

Data Deposit Agreement

SECONDARY ANALYSIS TO GENERATE EVIDENCE (“SAGE”), a research and data repository operating under the authority of POLICYWISE FOR CHILDREN & FAMILIES (“POLICYWISE”), a corporation organized and existing in the province of Alberta, with its principle office located on the 6th floor, 9925 – 109 Street, Edmonton, Alberta, T5K 2J8;

AND

[Data Producer/Data Producer’s Institution Name], [Address] (the “Data Producer’s Institution” or “Data Producer”);

Both SAGE and [Data Producer/Data Producer’s Institution] wish to enter into this Data Deposit Agreement effective as of the Effective Date specified on the final page hereof.

PRELIMINARY STATEMENT

The Data Producer has collected Data from individuals pursuant to applicable provincial legislation and policy, which include [Please type in the relevant legislation such as the Tri-Council Policy Statement 2 (TCPS), Health Information Act (HIA), Personal Information Protection Act (PIPA), and/or Freedom of Information and Protection of Privacy (FOIP)] (the “Data”).

The [Data Producer/Data Producer’s Institution] wishes to transfer the Data in its control and possession, from the Data Producer from the study approved by the Research Ethics Board on [Date], Ethics ID # [Number], entitled: [Title of Study] (the “Study”), to SAGE for storage and management of data for secondary data usage. Data related to this primary study may be routinely added to SAGE as it becomes cleaned, coded and available for transfer, at the discretion of the Data Producer.

NOW THEREFORE THIS AGREEMENT WITNESSETH that in consideration of the premises and of the mutual covenants herein set forth, the parties hereto have covenanted and agreed as follows:

I. DEFINITIONS AND INTERPRETATIONS

For purposes of this Agreement:

- A. “Agreement” refers to this Data Deposit Agreement.
- B. “Data” refers to the information included in Exhibit A, which has been collected and recorded from participants through methods described in Exhibit B.
- C. “FOIP” means the *Freedom of Information and Protection of Privacy Act*, R.S.A. 2000, c. F-25, as amended from time to time.
- D. “Study” refers to the project as described in Exhibit B.
- E. “Metadata” refers to all underlying data about the Data, without identifiers.

- F. “Data Producer” is the principal investigator, along with his or her research team, as listed on the Study outlined in Exhibit B. For the purposes of this agreement, the Data Producer will represent their research team and other co-investigators as the key contact and signatory.
- G. “Data Producer’s Institution” [Add definition. For example “a post-secondary institution, incorporate until the laws of.”]

1.2. Interpretation of Exhibits

This Agreement contains the following Exhibits, which form an integral part of this Agreement.

- A. Exhibit A – Data to be Deposited in SAGE
- B. Exhibit B – Study Description
- C. Exhibit C – Ethics Board Approval Certificate(s)
- D. Exhibit D – SAGE Overview of Repository Service
- E. Exhibit E – SAGE Data Access Process
- F. Exhibit F– Special Considerations for Data

II. AGREED TERMS AND CONDITIONS

1. **Research Work:** SAGE requests access to the Data at their sole risk and at no expense to the Data Producer solely for the non-commercial purpose outlined in Exhibit D.
2. **Non-transferability.** This Agreement is not transferable.
3. **Provision of Data.** The Data being transferred between the parties will be transferred via a secure file transfer protocol (SFTP). Upload access is limited to the Data Producer or their team through secure authentication.
4. **Control and Stewardship of Data.** The Data Producer is the steward of the Data (as defined in Exhibit A). Data stewards have specific responsibilities to manage data, to maximize its integrity through determining the authority to access, use, define, and control the quality of the data that pertains to their functional areas and/or is deemed to be under their purview.
 - A. As per this Agreement, the Data Producer agrees to share some of the Data steward responsibilities with SAGE, including determining the authority to use and access the data.
 - B. The Data Producer shall be responsible for providing applicable Ethics certificate(s) associated with the Data to be deposited.
 - C. The Data Producer shall provide a full copy of all the Data, defined in Exhibit A, to SAGE to allow it to perform the repository services defined in Exhibit D.
 - D. SAGE shall acknowledge the Data Producer’s stewardship of the Data in any Data Access Agreements, public statement or communications.
 - E. This transfer of Data does not constitute a transfer to SAGE of all stewardship rights or title to the Data. Except as provided in this Agreement, no express or implied licenses or other rights are provided to SAGE under any proprietary rights of the Choose an item to the Data.

