



Award for Research Data Stewardship Case Study:

Mayo Clinic & Optum

Research from Mayo Clinic and Optum: Predicting the Risk of Severe Hypoglycemic and Hyperglycemic Events in Adults with Diabetes

Rozalina McCoy MD, MS from Mayo Clinic and Mehwish Qasim from healthcare services company Optum are the recipients of the third-annual FPF Award for Research Data Stewardship. Their collaboration on a project called, *“Predicting the Risk of Severe Hypoglycemic and Hyperglycemic Events in Adults with Diabetes,”* was composed of addressing a critical gap in diabetes management in the United States by examining the epidemiology of severe hypoglycemia and hyperglycemia.

Robust privacy protections were built into the project – including the use of de-identified data sets, hashing methodologies, encryption, and administrative safeguards. The team utilized Optum data via an open, collaborative research and innovation center that follows a robust data policy that prevents the transfer of Protected Health Information, physically limiting access to labs and provided secure virtual desktop infrastructure. Optum systems are housed in a secure, 24-hour monitored facility with physical controls to limit access.

“Optum takes seriously the protection of health care member and patient information with comprehensive safeguards, practices, and policies,” said Dr. Mehwish Qasim, Senior Director of Academic & Advocacy Partnerships at Optum. “At the same time, Optum supports independent research that aims to accelerate ground-breaking research that translates into improved health, better patient experience, and reduced costs.”

The Research Project

Diabetes affects more than 37 million people across the United States, making it the country’s most common and costly chronic disease. Despite advances in the science of diabetes therapy, rates of diabetes complications have not improved over the past decade, especially as it relates to garnering glycemic control to avoid severe hypoglycemia and hyperglycemia. With little research about acute hypoglycemic and hyperglycemic diabetes complications available, many clinicians lack a practical and reliable means to identify patients at the highest risk for severe hypoglycemia and hyperglycemia, preventing the ability of the patient and their doctor to individualize glycemic goals and therapy. By developing risk prediction models from real-world data, the Mayo Clinic and Optum addressed the critical gaps in diabetes management to emulate clinical trials, improving glycemic control and reducing risks of severe hypoglycemia and hyperglycemia.

Data Protection Procedures and Processes in the Research by Mayo Clinic and Optum

1. **De-identifying Data.** Optum employs certified de-identified datasets, following HIPAA guidelines and a hashing methodology to match individuals from multiple data sources while preserving the statistical de-identification of the data. The de-identified status of the data is certified through an expert determination. Furthermore, Optum employs measures designed to ensure that no personnel ever attempts to re-identify data or link it against internal and external data sets.
2. **Establish Limits on Sharing and Use of Personal Health Information (PHI) Data.** Files containing any PHI are prohibited from being transferred, according to Optum's privacy policy. Users are, therefore, only allowed access to data under an agreement with Optum and only from approved locations. Each user is trained, authenticated, and authorized before gaining access to data, as well as being provided credentials with multi-factor authentication.
3. **Limiting The Transmission of Data.** Secure File Transfer Protocol (SFTP) accounts are required to move files in and out of Virtual Desktop Infrastructures (VDIs), and both inbound and outbound files are reviewed for compliance with the Optum Data Policy. No Optum Systems containing Optum data were transported or transmitted off of Optum premises unless encrypted and transported with the owner's approval. Data was only transmitted via encrypted channels such as a virtual private network (VPN), encrypted File Transfer Protocol (FTP), Secure Sockets Layer (SSL), Hypertext Transfer Protocol Secure (HTTPS), or SFTP.
4. **Physical and Technical Safeguards.** Optum Systems are housed in a secure facility with physical controls to limit access, such as locks or physical tokens, and are monitored 24-hours every day. Technical safeguards were employed, including encryption, identity management, anti-malware, and intrusion prevention/detection to prevent improper access, use, or disclosure of the data in their networks. There was a multi-party review of each research proposal for the appropriate use of the data before allowing researchers access to the data.

Lessons for Future Data-Sharing Projects

The data-sharing collaboration between the research team at Mayo Clinic and Optum highlights several valuable lessons that companies and academic institutions may apply to future data-sharing collaborations.

- **The Importance of Accessible, Private Data.** The ability of large, de-identified data sets to be published and made accessible for researchers and journalists have informed over 35 studies seeking to improve diabetes management and health outcomes. This work was uniquely possible because of the volume (over 140 million people, representing a diverse population), variety (claims, labs, and electronic health records), velocity (timely availability of data that allows rapid-cycle innovation), veracity (rigorous and standardized data cleaning, validation and linkages) and value of this data.

The Result

The availability of a large dataset aided researchers in identifying indicators for severe hypoglycemic and hyperglycemic events by providing thorough context for patients' health history. This work addresses critical gaps in diabetes management and helps provide patients and providers with reliable guidance.

The study discovered several key findings, including the fact that severe hypoglycemia and hyperglycemia are common among adults with diabetes in the United States, with no improvement over time. Additionally, the strongest risk factors for severe hypoglycemia are geographical and socioeconomic status, multimorbidity, and high-risk conditions like kidney disease. Researchers also found the risk of severe hyperglycemia increases with inadequate treatment, whereas the risk of severe hypoglycemia increases with excessive treatment.

Claims data (like those used in this study) can be utilized to emulate clinical trials to compare the effectiveness of medical interventions. In this study, data was used to compare the effectiveness and safety of second-line diabetes medications for maintaining glycemic control. For example, glucagon is an emergency treatment for severe hypoglycemia and should be given to all patients at risk of developing it. However, even after severe hypoglycemia, it is rarely prescribed, and patients from racial and ethnic minorities, as well as low-income patients, are the least likely to have access to glucagon.

The Selection Process

Nominees for the Award for Research Data Stewardship were judged by an Award Committee composed of representatives from FPF, academics, and industry leaders. The Award Committee evaluated projects based on several factors, including their adherence to privacy protection in the sharing process, the quality of the data handling process, and the company's commitment to supporting independent academic research.

To learn more about best practices for instituting research data-sharing programs between corporations and research institutions, download the Future of Privacy Forum's "[The Playbook: Data Sharing for Research](#)." This report addresses vital steps for data management, sharing, and program execution between companies and researchers while aiming to encourage safe, responsible data-sharing between industries and researchers.