Using routinely collected clinical data and advanced analytics to generate real-world evidence

Louisa Jorm, Centre for Big Data Research in Health, UNSW Sydney
Lumos Symposium 9 October 2024

CENTRE FOR
BIG DATA RESEARCH
IN HEALTH

Real-world data and real-world evidence

Real-world data: data relating to patient health or experience or care delivery collected outside the context of a highly controlled clinical trial.

Real-world evidence: evidence generated from the analysis of real-world data.

Real-world evidence

Covers a large array of evidence types:

- Disease epidemiology
- Comparative effectiveness research
- Health service research

Can be generated from a large range of study designs and analytical methods

May use routinely collected data, bespoke data collection, or both:

- Electronic health records
- Administrative data
- Claims data
- Patient registries

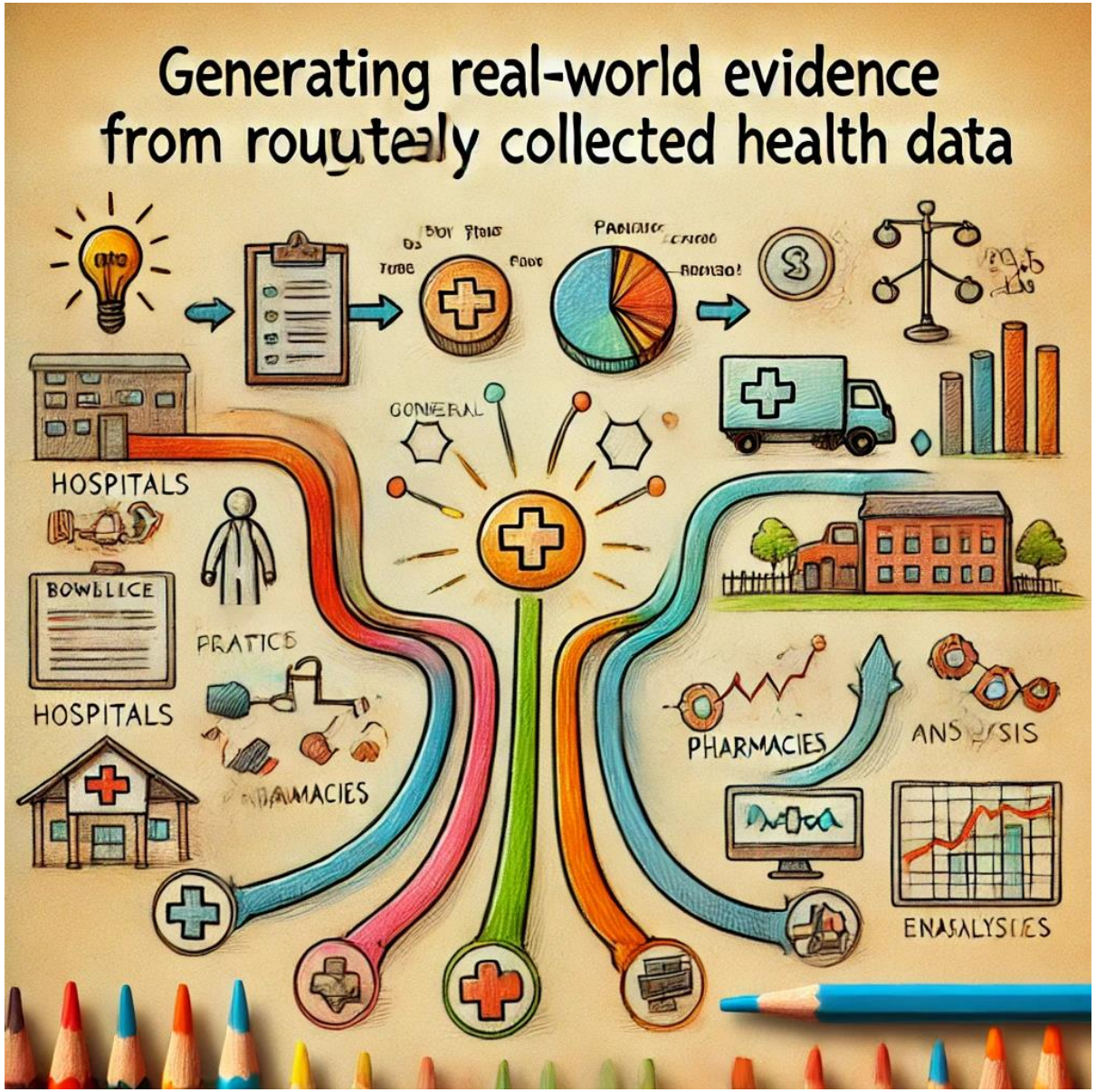
Focus for today

What has been possible?

What is now possible?

Promising possibilities!

What has been possible?



Generated using DALL·E 3

The Australia and patient o cohort study

Duong T Tran[1], Michael O Fal Jane Hall[4], Louisa Jorm[1]

Med J Aust 2024; 220 (7):

Abstract

Objective: To assess the impact of the Health Care Homes (HCH) primary health care initiative on quality of care and patient outcomes.

Design, setting: Quasi-experimental, matched cohort study; analysis of general practice data extracts and linked administrative data from ten Australian primary health networks, 1 October 2017 – 30 June 2021.

Participants: People with chronic health conditions (practice data extracts: 9811; linked administrative data: 10 682) enrolled in the HCH 1 October 2017 – 30 June 2019; comparison groups of patients receiving usual care (1:1 propensity score-matched).

