Riaz Alvi
Saskatchewan Cancer Agency Saskatoon Cancer Centre

Improving access to cancer data in Saskatchewan

Dr Riaz Alvi will design a data storage system to improve access to cancer surveillance data in Saskatchewan.

Most cancer-related data in Saskatchewan is collected and stored by the Saskatchewan Cancer Agency, which is responsible for ensuring quality of cancer care in the province. However, the way the data is currently stored and managed can make it difficult to access and analyse. This impacts projects such as the Canadian Strategy for Cancer Control (CSCC), which aims to reduce the burden of cancer in Canada. With funding from the Canadian Cancer Society, Dr Alvi is leading a project to make sure the data from Saskatchewan can be accessed more easily by those who need it, so that it can be better used to improve cancer care. The team will seek to understand the gaps in the current data storage and analysis system, so they can develop solution that makes the data more accessible. If successful, the team hope that their strategy can also be used to improve d use of other cancer surveillance data in Saskatchewan and Canada more widely.

Dr Darren Brenner
University of Calgary

Improving access to cancer data

Dr Darren Brenner will create an interactive online data dashboard system to allow more people to access and download cancer data in Canada.

The Canadian Cancer Statistics report is published annually and contains detailed information on cancer in Canada, including estimates of new cancers, cancer deaths and cancer survival as well as sex, age, location of people affected. The yearly data also includes estimates of cancer trends over time and the probability of developing or dying from the most common types of cancer in Canada. Currently, the data is published in a large document, which although detailed and comprehensive, is most suited to people with experience in statistics and not widely accessible to a broader audience. Canadian Cancer Society-funded researcher Dr Brenner will create an

innovative data dashboard to present the information in a much more accessible and flexible way, enabling more people to interpret and easily use the information. The dashboard will have the statistics already generated, but will also have new features, so that people can search by criteria such as age and cancer type and create their own visualizations, which they can then download for personal or professional use.

Dr Kathleen Decker
University of Manitoba

Adding data about cancer progression to cancer registries

Dr Kathleen Decker will develop a machine learning tool to extract information about cancer progression from medical records.

Cancer registries are databases of information about cancer cases which are frequently used for research and to evaluate various cancer control strategies such as cancer prevention campaigns. Cancer progression, which is when a tumour grows or spreads, is important clinical information that is not currently collected for these registries. This is because details about cancer progression are not automatically added to registries and are only found in electronic medical records. Currently, extracting this information requires a manual, time-consuming approach which may prevent researchers from effectively using this data. With support from the Canadian Cancer Society, Dr Decker will lead a team to develop a machine learning tool to process text within medical charts to extract information about progression. The researchers will use charts from individuals treated for breast or colorectal cancer and test the tool to see how effectively it captures progression information. The research team hope that this tool will help accelerate research that examines cancer progression by automating the chart review process for some cases. The team will make their tool available publicly, meaning it can be used all over Canada improving the amount and quality of data that registries contain.

Dr Philippe Després
Laval University

A database of lung cancer images to train artificial intelligence tools

Dr Philippe Després is creating a database of lung cancer scans and clinical information for training artificial intelligence tools to help radiologists screen people for lung cancer.

Lung cancer is the leading cause of cancer death in Canada and new ways of detecting it earlier are needed to improve outcomes. Screening programs for lung cancer exist in some jurisdictions but have been found to add a lot of pressure on healthcare systems, partly due to a high rate of false positive results. With Canadian Cancer Society funding, Dr Després will lead a team to develop a high-quality dataset of lung cancer images to assist in training artificial intelligence models. These models can help radiologists screen patients for lung cancer and assist clinicians with making decisions about patient care, potentially saving time, resources and increasing sensitivity and specificity of detection. The dataset will include information from 4,000 lung cancer cases and combine scan images with clinical information such as disease characteristics and also research data such as genetic analysis. It will be made available to the research community to aid in the development of artificial intelligence tools, ultimately leading to more lung cancers being detected at an earlier stage, improving patient outcomes.

Dr Jeff Dowden
Newfoundland Cancer Clinic

Improving the Newfoundland Cancer Care Registry

Dr Jeff Dowden is improving the way that cancer data is stored and accessed in Newfoundland.

The Newfoundland Cancer Care Registry is an important repository of data on cancer, containing information about screening for breast, colon and cervical tumours, as well as tumour surveillance and treatment information. However, the way that some of the data is currently stored can make it difficult to analyse. For example, there are many scanned documents, where data is not easily searchable. With funding from the Canadian Cancer Society, Dr Dowden will lead a team to vastly improve the registry information system for the province, using innovative artificial intelligence techniques. This will make the data much easier to access and analyse. This project will result in fewer hours of manual work to process and analyse the data, improving efficiency. This, together with improved access to the data will mean that Newfoundland can better analyse patient care, cancer screening and health system performance, ultimately leading to improvements for patients. The researchers hope that if successful, the method they develop can be expanded to other areas of healthcare, further improving access to information in the province and further afield.

