POLICY BRIEF #32

9 December 2019

Does it hurt? The sensitivity of health data

Ine van Zeeland, An Jacobs, Anouk Verhellen, Jonas Albert, Jo Pierson

This policy brief addresses challenges in smart health under the European Union's General Data Protection Regulation (GDPR). One of the main challenges is mitigating risks when sharing health data, biometric data and genetic data - categories of personal data that are merited special protections in the GDPR. Issues of unclear definitions, even for such basic concepts as 'health data' and 'research', complicate practical implementation of personal data protection to such an extent that risks to individuals cannot be effectively addressed. We therefore recommend that either legislators or data protection authorities issue clear guidelines that can remedy the lack of clarity and improve the protection of patients while stimulating innovative health research.

1. What is health data?

The GDPR defines health data 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". Does this preclude any other type of information that could contribute to determining someone's physical or mental condition? For instance, if location data shows regular visits to a maternity hospital, should these data then be classified as health data? So-called "lifestyle" data, which includes information on exercise, mood or nutrition, among other topics, are often used as indications - or 'proxies' - of a person's health status. Moreover, medical research[1] has shown that social determinants such as poverty, education and employment significantly influence health. While these types of data are often freely available or readily shared via social media or smartphone apps, similar data could just as well be part of a medical record. Lifestyle information extracted from smartphone use can demonstrate behaviour patterns that indicate mental illness or altered physical well-being[2], while the person could be unaware themselves of the insights that can be generated by other parties with access to this data.
Though the GDPR offers definitions of genetic data, biometric data, and 'data concerning health', it is not entirely clear how these concepts should be delineated in a data economy where behavioural data are often used as proxies for health data. Distinctions matter, as genetic data, biometric data and health data are under a stricter regime for processing: the individuals that these data relate to, need to explicitly consent to any use, except when the data are processed to protect their

[1] E.g. Marmot, M. (2005). Social determinants of health inequalities. *The lancet*, 365(9464), 1099-1104.

[2] Hays, R. et al. (2019). Assessing Cognition Outside of the Clinic: Smartphones and Sensors for Cognitive Assessment Across Diverse Psychiatric Disorders. *Psychiatr Clin North Am*. Dec;42(4):611-625.

vital interests.[3] The rationale behind this stricter regime is that misuse of these data is likely to have more severe consequences for a person's fundamental rights, such as privacy and non-discrimination, than misuse of less sensitive types of data. As the data protection authorities have pointed out, clear-cut delineation of the concept of health data has been difficult because different EU member states have different conceptions of what constitutes health information.[4]

Lifestyle data, on the other hand, can be collected and processed on other grounds than consent, as they are not considered to be sensitive. For example, activity data can be processed based on a contract with a sports outfit or a fitness app provider. Such lifestyle data can be invaluable to care providers. In an imec-SMIT research project centred around self-management, called ProACT[5], older adults self-measured various health parameters using digital tools to increase insight into the impact of their lifestyle choices and changes on their health. Within this project, patients autonomously chose whom to include in their digital care network, giving the network access to their collected health data. While during the course of the project not many caregivers made use of the access they had to the health data, our analysis showed that they would appreciate such access in situations where patients have become unfit to discuss their health in person.

One important question about lifestyle data collected by wearables and apps is whether beside the user, a commercially driven provider (e.g. a private health insurer or technology company) also processes the data for re-use or sale. Among the risks associated with such unforeseen use are unfair treatment and exclusion. There are also health-related risks: incorrect or unreliable advice may have similar long-term health consequences for people as misuse of data labelled as 'health data'. It is often unclear for lifestyle apps to what extent their advice is supported by clinical research or even which norms the app relies on.

In conclusion, to assess whether personal data are used as health data, it will not suffice to label the data as 'sensitive' or not - the *intended use* must be considered. Imec-SMIT is proficient in working and experimenting with different types of data that is not generated by health and care providers. Through projects such as ProACT and Nervocity[6], we actively investigate the implications of collecting sensitive data on behaviour, physical wellbeing, policy and society via proof-of-concept studies and larger scale trials.

2. Sharing health data and genetic data for public or research purposes

New machine learning techniques offer the promise to find unforeseen patterns in extensive datasets. Pooling health data and genetic data could thereby greatly contribute to a better understanding of health conditions, for instance of the causes of certain diseases or of curative and preventative approaches. This is especially true for the study of rare diseases. Collections of electronic health records can be indispensable for research, but how are they best shared between individuals and

[3] An example of a 'vital interest' situation is when a person is incapacitated and a care provider urgently needs to obtain any available health records. Article 9 GDPR provides several other exemptions for exceptional situations.

[4] Annex to the Letter from the ART 29 WP to the European Commission, DG CONNECT on mHealth https://ec.europa.eu/justice/article-29/documentation/other-document/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf

