Acknowledgement of country

I acknowledge the Traditional Custodians of the land upon which we meet today, the Gadigal people of the Eora Nation. I also pay respect to Elders past, present and emerging, and extend that respect to other Aboriginal and Torres Strait Islander people who are here today

The George Institute for Global Health

The presenter has advised that the following presentation is subject to no conflicts of interest and has nothing to disclose.

CONNECT

SSaSS
The China Salt Substitute and Stroke Study

resolve
Randomised Evaluation of SOdium dialysate Levels on Vascular Events

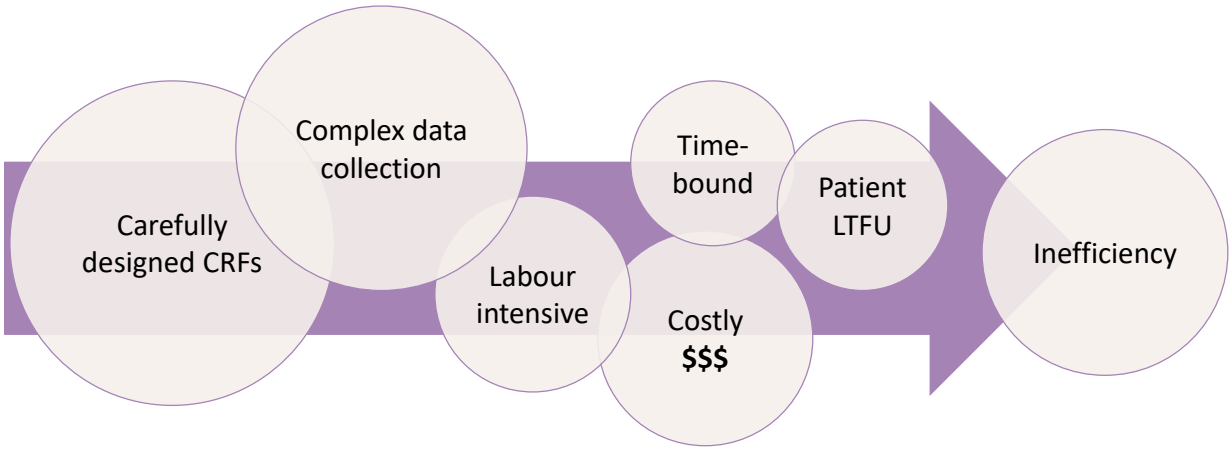
Routine data in randomised controlled trials: The George Institute for Global Health experience

The George Institute for Global Health

Carinna Hockham

On behalf of the Catalyst Program Working Group on Routine Data in Clinical Trials and trial investigators

Conventional RCTs are slow



We need ways to get around these issues

Data are being generated all the time

Data are being generated all the time

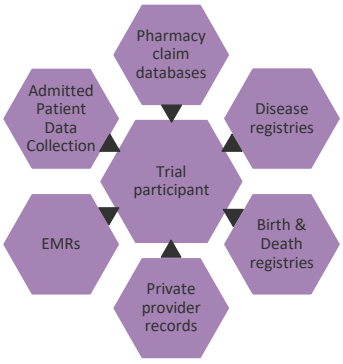
Advantages of using routine data in RCTs:

- Reduces the operational burden as compared to de novo data collection
- More cost-effective for long-term follow-up
- More reflective of real-world events or practice

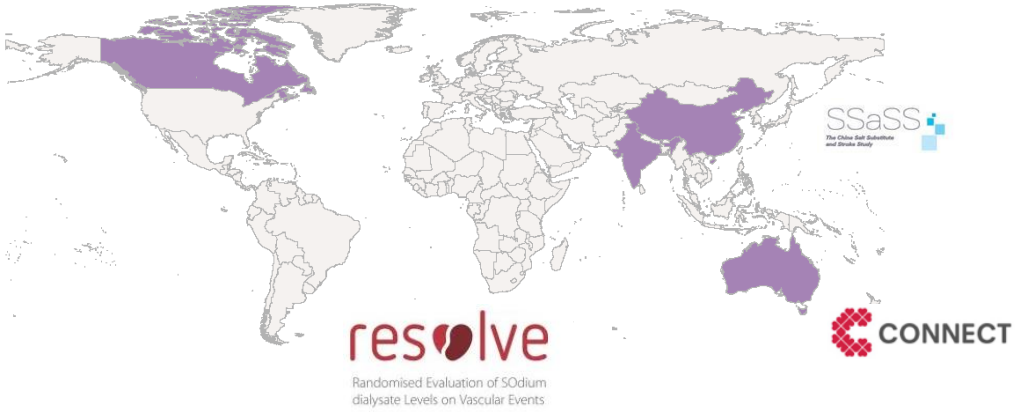
The challenge is how to best make use of this data.

