

The centre for health informatics: a novel approach to accelerating the field of health data science

Danielle A. Southern^{1,*}, Cathy A. Eastwood¹, Erin O'Connor¹, Chelsea Doktorchik², Marcello Tonelli³, Jon Meddings³, Hude Quan¹, and Tyler S. Williamson¹

Submission History

Submitted:	07/11/2023
Accepted:	19/02/2024
Published:	21/03/2024

¹Centre for Health Informatics, Cumming School of Medicine, University of Calgary, Calgary, Canada

²Faculty of Medicine & Dentistry, University of Alberta, Edmonton, Canada

³Cumming School of Medicine, University of Calgary, Calgary, Canada

Abstract

Precision Medicine and Precision Public Health are approaches to improve population health. Achieving these goals requires innovation in health informatics. The Centre for Health Informatics (CHI) within the Cumming School of Medicine (CSM) at the University of Calgary (UC), Canada, was created to respond to this need by fostering multidisciplinary collaborations, building capacity by recruiting and training outstanding faculty and students, and harnessing Alberta's rich health data to advance health informatics. To establish CHI as a health informatics leader, CHI has struck partnerships with stakeholders, including Alberta Health Services (AHS), Alberta Health (AH), and the Alberta Strategy for Patient-Oriented Research Unit (AbSPORU) among others. Through these close relationships, the CHI intake team facilitates access to Alberta's rich health data sources and educates researchers on the available health data in Alberta. The concept of a "One Stop Shop" for CSM and UC researchers encourages multidisciplinary collaboration, helps investigators access a wide range of datasets, and receive analytical support. Population-based data sets enable the development of methods to turn raw data into health information, improve health data collection, linkage, analysis, and quality, and applied studies creating clinical decision-support tools, prognostic tools, improved health surveillance methods, and health system performance indicators. CHI's ecosystem of diverse research expertise, cutting-edge technology, and embedded AHS analysts to support data access via a wide-ranging network of partnerships allows our provincial researchers, national and international collaborators tremendous opportunities for empirical research. It paves the way for implementing Precision Medicine in the real world.

Keywords

health data; Precision Medicine; Precision Health; health informatics; data science

*Corresponding Author:

Email Address: dasouthe@ucalgary.ca (Danielle A Southern)



Introduction

In 2019, the Centre for Health Informatics (CHI) [1] was established to support the University of Calgary (UofC) Cumming School of Medicine's (CSM) vision of Precision Health [2]. Precision Health includes Precision Medicine (determining disease risks and treatments that will work best for an individual) and disease prevention and health promotion activities [3]. Under the vision and direction of Dr. Hude Quan, the concept and infrastructure for creating CHI to be a "One Stop Shop" for data science was established. The 'shop' includes technology, faculty, and staff with varied expertise, access to secure data storage and cluster computing, data access via embedded AHS staff, and training programs to create the next generation of highly trained data scientists. The Centre for Health Informatics' vision is to improve health and healthcare for all through data-driven innovation and collaborative research.

The mission of CHI is to be a trusted resource to transform health data into evidence for the community through collaboratively developing innovative health informatics science for Precision Health and Precision Public Health. This mission is to be achieved through ensuring data validity, robust methods, unbiased analytics, and scientific presentation of data and maximizing data value, responding to societal needs, and providing opportunities for education and mentorship.

CHI moved into a physical location on the fifth floor of the recently renamed Cal Wenzel Precision Health Building at Foothills Medical Centre Campus, Calgary in January 2019 and held a Grand Opening May 1, 2019. The bright and modern space was designed to facilitate collaboration among multi-disciplinary investigators to advance health informatics for Precision Health through timely data access and data analytics, and lead the use of secondary data for research and clinical decision support tools. Secondary data in CHI refers to information that has been collected for a purpose other than the current study's specific research question. These data are already available and have been gathered by government agencies, healthcare organisations, or research institutions. Without having to collect new data, there is a potential for researchers to utilise secondary data to address their research questions. CHI is home to seven faculty members with permanent space and their research teams, two collaborative research teams (SPHERE and SeroTracker), and three faculty members in CSM with drop-in space for collaborative purposes. In addition, CHI has collaborated with and supported over 38 different researchers from across the University's faculties (academic and clinical research appointments) over its four operational years. CHI hosts many different research groups including the Alberta Strategy for Patient Oriented Research Unit (AbSPORU) Data Research Services team [4], the SPOR National Data Network [5], the University of Calgary World Health Organization (WHO) Collaborating Centre for Classifications, Terminologies, and Standards, SeroTracker [6], the Strategy for Precision Health in Breast Cancer (SPHERE) [7], and the Data Intelligence in Health Lab (DIH) [8] among many research programs for the various faculty or related faculty members in the Centre. CHI residents have grown to over 60 current researchers, including the faculty, support staff, students, trainees, and administrative personnel. CHI supports the CSM research

institutes and research groups by providing a physical space for embedded AHS analysts. CHI supports these groups by integrating their analysts in the Centre and providing a collaborative environment for interaction with other analysts, AHS oversight to assist with privacy concerns related to the data and training opportunities for analysts to learn more about the datasets.