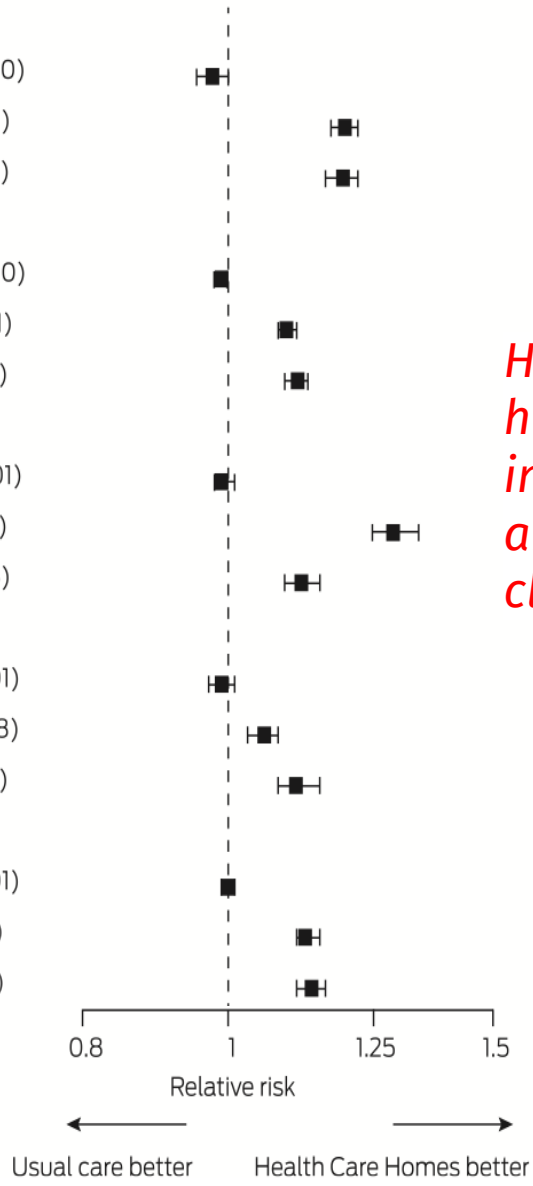
Intervention: Participants were involved in shared care planning, provided enhanced access to team care, and encouraged to seek chronic condition care at the HCH practice where they were enrolled. Participating practices received bundled payments based on clinical risk tier.

Main outcome measures: Access to care, processes of care, diabetes-related outcomes, hospital service use, risk of death.

UNSW SYDNEY

4 Chronic disease management processes of care for patients in the matched Health Care Homes (HCH) and usual care cohorts prior to enrolment and during first and second years after enrolment of HCH patients

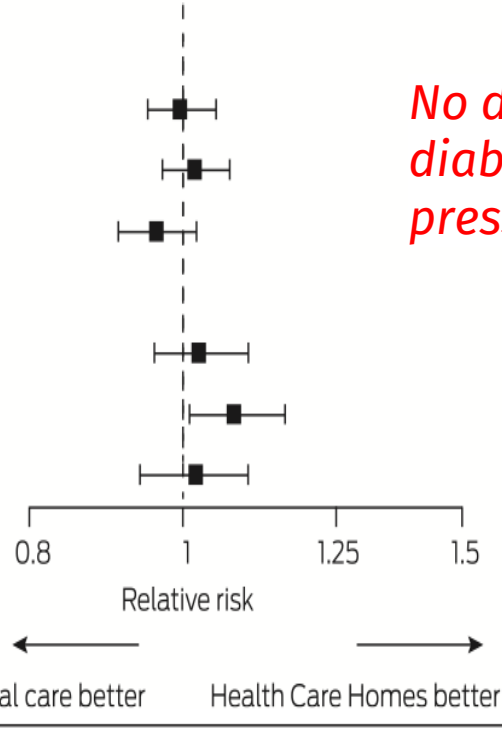
Outcome	Health Care Home: number (proportion)	Usual care:* number (proportion)	Relative risk (95% CI)
Influenza vaccination			
Pre-enrolment	5628 (57.4%)	5769 (58.8%)	0.98 (0.95–1.00)
First year	6241 (66.2%)	5221 (55.3%)	1.20 (1.17–1.22)
Second year	5541 (61.0%)	4644 (51.1%)	1.19 (1.16–1.22)
Blood pressure recorded			
Pre-enrolment	8254 (84.1%)	8353 (85.1%)	0.99 (0.98–1.00)
First year	7940 (84.2%)	7259 (77.0%)	1.09 (1.08–1.11)
Second year	6795 (74.8%)	6106 (67.2%)	1.11 (1.09–1.13)
Blood lipids recorded†			
Pre-enrolment	7408 (75.5%)	7473 (76.2%)	0.99 (0.98–1.01)
First year	6573 (69.7%)	5551 (58.8%)	1.19 (1.16–1.21)
Second year	5688 (62.6%)	5073 (55.9%)	1.12 (1.09–1.15)
HbA_{1c} recorded‡			
Pre-enrolment	2414 (85.7%)	2436 (86.5%)	0.99 (0.97–1.01)
First year	2282 (83.8%)	2155 (79.2%)	1.06 (1.03–1.08)
Second year	2058 (78.9%)	1854 (71.1%)	1.11 (1.08–1.15)
Kidney function recorded§			
Pre-enrolment	6044 (88.7%)	6048 (88.8%)	1.00 (0.99–1.01)
First year	5600 (85.6%)	4972 (76.0%)	1.13 (1.11–1.15)
Second year	4989 (79.0%)	4389 (69.5%)	1.14 (1.11–1.16)



HCH patients had higher rates of influenza vaccination and recording of clinical variables

4 Chronic disease management processes of care for patients in the matched Health Care Homes (HCH) and usual care cohorts prior to enrolment and during first and second years after enrolment of HCH patients