- F. This data transfer may be updated from time to time under the same conditions of ownership, use, control and access as defined in this contract.

5. Use of Data

5.1. Responsibilities of the Data Producer

A. Initial Data Assessment

- (i) The Data Producer will warrant that the Data is original and does not infringe copyright of a third party.
- (ii) The Data Producer will warrant that it has the right to transfer and share data with SAGE.
- (iii) The Data Producer will ensure that the appropriate Ethical Board approvals and any participant notification have been obtained, or that the Ethical Board has waived the consent and notification requirements for for the deposit of data in SAGE.
- (iv) The Data Producer shall ensure the quality and integrity of the Data being transferred to SAGE but does not warrant absolute accuracy.

B. Metadata Preparation

- (i) The Data Producer will provide the information and materials needed to prepare Metadata in consultation with SAGE personnel including items listed in Exhibit A.
- (ii) The Data Producer will allow SAGE to publicly post/circulate Metadata on the Study to potential users and the public.

C. Role in Access Processes

- (i) The Data Producer is eligible for review of data access requests as defined in SAGE's Data Access Process (included as Exhibit E).
- (ii) The Data Producer shall follow the de-identification recommendations proposed by SAGE and work with SAGE to ensure that they meet current technical standards/best practices.
- (iii) The Data Producer shall provide written consent on a case-by-case basis to SAGE with permissions for handling and modifying the Dataset for the purposes of preservation and the right to delete or remove the Data from the repository or update or add to the Data deposited in the repository. Any modification of the Data shall become Data and shall vest with the Data Producer Institution.

5.2. Responsibilities of SAGE

SAGE shall be responsible for Data management/user support processes.

A. Initial Dataset Assessment, De-identification and Confidentiality

- (i) SAGE personnel shall treat all information received pursuant to this Agreement in a confidential manner in conformity with this Agreement, applicable privacy legislation and policy.
- (ii) SAGE personnel shall administer an assessment of the initial Dataset and make recommendations to the Data Producer on any necessary changes to prepare it for use in relation to proper de-identification procedures.
- (iii) SAGE will embargo or otherwise limit the use of Data as guided by the Data Producer (as outlined in Exhibit F).

B. Metadata and Marketing

- (i) SAGE personnel will aid in preparation of Metadata in alignment with a standard format (consulting the Data Producer as needed).
- (ii) SAGE personnel shall ensure that the Metadata is reviewed and approved by the Data Producer before circulation to potential users and the public.
- (iii) SAGE personnel, in consultation with the Data Producer, will provide promotional services in relation to the Data, including but not limited to: making the research community and other relevant organizations aware of the availability of the Data through public posting of Metadata on websites, brochures, presentations, etc.
- (iv) SAGE shall respond to queries from potential users about the Dataset, the access process, and support services. SAGE shall consult the Data Producer as needed to provide timely and accurate responses.

C. Access Processes

- (i) SAGE shall provide a governance structure and administrative processes for formal access requests, approvals and release; including administering Data Access Agreements between SAGE and Secondary Data Users.
- (ii) SAGE personnel or the Data Producer may assist with secondary data use proposals (as specified in Exhibit F).
- (iii) SAGE shall provide assistance to approved Secondary Data Users for data management related services.
- (iv) SAGE shall provide reports back to the Data Producer on any active queries or processes that it is administering in relation to the Data Producer releasable Data upon request.
- (v) SAGE shall keep an inventory of known findings (such as published papers, abstracts, presentations from original and new uses of the Data and publish these on a public website with appropriate links or citations).
- (vi) For individual level data with unique identifiers, SAGE shall ensure that records are removed for participants who have requested withdrawal for all subsequent uses when the Data Producer notifies SAGE of these situations.

D. De-identified Data

- (i) SAGE shall administer Data Access Agreements with Secondary Data Users.
- (ii) SAGE shall monitor secondary use for released Data and access compliance.
- (iii) SAGE shall implement processes and policies for access and secondary use of data that mitigate the risk of infringement on copyright or intellectual property of the Data Producer and its institution and help ensure the use is in compliance with applicable Federal and Provincial legislative acts governing the use of the Data.
- (iv) SAGE shall provide appropriate security infrastructure, processes and policies for storage and protection of the Data, and access for secondary uses.
- (v) SAGE shall ensure compliance of policies and procedures of secondary use through Data Access Agreement with Secondary Data Users.

