Linkage of National and Commonwealth data

Dr Phil Anderson
Head, Data Linkage Unit
What's new

Australian Burden of Disease Study: impact and causes of illness and death in Australia 2011
media release: Less disease and injury burden since 2003– more improvements possible (10 May 2016)
Burden of disease (web page and presentations)

Back problems, associated comorbidities and risk factors (web report)
(17 Jun 2016)

Children's Headline Indicators (dynamic data display web update) (16 Jun 2016)

Alcohol and other drug treatment services in Australia 2014–15 (15 Jun 2016)
media release: 1 in 200 Australians seek treatment for alcohol and other drugs

Medication use by Australia's prisoners 2015: how is it different from the general community? (10 Jun 2016)
media release: Pain killer and antipsychotic medication use high among prisoners—particularly females

• More media releases
• Australian Institute of Health and Welfare Corporate Plan 2015–16 to 2018–19
• More publications
• Forthcoming publications

AIHW in the news

Almost 334,000 people used disability support services in 2014/15 under the National Disability Agreement (NDA) with another 4000 already transitioning to the National Disability Insurance Scheme (NDIS)

Australian Burden of Disease Study: fatal burden of disease in Aboriginal and Torres Strait Islander people 2010
10 Apr 2015

Alcohol and other drug treatment services in Australia 2013–14
18 June 2015

Hospital resources 2013–14: Australian hospital statistics
23 Jun 2015
A short history

• Mid-1990s: Linkages with the National Death Index, then Australian Cancer Database, internal and external projects
• Late 1990s: Linking service events for people. SLK851
• 2000s: Pathways in Aged Care and increasing linkages
• 2009: Dedicated Data Linkage Unit created
• 2011: First big Medicare linkage, AIHW joined PHRN
• 2012: AIHW approved as Integrating Authority
• 2015: Approval for Medicare data.
• 2016: Currently nearly 300 ‘active’ projects
AIHW data linkage

• Commonwealth
  – AIHW as an Integrating Authority

• National
  – Range of protocols for approvals

• State and Territory
  – AIHW works with state data linkage units

• Longitudinal surveys
  – Usually have consent for linkage

• Secondary use of linked data
Linkage Methodologies

• Probabilistic
  • Full name, date of birth and sex
  • Address/postcode
  • Date of last contact
• ‘Key-based linkage’
  • Community services statistical linkage key (SLK851)
  • Postcode
  • Event dates
• Event-based
  • Dates of admission and separation
  • Date of birth and sex
• GRHANITE
Recent Ethics Committee applications

- Linkage of the RFPCS and EOPCFS prostate cancer studies to the National Death Index
- Does continuity of primary care reduce demand on emergency department presentations and hospital admissions?
- Enhancing the accuracy of heart disease risk assessment in Indigenous Australians
- Vulnerable Youth: Pathways of homelessness through juvenile justice and child protection
- Cancer Survivors and Support Persons' Preferences for Care: A discrete choice experiment
- Optimal Pain Assessment in Joint Replacement Patients
- Utilising prescription patterns to identify increased risk of death after a diagnosis of epilepsy
- The Latrobe Early Life Follow-up Anonymised Data Linkage Study (Hazelwood Health Study)
- ACT Asbestos Health Study
- The safety of pharmacotherapies for smoking cessation
- Linkage of ACT Cancer Registry to the National Death Index
- Recent changes in IVF clinical practice: impact on fetal growth, birth defects & cerebral palsy
- Surveillance of hospitalizations, antiviral therapy, hepatocellular carcinoma and mortality among people diagnosed with hepatitis C or hepatitis B virus infection in New South Wales
- Hazelwood Adult Survey & Health Record Linkage Study
- Cancer Pathways project
National Death Index

- Agreements with State and Territory Registrars and National Coronial Information Service
- NDI set up specifically for data linkage
  - All deaths in Australia since early 1980s
  - Fact of death current to last month
  - Cause of death added annually (18 months lag)
  - Can update when available
  - One-step approval through AIHW EC
Australian Cancer Database

• AIHW has data for linkage.
• Projects need approval by all jurisdictions
  – Approval process reliable and becoming more efficient
  – Standard set of variables
• Regular linkages possible for both NDI and ACD
• Consent issues
National Aged Care Data Clearing House

- Permanent and respite residential aged care, community care, transition care, ACAP and HACC
- AIHW auspiced to link with EC approval
- Linkage began with Pathways in Aged Care project and now almost complete to 2013-14
National Diabetes Services Scheme

• Managed by Diabetes Australia (DA) for the Department of Health
• AIHW routinely links to National Death Index for DA and as part of updating AIHW National Diabetes Register
• Linkage undertaken for a number of projects
Community Services Data

- Disability Services NMDS
- Juvenile Justice NMDS
- Child Protection NMDS (new and not yet complete)
- Special Homelessness Services Collection
- Mental Health Services
- Drug and Alcohol Treatment Services
High Level Principles for Data Integration Involving Commonwealth Data for Statistical and Research Purposes

Australian Institute of Health and Welfare (AIHW)

Date accredited: 28 June 2012
Accreditation application summary: Australian Institute of Health and Welfare (AIHW)
Accreditation application summary (pdf): [pdf]
Contact details:
Visit the AIHW web site at www.aihw.gov.au
Email AIHW: disc@aihw.gov.au
Visit: 28 Thynne Street, Fern Hill Park, Bruce, ACT 2617

Entailed by: Portfolio Secretaries
3 February 2010

High Level Principles

- Treat data as a strategic asset
- Custodians retain accountability
- Integrating Authority to be nominated
- Public Benefit
- Statistical and Research Purposes
- Transparent Governance
Cancer risk and tomography scans

Is there a link?

