

# Is ‘Big Data’ the new epidemiology?

**“We know more than ever about the behaviors of our communities but what are the opportunities for linking this information effectively with action to improve health and with health services?”**

Simon Ross  
Manager, Analysis & Reporting  
Ministry of Health

[Simon\\_Ross@moh.govt.nz](mailto:Simon_Ross@moh.govt.nz)

---

# Overview

---

A bit about me and the work I do

Various contexts

- International – the times we live in
- Government / NZ Government– what is government to achieve with data?
- Health and social sector – what opportunities are there in our own sector?

Case studies

- Integrated Data Infrastructure (5 minutes)
  - Virtual Health Information Network (5 minutes)
  - Operational uses of health and social sector data (5 minutes)
    - PRM
    - Live NHI feedback (VDR and others)
-

## My background (for context)

---

- Manage a team of analysts at the Ministry of Health – Analytical Services
- We do three things:
  1. **Supply data** to researchers and operational users across the Ministry, government, the health sector, academia and beyond
  2. **Advise on the use and interpretation** of supplied data
    - Expertise
    - Metadata
    - Conduit for data quality advice and investigation
    - Conduit for coding advice and explanations
  3. **Protect patient privacy and confidentiality** and Ministry of Health information security through safe management of access to data at patient and health event level

Contact us: [data-enquiries@moh.govt.nz](mailto:data-enquiries@moh.govt.nz)

---

# Using data well is only becoming more critical

---

- What's happening globally? Information age, open government
    - Value
    - Efficiency
    - Outcomes
    - Breaking up of government and sectoral silos
  - What's happening in government in NZ?
    - NZ Data Futures Forum / Partnership
    - Open Data
    - Integrated Data Infrastructure
  - What's happening in health?
    - Health Strategy refresh, capability & capacity, funding models...
    - Better and more efficient use of data through sharing and collaboration
    - Virtual Health Information Network
    - Debate about how to use the outputs of analytic work when they point to individual level interventions
-

# Integrated Data Infrastructure

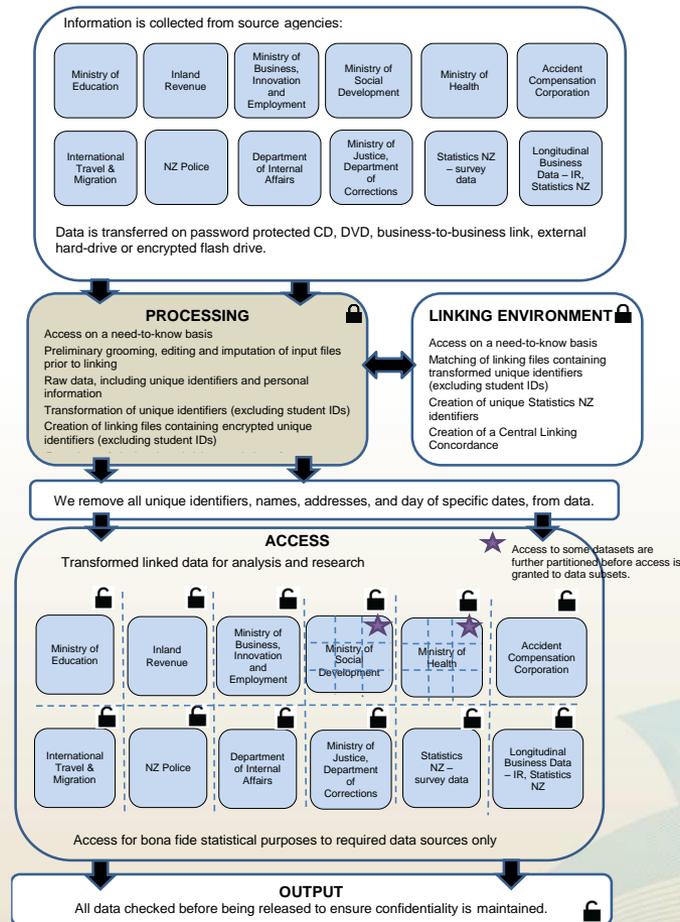
**NZ Government**

**Statistics New Zealand**

---

# Integrated Data Infrastructure (IDI) - overview

Figure 1: Flow of personal information within the Integrated Data Infrastructure



# Integrated Data Infrastructure – data in

Information is collected from source agencies:

Ministry of  
Education

Inland  
Revenue

Ministry of  
Business,  
Innovation  
and  
Employment

Ministry of  
Social  
Development

Ministry of  
Health

Accident  
Compensation  
Corporation

International  
Travel &  
Migration

NZ Police

Department  
of Internal  
Affairs

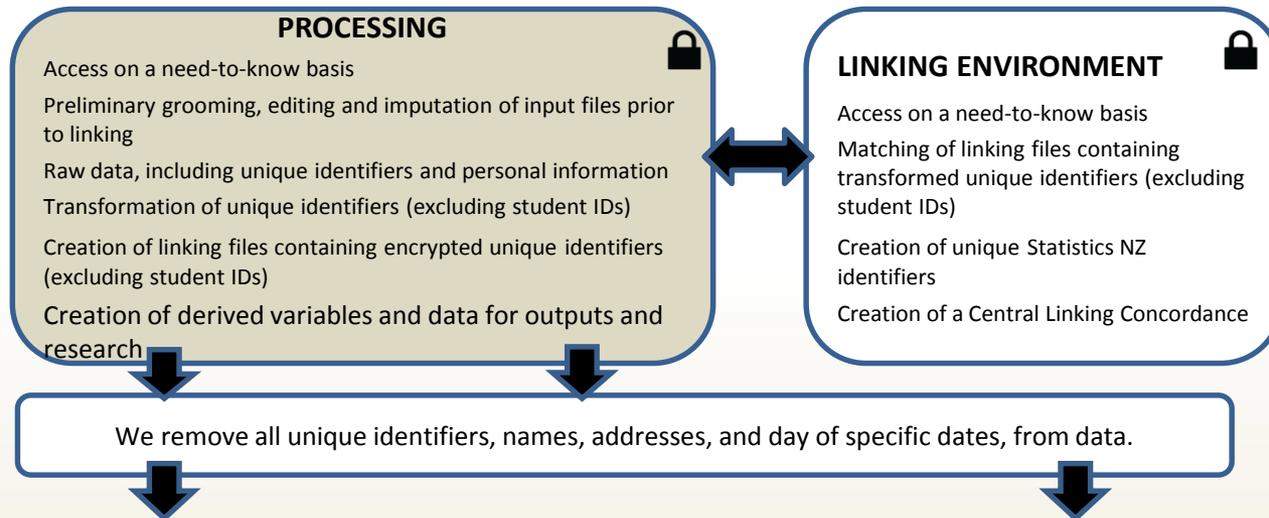
Ministry of  
Justice,  
Department  
of  
Corrections

