UNSW Centre for Big Data Research in Health

Distilling Impactful Health Insights from Big Data

Future Health Expo

Thursday March 14th 2024

Contents

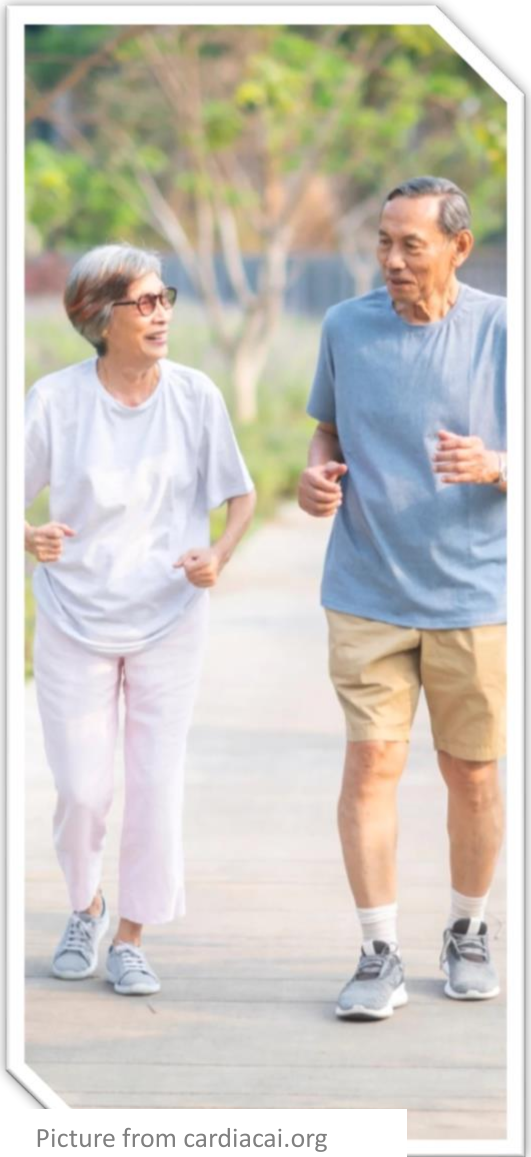
1. CardiacAI: Anticipating the needs of Cardiac patients using Artificial Intelligence

Key points

- The Cardiac Analytics and Innovation (CardiacAI) Data Repository is an independent, not-for-profit platform that provides a secure environment for Australian research groups to access and analyse real-time cardiovascular electronic medical record data.
- Using this platform, we have built AI-based algorithms for early identification of cardiac patients at high risk of readmission and death following discharge from hospital.
- We aim to anticipate the need for telehealth monitoring for patients at high risk of readmission, as well as the need for symptom management and meaningful discussions for patients at the end of life.
- We are building operational and clinician dashboards for heart failure management to improve outcomes and processes of care for these patients.

Key links

- https://www.cardiacai.org/



Picture from cardiacai.org

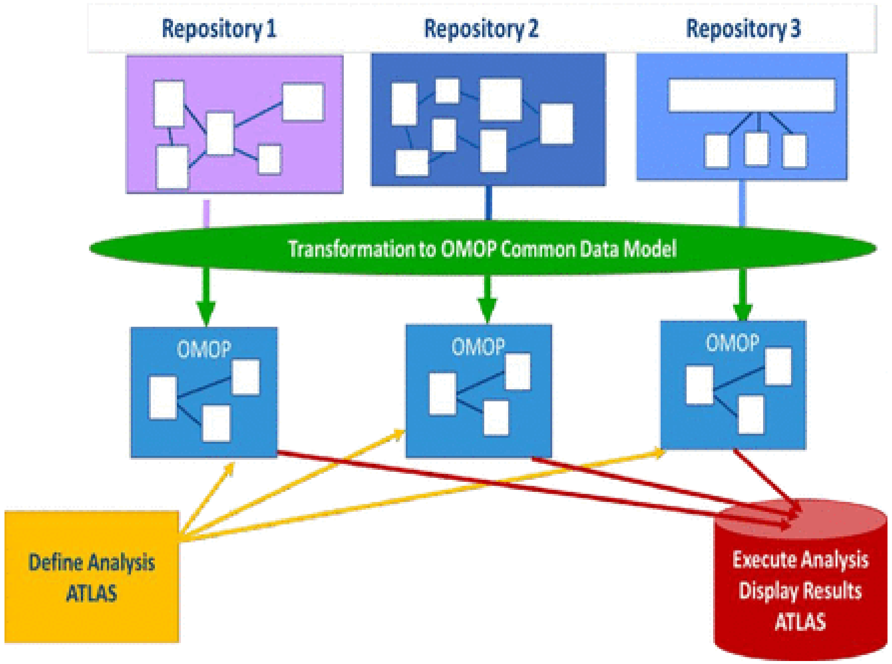
2. OMOP-EMR: Facilitating research by standardising electronic medical record data from multiple sites