Routine data

- Any data that are created and maintained externally to a trial
- As a by-product of patient care or service provision
- Typically linked, through an algorithm, to trial participants



Three trials



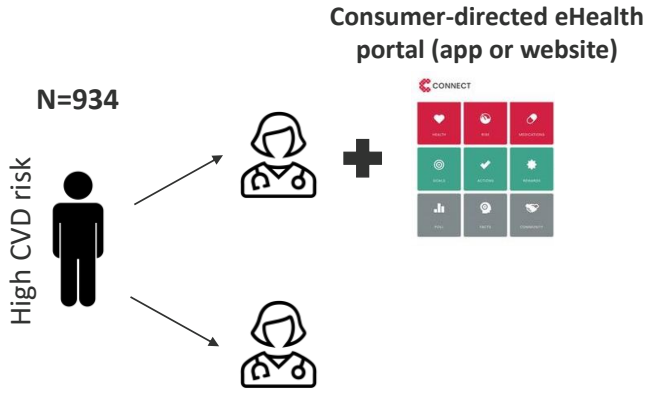
Consumer Navigation of Electronic Cardiovascular Tools

CONNECT

Objective: To investigate whether a consumer-focused eHealth strategy can improve CV health in individuals at high risk of CVD

Consumer Navigation of Electronic Cardiovascular Tools

CONNECT

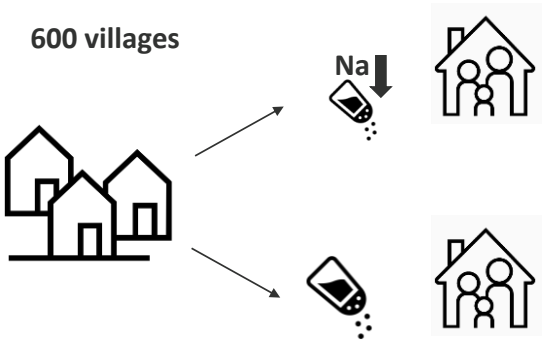


- Outcomes of interest:
  - Medication adherence
  - Clinical targets (e.g. BP, cholesterol)
  - Lifestyle risk factors
- Data collection through:
  1. Questionnaires at baseline, 12, and 24 months
  2. **Linked PBS data**

Salt Substitute and Stroke Study

Objective: To investigate whether a reduction in sodium intake through a low-sodium salt substitute reduces the risk of stroke.

Salt Substitute and Stroke Study



35 participants/village

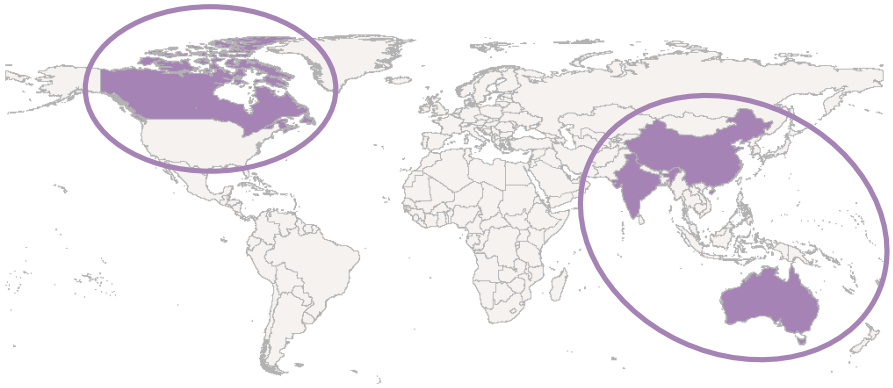
Total N=20,996

- Primary outcome: Stroke
- Suspected events through:
  1. 6mth questionnaires (first 24mth)
  2. **Linked NCMS data**
  3. **Linked national mortality surveillance**
- Independent adjudication of suspected events

Randomised Evaluation of Sodium Dialysate Levels on Vascular Events

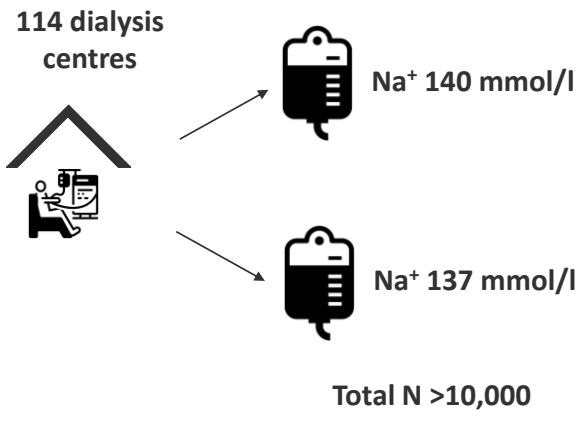
resolve
Randomised Evaluation of SOdium dialysate Levels on Vascular Events

Objective: To evaluate the effect of sodium dialysate levels on mortality and cardiovascular outcomes in haemodialysis patients.