Dr Khaled El Emam
CHEO Research Institute

Using synthetic data to enable sharing of clinical trial results.

Dr Khaled El Emam will use machine learning techniques to create synthetic data from real-world cancer clinical trial results to enable secure, private data sharing.

Multi-centre clinical trials have greatly improved outcomes for people with cancer and have generated significant quantities of data for research. Currently, however, making these datasets widely available for researchers to use in subsequent research studies has been difficult due to privacy concerns. With support from the Canadian Cancer Society, Dr El Emam will investigate the possibility of generating 'synthetic variants' of data, a technique which takes original data from trials and creates a synthetic dataset closely mimicking the findings and patterns of the original. This means that the synthetic data can be shared and analysed by other groups of researchers without concerns for privacy of individuals in the trial. Dr El Emam and team will evaluate 3 different techniques to produce the synthetic data from real-world breast cancer trial results and perform experiments to see if research questions using these synthetic datasets come to the same conclusions as with the original trial data. If successful, the use of synthetic datasets will greatly speed up sharing of data, allowing researchers to conduct further experiments=.

Dr Benjamin Haibe-Kains
Princess Margaret Cancer Center – UHN

Making complex cancer data more accessible.

Dr Benjamin Haibe-Kains will further develop a platform to make complicated and large datasets more accessible to researchers.

The complexity of biomedical data has increased substantially in recent years in tandem with technological advances in techniques like genomics. Data outputs from these analysis techniques often require specially trained scientists to process them into a form where they can be used to answer research questions. This creates a bottleneck to using the data in research. Canadian Cancer Society-funded researcher Dr Haibe-Kains will lead a team to further develop a data platform called ORCESTRA which will enable the curation of data in a way that makes it

accessible and ready to analyse for researchers. ORCESTRA currently contains results from almost 40 datasets. The team will use the new funding to expand this, including clinical trial results for patients treated with new targeted therapies and immunotherapies. By hosting the data in this ready to analyse way the team hopes to remove barriers and save time, enabling researchers to verify existing results and make new discoveries.

Dr Morag Park
McGill University

Sharing breast cancer data with the world

Dr Morag Park will develop a web portal to make breast cancer datasets available and accessible to researchers around the world.

Breast cancer is the most common cancer type and the second leading cause of cancer death in women in Canada. The genetic alterations that cause breast cancers and influence how they respond to treatment are highly varied, meaning some women have excellent prognoses while others die from the disease. Canadian Cancer Society-funded researcher Dr Park has been researching breast cancer for 2 decades and has generated a vast amount of data on over 600 breast cancer tumours. The data is publicly available but can only be analysed by researchers with specialized skills in bioinformatics. Dr Park will lead a team to develop an open-source web portal to make it easier to store, access, visualise and analyse the data. The portal will include anonymized information on the genetic alterations found in the breast tumours as well as clinical data and information about how the tumour responded to treatments. The researchers will continue to add to the portal as more samples become available and development of this open-access resource will enable researchers from around the world to access and use the data, contributing to further research projects and ultimately improving patient outcomes.

Dr Trevor Pugh
Princess Margaret Cancer Centre – UHN

A central storage solution for liquid biopsy data

Dr Trevor Pugh will develop and test a centralised storage solution for detailed liquid biopsy cancer data, driving innovation in earlier detection.

Diagnosing cancer often involves taking a piece of the tumour to analyse in a lab. But more recently “liquid biopsies” – which use samples of bodily fluids like blood or even urine to look for tiny fragments of DNA from tumour cells – have been a key area of interest in cancer research, with the potential to revolutionise diagnostics and early detection. The amount of data generated by liquid biopsy analysis is substantial and there is often no centralised repository to store the data, which slows down research progress. With support from the Canadian Cancer Society, Dr Pugh will lead a team to create an accessible storage solution for this data, allowing researchers to quickly and easily share it with their colleagues to enable greater innovation and collaboration. Creating a central repository for liquid biopsy data will enable more researchers to use the data, leading to more efficient research processes and quicker impact on diagnosing patients at an earlier stage.

Dr Sarah Quirk
University of Calgary

Automating radiation therapy data collection

Dr Sarah Quirk will develop a system to automate the collection of detailed radiation therapy data to help model treatment efficacy and side effects.

Radiation medicine systems used to deliver radiotherapy treatments to people with cancer collect lots of data including on the type of therapy given, dosing schedule and dose given. These treatments can come with side effects and the standard information collected currently misses lots of data that affect treatment toxicity and efficacy. Manually collecting this data from clinical systems and linking it up to other clinical data to investigate efficacy and treatment related side effects is inefficient and better systems are required to streamline this process. Canadian Cancer Society funded researcher Dr Quirk will lead a team to develop an automated system to collect detailed radiation medicine data, first on a cohort of 1,500 patients with prostate cancer treated with radiotherapy in Calgary. This will enable clinicians and researchers to gain a greater understanding of radiation treatment efficacy and also toxic side-effects, ultimately improving patient outcomes.