Key points

- The Observational Medical Outcomes Partnership Common Data Model (OMOP-CDM) standardises the structures and semantics of health datasets, enabling reproducibility and large-scale studies that leverage the data from multiple locations and settings.
- We have built an Extract-Transform-Load (ETL) framework for the conversion of health databases to the OMOP CDM that supports transparency of the mapping process, readability, and maintainability.
- We have used this framework to build a software tool that transforms Cerner Millennium (Cerner Corporation)) electronic health records used by Australian Local Health Districts to OMOP-CDM.
- This allows for the creation of shared end-to-end analysis packages using hospital electronic medical record data without the need for direct data exchange.

Key links

- Quiroz et al. Extract, transform, load framework for the conversion of health databases to OMOP. *PLOS ONE* (2022) https://doi.org/10.1371/journal.pone.0266911
-
- Hallinan et al. Seamless EMR data access: Integrated governance, digital health and the OMOP-CDM. *BMJ Health and Care* Informatics (2024) https://doi.org/10.1136/bmjhci-2023-100953



Hallinan et al (2024).

Figure 1. Observational Medical Outcomes Partnership Common Data Model (OMOP-CDM). Adapted from Standardised Data: The OMOP Common Data Model.

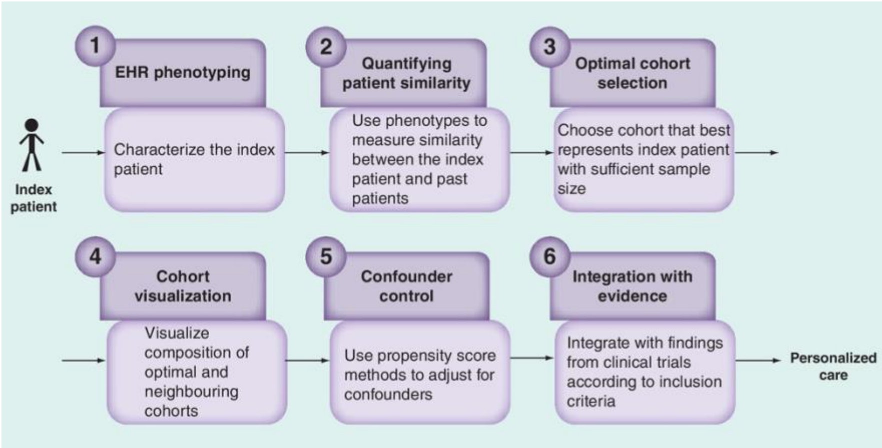
3. Causal AI: Guiding clinical decision-making by estimating heterogeneity of treatment effects from large routinely collected health data

Key points

- We have developed an analytics framework and methodology for measuring and reporting on the use and effectiveness of interventions in heterogenous populations as used in routine care.
- These methods have been used to analyse the effect of: (1) electroconvulsive therapy in patients with severe psychotic disorders; (2) percutaneous catheter ablation in patients with atrial fibrillation; and (3) vitamin D in patients with osteoarthritis.
- The tools generated by this project are being compiled into a CausalAI software package, which can be used by researchers and regulators to address comparative effectiveness questions.

Key links

- Gallego et al. Bringing cohort studies to the bedside: framework for a 'green button' to support clinical decision-making. *Journal of Comparative Effectiveness Research* (2015) https://doi.org/10.2217/cer.15.12
- Wendling et al. Comparing methods for estimation of heterogeneous treatment effects using observational data from health care databases. *Statistics in Medicine* (2018) https://doi.org/10.1002/sim.7820
- Jin et al. Estimating incidence rates of periprosthetic joint infection after hip and knee arthroplasty for osteoarthritis using linked registry and administrative health data. *The Bone and Joint Journal* (2022) https://doi.org/10.1302/0301-620X.104B9.BJJ-2022-0116.R1
- Zhu et al.Targeted estimation of heterogeneous treatment effect in observational survival analysis. *Journal of Biomedical Informatics* (2020) https://doi.org/10.1016/j.jbi.2020.103474



Gallego et al (2015) **Figure 1.** Proposed process for generating real-time cohort studies at the point of care.

