Confidentiality ensures that information remains protected from unauthorized access. Integrity ensures data remains unaltered and accurate. Meanwhile, availability guarantees that data and services are readily accessible whenever and wherever required.[70]

x. Accountability

The principle of accountability[71] mandates that controllers should not only ensure but also demonstrate their adherence to the regulations set forth within the GDPR. In practical terms, this entails that controllers are obligated to maintain detailed documentation that validates their compliance. The DPIA is a key accountability mechanism in itself; but it should further specify other means of ensuring compliance can be demonstrated, such as specifying a systematic record of data processing activities. This should capture information about data categories, the purposes of processing, and details of data recipients, among other things. The activities of an organization's DPO, which is required for entities involved in *inter alia* large-scale of processing sensitive data as most healthcare research does, provides a further means of demonstrating accountability. The DPO is tasked with independently overseeing and providing guidance on GDPR compliance within an organization,[72] and its mandate includes "providing advice where requested as regards the data protection impact assessment and monitor its performance."[73]

b. Assess Controls to Meet Data Security

Once compliance with the GDPR obligations has been assessed, the next step is to assess the controls regarding data security. A failure of "data security" can be equated to the GDPR's definition of a "personal data breach", i.e., "a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed".[74]

The risks of not maintaining adequate data security are very specific to the processing operation and reflect the potential privacy harms data subjects may experience. For example, a data breach concerning a data subject's health data will pose vastly different risks for the data subject compared to one where geolocation data is impacted.

Controllers must consider carefully and objectively what risks their specific processing may entail, the likelihood of the risks materializing and the severity of the harm if the risk occur. A risk that will materialize with a low likelihood might still warrant very high controls because the severity if it happens is very high. The following table[75] provides a visualization of the graduation between severity (i.e., impact) and likelihood.

As with the controls to meet the data protection obligations, each risk assessment needs to include an assessment of the technical and organizational measures to mitigate the risks. In a health context, consider a study that involves the processing of genetic and involving multiple clinical institutions across the EU, with the purpose to gather EU-level insights on childhood and adolescent obesity.[76] Such patient data is highly sensitive, with severe risks to data subjects if information is not properly secured. Moreover, the sharing of information between organizations, across the EU, is inherently problematic from both a broader data protection context[77] and narrower security context. On the latter, multiple state-of-the-art measures may need to be implemented in order to reach the GDPR's requisite standard of "appropriate technical and organisational measures".[78] These might include advanced cryptographic methods to reduce the risk of unauthorized access or to

70 See Marcelo Corrales Compagnucci, *Big Data, Databases and "Ownership" Rights in the Cloud* (Springer, 2019), p. 289.

71 Article 5 (2) GDPR.

72 For an outline of the DPO's responsibilities see Art. 39 GDPR.

73 Article 39(1)(c) GDPR.

74 Article 4(12) GDPR.

75 © Marcelo Corrales Compagnucci. Used with permission. See Marcelo Corrales Compagnucci, Big Data, Databases and "Ownership" Rights in the Cloud (Springer, 2019), p. 284.

76 This example is based on the Bio-Streams project; see <https://www.bio-streams.eu> accessed 5 September 2023.

77 On this conundrum, see further Jennifer Viberg Johansson, et al. Governance mechanisms for sharing of health data: An approach towards selecting attributes for complex discrete choice experiment studies. *Technology in Society* 66 (2021): 101625.

78 Article 32 GDPR; see also Article 25 GDPR (concerning data protection by by design and by default).

Likelihood	Impact				
	Negligible	Minor	Moderate	Major	Extreme
Rare	Low	Low	Low	Medium	Medium
Unlikely	Low	Medium	Medium	Medium	High
Possible	Low	Medium	Medium	High	High
Likely	Medium	Medium	High	High	Very high
Almost certain	Medium	High	High	Very high	Very high

Table 1: *Graduation between severity (i.e., impact) and likelihood*

anonymize or pseudonymize personal data,[79] localization of certain data, the implementation of access controls, training of staff, and maintenance schedules to ensure vulnerabilities are discovered and patched.[80] Such measures ensure both the protection of patient privacy and compliance with data protection regulations. Incorporating a risk assessment tool further ensures that any potential risks are addressed.

Further data security risks may demand specific attention. For example, consider the risk of data destruction, data loss or data alteration.

i. Data Destruction

Data destruction refers to the irreversible removal or corruption of data on storage devices. Cloud services, like Amazon, utilize Virtual Machines (VMs). VM data can be more unstable compared to traditional IT settings. VM data loss can occur if VMs crash or shut down. Also, if VMs are stored on a single physical server, a server crash can result in the loss of all VMs. VMs' ease of erasure and potential single point of failure risks are mitigated by the replicating VMs in different locations, enhancing data resilience.[81]

ii. Data Loss

The GDPR emphasizes safeguarding both logical and physical data availability against unforeseen events, such as natural disasters or hardware malfunctions.[82] Keeping backups is a recommended safeguard. The advantage of using cloud computing services in this context is its inherent redundancy; data is dynamically stored across different locations, and VM recovery can be faster than physical servers.[83]

iii. Data Alteration

Data alteration refers to any change made to existing data. Data should be protected against unauthorized changes to maintain its integrity. Systems should track and log data access, modifications, and the intent behind such modifications. Ensuring the integrity of data requires tracking systems that document data entry, alterations, and access. Log creation is essential for audits, certifications, security incidents, and potential unlawful data modifications.[84]

Tables 2[85] and 3[86] below illustrates an example of the key risk assessment elements in both legal and

79 Article 32(1)(a) GDPR.

80 See further European Data Protection Board, Guidelines 4/2019 on Article 25 Data Protection by Design and by Default Version 2.0, Adopted on 20 October 2020, pp. 26-28.