Dr Danielle Rodin
Princess Margaret Cancer Centre – UHN

Evaluating the effectiveness and cost of lymphoma therapies

Dr Danielle Rodin will use design a database to evaluate the cost and effectiveness of a type of cellular therapy used in lymphoma treatment.

Approval of new cancer drugs in Canada is based on clinical trial results. But not all patients are eligible for clinical trials and those who do enrol tend to be younger and healthier. This means that, sometimes, drugs that perform well in trials don't work as well when applied to a wider population. In lymphoma, this is of particular interest because there are over 60 different subtypes of lymphoma and the 'real-world' effectiveness of drugs and the cost to the healthcare system is not currently well known. With funding from the Canadian Cancer Society, Dr Rodin will lead a team to investigate which treatments are effective for patients and also good value for the healthcare system. The researchers will create a new database that includes molecular information about a person's disease, and merge this with clinical information on outcomes, treatments and hospitalizations. The team will test the database initially by looking at the effectiveness and cost of a cellular therapy called CAR T-cells. The results will be used to help inform patients, clinicians and policymakers. The researchers hope that this project will provide new information about the cost and efficacy of drugs for lymphoma and better inform decision making for treatments.

Dr Lincoln Stein
Ontario Institute for Cancer Research

Streamlining procedures allowing researchers to access patient data

Dr Lincoln Stein will lead a project to simplify access to cancer datasets, making it easier for researchers to use the information.

Many types of cancer research rely on the use of large datasets containing information about thousands of patients and their samples. One way to do this is to combine multiple smaller cohorts, but accessing and compiling different studies is often difficult. Institutions that hold and manage the data frequently having different access rules and procedures, which slows research

and limits use of the data. Canadian Cancer Society-funded researcher Dr Stein will lead a team to create a new system to improve and streamline the way in which these datasets are accessed by designing new models for reviewing and approving applications from researchers who apply to access the data, while maintaining important ethical and legal requirements.

Dr Robin Urquhart
Dalhousie University

Linking national health data to improve cancer research

Dr Robin Urquhart will combine large health datasets to create an impressive resource for cancer research in Canada.

Canada has a wealth of health care and health research data, but numerous barriers prevent efficient sharing of these datasets between provinces. This, combined with further difficulties with linking datasets together, causes major limitations in using the data productively to improve the health of Canadians. With support from the Canadian Cancer Society, Dr Urquhart will lead a team to connect several cancer registry and administrative health data repositories, creating a unified resource for cancer research in Canada. One of the datasets, CanPath, contains data on 1 in 100 Canadians, including lifestyle, family histories, genetics and environment and this, combined with other datasets, will be analysed to understand the causes of cancer and what happens to people with cancer in the short- and long-term. The researchers will also be able to ask important questions about the impact of the COVID-19 pandemic on people who have been diagnosed with cancer during this time. Linking these huge datasets will provide a very important resource for cancer research in Canada.

Dr Yuan Xu
University of Calgary

Transforming patient data into evidence to improve cancer care

Dr Yuan Xu will develop and use algorithms to extract important information from medical records which is often missed.

Detailed information about people with cancer and their clinical care is currently stored in large databases, but is not often easily accessible by researchers. This significantly limits researchers' ability to analyse the effectiveness of healthcare delivery and improve cancer outcomes. Canadian Cancer Society-funded researcher Dr Xu will lead a team to develop algorithms that combine administrative and medical record data and identify important information such as recurrence of disease, which is often missed in current databases. The research team will look at historical data from over 70,000 people with solid tumours in Alberta to develop their algorithm. The team hopes that this will make it easier for researchers to access this important data to conduct studies on cancer outcomes, accurately allocate healthcare resources for treatment, and inform cancer prevention strategies.

Dr Katherine Zukotynski
London Health Sciences Centre Research Inc

A database of prostate cancer scan images

To support research and clinical education, Dr Katherine Zukotynski will compile a database of scan images of prostate cancers imaged using a new radioactive tracer technique.

A new imaging technique is increasingly being used to help identify the stage of prostate cancer in men. The technique involves a radioactive drug that binds to a protein found on prostate cancer cells. A PET scanner can detect the cells that have bonded to the drug, making it easier for medical staff to see the tumour and determine treatment. Canadian Cancer Society-funded researcher Dr Zukotynski will lead a team to create a database of scan images using this new technique to be used for research and education. The database will contain images from patients with several different stages of prostate cancer and will also be used to develop artificial intelligence tools that can help clinicians stage and monitor the tumours. The researchers hope that this resource will also play an important role in educating clinical professionals working with people with prostate cancer, ultimately improving the staging of their tumours and management of their disease.