Statistics NZ  
– survey  
data

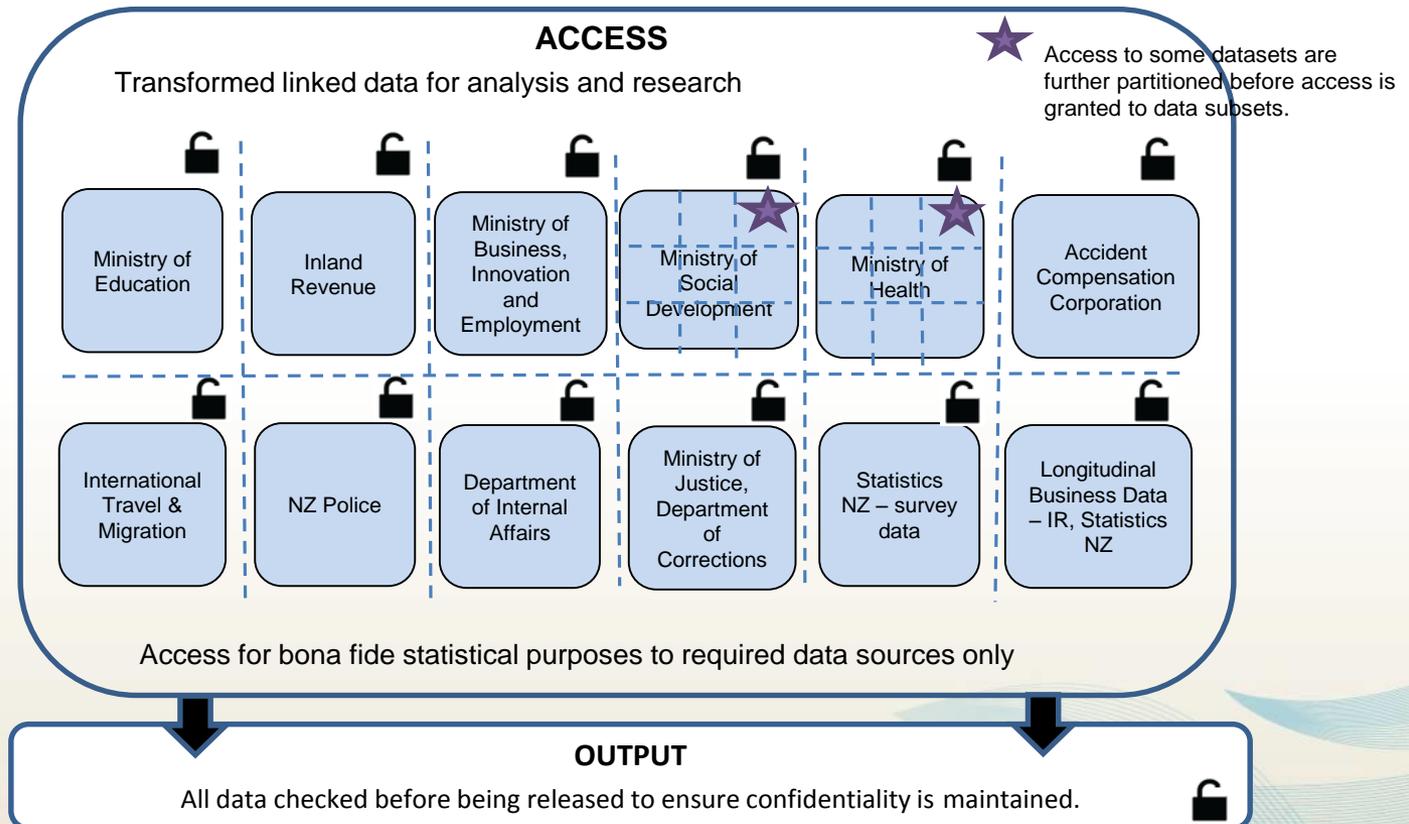
Longitudinal  
Business  
Data – IR,  
Statistics NZ

Data is transferred on password protected CD, DVD, business-to-business link, external hard-drive or encrypted flash drive.

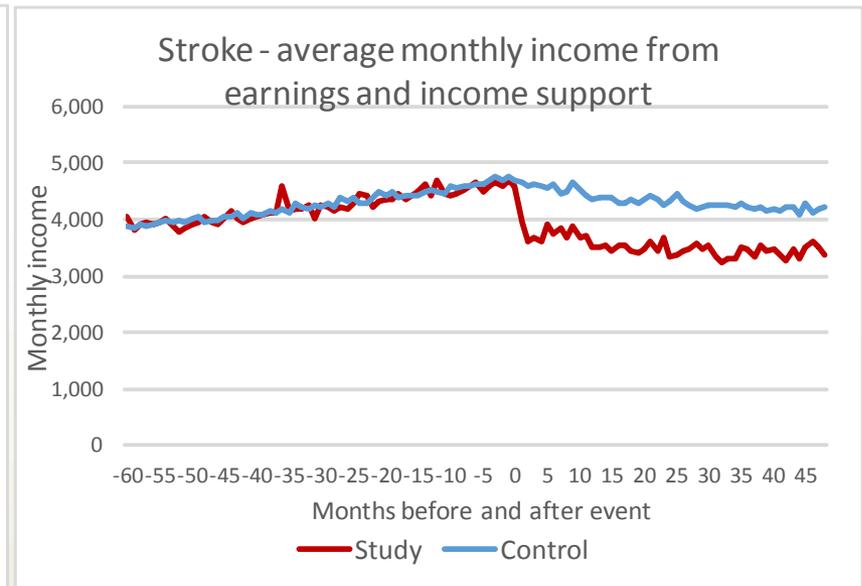
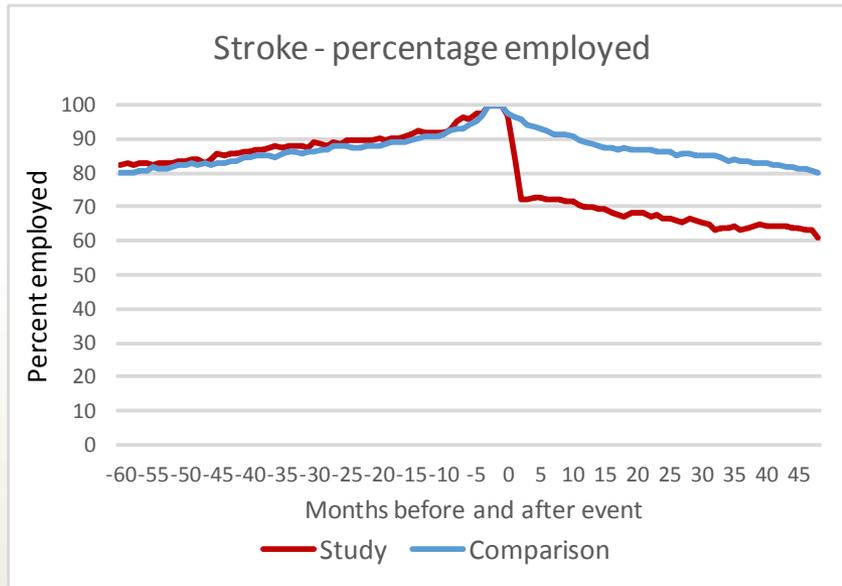
# Integrated Data Infrastructure – data processing