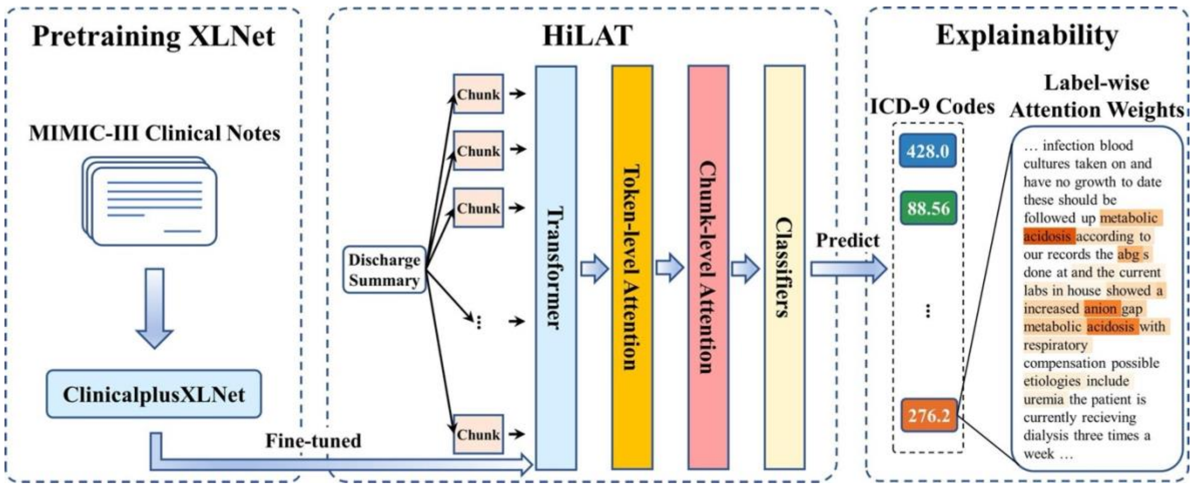
4. Automating ICD coding from clinical text

Key points

- A hierarchical label-wise attention transformer (HiLAT) model was proposed for automating ICD coding process and employed a two-level hierarchical label-wise attention mechanism that creates label-specific document representations.
- HiLAT achieved state-of-the-art performance on the top 50 most frequent ICD-9 codes from MIMIC-III and presents a potential explainability tool for checking the face validity of ICD code predictions by visualising label-wise attention weights.
- Several Transformer-based models were explored to address the extreme label set and long text classification challenges that are posed by automated ICD coding tasks.

Key links

- Liu et al. Hierarchical label-wise attention transformer model for explainable ICD coding. *Journal of Biomedical Informatics* (2022) doi.org/10.1016/j.jbi.2022.104161
- Liu et al. Automated ICD coding using extreme multi-label long text transformer-based models. *Artificial Intelligence in Medicine* (2023) doi.org/10.1016/j.artmed.2023.102662

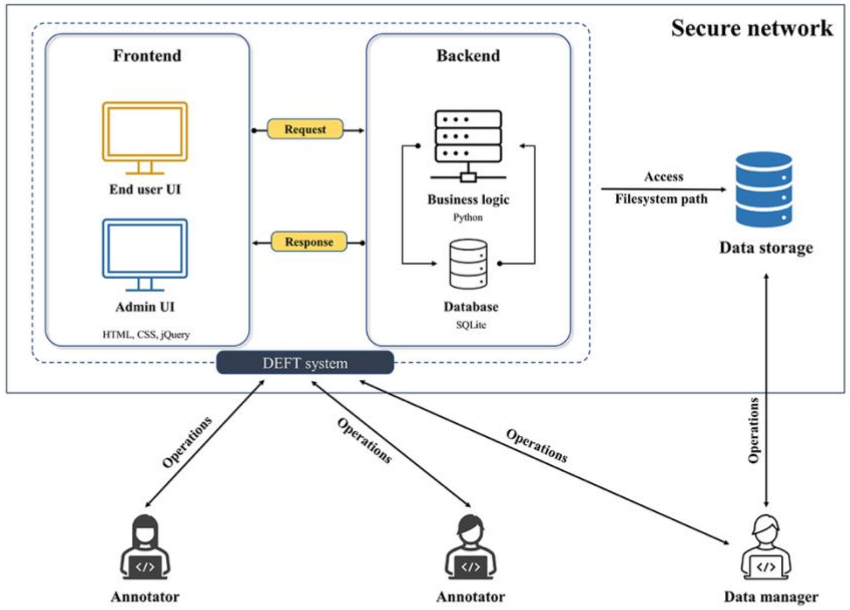


Liu et al. (2022). **Graphical abstract**

5. Privacy protection from clinical text

Key points

- An end-to-end framework was designed to help health data custodians protect privacy in clinical text.
- The framework uses an ensemble approach which combines several deep learning models to provide a robust solution for de-identifying clinical text.
- A web-based application of the framework was developed with a human-in-the-loop pipeline to speed up the deidentification process and increase its performance over time.
- The application has been employed to remove personal identifiers from clinical text for the Cardiac Analytics and Innovation (CardiacAI) project.