- (vi) SAGE will support or broker the development of research collaborations.
- 6. Non-Identification.** SAGE agrees that Data will not be used, either alone or in conjunction with any other information, in any effort whatsoever to establish the individual identities of any of the participants from whom Data were obtained.
- 7. Compliance Monitoring and Investigations,** SAGE shall record and monitor access to the Data in its stewardship, in order to establish a chain of responsibility, as follows:
- A. SAGE will investigate all reported cases of:
 - (i) Unauthorized access to or modification of the Data in its stewardship;
 - (ii) Unauthorized use of the Data in its stewardship;
 - (iii) Unauthorized disclosure of the Data in its stewardship;
 - (iv) Unauthorized research on the Data provided to Secondary Data Users
 - (v) Breaches of privacy or security with respect to the Data in its stewardship or with respect to any computer system in its custody that is used to access the Data.
 - B. If requested, SAGE shall report to the Data Producer the results of any such investigations and the steps taken to address any remaining issues or concerns about the security of the Data or computer systems, or the privacy of individuals to whom the Data relates.
- 8. Notice of Breach.** SAGE shall advise the Data Producer immediately of any circumstances, incidents or events which to its knowledge have jeopardized or may in future jeopardize; (i) the privacy of individuals, (ii) the security of any computer system in its custody that is used to access the shared Data, and/or (iii) any suspected or apparent risk of a breach, or actual breach, of any term of this Agreement.
- 9. Termination for Unremedied Breach.** The Data Producer may give SAGE notice of breach of this Agreement and, if the breach has not been addressed to the Data Producer's reasonable satisfaction (including by retrieving research information that has been used or disclosed contrary to this Agreement) within thirty (30) days after the notice is given, the Data Producer may give notice of termination of this Agreement, which becomes effective fifteen (15) days after it is given. Within thirty (30) days after termination of this Agreement SAGE must return all copies of Data in any medium or must securely destroy all copies to the Data Producer's reasonable satisfaction, as the Data Producer directs in a notice given to SAGE. SAGE's obligations under this Agreement respecting use, disclosure and security of Data remain in force despite termination of the rest of this Agreement.
- 10. Non-Data.** Notwithstanding the definition of "Data" or the agreed Terms and Conditions of this Agreement, if the Data Producer transfers written confidential information concerning the Data along with the Data, then to the extent permitted by law, SAGE agrees to treat in confidence, any of Data Producer's said confidential information. SAGE's obligations of confidentiality under this Agreement shall not extend to any information:
- A. that can be demonstrated to have been publicly known at the time of disclosure; or
 - B. that can be demonstrated to have been in the possession of or that can be demonstrated to have been readily available to SAGE from another source prior to the disclosure; or

- C. that becomes part of the public domain or publicly known by publication or otherwise, not due to any unauthorized act by SAGE; or
- D. that can be demonstrated as independently developed or acquired by SAGE without reference to or reliance upon Data provided under this Agreement; or
- E. that is required to be disclosed by law, provided SAGE takes responsible and lawful actions to avoid and/or minimize such disclosure.

- 11. Intellectual Property.** Existing intellectual property, inventions, and technology (collectively “IP”) that formulate and constitute SAGE are the sole property of SAGE. No other Party shall have any claims to rights in such existing IP. IP rights to the Data generated through the course of the Study and collected under this Agreement are the property of the Data Producer.
- 12. Acknowledgments.** SAGE agrees to require Secondary Data Users to acknowledge the contribution of the Data Producer in any and all oral and written presentations, disclosures, and publications resulting from any and all analyses of Data. Where appropriate, and in accordance with academic standards, this acknowledgement may be in the form of authorship.
- 13. Limitation of Liability and Indemnity.** SAGE assumes all liability for loss or damages arising from the use, storage or disposal of the Data and further agrees to indemnify, defend and hold harmless the Data Producer and Data Producer Institutions from all claims, actions and damages whatsoever, including legal fees, resulting from or in connection with the use, storage or disposal of the Data. The Data Producer shall in no event be liable for any loss of profits, be they direct, consequential, incidental, or special or other similar or like damages arising from any defect, error or failure to perform with respect to the Data, even if the Data Producer has been advised of the possibility of such damages.
- 14.** Any notice, direction or waiver which may be or is required to be given under this Agreement must be in writing to be effective and must be delivered or sent by fax transmission or e-mail using the address information set out below. Any notice, direction or waiver that is delivered is considered to have been given on the next business day after it is dispatched for delivery. Any notice, direction or waiver that is sent by fax transmission or e-mail is considered to have been given on the day it is sent if that day is a business day, but if that day is not a business day, it is considered to have been given on the next business day after the date it is sent. If a party changes its address, fax number, e-mail address or all of them, it must immediately give notice of the new address, fax number or e-mail address to the other party as provided in this section.