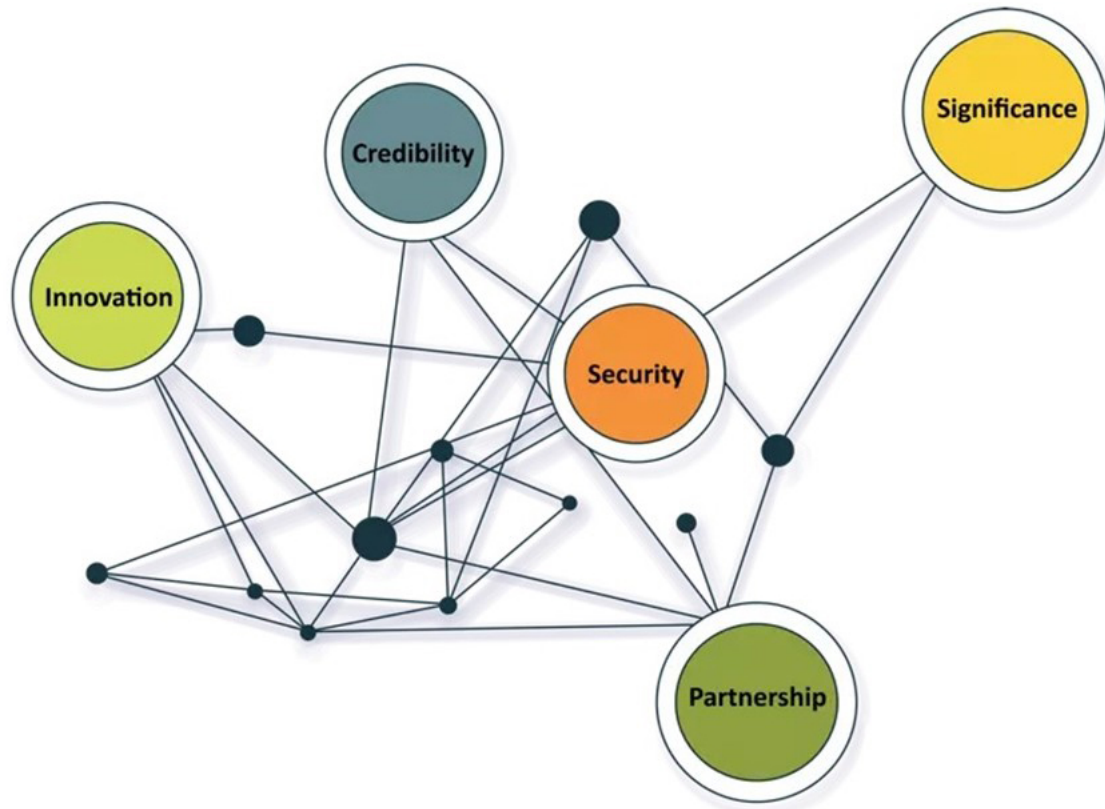
Centre for Health Informatics' faculty and support staff consists of many individuals from varying disciplines like biostatistics, epidemiology, computer engineering, physics, clinical coding, and nursing. By providing a space for socialisation, CHI fosters inter-disciplinary collaborations for cutting-edge new projects. The physical CHI space was specifically designed to foster a collegial environment for faculty, research support team members, and analysts to learn from each other, work together through the various backgrounds to further Precision Health, and improve efficiency through standardised approaches. In addition, the CHI culture has purposely been built to encourage collaboration with weekly internal newsletters, informal social gatherings, and knowledge-sharing opportunities.

CHI has worked toward connecting researchers, industry leaders, analysts, and students on several projects since opening. The impact of CHI in the research community goes beyond those physically located in CHI and has created an impact throughout CSM and its Institutes, the main campus of the University of Calgary, the industrial community, the general population in Alberta, the research community in Calgary, Canada, and globally. CHI looks to grow connections with academia, industry, and the public.

Centre for health informatics values

The core principles of CHI are *Innovation*, *Significance*, *Credibility*, *Security*, and *Partnership* (Figure 1). CHI leadership, faculty, and support staff who are physically located in CHI believe in *innovation* as a mindset. CHI strives to redefine the standard of excellence in our healthcare system by fuelling young minds and novel research. Conducting research that is relevant and *significant* ensures that our work positively impacts the lives of Albertans and Canadians. At CHI, we aim to produce trustworthy results, dependable solutions, and become a reliable, *credible*, long-lasting source of healthcare information. The ability to handle, use and transmit healthcare data is recognised as an invaluable resource that aids in creating a seamless pathway for data collection, analysis, and collaboration directly within our Centre. CHI employs a proactive approach to implementing data *security* measures including administrative controls, physical security, logical controls, and organisational standards [9]. To accommodate the complexity and urgency of the challenges presented in the health care system today, *partnership* and collaboration among government, health services researchers and academic health science networks is vital. CHI believes that partnership is essential to the implementation of flexible and sustainable healthcare solutions. Building partnerships has involved reaching out to various professionals and groups such as the faculties of nursing, science, and engineering, among others to create a network within the health informatics and data science communities. Early discussions with partners include exploring

Figure 1: Centre for health informatics values



other data sources such as social determinants of health, data related to justice and other sources not traditionally used in health research.

Approach

The Centre for Health Informatics focuses on the transformation, analytics, and application of health data. Each area requires expertise in several domains and collaboration for transition from one to the other. Overall, the Centre has three main priorities of Research and Design, Training and Education, and Collaboration, Service and Community.

Research and design

A primary goal of CHI is to build a strong academic community in data science. A network of experts specialised in data science have been recruited to establish CHI as a recognised hub for health informatics research. Academics from diverse healthcare and scientific disciplines reside in CHI, including experts in health systems research, epidemiology, public health, and clinical research. The variety in expertise encourages innovative research, facilitating the creation of multi-disciplinary teams for project development and co-supervision of students.