[5] http://proact2020.eu/

[6] https://www.imec-int.com/en/nervocity

researchers - be they associated with commercial or public and non-profit organisations?
In many European countries, we currently encounter systems that enable electronic health data to be stored and transferred. Many of these systems appear to be mainly designed to share data that is recorded by healthcare professionals with other authorised healthcare professionals. In these cases, the patient usually has to give explicit consent to share his or her data with a healthcare professional who is not a direct care provider.
Research occupies a privileged position within the framework of the GDPR. Entities that process personal data for research purposes are exempted from certain restrictions on secondary processing, such as an individual's right to object to data processing. Exemptions differ between EU member states and 'research' is broadly defined, enabling a wide array of activities to be classified as research and thereby being eligible for exemptions. For research projects within the H2020 framework of the European Commission, health data is shared between the partners in research consortia with clear delineations on who may access data and who is able to view the data before they are de-identified. A data governance plan forms part of the application to an ethical commission that is needed for any type of research involving data from individuals. Participants are presented with an informed consent request form that needs to set out the limitations of use and re-use of the data.
One interesting insight that we gained from the ProACT project, which was also executed under the H2020 umbrella, was that participants tended to be dismissive about the protection of their personal data. This meant that researchers needed to be extra vigilant in addressing their rights and in explaining what they were consenting to, as well as how the regulations and standards protect them. In instances such as these, imec-SMIT has been able to collect first-hand experience in implementing GDPR-compliant trials, but also has been conducting research on the shortcomings on national and European legislation on the topic.

3. What are the risks of processing health data?

The main risks related to the processing of health data have to do with unlawful or unforeseen use or access. Risks to individuals that the GDPR mentions as a consequence of such misuse or unintended access are: discrimination, identity theft or fraud, financial loss, damage to the reputation, loss of confidentiality of personal data protected by professional secrecy, unauthorised reversal of pseudonymisation, or any other significant economic or social disadvantage.[7]
In early November 2019, news came out of a research project between Google and Ascension, a hospital chain and health insurer in the United States, that resulted in the transfer of 50 million medical records to one of the world's biggest advertisers. The individuals whose medical records were transferred to Google were neither asked nor informed. In addition, the data was not de-identified. While both Ascension and Google assured that the project's goals were benign, patients reportedly felt uncomfortable with Google knowing about such conditions as mental health issues or venereal diseases. Previously, Google had also closed a deal with the United Kingdom's National Health Service (NHS) to access millions of patients' hospital records, again with many patients unaware, prompting intervention by the UK's supervisory authority on data protection, the Information Commissioner's Office.[8] People's discomfort aside,

[7] Recital 75 GDPR

[8] Royal Free NHS Foundation Trust update, July 2019, https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/07/royal-free-nhs-foundation-trust-update-july-2019/

risks of manipulation[9], discrimination[10] and deficient data management practices by technology companies are real.
The government eHealth portals in Belgium, such as mijngezondheid.be, base the processing of health data within these portals on consent. Patients are asked to consent to the portal's processing of their medical data in order for the patients themselves to have access to it. The patient cannot, in theory, see any of their medical data online if they have not given their consent, as it is only locally stored with the respective healthcare professional, similar to the paper-based recording of healthcare data. Remarkably, giving your consent as a patient to sharing your data with the portal also entails that you consent to the exchange of the records between healthcare professionals - a "dual consent". There are three ways this consent can be given, namely 1) via the eHealth portal or eHealth platform that the patient uses, 2) via their general practitioner, pharmacist or hospital staff, or 3) via the patient's health insurance provider. The objective of this procedure is that when the patient consents, it is clearly explained what they are consenting to. There are plans to change this dual consent to two different consent requests where a patient can agree to see their own data but disagree to share it with other healthcare professionals.
Unforeseen access to health data makes patients uncomfortable and comes with risks, such as unintentional disclosure of a person's health status to their social circle. A consequence may be that patients withhold vital information from their care providers for fear of loss of confidentiality, with potentially detrimental effects.[11]

4. Recommendations

Based on the above, we have the following recommendations:

1. ***For data protection authorities and legislators: Provide guidance on the delimitations of "health data", "lifestyle data" and related categories***
   In order to have a differentiated discussion and to properly apply national and European legislation to all types of data which are relevant to health, universally agreed-upon delineations have to be pursued. Alternatively, strict guidance on distinctive criteria can support practical implementations.
2. ***For legislators: Provide clear rules for transparency on norms underlying advice in lifestyle apps with health claims***
   We need to acknowledge the value of lifestyle advice for health improvement, as well as the actual use of lifestyle advice by individuals who believe to improve their health. For any health-related claim it should be clear what the advice is based on and which data lead to which conclusions.
3. ***For eHealth portals: disentangle consent requests***
   As the GDPR requires granularity and specificity for consent to be valid, consent for data processing by different parties should be disentangled, i.e. dual consent should be replaced by two different consent requests.

[9] European Commission (2017), Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service, https://ec.europa.eu/commission/presscorner/detail/en/IP_17_1784
[10] Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. nyu Press.
[11] Heaton J. (2019). Why patients are picky about what health data they're willing to share. *HealthData Management* (5 December 2019), at: https://www.healthdatamanagement.com/opinion/why-patients-are-picky-about-what-health-data-theyre-willing-to-share

About the authors:

Ine van Zeeland is a PhD student within the VUB research chair on Data Protection On The Ground.

An Jacobs is the Programme lead of the "Data & Society" programme at imec-SMIT. She is also an associate professor in the Department of Media and Communication Studies at the Vrije Universiteit Brussel.

Jonas Albert is in charge of the research unit "Digital Health & Work Living Labs" at imec-SMIT.

Anouk Verhellen is a Digital Health Researcher at imec-SMIT.

Jo Pierson is in charge of the research unit 'Privacy, Ethics & Literacy' at imec-SMIT and associate professor in the Department of Media and Communication Studies at the Vrije Universiteit Brussel. He holds the VUB research chair on Data Protection On The Ground.

SMIT (Studies in Media, Innovation and Technology) is an imec research group at Vrije Universiteit Brussel.