If to the Data Producer: ATTN: [Data Producer’s Name] [Institution name, if applicable] [Address] [City, Province] [Postal Code] Telephone: [Phone Number] Fax: [Fax] [E-mail]	Technical Notices: [Name] [Institution] [Address] [Phone Number] [Fax] [E-mail]
If to SAGE: ATTN: SAGE PolicyWise for Children & Families	

601, 9925 -109 Street Edmonton, AB T5K 2J8 Telephone: (780) 944-8630 Fax: (780) 944-8637 data@policywise.com	
--	--

- 15. Designated Representatives.** The parties agree the individuals signing this Agreement are their designated representatives for all purposes of this Agreement. A party may give notice to the other designating a new representative, and must do so promptly after its representative changes.
- 16. Binding on Successors.** This Agreement ensures to the benefit of and is binding upon the parties and their respective successors, contractors, trustees, administrators and receivers, despite any rule of law or equity to the contrary. If SAGE ceases to exist, the Data Producer reassumes total responsibility for the Data and further use of the Data must be approved by the Data Producer.
- 17. Whole Agreement.** This Agreement and its Exhibits, is the entire Agreement between the parties and it terminates and supersedes all previous communications, representations, warranties, covenants and Agreements, whether verbal or written, between the parties with respect to the subject matter of this Agreement.
- 18. Waiver.** Waiver of any breach of this Agreement must be expressed and in writing to be effective and a waiver of a particular default does not waive any other default.
- 19. Counterparts.** This Agreement may be executed in counterparts and when each of the parties to this Agreement has executed a counterpart, each of those counterparts is considered to be an original and all of those counterparts when taken together constitute one and the same Agreement.
- 20. Severance.** If any portion of this Agreement is held invalid by a court of competent jurisdiction, the invalid portion must be severed and neither the decision that it is invalid nor the severance affects the validity of the remainder of this Agreement.
- 21. Amendments.** This Agreement may only be amended by a document signed by the authorized representatives of the parties and no other purported amendment is effective.

As evidence of their Agreement to be bound by the above terms, the authorized signatories of the parties have executed and delivered this Data Deposit Agreement, to be effective as of [Date] (“Effective Date”)

Signed for and on behalf of [Data Producer’s Institution Name] by its duly authorized officer:

Name:
Title:
Date:

Read and Acknowledged by:

Name: [Data Producer’s Name]
Title:
Date:

Signed for and on behalf of SAGE Executive by its duly authorized officer:

Name:
Title:
Date:

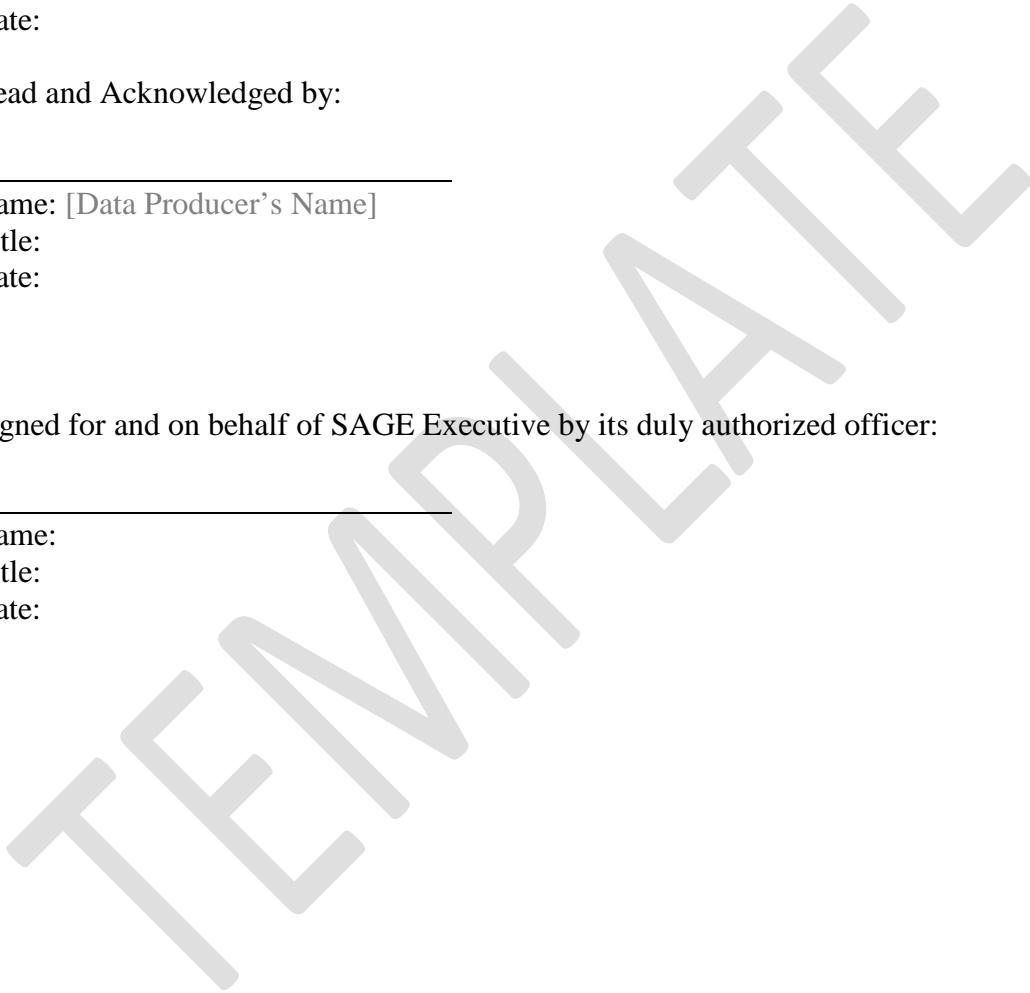


Exhibit A -

Data to be Deposited in SAGE

<u>Dataset Name</u>	<u>Dataset Description</u> <i>Please list time of collection, if applicable</i>
<i>i.e. CBIRTH</i>	<i>Child's Birth Record</i>

TEMPLATE

Exhibit B –

Study Description

TEMPLATE

Exhibit C-

Ethics Board Approval Certificates

Attach REB Approval Here.

TEMPLATE

Exhibit D-



SAGE Overview of Repository Services

What is Secondary Analysis to Generate Evidence?

SAGE is a data and research platform where research data, service delivery data, and administrative data related to child and youth development, health and well-being are stored, cleaned, catalogued and managed for secondary research and policy uses. SAGE will have highly secured computing systems and staff at PolicyWise for Children & Families (PolicyWise) in Edmonton and Calgary. PolicyWise will provide state-of-the-art technical infrastructure and governance processes for secondary use of data while ensuring privacy.

Outcomes for Alberta's children, youth, families and communities will be improved through using and combining data in *new ways* to inform social, educational and health policy, practice and service planning. PolicyWise will build and leverage research capacity, (including infrastructure and both data-related and child health and development-related content expertise) in an unprecedented way.

What are its Vision and Mission?

The VISION of SAGE is to *'maximize the potential of Alberta's children and youth'*. The MISSION of SAGE is *'to enable evidence-informed policy and practice through collaborative, optimal use of data'*.

Why is it important?

There are many benefits of sharing data, including that it:

- Encourages quicker generation of new knowledge for greater impact;
- reduces costs because similar data don't have to be collected repeatedly;
- enables richer answers to complex policy and research questions by combining data from multiple sources and of multiple types;
- promotes collaboration and cooperation across diverse disciplines;
- allows new questions on existing datasets and new dataset combinations;
- facilitates independent verification of findings;
- encourages quality data management practices;
- reduces unnecessary/duplicate participation of research participant or clients;

and perhaps most importantly; it -

- uses "to best effect the gift of data made by study participants"

Recently, the Social Sciences and Humanities Research Council of Canada (SSHRC) set a policy encouraging data sharing and allowing associated costs in research budgets.

