Patient data in the cloud

As growing volumes of patient data are stored on externally hosted platforms, worries are mounting among patient advocates that patient data might be at risk of privacy breaches.

In some cases, government and private health systems alike are entrusting their patient data to cloud data service providers such as Amazon, Apple, Google, and Microsoft. In other cases, health systems are entrusting patient data to vendors of electronic records systems to aid clinical decision making and record management.

Not surprisingly, the scale and speed of the patient data rush among data-hungry entities worries many data security, confidentiality, and fair-use advocates.

On September 23, in an open letter to Congress's Committee on Energy and Commerce, the American Health Information Management Association, the American Medical Association, the American Medical Informatics Association, the College of Healthcare Information Management Executives, the Federation of American Hospitals, and the Medical Group Management Association, warned that within pending legislation designed to improve information sharing, "it is imperative that policies be put in place to prevent inappropriate disclosures to third-parties and resultant harm to patients."

After stating support for "patients using apps to access their information," the six groups recommended that the US federal Government should "adopt a holistic and coordinated approach to addressing the access, exchange and use of health information by third parties... including the sale and commoditization of data not intended by patients."

In a report on digital health and artificial intelligence (AI) for the UK National Health Service in February, Eric Topol (director and founder of Scripps Research Translational Institute in San Diego, California) noted that advances in mathematics, computing power, cloud computing, and algorithm design have accelerated the development of methods that can be used to analyse, interpret, and make predictions using these data sources. Topol also emphasised that AI has the potential to transform the delivery of healthcare in the NHS, from streamlining workflow processes to improving the accuracy of diagnosis and personalising treatment, as well as helping staff to work more efficiently and effectively.

Topol's report also warned that "without a legally enforceable and effective system of data governance, which the British public regard as ethical, respectful of rights, and secure and trustworthy, the promises of digital healthcare technologies could be undermined."

Heed this warning, says Phil Booth, coordinator of UK-based Med Confidential, a UK based patients' advocacy group. "Governments are losing control of patient data to the extent we don't have proper audit systems and requirements governing patient data," Booth cautions.

Booth and Topol both point to a controversy in which the Royal Free NHS Foundation Trust, a London hospital group, failed to comply with the UK Data Protection Act when it provided patient details to Google DeepMind, a London-based unit of Google Health that plays a leading role in Google's cloud platform businesses. The Trust provided personal data of around 1·6 million patients to DeepMind as part of a trial to test an alert, diagnosis, and detection system for acute kidney injury.

In July, 2017, an investigation by the UK Information Commissioner's Office found several shortcomings in how the data was handled, including how patients were not adequately informed that their data would be used as part of the test.

"There's no doubt the huge potential that creative use of data could have on patient care and clinical improvements, but the price of innovation does not need to be the erosion of fundamental privacy rights," UK Information Commissioner Elizabeth Denham said about the matter in July 2017. Denham recommended the Trust "establish a proper legal basis under the Data Protection Act for the Google DeepMind project and for any future trials."

Changes in technology "mean that vast data sets can be made more readily available and can be processed faster and using greater data processing technologies," Denham added in a blog post, "but just because evolving technologies can allow you to do more doesn't mean these tools should always be fully utilised, particularly during a trial initiative." In the wake of the controversy, Booth says it remains difficult to know what the NHS is doing with the patient data in its vast repositories. The advent of cloud storage repositories that are accessible to multiple users for an almost infinite variety of purposes—medical, commercial, and possibly even for state surveillance usage—has severely destabilised traditional health records management customs, he warns. "We've uncovered all sorts of instances where basic data control requirements were being breached," Booth charges.

"We don't want this to happen again," Topol told *The Lancet Digital Health* in an interview about the Google DeepMind controversy. "We can all learn from it."

Topo highlights that a contract announced in September between Google and Mayo Clinic, one of the most prestigious US hospital networks, indicates that health systems are intent on entrusting large volumes of patient data to external data platforms. But while there is the potential for trouble to arise from some such arrangements, he added, there are also great potential benefits for patients. "Deep depth of talent is needed for data analysis and

For the **open letter to Congress's Committee on Energy and Commerce** see https://chimecentral.org/wp-content/uploads/2019/09/HIT-Provider-Group-Letter-to-Senate-EC-FINAL.pdf

For the **UK NHS report on digital health and artificial intelligence** see https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf

For more on the **Royal Free NHS Foundation Trust controversy** see https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2017/07/royal-free-google-deepmind-trial-failed-to-comply-with-data-protection-law/

For more the **Google and Mayo Clinic contract** see https://cloud.google.com/blog/topics/customers/how-google-and-mayo-clinic-will-transform-the-future-of-healthcare