Exposure Data

Incidence Data
Medicare data files at AIHW

• Medicare enrolment file for data linkage
  – From 1984 with history
  – 30.6 million people with 43.2 million names
  – 36 million addresses

• MBS and PBS data files 2010-2015
Application and approval process

1. Research question/proposal
   Data Users

2. Preliminary discussions
   Data Users, IA, Data Custodians*

3. Assess project feasibility
   Integrating Authority

4. In principle project approval
   Data Custodians

5. Privacy and ethics approvals
   (as required)

6. Final project approval
   Data Custodians

7. Formal agreement
   IA & Data Custodians

8. Formal agreement re: release of data
   IA and Data Users
Arrangements with DoH

• Approval still necessary from Department of Health for data prior to 2010
  - facilitated by AIHW
  - if cohort from MBS/PBS has to happen first
• AIHW will also carry out risk assessment
• Expect further historical data to be provided in time, which will simplify process
• Currently still project by project
Current application process

• One application form
  – AIHW Data Integration Services Centre application
  – PHRN Online Application System
  – State data linkage unit application

• AIHW ethics approval

• Home ethics approval (and other custodian approvals)

• Use of SURE
Other data sets

- Australian Childhood Immunisation Register
- Cancer screening registries
- Australian Early Development Census
- Centrelink data  
  - First project approved
- Others
Linkage with state data

• Can be done directly or through State data linkage units
  – Data can be linked at State level and then to national data sets
  – Currently most projects involve one or a few states

• Working towards a better national system particularly with hospital data
Better Cardiac Care Data Linkage (NSW)

**AIHW**

**Linkage Domain**
AIHW undertakes separate linkages to identifier data of each approved dataset, using the Cohort file.

**Merging Domain**
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

**Secure Unified Research Environment**
De-identified Content data and PPNs

Only Content data from MBS, NDI, NSW and ACT Hospital and Emergency datasets including PPNs are accessible.

**CHeReL (NSW Linkage Unit)**

- Creates the ‘Hospital Cardiac Cohort’ file from the Master Linkage Key and assigns a PPN to each person on the Cohort file.
- Provides IDs and PPNs of Cohort records to state data custodians to extract Content data for Cohort records.

**NSW APDC**
- APDC: Admitted Patient Data Collection

**NSW EDDC**
- EDDC: Emergency Department Data Collection

**ACT APC**
- APC: Admitted Patient Collection

**ACT EDIS**
- EDIS: Emergency Department Information System

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.
Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

Only Content data and PPNs are released to SURE

Cancer Pathways Study

Study Cohort file
Contains identifier data and project-specific person number (PPN) of colorectal and breast cancer patients

Victorian Data Linkage Group (VDL and DHHS)
- Creates the Study Cohort file and assigns a PPN to each person on the Cohort file.
- Provides IDs and PPNs of Cohort records to the data custodian (Dept. of Health and Human Services) to extract Content data for Cohort records.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data from MBS, PBS and specific Victorian datasets including PPNs are accessible

Secure access

Cleared output

AIHW

Linkage Domain
AIHW undertakes separate linkages to identifier data of each approved dataset, using the Cohort file.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.

MBS

PBS

Merging Domain
AIHW obtains Content data separately from each approved data file and creates de-identified Content datasets with PPNs.

Secure Unified Research Environment

De-identified Content data and PPNs

Only Content data and PPNs are released to SURE

Cleared output

Content data includes services, diagnostic, clinical and other data the researcher has approval to access.
Commonwealth initiatives

• Public Sector Data Management Report
• Australian Government Public Data policy Statement
  – Securely share data across Australian government
  – Make high value data available
  – Build partnerships with public, private and research sectors
Commonwealth initiatives (cont)

- Senate Select Committee on Health hearings on data
- Productivity Commission enquiry on Data Availability and Use
- OAIC consultation on Big Data Guidance
- National Centre for Longitudinal Data (Department of Social Services)
Future prospects

- Moving from project-by-project linkage to ongoing system
  - Working towards regular linkages of key datasets
- Increasing collaboration between Commonwealth and States and Territories
- Greater availability of data
  - Increased efficiency in approval and provision
- Better geocoding
Thanks

• www.aihw.gov.au

• phil.anderson@aihw.gov.au

• linkage@aihw.gov.au