Randomised Evaluation of Sodium Dialysate Levels on Vascular Events

resolve
Randomised Evaluation of SOdium dialysate Levels on Vascular Events



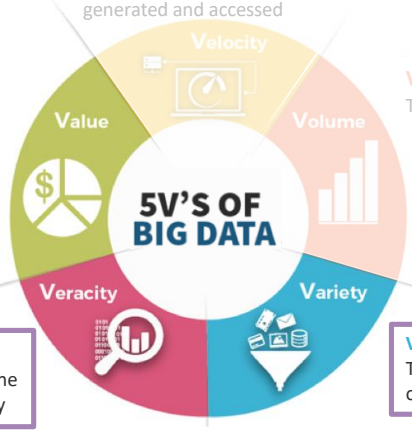
- Outcomes of interest:
  - Major CV events
  - All-cause mortality
- Data collection through:
  1. **Linked dialysis registry data OR EHRs OR private provider network records**
  2. Active data collection through eCRFs

Five V's of Big Data

Velocity
The speed at which the data is generated and accessed

Volume
The size of the data

Value
The value that can be derived from accessing and analysing big data



Veracity (accuracy)
The trustworthiness of the data in terms of accuracy

Variety
The different types of data

Variety in data and processes

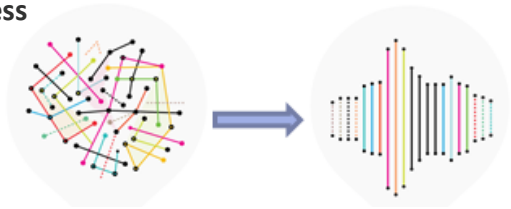
SSaSS:
- 10 counties in northern China
- 1 county = 1 health insurance provider
  → Data vendors
  → Data platforms

RESOLVE:
- Linked data from >1 countries and in multiple jurisdictions
  → Different data custodians
  → Different privacy laws

Processes for data access

File formats

Data variables



Variable formats

Identifying information

Data transfer restrictions

Variety: lessons learned

- Speak to key stakeholders and vendors up front
- Relationships and buy-in are key
- Involve government stakeholders
- Know their data – data dictionaries, in-depth discussions
- Understand their internal processes & arrange timelines around them

Variety: lessons learned

- Speak to key stakeholders and vendors up front
- Relationships and buy-in are key
- Inv
- Kn
- Un round
the

SSaSS	RESOLVE

Veracity (accuracy)

- Two levels to think about:
  1. Uncertainty in the raw data
  2. Uncertainty in using this data to derive outcomes of interest

Veracity (accuracy)

- Two levels to think about:
  1. Uncertainty in the raw data
  2. Uncertainty in using this data to derive outcomes of interest

SSaSS:

- Standardised mortality reporting across surveillance sites
- Underreporting is a problem
- Validation study: Cause of death in death registry compared to SSaSS-adjudicated data collected in first 24 months

Veracity (accuracy)

- Two levels to think about:
  1. Uncertainty in the raw data
  2. Uncertainty in using this data to derive outcomes of interest

CONNECT:

PBS dataset (original variables)	Clinical knowledge & assumptions	Statistical programing	Derived adherence
Date of supply Packet size	Typical dosage Medicine switching	→	**Proportion of Days Covered**

Value

- Value of routine data can be assessed in terms of cost or the information it provides

All trials

- Reduces trial burden on participants

CONNECT

- Reduces recall bias
- Limitations to self-reported adherence scales → routine data provides more detailed information

Value

SSaSS

- Huge undertaking to actively collect data for ~21,000 participants → routine data is faster and costs 40% less
- Also provides additional medical information not always available from self-report, and over a long time period

RESOLVE

- Pragmatic approach to evaluating the comparative effectiveness of two default dialysate sodium concentrations in real-world conditions

Value

Pros	Cons
(Often) Cheaper	Not always in your control
Reduces participant burden	Requires participants to be accessing healthcare
Reduces burden on site staff & monitors	Increases burden on data managers & trial statisticians
Can provide richer information	Can be difficult to validate

Up to researchers to weigh up the benefits and costs of using routine data

Summary

- Routine data has the potential to transform the way we do randomised controlled trials
- It shifts the burden from research nurses and monitors to statisticians and data managers who need to wrangle the data → this should not be underestimated
- Not a silver bullet and decisions need to be made about whether the use of routine data is appropriate

- **Thoughtful development of rigorous processes is needed to maximise trial quality and integrity**

Thank you

The George Institute for Global Health

For help with this presentation
A/Prof Meg Jardine
A/Prof Kris Rogers
Ms Erika Dempsey
Ms Liping Huang
Ms Yishu Liu

Trial Investigators
Prof David Peiris
Dr Maoyi Tian
Prof Bruce Neal
Prof Yangfeng Wu
Ms Manuela Armenis

Catalyst Program Working Group on Routine Data in Clinical Trials
Ms Lyndal Hones
Mr Thomas Bradbury
Dr Min Jun
Ms Sindhu Prasad
Dr Emily Atkins
Dr Thomas Lung
Dr Cecilia Ng
Ms Helen Monaghan