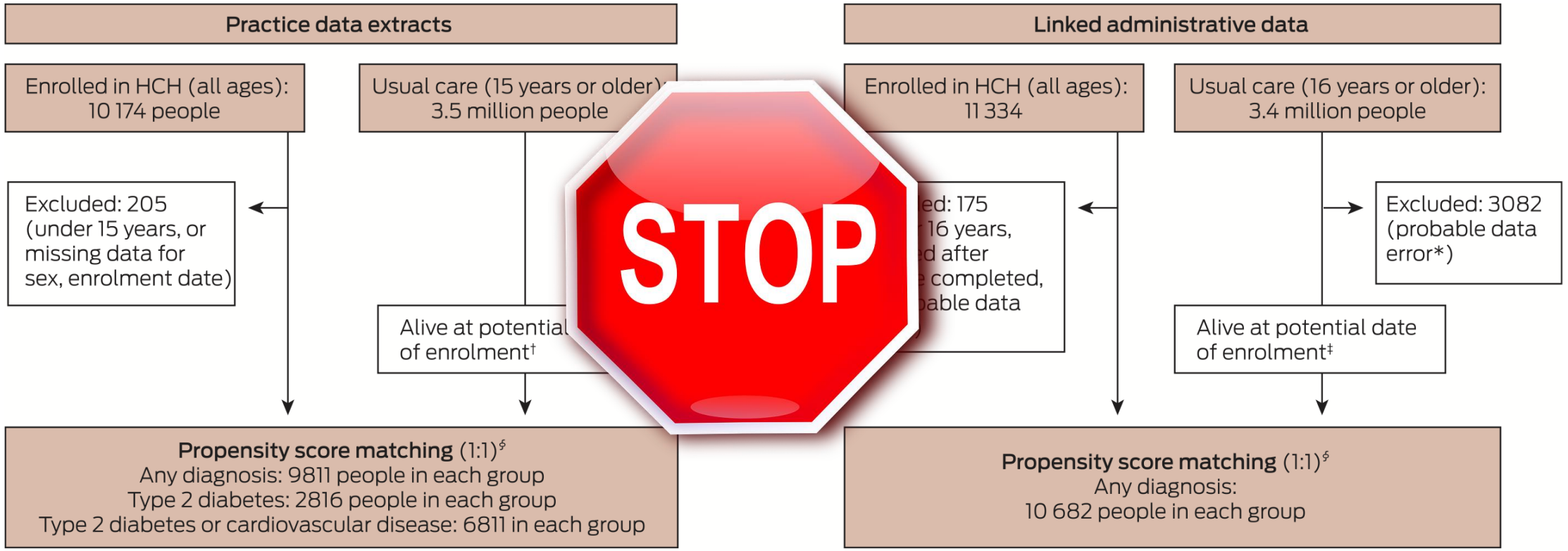
Outcome	Health Care Home: number (proportion)	Usual care:* number (proportion)	Relative risk (95% CI)
HbA_{1c} ≤53 mmol/mol[¶]			
Pre-enrolment	1355 (56.1%)	1371 (56.3%)	1.00 (0.95–1.05)
First year	1315 (57.6%)	1219 (56.6%)	1.02 (0.97–1.07)
Second year	1051 (51.1%)	984 (53.1%)	0.96 (0.91–1.02)
Blood pressure ≤130/80 mmHg[¶]			
Pre-enrolment	1017 (40.2%)	1005 (39.3%)	1.02 (0.96–1.10)
First year	1049 (41.6%)	890 (38.6%)	1.08 (1.01–1.16)
Second year	832 (37.5%)	718 (36.8%)	1.02 (0.94–1.10)



No differences in diabetes or blood pressure control

No differences in emergency department presentations, hospital admissions or mortality

1 Propensity score matching of the two cohort pairs for assessing quality of care and patient outcomes in the Health Care Homes (HCH) trial

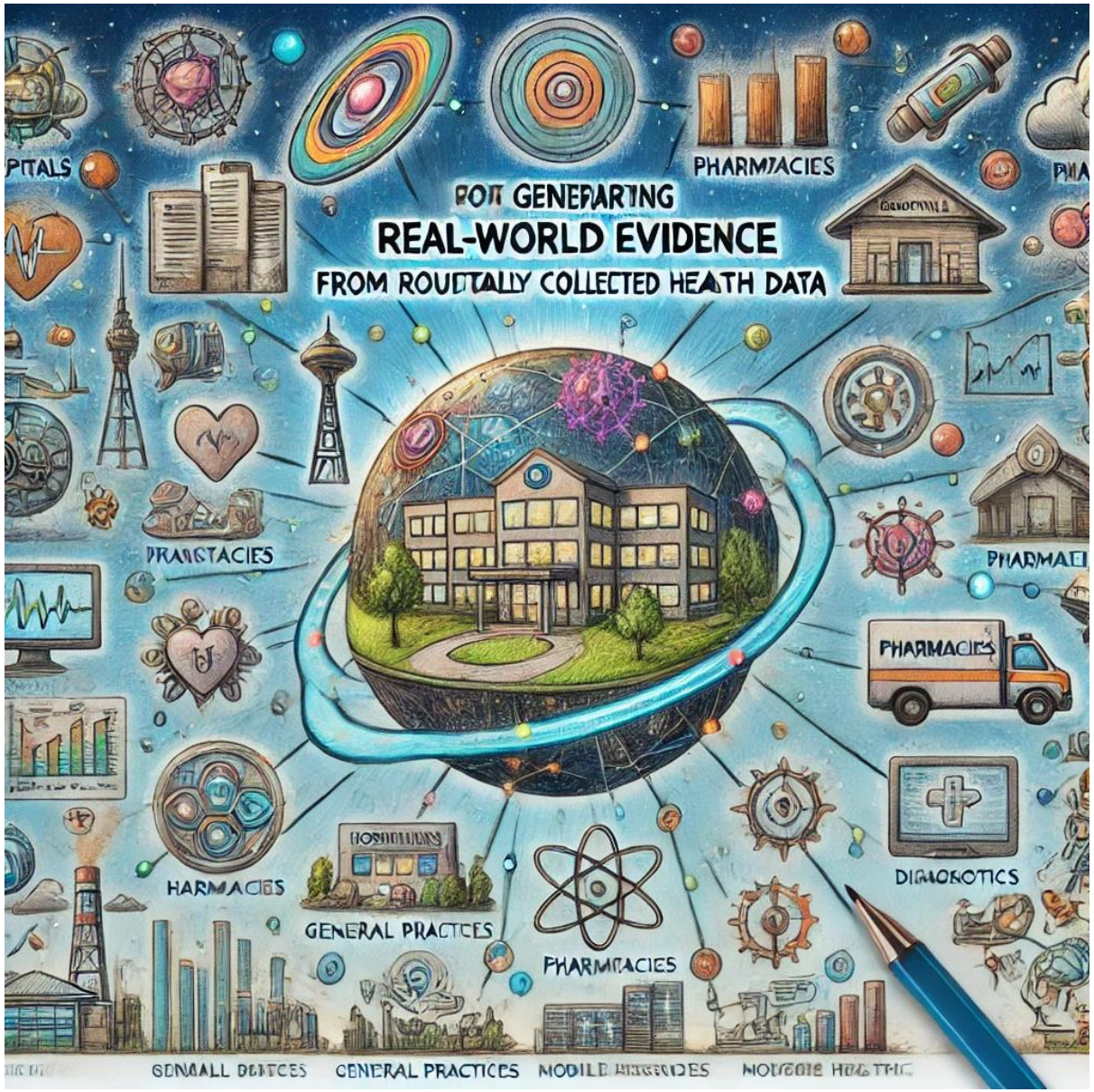


* Recorded date of health service use was later than recorded date of death. † Year of death was later than year of enrolment or not recorded. ‡ Date of death was later than date of enrolment or no record of death. § Matched by propensity score and year and month of enrolment. ◆

No linkage between practice data and administrative data (MBS, PBS, hospital, death)!