Liu et al (2023). **Figure 1.** Overview of the DEFT system architecture. DEFT: deidentifying free text; UI: user interface.

Key links

- Liu et al. De-identifying Australian hospital discharge summaries: An end-to-end framework using ensemble of deep learning models. *Journal of Biomedical Informatics* (2022) doi.org/10.1016/j.jbi.2022.104215
- Liu et al. Web-Based Application Based on Human-in-the-Loop Deep Learning for Deidentifying Free-Text Data in Electronic Medical Records: Development and Usability Study Interactive. *Journal of Medical Research* (2023) doi:10.2196/46322

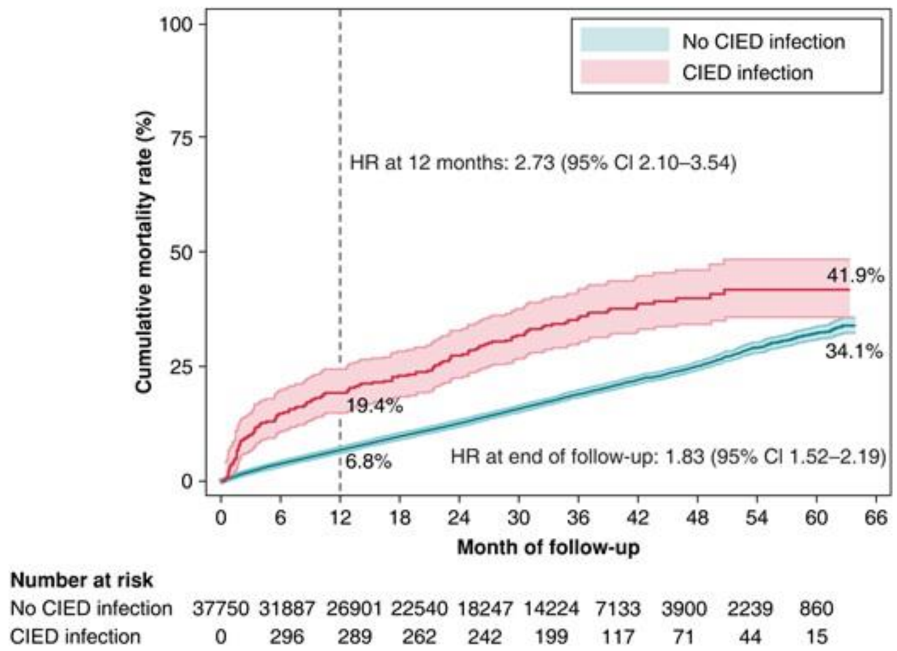
6. Evidence about the use, safety, effectiveness, and costs of new medical devices and technologies

Key points

- CBDRH and Medtronic Australasia established a partnership to leverage real-world evidence.
- We aimed to overcome data source limitations to improve coverage of Australian patients and ensure long-term follow-up.
- Employed longitudinal NSW administrative data from 2001 for patients with cardiovascular diagnoses.
- The projects undertaken were:
    1. "Cost of Cardiac Implantable Electronic Device (CIED) Infections: a Non-interventional Study using Linked Secondary Data."
    2. "Trends and Patterns in Transcatheter Aortic Valve Implantation (TAVI) Access and Utilization in New South Wales: a Data Linkage Study."

Key links

- Shawon et al. Real-world evidence on the association between cardiac implantable electronic device infection and all-cause mortality. *EP Europace* (2023) https://doi.org/10.1093/europace/euad274
- Challis et al. What Does Real World Evidence (RWE) Offer Health Technology Assessment (HTA) Procedures In Australia? *International Journal of Technology Assessment in Health* (2023) https://doi.org/10.1017/S0266462323002519
- Shawon et al. Association between cardiac implantable electronic device infection and mortality: long-term follow up of a complete, state-wide cohort in Australia. *European Heart Journal* (2023) https://doi.org/10.1093/eurheartj/ehad655.698



Shawon et al (2023)
Figure 1. Kaplan–Meier all-cause mortality curves. Hazard ratios (HR) are estimated from an adjusted Cox proportional hazard regression model.