To build our diverse network of researchers, CHI leaders made a deliberate effort to enhance faculty and student recruitment strategies from data-science related fields. These fields include data mining and machine learning, artificial intelligence, health data visualisation and computer science, implementation science, and mathematical prediction and modeling. To attract such high-calibre recruits, CHI helps

facilitate access to high quality provincial and national-level data, including registry data (for example the Alberta Provincial Project for Outcomes Assessment in Coronary Heart Disease (APPROACH) and the Canadian Primary Care Sentinel Surveillance Network (CPCSSN,) administrative data (including hospital discharge abstracts, ambulatory care records, pharmaceutical dispenses, practitioner claims and demographic information) [10], imaging data, province-wide electronic medical record data, and national survey data by linking researchers with the data custodians. More information about available data for research in Alberta and how to request access to data extracts can be found in our prior publication related to governance [9].

Training and education

Regular communication and presentation of research is encouraged within our faculty and staff community, typically in the forms of weekly, biweekly, and monthly roundtable discussions. Student and staff-led Health Informatics Knowledge Club provides a platform for all students and staff to share relevant articles, innovative methods, or projects with the team. A Health Informatics Speaker Series [11] is held once per month, where CHI invites external experts and researchers from local, national, and international universities and healthcare organisations to share their work and build collaborations, and CHI shares events and team updates on social media (<https://linktr.ee/ucalgarychi>).

Dr. Tyler Williamson, the current Director of the Centre for Health Informatics, serves as the Director of the Health Data Science and Biostatistics Diploma Program [12], a specialisation in the Data Science and Analytics Program,

jointly offered by the faculties of Science, Business, and the CSM. This is the first Health Data Science-focused program in Alberta, that launched in January 2020 and provides a pipeline of well-trained students into both research careers in the CSM, AHS, and throughout Alberta's growing health tech economy.

CHI also hosts a summer studentship program each year. Students at different stages in their degrees are accepted (high school, undergraduate, graduate) for a summer term to work on a project related to health informatics, under the mentorship of both a faculty supervisor and a CHI staff member. This program benefits both the students, who are exposed to data science methods and projects, and the staff who can expand mentorship opportunities and advance leadership skills.

Collaboration, service, and community

CHI works directly with various research institutes and groups from across the CSM to co-design and conduct research projects. CHI's goal is to orient UC investigators to the data that are available and support the development of analytical plans (including assistance with applications for ethics approval and data access). Within a cost recovery model, skilled analysts support the extraction of defined variables according to the requirements of UC investigators, assembly, and delivery of appropriately sized datasets to UC investigators (after appropriate ethics and data governance approvals) and complete the analysis of the data.

As an example, CHI has developed a strong working relationship with the Libin Institute [13], and more specifically, the Libin Precision Medicine Initiative [14]. This initiative aims to harmonise and combine Alberta's world-renowned data sources into a single, 'super' database which will allow researchers to conduct life-changing cardiovascular research. Within Libin, hospital administrative claims data and cardiovascular clinical data sets are being linked and used to monitor population health, implement Precision Medicine for physicians and patients, and aid in the design and development of community cardiovascular health interventions more closely.

CHI works closely with the O'Brien Institute for Public Health (OIPH). In 2020, leadership within the OIPH requested support to monitor the ongoing surveillance information about the spread of COVID-19. CHI developed a COVID-19 dashboard or "Tracker". COVID Tracker has been used to inform policy decisions for the City of Calgary's pandemic response (e.g., closing certain businesses early in the pandemic, supporting decision-making about relaxing restrictions, decisions around mandating mask use in all public spaces), and was the model for the later developed provincial COVID-19 dashboard. COVID Tracker was a valuable and trustworthy source of information to all Albertans, both the public and decision-makers [15]. The easily accessible design ensured that all Albertans could understand the data and interact with the data through basic visualisations, enabling them to make informed decisions for their own lives and activities. In 2022, COVID Tracker was viewed over 670,000 times [15] and is now the primary source of ongoing provincial surveillance information for COVID-19 in wastewater. Through COVID Tracker, CHI was able to demonstrate credibility, responsiveness, and relevance through the partnerships developed.

As previously mentioned, CHI hosts several research groups on the floor to encourage collaboration. SPHERE, which focuses on improving breast cancer outcomes and treatments through Precision Medicine, big data, and machine learning, has been able to leverage the CHI environment, becoming a collaborative health services research environment that informs best practices for Precision Medicine in breast cancer, ultimately improving outcomes for patients across Alberta. SPHERE's research focuses on understanding how and why breast cancer affects different cohorts in different ways. Study results are used to identify, inform, and create innovative strategies to improve breast cancer care quality, efficiency, and outcomes [7]. DIH is a group of health data scientists interested in applying data science, machine learning, artificial intelligence, natural language processing, software engineering, visualisation, and mobile technology to challenging health problems. Their particular focus is on unconventional digital health data sources. Some projects include leveraging machine learning to enable precision coronary artery disease patient care and using AI to protect children from unhealthy food and brand marketing [8].