# Integrated Data Infrastructure – data access



# Treasury's work on the employment and income effects of eight chronic health conditions



## What does the IDI provide for health analysts and researchers?

---

- A secure research environment
  - A single environment accessible from multiple locations
  - A collaborative space
  - Access to lots of contextual information not usually available for health research / epidemiological work
  - An enhanced ability to adjust for confounding factors
  - The opportunity to enhance and improve the environment overtime limiting the need for rework following data refreshes
    - Persistent derived variables
    - Some data management capacity
-

# Virtual Health Information Network

[WORKING TITLE]

**Ministry of Health analysts**

**Health sector analysts**

**Health academics**

---

## Virtual Health Information Network - Opportunities

---

- New Zealand is building from a strong base of consistently collected health administrative data with a strong unique patient identifier
  - However, a fragmented and inefficient approach to using this resource is holding us back
  - The IDI provides a platform – new infrastructure
  - The network approach seeks to maximise the value of the IDI service
  - A more efficient approach to the nuts and bolts of data linkage projects – linking, data management, data quality, data cleaning, metadata, derived variable creation, etc. – should lead to more competitive and efficient research
- 
- 

## Virtual Health Information Network – Vision

---

- to create and sustain an environment that captures value from linking health data collections and related social and economic data, through world leading health research, policy development and service planning
  - a joint initiative between the Ministry of Health, health sector participants and several New Zealand Universities
  - a collaborative and cooperative environment where participants contributions achieve their own goals but also add value to the work of colleagues and future data users
- 
- 

## Virtual Health Information Network – Objectives

---

To improve the health of New Zealanders, reduce health inequalities, and improve resource allocation by:

- more efficiently and effectively use routinely collected data
  - enhancing data integration
  - developing quantitative analysis methods
  - improving collaboration between health researchers and analysts through sharing
    - metadata
    - code
    - derived variables
    - knowledge of strengths and weaknesses of the data
-

## Virtual Health Information Network – Progress so far

---

- Formation of an Executive to take the project forward
  - Securing seed funding
  - Identifying the confirming catalyst projects
  - Recruiting / seconding contributors
    - Simon Ross, Interim Manager, Development Phase
    - An academic/technical resource (research fellow) to progress the catalyst projects
  - Promoting the vision to data users – academics and health sector professionals <https://www.facebook.com/groups/1658066281071881/>
-

# Operational uses of 'big data'

---

# Potential operational uses of 'big data'

---

Such operational uses are often contentious from a privacy perspective

## Examples

- Predictive Risk Modelling
- Virtual Diabetes Register
- Live NHI feedback

Information is being used for a different purpose than it was collected for

Quality, accuracy and completeness of becomes more critical

---

## Diabetes prevalence by ethnic group – source: VDR, MoH.

◆ European/Other   
 ◆ Indian   
 ◆ Māori   
 ◆ Pacific people

