Clinical Research Seminar Series
November 22, 2017

SAGE: Collaborations through Data Sharing
Hannah Lloyd-Jones

Outline

- PolicyWise for Children & Families
- Data Sharing
- SAGE
- Lessons Learnt and Future Plans
- Questions and Discussion
About Us

PolicyWise for Children & Families exists to improve child, family and community well-being by leading, creating, enabling and mobilizing research and evaluation for evidence-informed policy and practice.

Our Five Core Functions

• Undertake research, evaluation, communication, and knowledge mobilization
• Conduct, fund, and build policy-relevant research capacity
• Link, analyze, and manage data
• Manage resources and strategic relationships
• Measure and communicate our impact and value

SAGE
(Secondary Analysis to Generate Evidence)

Enable and support use to inform policy and practice
Data Definitions

- Research data: “The recorded factual material commonly accepted in the scientific community as necessary to validate research findings.” (OMB Circular 110 - US).
- Administrative data: Collected for administrative purposes such as registration, transaction and record keeping (Adapted from ADLS - UK). Example: data collected by government ministries.
- Community service data: Data collected for the purpose of delivering services. Example: data collected by food banks, emergency shelters.

Question 1

Do you currently use secondary data in your research?

Answers:
1. Yes, research data
2. Yes, administrative data
3. Yes, Statistics Canada data
4. Yes, other data (or a combination of the above types of data)
5. No

Researcher Survey 2016

Do you currently use secondary data in your research?
Question 2
If you collect data for research purposes, do you currently make it available for secondary use?

Answers:
1. Yes
2. No

Question 3
1. If you do share data, please tell us why.
2. If you don't share data, please tell us why.

What are the benefits of sharing data?
Shout out!
Benefits of Secondary Data Use 1

• Expedites scientific discovery and knowledge mobilization
• Encourages interdisciplinary research and collaboration between different sectors
• Greater return on investment
• Avoids duplication of primary collection
• Reduces burden on clients/participants/individuals (especially vulnerable communities)

Benefits of Secondary Data Use 2

• Ethical (especially for publicly funded data collection)
• Expands the use and value of existing data
• Gain systemic intelligence on performance, outcomes, and impact.
• Potential for linkage and cross-disciplinary analysis with other existing datasets

Funders and publishers want data sharing
Current Environment

Methods of Data Sharing

Open Access    Facilitated Access

Data sharing

What are the barriers to sharing data?

Shout out!
Secondary Data Use Issues

- Ethical and legal issues
- Costs: of cleaning, creating metadata, making data understandable and reusable
- Often larger dataset than for-purpose data collection, but not always the exact data point needed

Ethical Considerations

- Need to consent/re-consent
  - Covered in FOIP, HIA, PIPA
  - TCPS 2: Chapter 5, Section D
- Potential harm to individual
  - Misinterpretation/Misuse of data
  - Re-identification
  - Disclosure of sensitive information

Methods for ethical use

- Secondary use by design
- Collaboration
- Anonymization
- Access Control
Secondary Use by Design

- Client/participant consents
  - Informed about collection, use and disclosure
  - Provisions for sharing and future use of data
- Secondary use incorporated in proposal and grants

Collaboration

- Data cleaning and documentation
  - Metadata standards such as DDI, DCMI, CORMAP
- Open communication between primary researcher (i.e. data producer) and secondary user
- Proposal review process
  - Inclusion of data producer and external reviewer

Anonymization

- Limit upper and lower ranges
  - Hide outliers
- Aggregation/Categorization
  - Reducing the precision of variables
- Suppression
- Extreme methods
  - Dummy data sets
  - Perturbative methods (adding noises, swapping records, etc.)
Access Control

- Data Access Agreement
  - Data use is limited to the proposed research
  - Clauses: No attempts to establish the identity of individuals
- SAGE Analytics Environment
  - Two-factor authentication of authorized use
  - No internet access
- Vetting Rules
  - Result output: aggregated results, suppression of cell sizes

How can we help you?

1. Wording for consent forms

- Template wording around consent:
  - Informed about collection, use and disclosure of personal information
  - Provisions for sharing and future use of data

How can we help you?

2. Wording for ethics applications

- Template wording for ethics applications:
  - Data storage
  - Data privacy and security
  - Data access
How can we help you?

3. Data management planning

- Consider your research in the context of the research data lifecycle – you want it to be reproducible and reusable (by yourself and/or others)


How can we help you?

4. Data Governance

- Template data deposit and data access agreement
- Governance: access process, de-identification, security etc.
How can we help you?

5. Advice on metadata

Metadata standards:
- Data Documentation Initiative (DDI), Dublin Core
- Metadata Initiative (DCMI), Government of Alberta Core
- Content Standard (CORMAP)
- Ensure that data is understandable and usable by other researchers - prevents misinterpretation of data

How can we help you?

6. Collaboration

- Our goal is to help build collaborations, and connect you to a broader network of researchers, evaluators and community organizations

Calgary Thrives Project: January – December 2017

- Calgary Food Bank
- Families Matter
- Calgary Counselling Centre
- South West Community Resource Centre
- Sunrise Link
- CUPS
Lessons Learnt: Sharing Research and NFP Data

- Cultural change is slow
- Need to budget and plan for resource to make data useful for secondary use
- Need to plan for and prioritize data management and sharing at an early stage

Lessons Learnt: Sharing NFP data 1

- Need for capacity building
- Need to value data
- Very different levels of “data maturity” (including capacity, governance, resource and expertise)
- Specific and diverse needs

Lessons Learnt: Sharing NFP Data 2

Unclear legal situation
- Often fall in gaps of privacy legislation and ethical oversight
Risk averse
- Protection of vulnerable populations
- Prioritization of meeting immediate needs over long-term benefits of sharing data
- Special considerations for Indigenous communities
Future Plans 1: Capacity Building

• Clear advice, expertise and tools
• Partner with existing sources of expertise where possible

Future Plans 2: Knowledge Mobilization

• Currently track knowledge mobilization activities and outputs
• Investigating if we can offer a knowledge mobilization service for our users

Any questions?

data@policywise.com
Datasets at SAGE

Surveys
• Alberta Benchmark Survey 2008: What Adults Know about Childhood Development
• Second Provincial Benchmark Survey 2013: What Adults Know About Child Development
• Alberta Adverse Childhood Experiences Survey 2013
• What Albertans Know about Youth Mental Health 2010 Survey
• What Albertans Know about Fetal Alcohol Spectrum Disorders 2011 Survey
• What Albertans Know about Maternal Mental Health 2012 Survey

Longitudinal Pregnancy Cohort
• Alberta Pregnancy Outcomes and Nutrition (APrON) Study
• All Our Babies/Families (AOB/F) Study

Data Deposit

Data Access