Alberta real world evidence consortium

In partnership with the University of Alberta and the Institute for Health Economics at the University of Calgary, CHI supports the Alberta Real World Evidence Consortium (RWE) [16]. The RWE enables academic researchers with primary affiliations with the University of Calgary, the University of Alberta, or the Institute of Health Economics to study important population-based questions with industry sponsorship. The RWE engages key provincial stakeholders involved in Real World Evidence generation and utilisation to advance provincial capabilities, to ensure high-quality projects that are managed to timelines and budget, and to help focus projects and provincial resources on joint priorities between sponsors and the province. CHI has a dedicated team to develop these initiatives further, set up processes, protocols, and financials and carry out projects for these industry partners. Beyond legal contracts, the same protocols are followed for accessing data for research which involve ethics and data disclosure agreements [9].

Public and patient involvement with research

CHI works closely with AbSPORU which specialises in patient-oriented research and provides several ways for patients and the public to be involved in research [17]. Researchers are encouraged to involve patients and referred to AbSPORU for information and tools for doing so.

Governance, legislation, and management

The Centre facilitates researchers to access data through an established procedure for requesting, accessing, storing, and computing health data resources within Alberta [9]. The Centre for Health Informatics hosts members of AbSPORU Data and Research Services (DRS) team that provide the knowledge and expertise around data access governance, legislation, and extraction of AHS, AH, and other custodian-owned data sources. In general, research studies require ethics,

data custodian review and approval, researcher agreement and a data disclosure agreement with AHS and a data extract submission [9]. Data is accessed through AbSPORU and is released to researchers with non-identifying health information as appropriate, for example removing or encrypting patient identifiers such as personal health numbers and postal codes [18]. AbSPORU is funded by Alberta Innovates and the DRS team consists of AHS analysts that support access to provincial health data. CHI provides a central location for these analysts within the University setting which enables close collaboration between academics and AHS, open sharing of expertise, and safe, legal data access for healthcare researchers. This directly benefits the research community by having data experts co-located with those that frequently use the data, but it also benefits the health system directly. Questions of high priority can be given appropriate attention, mutual training occurs, and the principles of a learning health system are implemented. CHI utilises the University of Calgary's High-Performance Computing infrastructure which was created in partnership with AHS and the Secure Computing Platform allowing storage and sharing of level 4 (restricted) data in a secure environment and the Medical Advanced Research Computing (MARC) cluster for processing Level 4 data [9].

CHI also hosts a secure Statistics Canada Research Data Centre. Data users who become deemed employees of Statistics Canada can be granted access to the data room and access a wide variety of data, including social and business surveys, administrative data, and linked data. The confidentiality of respondents is protected using policies and procedures managed by Statistics Canada. Becoming a "deemed employee" includes attaining a security clearance, completing a mandatory training, and swearing or affirming the Oath of Office and Secrecy to Statistics Canada [19]. Statistics Canada staff in this unit help facilitate data access to national statistics datasets for researchers [20].

Secure data storage and cluster computing

Our partnership with the Research Computing Services (RCS) of the University of Calgary has enabled advanced computing storage and power. An Information Management Agreement between AHS and CHI enables the use of a secure, high-performance computing (HPC) GPU/CPU infrastructure environment [21, 22]. The advanced computing options of UofC make large data analysis feasible, including analytical techniques such as deep learning. For example, one study involves the linkage of clinical registry, Electronic Medical Record (EMR), and hospital administrative claims data and includes predictive analytics. However, with big data comes challenges with the analytical capacity of networks and computer resources. The University of Calgary RCS provides a secure cluster computing environment to facilitate such analyses and secure data storage for PIs to remain compliant with the Health Information Act.

Key activities & noteworthy outputs

Key activities

The Centre for Health Informatics is at the front of the imminent data wave and connects health data for

transformation, analytics, and application in our society (Figure 2).

Transformation of data

CHI strives to understand data and its complexities. To achieve high-quality research, one requires quality data. According to the Canadian Institute for Health Information (CIHI), the five dimensions of quality are accuracy and reliability, relevance, timeliness and punctuality, accessibility and clarity, and comparability and coherence [23]. CHI strives to help researchers understand data quality and explore how data is collected and how valid it can be for research purposes. Objectives of this work include standardising and refining the basic terminologies, classifications, and standards used to catalogue the available data elements; developing innovative data collection and linkage methods of multiple data sources; transforming the data into analysable formats; developing algorithms to define conditions and indicators of the health system and population health; improving data quality; and in turn devising novel methods for statistical prediction and modelling. CHI links researchers with the UC's WHO Collaborating Centre for Classifications, Terminologies and Standards, and AbSPORU for data translation expertise to reach these goals.

Transformed data remain available to researchers; however, access to the raw data is available through analysts within AHS in the Centre. Knowledge Club seminars provide all analysts with shared techniques and issues identified through analysis of data. The Centre's partnership with data custodians on these data issues enables transferable learning, reproducible research and research at scale.

Analytics and visualisation

Machine learning methods capitalise on the rich clinical data available through EMR and other data sources. CHI boasts a high-tech visualisation room allowing researchers from the CSM and across the UC campus to investigate novel ways to represent and visualise data. The overarching aim is to use state-of-the-art visualisation methods to transform data into explainable information in partnership with clinicians and patients, facilitate communication with policymakers and the public, create new insights about improving care or reducing costs, and develop new methods of data visualisation.

