

## A Framework for Using Electronic Intraoperative Anesthesia Records for Genomic Discovery

**Author:** Matthew A. Levin, MD, Icahn School of Medicine at Mount Sinai, New York NY

**Co-Authors:** Thomas T. Joseph, MD, PhD Icahn School of Medicine at Mount Sinai, New York NY; Janina Jeff, PhD, Icahn School of Medicine at Mount Sinai, New York NY; Eimear E. Kenny, PhD, Icahn School of Medicine at Mount Sinai, New York NY

**Introduction:** During the perioperative period, patients are routinely monitored to ensure their safety and well-being. This monitoring generates a wealth of clinical and physiological data. To date, few if any researchers have fully utilized this rich source of data for genetic discovery. We have developed a framework that integrates physiologic data from our institution's electronic intraoperative anesthesia record with data from our genetic repository.

**Methods:** The framework integrates data from two primary sources: The Charles F. Bronfman Institute for Personalized Medicine (IPM) Biobank (*BioMe*) and our operating room data warehouse (ORDW). *BioMe* is an ongoing, consented Electronic Health Record linked bio- and data repository that integrates phenotypic data from clinical systems with detailed ancestry information obtained via questionnaire, and genomic data derived from banked blood samples taken at the time of patient enrollment. ORDW is a comprehensive data warehouse that contains detailed physiological data from our Anesthesia Information Management System (CompuRecord, Philips Healthcare, Andover MA) along with perioperative clinical, laboratory and administrative data. The pipeline for integrating these two data sources is implemented in SQL and Python, and automated using *chron*. Key design goals were to create an idempotent, automated, secure, unidirectional data pathway with robust logging and a clear audit trail. Briefly, a list of patient identifiers for current *BioMe* participants is pushed on a bi-weekly basis to a secure separate database within ORDW. A Python job uses these identifiers to find anesthetic records. Each *BioMe* participant may have had zero, one, or more anesthetics. An opt-out flag set by *BioMe* identifies patients who should be removed. Simple set arithmetic is then used to identify new, old, and disappeared cases, and a timestamp and status set accordingly. Once per quarter a list of HIPAA compliant identifiers is pushed from ORDW back to *BioMe*, and detailed phenotype and ancestry data as well as flags indicating the availability of genotype data are extracted. This data set is pushed back to ORDW. Subsequently, detailed intraoperative physiologic and clinical data is extracted using standardized Python scripts and combined with genotype data on a per project basis.

**Results:** The pipeline went live in June 2015. At that time, there were 31,797 patients in *BioMe*; 15,891 had at least one anesthetic (49.9%). Total number of anesthetics administered to these patients was 42,260 (mean 2.7, median 2 anesthetics per patient, range 1 – 130 anesthetics). Collectively there were 38,149,718 vital signs and 414,709 bolus drug administrations recorded. As of October 2015, an additional 521 patients were enrolled in *BioMe*, bringing the total number of *BioMe* patients to 32,318; 16,164 (50%) had at least one

anesthetic. The total number of anesthetics increased by 479 to 42,739, of which 248 were on new patients, and 231 were on existing patients. 12,757 (80.3%) of patients that have had an anesthetic have genotype data. Historical analysis indicates an anticipated yearly growth rate of approximately 5% for both patients and anesthetics.

**Conclusion:** We have implemented a robust, scalable pipeline to find and extract intraoperative anesthetic data for participants and merge it with genetic data from our institutional biobank. This pipeline will enable us to investigate the relationship between genotype and the response to anesthetic and surgical stress.