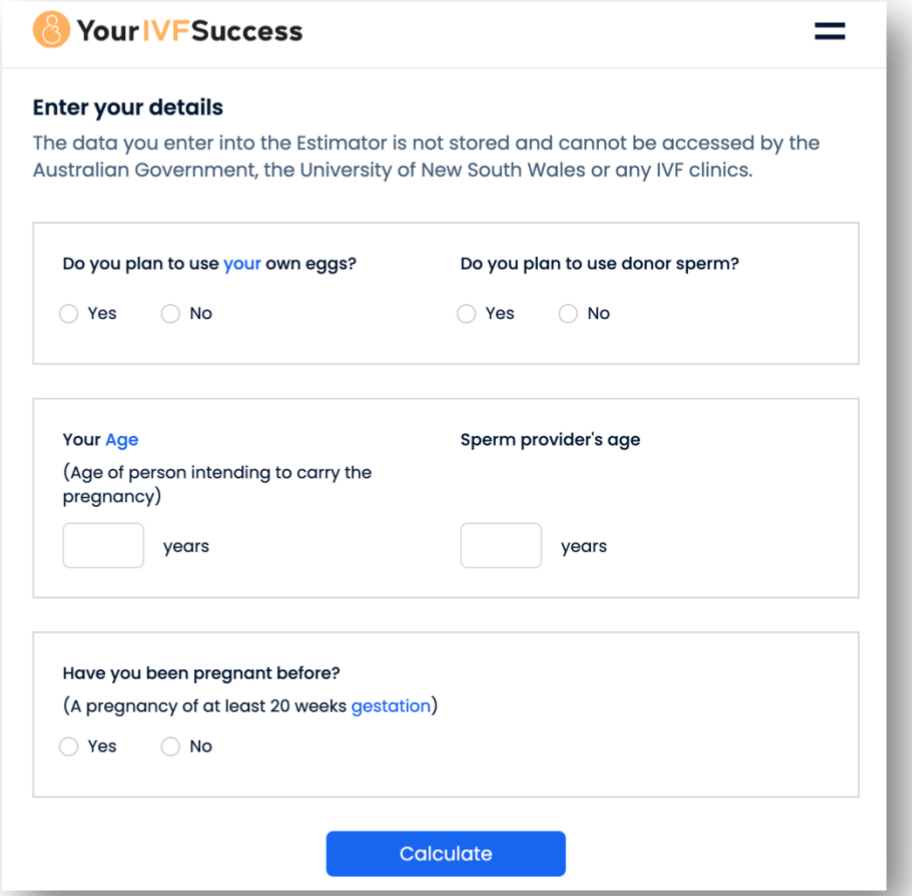
7. YourIVFSuccess: Supporting hopeful parents navigate the difficult and complex IVF journey

Key points

- Funded by the Australian Government, the YourIVFSuccess website was created to help people make informed decisions about IVF treatment.
- The website provides a searchable database of all fertility clinics in Australia, and includes comparable information on the types of patients, their treatment and their success rates compared to national averages.
- It also includes an IVF Success Estimator which allows patients to enter their own characteristics and receive an individualised estimate their chances of having a baby using IVF.
- The IVF Success estimates are based on the analysis of over 800,000 IVF cycles performed in Australia using AI algorithms and best practice in predictive modelling.
- The YourIVFSuccess website was awarded the Research Australia Data Innovation Award.

Key links

- UNSW Newsroom
https://www.unsw.edu.au/newsroom/news/2023/02/new-national-ivf-clinic-success-rates-released
- YourIVFSuccess website and online calculator
https://yourivfsuccess.com.au/



The YourIVFSuccess Estimator
https://yourivfsuccess.com.au/estimate

8. Educating the next generation of Health Data Scientists

Key points

- Our master's degree, graduate diploma and graduate certificate in health data science are pioneering programs that examine data-driven solutions to complex health problems.
- We also offer continuous professional development in areas such data management, statistical modelling, machine learning, data visualisation and clinical artificial intelligence.
- We supervise post-graduate PhD and research master's students.

Key links

- UNSW website
  https://www.unsw.edu.au/medicine-health/study-with-us/study-areas/health-data-science
- Health Data Science Student Hub
  https://hds-hub.cbdrh.med.unsw.edu.au/

Over 100 students have already graduated from our Master of Science in Health Data Science Program