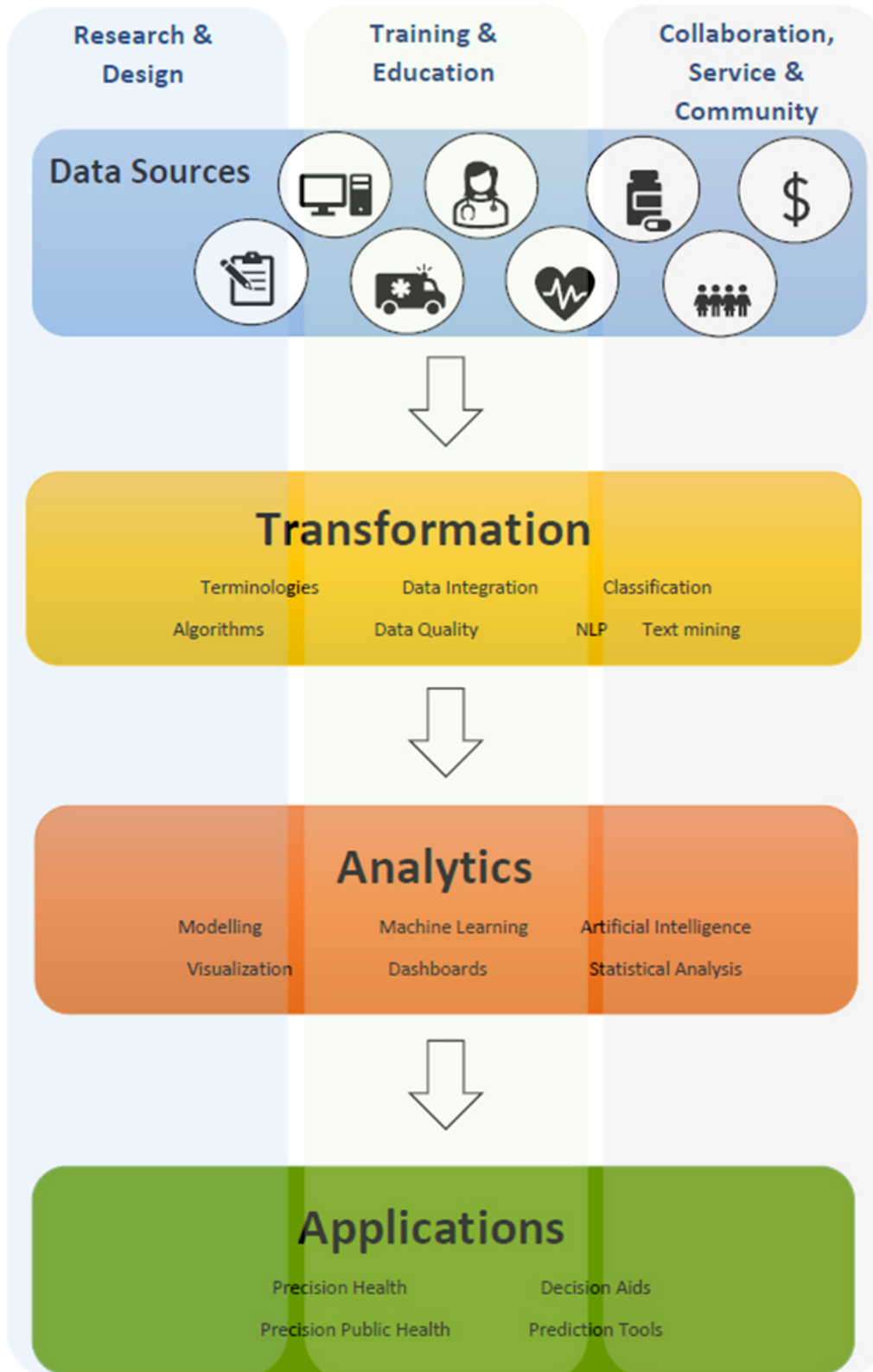
Applied studies

Some of the applications CHI researchers are working on are Precision Medicine applications (e.g., clinical decision support; prognostic tools) [24]; precision public health applications (e.g., health promotion and surveillance) [25, 26]; RWE analyses [27]; evaluation of health system performance [28, 29]; and identification of potential policy tools [30].

Training & capacity building

Since novel methods and new data sources are constantly emerging, sustaining research activities in CHI provides and participates in ongoing training programs for faculty and professional research staff. For the reasons given above, AHS is working to upgrade the skills of its analytical staff to maximise

Figure 2: Overview of centre for health informatics



utility of the EMR data – especially those stored as free text or as inconsistently formatted data (e.g., imaging or pathology reports).

The CHI collaborates with AbSPORU and AHS to develop advanced training programs focusing on hands-on practice and applying new data management methods to complex data. For example, CHI analysts use machine

learning to mine EMR data to define conditions and predict events.

Community and government support

With the COVID-19 Pandemic, CHI’s diverse team of experts used its unique infrastructure, close partnerships with AH

and AHS, and ability to facilitate direct AHS data access to create a COVID-19 “Tracker” to inform evidence-based policy decisions. CHI plans to continue transforming data into information for other topics to help inform the public, aid in policy decision-making, and support clinical decisions with expertise and timely data access to drive forward Precision Medicine and Precision Health. CHI has established an Associate Director of *Research Strategy and Partnerships* role who is responsible for establishing strong partnerships both internally at the University of Calgary and externally with health system, academic, provincial, national and international partners.