For more on **Xuefeng Jiang's study** see https://www.ncbi.nlm.nih.gov/pubmed/31546251

For more on the **Cosmos Initiative** see https://www.beckershospitalreview.com/ehrs/epic-unveils-patient-data-research-initiative-new-software.html

For more on **Microsoft Azure** see https://azure.microsoft.com/en-ca/resources/videos/ochsner-health-system/

For more on the **Cerner and Amazon partnership** see https://www.cerner.com/blog/cerner-leads-new-era-of-health-care-innovation

that can come when health systems link with organizations with powerful artificial intelligence capabilities. If there is good complementarity, it can be a good thing."

Mayo Clinic's aim in partnering with Google, explained Christopher Ross, Mayo Clinic's chief information officer, in an interview with *The Lancet Digital Health*, is in large part to probe whether combining genetic information with patients' clinical histories and social and economic data could help to deliver more effective treatments for patients with rare and complex diseases.

"Our datasets are some of the most robust sources of clinical insights in the world," Ross said. "This partnership will help propel a multitude of artificial intelligence projects by our scientists and physicians. Working with Google will provide us with the technological tools to deliver answers on a vastly increased scale."

As part of the partnership, Ross said, Mayo Clinic could decide to share de-identified patient data with Google and other parties for specific research projects. But Mayo Clinic "retains complete control" over its patient data within its arrangement with Google, Ross emphasised. The data, he added, "can only be used for the purposes of research dedicated to improving patients' healthcare," and will not be used for non-medical purposes or the development of non-therapeutic products.

"It's Mayo-controlled private data that we keep on behalf of our patients," he stressed. "Google doesn't have any right or ability to get to those data."

Patients can also take comfort from knowing that the deal with Google might possibly enhance protections against theft and misuse of data, Ross indicated. "Google has vast resources dedicated to data security," he noted. "It's at the core of their business."

Xuefeng Jiang, a health data security specialist at University of Michigan who recently published a study that reviewed records from the US Department of Health and Human Services from 1461 data breaches at 1388 entities that affected 169 million patients in aggregate, agrees that data exported to external cloud hosts from health organisations, like Mayo Clinic, are likely to be better protected against external and internal theft and misuse than data stored internally by health organisations.

"That's the inference we drew from talking with computer scientists and seeing the mistakes made by healthcare providers," Jiang told *The Lancet Digital Health*.

Mayo Clinic's compact with Google is part of a patient data-gathering drive spearheaded not just by cloud services providers such as Amazon, Apple, Google and Microsoft, but by electronic medical records vendors as well.

In August, US-based electronic medical records software vendor EPIC announced that its Cosmos Initiative now aims to compile records from 25 million patients in order—like the Mayo-Google plan—to widen its information base for research into medical practices.

So far, nine health-care systems are contributing patient data to Cosmos, Sumit Rana, senior vice president of research and development for Epic, said at a gathering of 17 000 health-care IT specialists convened by EPIC in late August. The company is in discussions with more than 30 other customers that could contribute data, which would result in the data of about 25 million patients being included. EPIC has also announced the development of a new Cognitive Computing Platform, built on the Microsoft Azure cloud, that Microsoft says will drive predictive analytics and AI into a wide range of workflows.

"Louisiana's largest health system, Ochsner Health System, used advanced machine learning algorithms to create a predictive model leveraging Epic's machine learning platform powered by Microsoft Azure to accurately predict patient deterioration hours before an adverse event," the Microsoft Azure website explains. "This early warning system was tightly integrated into Epic, enabling Ochsner's Rapid Response team to intervene on patients proactively, rather than reactively."

In June, US-based Cerner, which, like EPIC, controls a substantial share of the global electronic medical records software market, announced a partnership with Amazon web services.

Within this deal, Amazon web services "is providing Cerner with the broadest portfolio of innovative analytics and machine learning services that will empower them to gain new clinical and business insights that have the potential to transform patient care delivery," Andy Jassy, Amazon web services CEO, explained in a media release posted by Cerner.

With cloud service providers announcing new deals almost daily to secure vast troves of patient data for advanced analytical probes, Eric Topol warns that there is a potential for trouble alongside powerfully positive possibilities. "We need to have extensive data security surveillance as well as guardrails," he says.

One solution to averting trouble could lie in a new AI form called "federated AI", in which a patient's data never leaves the health system and is analysed locally, Topol says.

"That is a better model than one in which the data is given or shared with a company or tech entity. I'd like to see more movement toward this kind of safeguarded approach in which the data is not leaving the premises."

Paul Webster