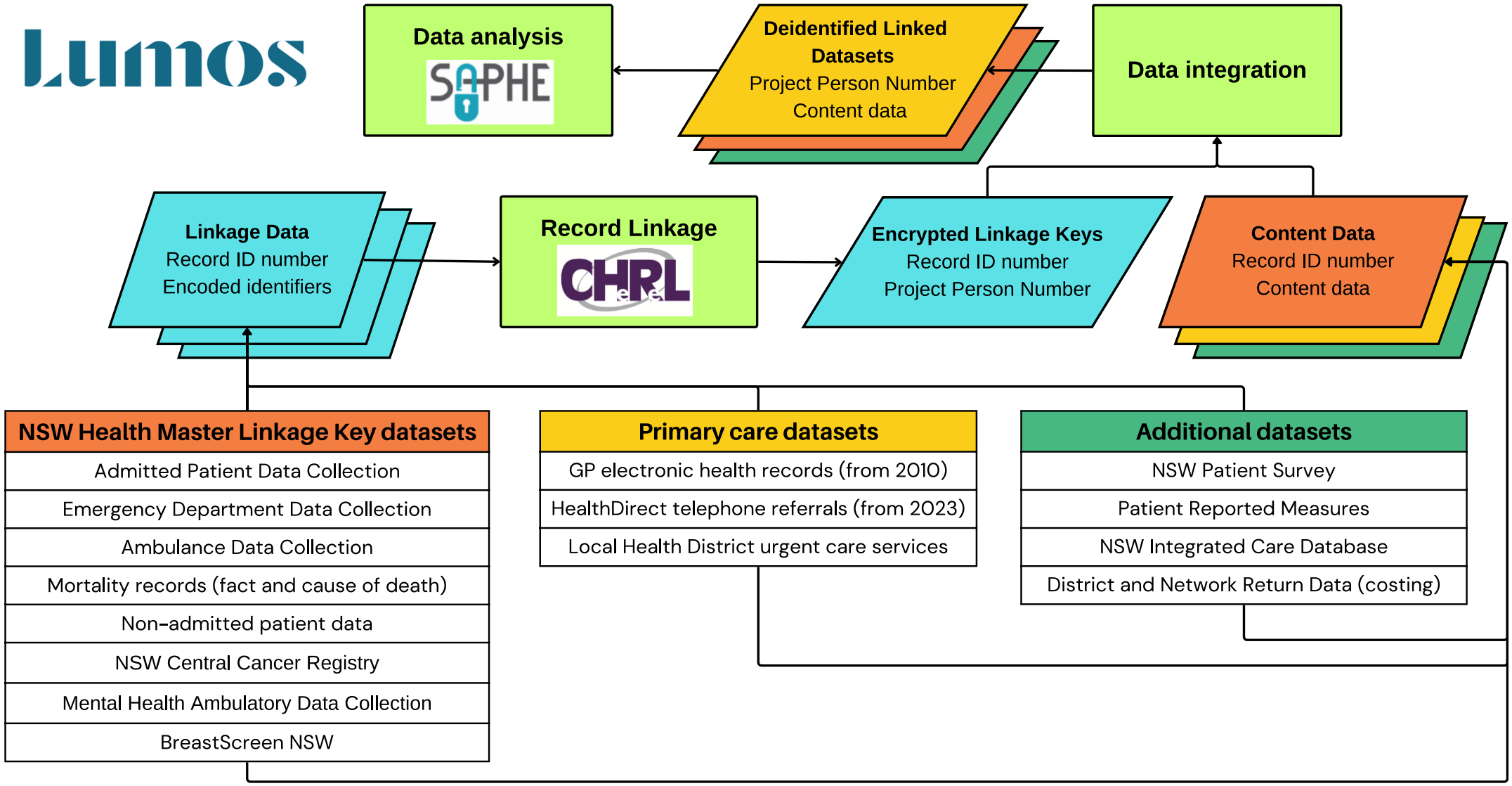
UNSW SYDNEY

What is now possible?



Generated using DALL·E 3

Lumos



UNSW SYDNEY

Estimating Five-year Absolute Risk of Cardiovascular Disease Using Routinely Collected Electronic Medical Records from Australian General Practices

Nic Kuo, Sebastiano Barbieri, Clare Arnott, Blanca Gallego-Luxan, Ziba Zandomkar, Shahana Ferdousi, Kirsty Douglas, Mark Woodward, Louisa Jorm

Funded through MRFF Cardiovascular Mission Grant 2020-2025

Partnership with WentWest PHN

CENTRE FOR BIG DATA RESEARCH IN HEALTH

UNSW SYDNEY

Motivation….

- ~70% of Australian adults have 3+ modifiable risk factors for CVD[1]
- From 2019, MBS funds free annual Heart Health Check for people aged 45+ (30+ for Aboriginal and Torres Strait Islander peoples)
- In 2023-24, 1.7% of women and 1.4% of men aged 45+ had a GP claim for this item (699)[2]
- Australian CVD risk calculator (AusCVDRisk)[3] launched mid-2023
  - Based on PREDICT[4] equations developed for NZ
  - Not validated using Australian data

[1]National Heart Foundation. HeartWatch Survey, customised data. 2019.
[2]http://medicarestatistics.humanservices.gov.au/statistics/mbs_item.jsp
[3]Australian CVD Risk Calculator, https://www.cvdcheck. org.au/calculator
[4]Pylypchuk R, et al, The Lancet, 2018

UNSW SYDNEY

Methods: study population and outcome

- Data from the New South Wales Health Lumos program
  - Primary care EMRs from 680 general practices, linked with administrative records for hospitalisations and deaths
- Study cohort: individuals aged 30-74 years on 1 January 2017
  - no prior history of CVD
  - at least one prior EMR record for a measurement or pathology test
- Primary outcome: first fatal or non-fatal CVD event in the period up to 31 December 2021 (APDC records, deaths)

UNSW SYDNEY

Methods: predictors

- age
- socioeconomic deprivation
- smoking status
- chronic conditions
  - diabetes, hyperlipidemia, atrial fibrillation, asthma, COPD, CKD, depression, serious mental illness, other mental illness
- measurements
  - BMI, SBP
- pathology results
  - TC/HDL, eGFR, triglycerides, HbA1c
- medications
  - antihypertensives, lipid-lowering agents, antithrombotics

UNSW SYDNEY

Methods: models

- Sex-specific Cox proportional hazards models
- 5x2 cross-validation
- Two models presented here:
  - Full model using all available predictors (“Full”)
  - Parsimonious model using least absolute shrinkage and selection operator (LASSO) regression (“LASSO”)

Next steps – implementation?