Noteworthy outputs

Responding to the COVID-19 pandemic to provide a surveillance dashboard, predictive modeling, and a novel database

The World Health Organization announced COVID-19 as a global pandemic on March 11th, 2020. CHI was tasked by the City of Calgary and CSM to provide information on COVID-19 developments in Alberta and characterise the Canadian and international pandemic. To create a surveillance dashboard, CHI quickly developed an internal team, including software engineers, data scientists, visual design experts, statisticians, and epidemiologists [1]. The dashboard provided information such as new cases and total case counts, testing efforts, mortality rates, cases stratified by age groups, and hospitalisation rates. Upon developing a “COVID-19 Task Force,” CHI became involved in several other projects, including predictive modeling for COVID-19 outbreaks and trends [15, 31–33]. Our dashboard is a credible and reliable source for researchers, decision-makers, and the public. The dashboard project enabled a partnership with the Pan-Alberta Wastewater Monitoring Group and led to several grants and publications for that group [34]. Wastewater samples are taken three times a week and results are shared on the COVID Data Tracker website, which allows members of the public to view recent data points in different communities to see how cases are trending. CHI provides regular online reporting of the wastewater results from Calgary since 2020 with publicly available real-time data.

Predicting hospitalisations in patients with hypertension using linked data and advanced modeling

Precision Health relies on accurate prediction. Therefore, CHI undertook a study exploring machine learning (ML) methods for predicting cardiovascular disease (CVD) outcomes in patients with hypertension [35]. The study used routinely collected administrative health data (hospital claims, physician claims, laboratory, and pharmacy medication data) to determine whether it can predict individual patient outcomes for heart failure, myocardial infarction, and stroke in hypertensive patients. In addition, the study included demographic data, lab test results, medication information, and comorbidity status to improve prognosis assessment and guide treatment decisions for hypertensive patients. The researchers found that all four models exhibited high discrimination and calibration in predicting CVD outcomes,

with the NMTLR model having the best individual survival prediction and the RSF model having the best population survival prediction.

This precision health study suggests that improved prediction of outcomes can help clinicians make more meaningful treatment decisions for hypertensive patients. Additionally, the study highlights the potential of administrative health data, routinely collected but underutilised by clinical health systems, to be used in predictive models for CVD outcomes. Overall, this study is one example of the growing body of research at CHI exploring the potential of ML in healthcare and highlights the importance of using routinely collected data to improve patient outcomes [35].

Data resource to support cardiac precision health

The CREATE database is a collection of health data from various sources, including the Sunrise Clinical Manager (SCM) Electronic Medical Record (EMR), the Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease (APPROACH) clinical registry, and the Discharge Abstract Database (DAD) and National Ambulatory Care Reporting System (NACRS) administrative databases. The data elements captured include clinical information, demographics, diagnoses, and procedures related to cardiac catheterisation patients in Calgary, Alberta between April 1, 2011, and March 31, 2017. The data were linked through personal health numbers, last name, first name, sex, and date of birth, and data exploration and cleaning were conducted to ensure data accuracy. The data has been used for research purposes to advance the understanding of cardiovascular disease and improve patient outcomes. The CREATE database is housed in the joint AHS and UofC environment [36]. The CREATE database is part of an initial effort to link EMR data to clinical registries and the administrative data for Precision Health research innovations. We anticipate similar projects will be undertaken within the Centre for Health Informatics within the next few years.

Sustainability and future work

The vision, goals, and identity of CHI became clear over the first operational year. The CHI team fully believes in the future of the Centre as a key investment for the CSM in driving forward Precision Medicine and Precision Public Health. In the infancy of CHI, investment funding from the Cumming School of Medicine was needed to build CHI into an operation that can create enough revenue to sustain staff and operations and continue to grow in the future. CHI aims to be financially self-sustaining through revenue generated by contracted research services, Alberta Real World Evidence Consortium projects, and philanthropic donations.

Launching a unique data science diploma program

The cross-campus collaborative project of establishing a Master in Data Science and Analytics program at the University of Calgary has been a success. The program has grown to have four specialisations from different faculties across the University including the CSM with additional specialties being planned. The Health Data Science and

Biostatistics (HDSB) Diploma program, as part of the ladder Master of Data Science and Analytics program facilitated by the Faculty of Science, is hosted by CHI with students being integrated into the Centre throughout their semester. Students can interact with staff working in the field, collaborate with other students in different programs with different expertise, and build relationships with CHI faculty and their instructors which are often part of the CHI community. This program has been successful in creating a pipeline of students to working professionals in the field of health informatics. CHI and related research groups within the Centre, have hired 3 previous students from the program that now work in various health specialisations and are employed by AHS or UofC. Students have gone on to work in the health field, elsewhere in industry, or founded a thriving health start-up venture. The program has had 56 graduate students complete the Health Data Science and Biostatistics Diploma at CHI and is preparing for future cohorts to reach up to 50 students per year.

Building a pathway for secondary use of EMR for research

As Alberta evolved to a trans-provincial connected computer information system in 2020 (Connect Care), CHI is building a pathway involving training staff to use EPIC data for secondary research purposes. CHI is focused on developing methods to facilitate the use of EPIC data and building a foundation for using EMRs, in general, for clinical decision support. EPIC is the name of the new electronic health record system for patients that rolled out under Connect Care [37]. Our team, with CHI faculty Dr. Shawn Xu and Dr. Cathy Eastwood leading the project, was awarded CIHR funds for Developing EMR-based algorithms to identify hospital adverse events for health system performance and evaluation. This will provide natural language processing algorithms to enhance investigations of hospital adverse events.

Conclusion

In conclusion, CHI is an established hub at the University of Calgary for innovation, collaboration, and training on data science techniques and health informatics. Through close relationships with AHS and AbSPORU, we are advancing data analysis skills, linking novel data sets, and serving the public with precision health and information transformation projects. Future plans include supporting complex linked data projects with novel data sets such as ECG waveforms, and outpatient EMR text.

