When data protection becomes dangerous

Thousands of medical studies are in jeopardy because personal information of participants may no longer flow from Europe to America. ***By Piotr Heller***

At first glance, Robert Eiss doesn't exactly have the most spectacular job at the National Institutes of Health (NIH). As an advisor to the U.S. research agency, he sometimes has to grapple with data protection issues. If a medical study stalls because European data are not allowed to flow to America, the case lands on his desk. But as dull as it sounds, thanks to this task he looks through a burning glass at a problem that is currently challenging international biomedical research.

One case is etched in his memory. It occurred at the NIH's in-house clinic. There, a patient was being treated for lymphoma. Proven therapies were unsuccessful, and his only chance was an investigational immunotherapy that required a stem cell donation. Doctors found a match in a German donor registry. Because the therapy was a research project, they also needed personal information and samples from the potential donor. But because of the European General Data Protection Regulation (GDPR), which has been in effect since May 2018, this data was not allowed to reach the U.S. agency. The German donor center refused to cooperate.

"Imagine the healthcare experts, who were caring for the patient and were not able to provide therapy because of a data privacy law", says Eiss. After weeks of negotiations, the NIH did manage to convince the German side to get permission from the donor and send the data. "Our German counterpart indicated to us, that this is a one off exception", he says. That's a problem because 40 percent of the donors the NIH accesses are in Europe. The next similar case was not long in coming.

This story stands out because here the GDPR had a direct impact on a patient's treatment. Much more often, purely scientific studies are affected. A diabetes study that had been running for 25 years had to be paused for 18 months. At least

40 research projects on cancer risk factors are on hold. The NIH has invested $1.5 million to collect data from a Danish birth cohort. Scientists wanted to use it to explore genetic factors for gestational diabetes. "Following the GDPR our Danish counterpart asked us for the data and bio samples to be returned to Copenhagen", Eiss explains.

Other European databases containing blood samples or genetic material have also stopped their data flows with U.S. institutions. NIH supports a over 5000 medical studies in Europe. "We estimate that most of them will be affected by the regulation," Eiss says. Meanwhile, three major academic networks from Europe have weighed in on the issue. Recently, they published a report together for the first time. In it, they appeal to the European Commission to finally take up the cause. "Less global exchange of health data for research hurts everyone," the authors write.

At the heart of the problem is that, unlike Switzerland or Japan, for example, the United States is not considered a trustworthy partner for data protection from the EU's perspective. They lack an "adequacy decision," as it is called in bureaucrat jargon. That's why personal data is not allowed to cross the Atlantic from Europe without further ado - whether it's commercial data from Amazon or Facebook or information from medical studies. "I think the people who drafted the GDPR and implement the GDPR at the EU member state level now, they think with great concern about Amazon, Google, Facebook – that's their frame of reference," Mark Barnes explains. The Boston attorney represents major U.S. universities and also does pro bono work for international medical collaboration. He describes biomedical research as a little caboose on the train, that is pulled along by those forces. "They're actually not related to us at all. We are trying to do better science by having more data," he complains. But to get the data from Europe, U.S. institutions must sign contracts with non-negotiable standard clauses. They require them to open U.S. data systems for audits or submit to European jurisdictions in

privacy disputes. NIH, as a federal agency, is not allowed to do that, and state universities also have statutes prohibiting them from doing so.

It would be too short-sighted to see data protection as a mere obstacle. Medical studies rely on sensitive information such as disease diagnoses or drug use details. Even if they don't include names, the people behind them can be tracked down using Big Data and public gene databases for genealogical research. Back in 2013, Harvard professor Latanya Sweeney proved she could identify 241 of 1130 people from a DNA study based on their zip codes, birthdays and gender. In addition, a 2019 study by Dutch researchers shows that volunteers want to share their health data only on the condition that their privacy is respected. So privacy is in the interest of scientists, including U.S. scientists. Barnes asserts that there are no problems with medical research in the United States. The NIH, the National Science Foundation and state universities have a sterling track record of protecting privacy, he says. "It's no more risky for European data to be there in the NIH versus to be at an European University", he adds. So the problem is purely formal. European researchers, not least the report's authors, are also insisting that the matter be addressed here in Europe.

After all, it's about time. Giske Ursin is feeling the consequences of the wrangling that has been going on for three years now. As head of the Norwegian Cancer Registry, she oversees one of the largest blood serum banks in the world. Samples from 300,000 Norwegians over several decades are stored there. 80,000 of them have developed cancer over time. This data can be used to study such things as risk factors and early signs of cancer. "Despite this large database, we rely on international collaboration for rare diseases like gallbladder carcinoma," Ursin says. When American colleagues ask her for samples for their studies, she is no longer allowed to comply. Like many others, she resorts to workarounds. One of them is to do all the analysis in Europe - after all, American data may still be sent to Europe. "But we can't do all transatlantic

cooperation at our site," Ursin says. An alternative is to analyze the European data in Europe and the American data in America and then compare the results. But that's expensive, lengthy and for statistical reasons, very rare phenomena cannot be identified in such frayed data sets.

The problem is not limited to federal institutions in America. It's true that the GDPR offers opportunities to share personal data with private universities, but their design is so complex, Ursin says, that four cancer registry studies are currently waiting for lawyers to hammer out the details. "At some point, Americans won't even consider us because they know how complicated it all is," Ursin says. According to Mark Barnes, that situation has already occurred in many cases.

Gerard Schellenberg of the University of Pennsylvania also sees privacy frustrating his colleagues. He researches the genetic basis of Alzheimer's. When he studies rare gene variants, he too needs large samples from multiple countries. To solve the problem, he hopes to collect all the information on a European database. American researchers can then perform analyses there and download the results without seeing the data themselves. But it's questionable whether the necessary infrastructure even exists in Europe. "We're wasting the talent of young researchers here who have good ideas about how to analyze the data but just can't do it because the information isn't in one place."

The authors of the report and other experts propose a number of solutions to the European Commission. One would be to include exceptions for research data in the standard clauses so that government agencies like those in the United States can sign them. Another would be an international code of conduct in which research institutions commit to data protection. The EU proved in April 2020 that it can in principle allow exceptions. At that time, the issue was the transfer of data for Covid-19 research, but here, too, experts criticized the strict design. For example, the data had to be deleted again after a certain time. Robert Eiss says that sixty institutions in Europe refused to

participate in studies on monoclonal antibodies for Covid-19 therapy because of the GDPR.

The European Commission does not want to comment on the report. It points out that the rules on data transfer have not changed at all with the GDPR. They have merely become clearer. In addition, the regulation offers tools for transferring personal data, which also apply to research. The Commission has investigated cases in which data protection allegedly stood in the way of the flow of information to American authorities. Often, privacy was not the problem after all. The European Data Protection Board seems to perceive the matter as somewhat more urgent. Upon request, the committee stated that it was aware of the report and had forwarded it to its own experts for discussion. This year and next year, they plan to develop further guidance for scientific research.

Robert Eiss has the feeling that after three years the matter is slowly reaching Brussels. He says this was seen, for example, in the fact that the GDPR was on the agenda during the recent visit of U.S. Secretary of State Antony Blinken. But time is pressing, as illustrated not least by the second case of a cancer patient at the NIH clinic. Eiss describes this case carefully, without assigning blame, and says it is unclear exactly what factors had an impact on this individual's health. But the facts are these: Again, an experimental treatment with stem cells was planned. Again, the donor was in Germany. Again there were long negotiations because of data protection. Before they came to a conclusion, the patient's condition had worsened to such an extent that he was no longer eligible for the therapy.