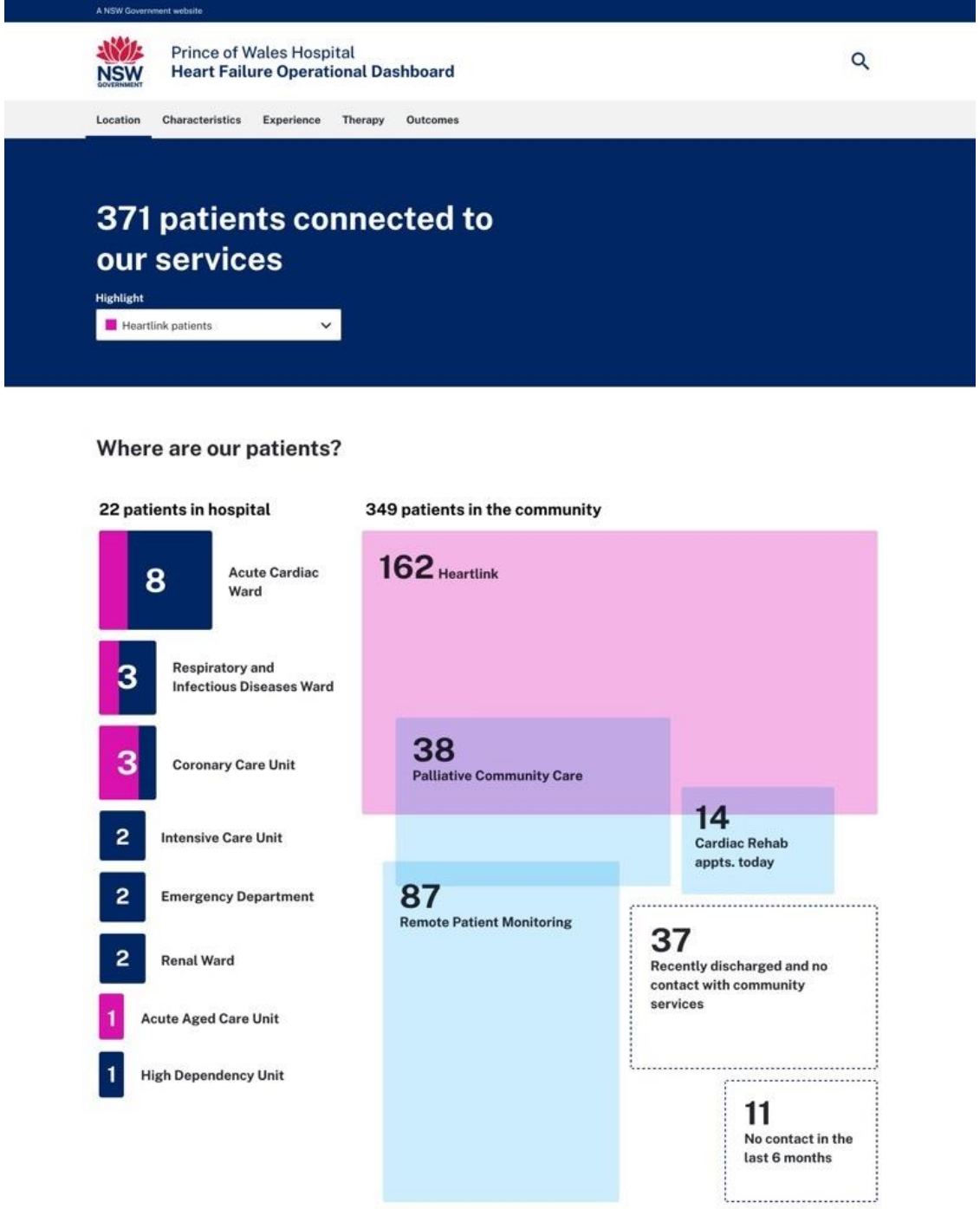
Design implementation pilot with Western Sydney PHN

- Implement algorithm within PHN data systems
- Produce patient lists, dashboards
- Test approaches to patient recall