Ethics statement

Ethics approval was not sought for this manuscript as its sole purpose is to describe the Centre for Health Informatics.

Statement on conflicts of interest

The authors have no conflicts of interest to declare.

References

1. Centre for Health Informatics. [Available at: <https://www.chi-csm.ca/>, last accessed February 28, 2023].
2. University of Calgary. Precision Health: Precision Medicine Specialization. [Available at: <https://cumming.ucalgary.ca/gse/about/programs/precision-health/precision-medicine>, last accessed February 28, 2023].
3. Centers for Disease Control and Prevention. Precision health: Improving health for each of us and all of us. [Available at: https://www.cdc.gov/genomics/about/precision_med.htm, last accessed April 26, 2023].
4. Alberta Strategy for Patient Reported Outcomes Support Unit. [Available at: <https://absporu.ca/>, last accessed February 28, 2023].
5. Health Data Research Network Canada. [Available at: <https://www.hdrn.ca/>, last accessed February 28, 2023].
6. SeroTracker. [Available at: <https://serotracker.com/>, last accessed February 28, 2023].
7. Strategies for Precision Health in Breast Cancer. [Available at: <https://spherecalgary.ca/>, last accessed February 28, 2023].
8. Data for Intelligence Health Lab (DIH). [Available at: <https://cumming.ucalgary.ca/lab/dih/dih>, last accessed April 25, 2023].
9. Sandhu N, Whittle S, Southern DA, Li B, Youngson E, Bakal JA, Mcleod C, Hilderman L, Williamson TS, Cheligger C, Walker RL, Kaul P, Quan H, Eastwood CA. (2023) "Health Data Governance for Research Use in Alberta", *International Journal of Population Data Science*, 8(4). <https://doi.org/10.23889/ijpds.v8i4.2160>.
10. Alberta Health, Analytics and Performance Reporting Branch. Overview of Administrative Health Datasets. 2017. [Available at: <https://open.alberta.ca/dataset/657ed26d-eb2c-4432-b9cb-0ca2158f165d/resource/38f47433-b33d-4d1e-b959-df312e9d9855/download/research-health-datasets.pdf>, last accessed April 25, 2023].
11. University of Calgary. CHI events. [Available at: <https://events.ucalgary.ca/centre-health-informatics/>, last accessed February 7, 2024].
12. University of Calgary. Health Data Science and Biostatistics Diploma Program. [Available at: <https://cumming.ucalgary.ca/centres/centre-health-informatics/education/graduate-training/health-data-science-and-biostatistics-diploma-program>, last accessed February 28, 2023].
13. Libin Institute. [Available at: <https://libin.ucalgary.ca/>, last accessed February 28, 2023].

