

ENABLING FUTURE CARE NOW

The Princess Margaret Data Science Program

Cancer is a complex group of diseases. As we expand our understanding of these diseases and integrate new technology into care and research, the volume of data generated continues to grow exponentially. The progress in computerization and data science presents a unique opportunity to integrate, analyze and extract meaning from rich data assets.

The Princess Margaret Data Science Program (PMDSP) aims to be at the forefront of applying and discovering novel analytical approaches and predictive models that could be leveraged to inform care, fuel new research, optimize operations, and enhance learning. The program engages a diverse team of clinicians, scientists, analysts and technology developers and is co-led by Tran Truong - Program Director, Dr. Benjamin Haibe-Kains - Scientific Lead, Dr. Alejandro Berlin - Clinical Lead, and Luke Brzozowski - Partnership Lead. The PMDSP is aligned with the UHN-wide data science strategy and supports local, national and international data science collaborations.

Leveraging data driven tools in patient care can push the boundaries of delivering personalized medicine and maximize our ability to predict outcomes, identify the

most effective individualized therapies, monitor treatment response, and prevent toxicities. "There is no doubt that access to multidimensional and high-quality data can empower providers and researchers to new heights, in direct benefit of our current and future patients," says Dr. Berlin. Data science also holds the promise of enabling transformational discoveries through basic and translational research that can uncover previously undetectable trends. There is untapped potential that can arise from integrating heterogeneous, seemingly disconnected sources of data such as clinical and research data sets, and external data from community, environmental, social and patient-reported sources. "We are at a point where the dimensionality of the data collected during the patients' journey and in research laboratories exceeds by far what the human brain can grasp. The development and adoption of new computational tools will be crucial to assist clinicians and researchers in their work," says Dr. Berlin.

To realize this potential, the PMDSP is working to develop an integrated cancer data platform, utilizing UHN's advanced data lake technology to centralize multi-modal data sources including clinical documentation, medications, imaging, multi-'omics' (e.g., genomics, transcriptomics, proteomics), patient recorded outcomes and research registries. In



addition, the program is working to develop an oncology data catalog and a platform for cancer centre-wide outcomes collection that will improve data capturing and enhance documentation.

A series of catalyst projects were launched in 2019 to foster engagement across the centre and inform the development of a data science framework at the Princess Margaret. The projects were selected based on scientific quality, opportunities for large-scale data integration and the potential to build infrastructure and enhance our collective expertise. The ultimate results of these projects will have broader impact and contribute to the development of platforms to support delivery of value-based care and research, automated processes for integrating and analyzing institution-wide data routinely, and computational platforms to jointly analyze multimodal molecular, imaging and clinical data.

Opportunities for learning and collaboration between data scientists, clinicians and researchers are vital to deepen our understanding of the benefits, applications and limitations of AI and deep machine learning in cancer. The team is poised to collaborate with leaders in AI research as well as identify opportunities for internal collaborations across programs, departments and professions. “Collecting complex biomedical data, developing advanced machine learning algorithms and deploying new computational models in research and clinical settings require a highly multidisciplinary team,” says Dr. Haibe-Kains. “This is why the PMDSP is working with the broad community of the Princess Margaret towards a common goal: improve cancer care for the benefit of our patients.” 🤖

The Princess Margaret Data Science Leadership Team



Tran Truong,
Program Director



Benjamin Haibe-Kains,
Scientific Lead



Luke Brzozowski,
Partnerships Lead



Alejandro Berlin,
Clinical Lead

Read the full report to learn more about our progress and activities in clinical care, research and education in 2019.

[View Annual Report](#)