Also potential for online tool and/or app

Also potential for incorporation into practice software

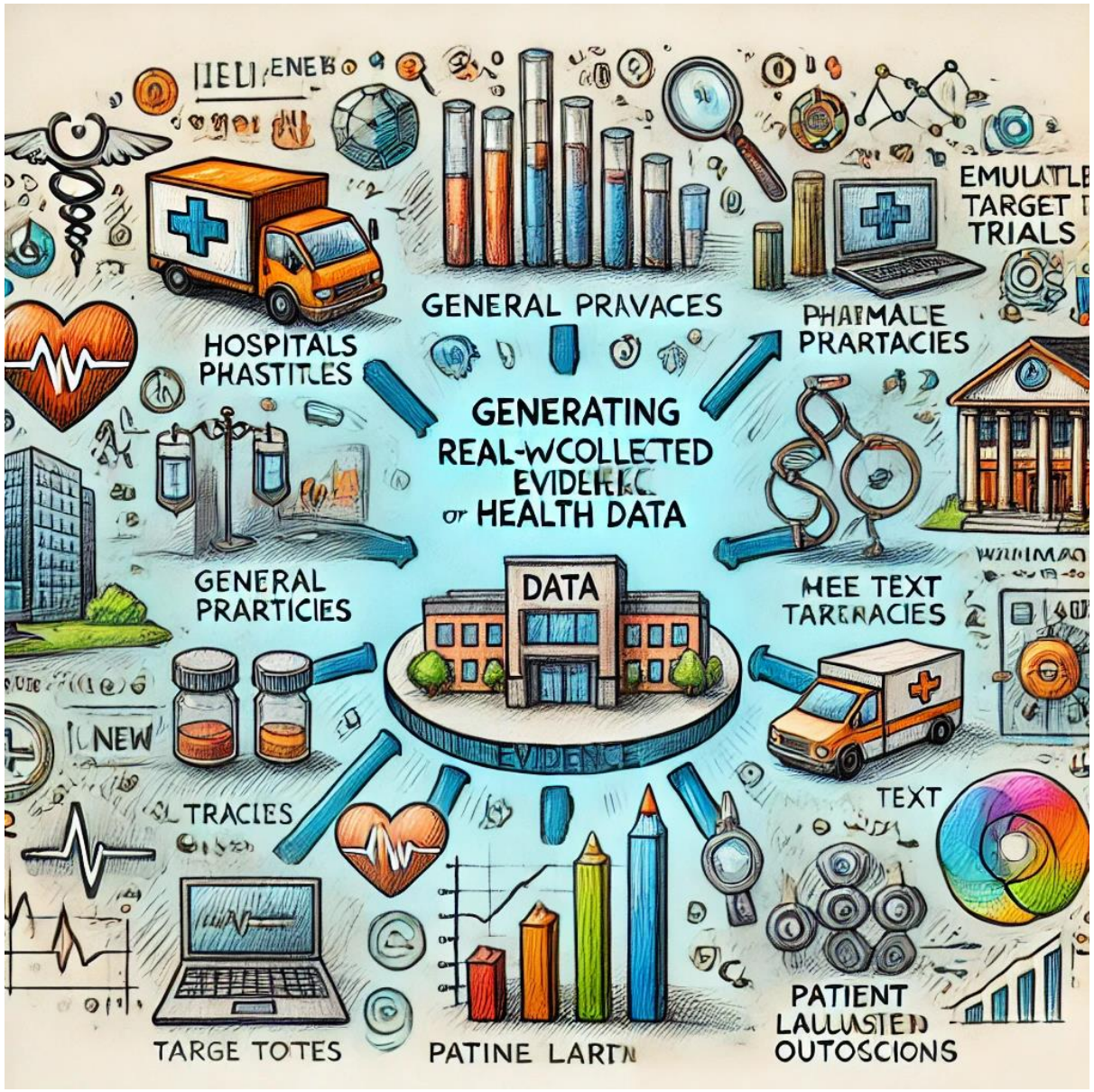
Heart failure operational dashboard co-design



Cardiac AI

UNSW SYDNEY

Promising possibilities



Generated using DALL·E 3

Use of emulated target trial methodology

Weight Change and the Onset of Cardiovascular Diseases: Emulating Trials Using Electronic Health Records

Katsoulis, M, et al. Epidemiology32(5):744-755, September 2021.
doi: 10.1097/EDE.0000000000001393

Background:

Cross-sectional measures of body mass index (BMI) are associated with cardiovascular disease (CVD) incidence, but less is known about whether weight change affects the risk of CVD.

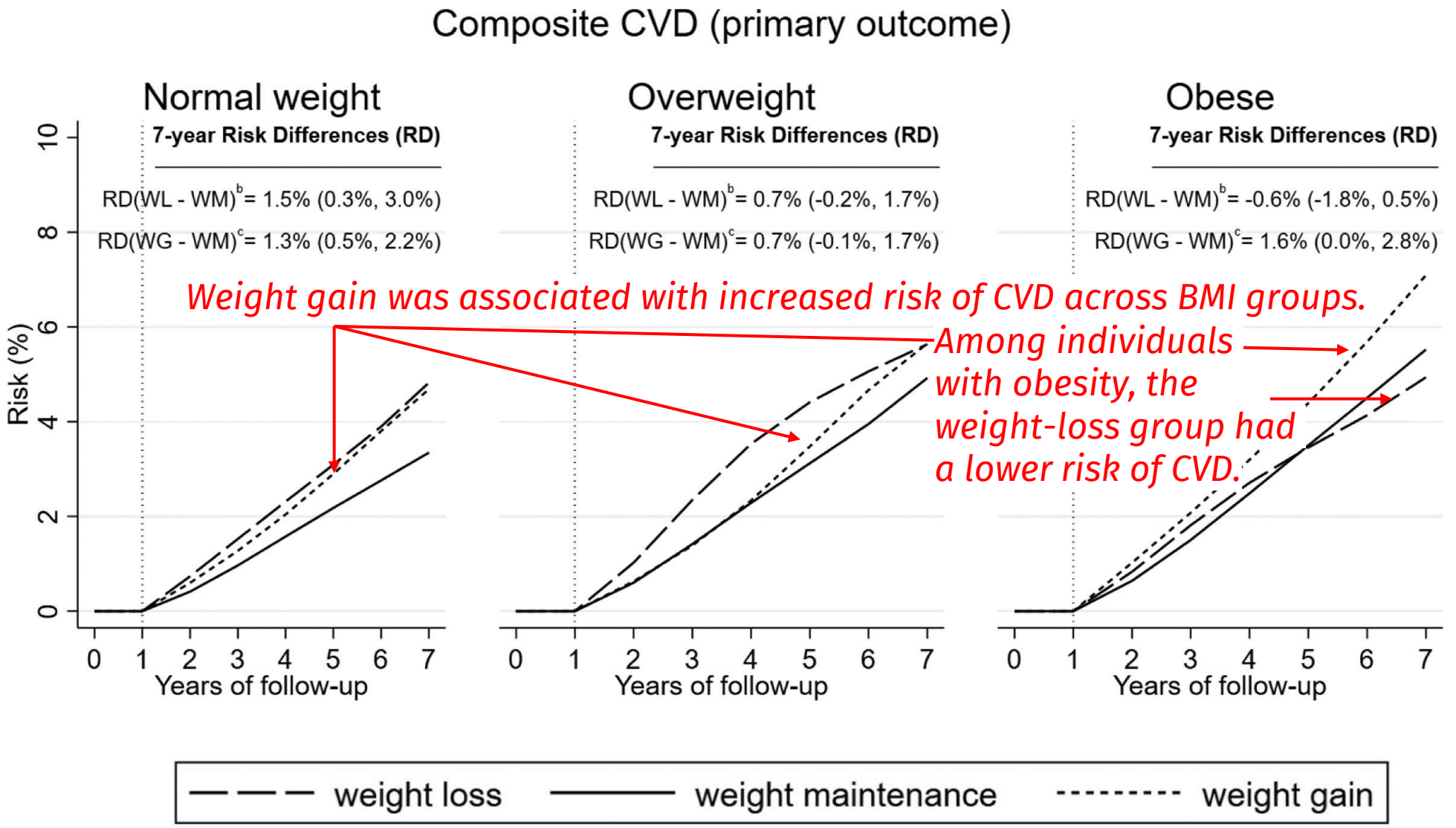
Methods:

We estimated the effect of 2-y weight change interventions on 7-y risk of CVD (CVD death, myocardial infarction, stroke, hospitalization from coronary heart disease, and heart failure) by emulating hypothetical interventions using electronic health records. We identified 138,567 individuals with 45–69 years of age without chronic disease in England from 1998 to 2016. We performed pooled logistic regression, using inverse-probability weighting to adjust for baseline and time-varying confounders. We categorized each individual into a weight loss, maintenance, or gain group.

UNSW SYDNEY

CALIBER links anonymized coded EHR from primary care (Clinical Practice Research Datalink), hospital care and death registry.

138,567 individuals 45–69 years of age without chronic disease in England from 1998 to 2016.



Katsoulis, M, et al. Epidemiology32(5):744-755, September 2021.
doi: 10.1097/EDE.0000000000001393

Annals of Internal Medicine RESEARCH AND REPORTING METHODS

Target Trial Emulation for Evaluating Health Policy

Nicholas J. Seewald, PhD; Emma E. McGinty, PhD; and Elizabeth A. Stuart, PhD

Target trial emulation is an approach to designing rigorous nonexperimental studies by "emulating" key features of a clinical trial. Most commonly used outside of policy contexts, this approach is also valuable for policy evaluation as policies typically are not randomly assigned. In this article, we discuss the application of the target trial emulation framework in a policy evaluation context. The policy trial emulation framework includes 7 components: the units and eligibility criteria, definitions of the exposure and comparison conditions, assignment mechanism, baseline ("time zero") and follow-up, outcomes, causal estimand, and statistical analysis and assumptions. Policy evaluations that emulate a randomized trial across these dimensions can yield estimates of the causal effects of the policy on outcomes. Using the policy trial emulation framework to conduct and report on research design and methods supports transparent assessment of threats to causal inference in nonexperimental studies intended to assess the effect of a health policy on clinical or population health outcomes.

Ann Intern Med. doi:10.7326/M23-2440
For author, article, and disclosure information, see end of text.
This article was published at Annals.org on 8 October 2024.

Conversion to common data models

PLOS ONE

OPEN ACCESS PEER-REVIEWED

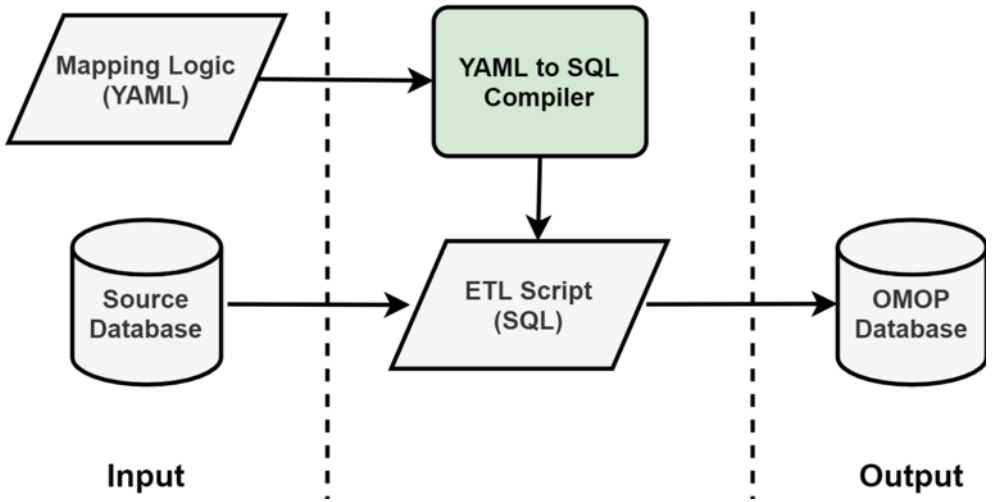
RESEARCH ARTICLE

Extract, transform, load framework for the conversion of health databases to OMOP

Juan C. Quiroz, Tim Chard, Zhisheng Sa, Angus Ritchie, Louisa Jorm, Blanca Gallego

Published: April 11, 2022 • https://doi.org/10.1371/journal.pone.0266911

See the preprint



OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

UNSW SYDNEY

Extracting value from clinical text

Clinical text

CHF due to rapid atrial fibrillation and systolic dysfunction. Patient is not diabetic.

Nuclear stress test showed moderate size, mostly fixed defect involving the inferior wall with a small area of peri-infarct ischemia. Ejection fraction is 25%. The patient is on beta-blockers and digoxin. Continue Coumadin. Deep venous thrombosis prophylaxis. The patient is on heparin drip. Monitor INR. Tobacco smoking disorder. The patient has been counselled.