14. Libin Institute. Precision Medicine Initiative. [Available at: <https://libin.ucalgary.ca/about-us/our-initiatives/precision-medicine-initiative>, last accessed February 28, 2023].
15. Krusina A, Chen O, Otero Varela L, Doktorchik C, Avati V, Kundsen S, Southern D, Eastwood C, Sharma N, Williamson T (2020) Developing a Data Integrated COVID-10 Tracking System for Decision-Making and Public Use. *IJPDS* 5(4). <https://doi.org/10.23889/ijpds.v5i4.1389>.
16. Alberta RWE Consortium. [Available at: <https://www.albertarwe.ca/>, last accessed April 25, 2023].
17. AbSPORU. Patient Engagement (PE) Definitions. [Available at: <https://absporu.ca/patient-engagement/definitions/>, last accessed January 29, 2024].
18. Alberta Health Services. Non-identifying health information. 2014. [Available at: <https://extranet.ahsnet.ca/teams/policydocuments/1/clp-ahs-privacy-non-identifying-health-information-standard-ipo-2013-0004.pdf>, last accessed January 29, 2024].
19. Statistics Canada. Research Data Centres. [Available at: <https://www.statcan.gc.ca/en/microdata/data-centres>, last accessed February 7, 2024].
20. Centre for Health Informatics. Statistics Canada RDC @ CHI. [Available at: <https://cumming.ucalgary.ca/centres/centre-health-informatics/RDC-access>, last accessed, September 13, 2023].
21. University of Calgary Information Technologies. Secure Computing Platform. [Available at: <https://it.ucalgary.ca/secure-computing-platform/store-your-data>, last accessed February 28, 2023].
22. University of Calgary Information Technologies. Research Computing Services. [Available at: <https://it.ucalgary.ca/research-computing-services/our-resources/high-performance-computing-hpc>, last accessed February 28, 2023].
23. Canadian Institute for Health Information. Data Quality Framework. [Available at: https://www.cihi.ca/sites/default/files/document/iqf-summary-july-26-2017-en-web_0.pdf, last accessed February 28, 2023].
24. Basmadjian RB, Kong S, Boyne DJ, Jarada TN, Xu Y, Cheung WY, Lupichuk S, Quan ML, Brenner DR. Developing a prediction model for pathologic complete response following neoadjuvant chemotherapy in breast cancer: a comparison of model building approaches. *JCO Clin Cancer Inform*, 6, e2100055(2022). <https://doi.org/10.1200/CCI.21.00055>
25. Wu G, Khair S, Yang F, Cheliger C, Southern D, Zhang Z, Feng Y, Xu Y, Quan H, Williamson T, Eastwood CA. Performance of machine learning algorithms for surgical site infection case detection and prediction: A systematic review and meta-analysis. *Ann Med Surg*, 2022 Nov 23;84:104956. <https://doi.org/10.1016/j.amsu.2022.104956>
26. Nurmambetova E, Pan J, Zhang Z, Wu G, Lee S, Southern DA, Martin EA, Ho C, Xu Y, Eastwood CA. Developing an Inpatient Electronic Medical Record Phenotype for Hospital-Acquired Pressure Injuries: Case Study Using Natural Language Processing Models. *JMIR AI* 2023;2: e41264. <https://doi.org/10.2196/41264>
27. Alberta Real World Evidence Consortium. [Available at: <https://www.ucalgary.ca/research/alberta-rwe>, last accessed February 28, 2023].
28. Eastwood CA, Quan H, Howlett JG, King-Shier KM. Factors Associated With 7-Day Rehospitalization After Heart Failure Admission. *J Cardiovasc Nurs*. 2017 Jul/Aug;32(4):339–347. <https://doi.org/10.3389/fphar.2021.732760>
29. Na Li, Fei Chiang, Douglas G. Down, Nancy M. Heddle. A decision integration strategy for short-term demand forecasting and ordering for red blood cell components. *Operations Research for Health Care*. 2021; 29. <https://doi.org/10.1016/j.orhc.2021.100290>
30. Olstad DL, Beall R, Spackman E, Dunn S, Lipscombe L, Williams K, Oster R, Scott S, Zimmermann G, McBrien K, Steer K, Chan C, Tyminski S, Berkowitz S, Edwards A, Saunders-Smith T, Tariq S, Popeski N, White L, Williamson T, L'Abbe M, Raine K, Nejatnamini S, Naser A, Basualdo-Hammond C, Norris C, O'Connell P, Seidel J, Cabaj J, Campbell D. (2022) Healthy food prescription incentive programme for adults with type 2 diabetes who are experiencing food insecurity: protocol for a randomised controlled trial, modelling and implementation studies. *BMJ Open*. 2022 Feb 15;12(2): e050006. <https://doi.org/10.1136/bmjopen-2021-050006>
31. Mahsin M, Lee S, Vickers D, Guigue A, Williamson T, Quan H, Quinn RR, Ravani P. The Coronavirus 2019 Pandemic in Canada: The Impact of Public Health Interventions of the Course of the Outbreak in Alberta and Other Provinces. *medRxiv* 2020.05.09.20096636. <https://doi.org/10.1101/2020.05.09.20096636>
32. Vickers DM, Baral S, Mishra S, Kwong JC, Sundaram M, Katz A, Calzavara A, Maheu-Giroux M, Buckeridge DL, Williamson T. Stringency of Containment and Closures on the Growth of SARS-CoV-2 in Canada Prior to Accelerated Vaccine Roll-Out. *Int J Infect Dis* 2022 May; 118:73–82. <https://doi.org/10.1016/j.ijid.2022.02.030>
33. Wu G, D'Souza AG, Quan H, Southern DA, Youngson E, Williamson T, Eastwood C, Xu Y. Validity of ICD-10 codes for COVID-19 patients with hospital admissions or ED visits in Canada: a retrospective cohort study. *BMJ Open*. 2022 Jan 21;12(1): e057838. <https://doi.org/10.9778/cmajo.20230033>

34. Acosta N, Bautista MA, Waddell BJ, McC Calder J, Beaudet AB, Man L, Pradhan P, Sedaghat N, Papparis C, Bacanu A, Hollman J, Krusina A, Southern DA, Williamson T, Li C, Bhatnagar S, Murphy S, Chen J, Kuzma D, Clark R, Meddings J, Hu J, Cabaj JL, Conly JM, Dai X, Lu X, Chekouo T, Ruecker NJ, Achari G, Ryan MC, Frankowski K, Hubert CRJ, Parkins MD. Longitudinal SARS-CoV-2 RNA wastewater monitoring across a range of scales correlates with total and regional COVID-19 burden in a well-defined urban population. *Water Res.* 2022 Jul 15; 220:118611. <https://doi.org/10.1016/j.watres.2022.118611>
35. Feng Y, Leung AA, Lu X, Liang Z, Quan H, Walker RL. Personalized prediction of incident hospitalization for cardiovascular disease in patients with hypertension using machine learning. *BMC Med Res Methodol.* 2022 Dec 17;22(1):325. <https://doi.org/10.1186/s12874-022-01814-3>
36. Lee S, Li B, Martin EA, D'Souza AG, Jiang J, Doktorchik C, Southern DA, Lee J, Wiebe N, Quan H, Eastwood CA. CREATE: A New Data Resource to Support Cardiac Precision Health. *CJC Open.* 2021. 3(5): 639-645. <https://doi.org/10.1016/j.cjco.2020.12.019>
37. Alberta Health Services. Connect Care. [Available at: <https://www.albertahealthservices.ca/cis/cis.aspx>, last accessed February 7, 2024].

Abbreviations

CHI:	Centre for Health Informatics
CSM:	Cumming School of Medicine
AHS:	Alberta Health Services
AH:	Alberta Health
AbSPORU:	Alberta Strategy for Patient-Oriented Research Unit
UofC:	University of Calgary
SPHERE:	Strategy for Precision Health in Breast Cancer
CIHI:	Canadian Institute for Health Information
WHO:	World Health Organization
HPC:	High-performance computing
RWE:	Real World Evidence Consortium
EMR:	Electronic Medical Record
RDC:	Research Data Centre