81 Karim Djemame et al., 2013, Legal Issues in Clouds: Towards a Risk Inventory, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371(1983): pp. 1-17.

82 Article 32(1)(c) GDPR refers to "the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident".

83 Ibid.

84 Ibid.

85 © *Marcelo Corrales Compagnucci. Used with permission.* See Marcelo Corrales Compagnucci, *Big Data, Databases and "Ownership" Rights in the Cloud* (Springer, 2019), p. 291.

86 ibid.

Table 2: Example of a risk assessment taking into account the graduation severity (i.e., impact) and likelihood

Risk category	Technical/Data security
Asset identified	Availability of data and databases
Vulnerability of asset	Lack of maintenance
Threat to asset	Database server failure
Risk likelihood	Rare
Risk impact	Moderate
Resulting risk level	Risk likelihood and risk impact = Low
Risk event	Data unavailability due to unexpected server failures occurring during the cloud service's operation.
Resulting risk mitigation	Fault-tolerance solutions provision

technical domains taking into account the graduation severity (i.e., impact) and likelihood.

3. Draft DPIA Report

After completing the assessment, the next step is to draft the DPIA report. Taking note of the outcome of the DPIA is part of the accountability process and paves the way for its approval and subsequent reviews.

The GDPR specifies the following necessary requirements:[87]

1. A systematic description of the planned processing activities and their purposes, including any legitimate interests pursued by the controller, if applicable.
2. An evaluation of the processing's necessity and proportionality concerning its purposes.
3. An examination of potential risks to data subjects' rights and freedoms referred (as referred to in section 3.2.1).
4. Proposed measures to mitigate these risks, highlighting safeguards, security protocols, and mechanisms designed to uphold data protection and demonstrate GDPR compliance, considering both the rights of data subjects and the interests of other relevant parties.

The DPIA report can adopt various formats, provided all elements are detailed and adequately addressed.

4. Approve and Implement DPIA

Having drafted the DPIA, the controller must obtain formal approval before proceeding with its implementation.[88] This entails presenting the DPIA to the DPO for insights, in line with Article 35(2) of the GDPR. The controller should meticulously document the rationale behind the DPIA's acceptance or rejection, adhering to the accountability mandate of Article 5(2) of the GDPR. Should the DPIA be declined, the intended processing activity must either be halted or the DPIA process must be revisited, bearing in mind the reasons for its initial rejection.[89]

87 Article 35(7) GDPR.

88 See e.g., CNIL, Privacy Impact Assessment (PIA), Methodology, February 2018 edition, p. 8; German Datenschutzkonferenz (DSK), Kurzpapier Nr. 5, Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO, 17.12.2018, p. 4.

89 Ibid.

Table 3: Example of a risk assessment taking into account the graduation severity (i.e., impact) and likelihood

Risk category	Technical/Data security
Asset identified	Availability of data and databases
Vulnerability of asset	Data center infrastructure (servers)
Threat to asset	Force majeure (such as floods, earthquakes, etc.).
Risk likelihood	Rare
Risk impact	Major
Resulting risk level	Risk likelihood and risk impact = Medium
Risk event	Data unavailability arises from server failures during service operation, especially when unforeseen events or force majeure incidents occur while the cloud provider is actively running the service.
Resulting risk mitigation	Employ redundancy and backup servers situated in diverse locations (cities). Continuously replicate data using databases and backup systems throughout the entire lifecycle of the cloud computing service.

Once the DPIA is approved, its rollout should follow a structured plan with specific milestones, ensuring documentation at every stage, and be in place before initiating the intended processing activity.[90]

5. Review

While the DPIA serves as a momentary snapshot, it is essential for the controller to periodically assess if any elements of the processing have changed. Factors such as emerging threat actors or advancements in technology could alter the risk landscape, requiring new or revised measures. Thus, the controller ought to integrate a routine review of the DPIA, for instance, within its annual data protection audit.[91]

IV. Conclusion

This article offers a step-by-step overview of the crucial processes integral to performing a DPIA, specifically within the realm of health science research projects. The purpose is to offer researchers and stakeholders with the necessary knowledge and practical tools to enhance the efficiency and robustness of DPIAs. This not only advocates responsible data management practices but also underscores the importance of strict adherence to data protection legislations and standards. In a world increasingly reliant on data, the insights presented here aim at fostering a culture of transparency, responsibility, and legal compliance in health science research.

Acknowledgment: This research was supported, in part, by a Novo Nordisk Foundation Grant for a scientifically independent Collaborative Bioscience Innovation & Law Program (Inter-CeBIL program – grant no. NNF23SA0087056), and by the European Union projects: CLASSICA, Horizon Europe (grant no. 101057321) and Bio-Streams, Horizon Europe

90 Ibid.

91 See Article 35(11) of the GDPR and e.g., CNIL, Privacy Impact Assessment (PIA), Methodology, February 2018 edition, p. 8; German Datenschutzkonferenz (DSK), Kurzpapier Nr. 5, Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO, 17.12.2018, p. 4.

(grant no. 101080718). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the NNF, the European Union or the Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.