Traditional and New NLP Tools

SNOMED International

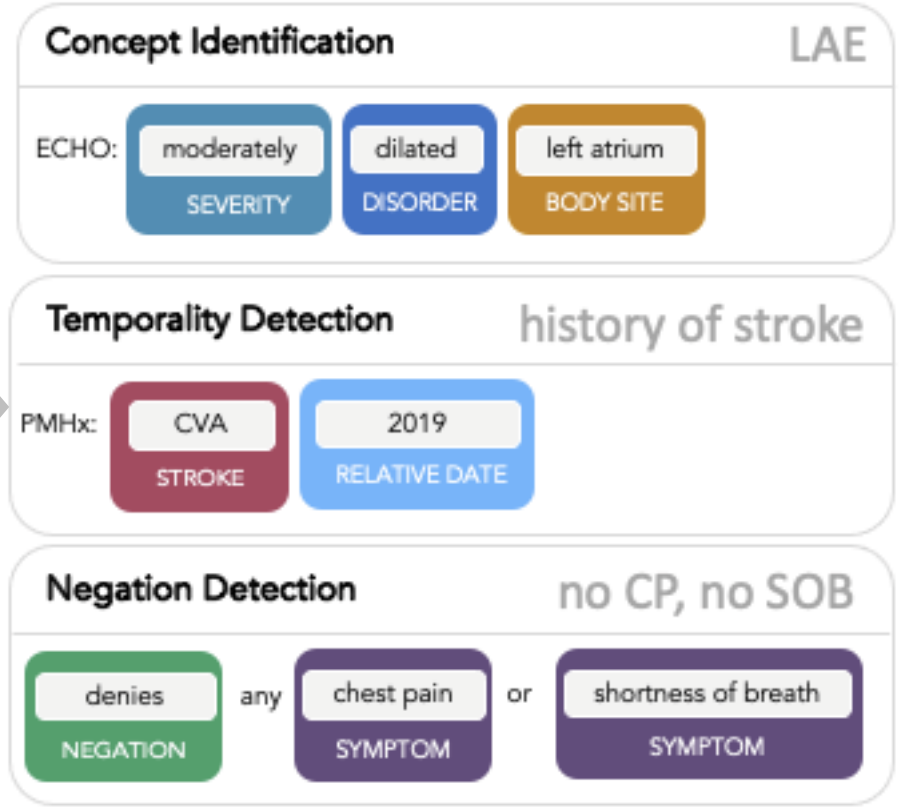
Biomedical Ontologies

Large Language Models

Automatic Text Labelling

Rule-Based NLP

Information Extraction



UNSW SYDNEY

De-identification of clinical text

ELSEVIER

Journal of Biomedical Informatics
Volume 135, November 2022, 104215

Original Research

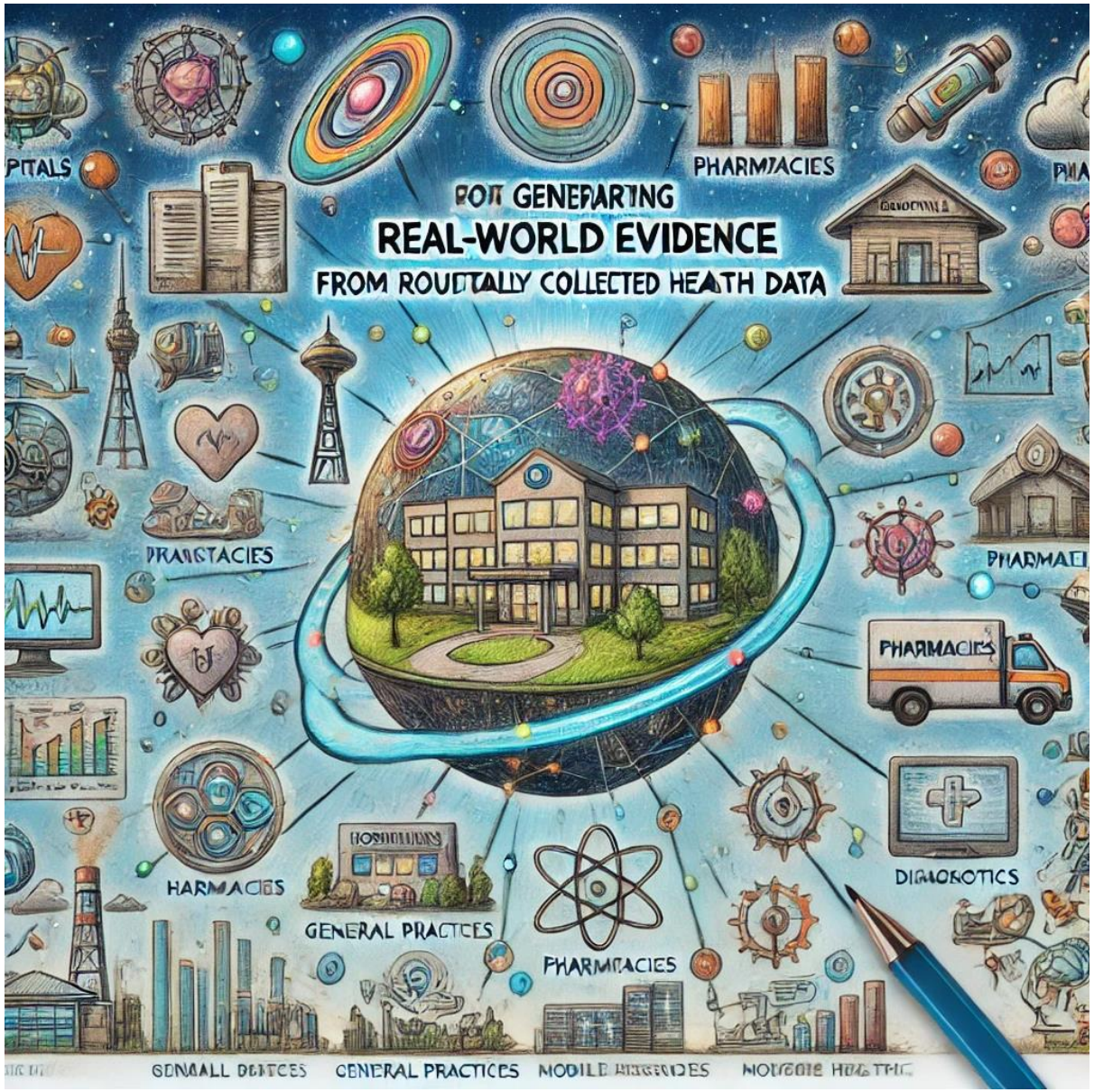
De-identifying Australian hospital discharge summaries: An end-to-end framework using ensemble of deep learning models

Leibo Liu [a], Oscar Perez-Concha [a], Anthony Nguyen [b], Vicki Bennett [c], Louisa Jorm [a]

	PII Category	Description
1	PERSON	Patient names, Doctor names, Family member names and other names
2	ADDRESS	Patient address, General Practitioner address
3	DOB	Date of Birth
4	IDN	Medical Record Number (MRN), Financial Identification Number (FIN), Doctor pager number
5	PHONE	Phone number, fax number

UNSW SYDNEY

Questions?



Generated using DALL·E 3