Governance, Ethics and Privacy Protection

SAGE has a three-pronged approach to ethics and privacy protection – guided by CIHR best practices, Canada’s Tri-Council Policy Statement 2 on the ‘Ethical Conduct for Research Involving Humans’ and the relevant literature placed in context of relevant legislation in Alberta. The first prong is **governance**. There is a Strategic Advisory Committee made up of representatives of key stakeholder groups that use an adaptive governance model to develop fair and accountable access, deposit, ethics review, privacy protection, and engagement processes. In addition, an internal SAGE Team will oversee the management of a systematic and efficient access, release and follow-up process.

The second prong to the ethics and privacy protection approach is **formalized connections** to Ethics Board review processes and the Office of the Information and Privacy Commissioner of Alberta (OIPC) for advice and, when necessary, privacy impact assessments. Data requiring the greatest level of protection will be analyzed on-site only. Those at a second level may be released to qualified researcher/policy analyst teams or analyzed by remote access after completion of an approval process (including ethical approval) under the terms of a formal Data Deposit Agreement. Datasets that can be completely anonymized will be more readily available. In all cases, only the minimum amount of data will be provided to the secondary user to meet the objectives of the research.

The third prong to the ethics and privacy protection approach is **technology**. Staff will have special training, and state of the art identity protection and de-identification software will be used. The approach to physical data security will integrate learnings from the Child and Youth Data Laboratory (CYDL) model which follows a strict code of practice for information security.

Functions and Operations

SAGE operates according to the following principles:

- Best practices for research repository operations internationally
- Best practices for research ethics in Canada
- Principles for data stewardship (Research Data Canada)
- Best practices for governance for biobanks
- A commitment to consult with and involve all stakeholders e.g., Research Ethics Boards, OIPC of Alberta, research participants, and the public in implementation and governance.

The services SAGE provides include core supports for data users that are known best practices for research repositories such as access processes, dataset preparation, metadata development, as well as value added services such training, and knowledge mobilization. Researchers and policy analysts with specific interests will be able to search for information on-line about the data holdings (Metadata).

Revision History

Version	Date	Produced by	Change
Jan 26	January 2016	Amanda Lau	Draft items to be included in Manual
V 1.0	June 16 2016	Amanda Lau	Draft with full structure of Manual
V 1.1	June 24 2016	Amanda Lau	Major revisions
V 1.2	June 30 2016	Lucie Richard	Formatting and minor revisions
V 1.3	October 2 2017	Hannah Lloyd-Jones	Minor revisions and changes to review process
V 1.4	November 23 2017	Hannah Lloyd-Jones	Minor revisions following SAGE Implementation Meeting discussion

TEMPLATE

Contents

Revision History	15
Section I: Introduction	17
Overview	17
Why access data through SAGE?	17
Section II: Data Access Process	18
Figure 1- Procedure overview	18
Step 1: Secondary data user submits Data Inquiry Form	19
1.1. Eligibility	19
Step 2: SAGE staff consults with Secondary data user	19
Step 3: Secondary user data submits Data Access Request form	19
Step 4: Data Access Request Review	20
4.1 Review Criteria	20
4.2 Review timeline	20
4.3 Review outcome	20
Step 5: SAGE staff provides a preliminary cost estimate	21
Step 6: Secondary data user applies for ethical approval from a recognized research ethics board	21
Step 7: SAGE staff and Secondary Data User sign Data Access Agreement	21
7.1 Intellectual property	21
7.2 Repatriation of derived variables	21
7.3 Privacy of Participants	21
Step 9: Secondary data user completes analysis by accessing customized de-identified dataset, analytic tools, and research supports through SAGE'S Analytic Environment	22
8.1 Alternate methods of data access	22
8.2 Research Output	22
Step 10: SAGE evaluates research impact and tracks publications	22
Contact Information	23

Section I: Introduction

Overview

SAGE (Secondary Analysis to Generate Evidence) is a platform where research data, service delivery data, and (at a later stage) administrative data are catalogued and managed for secondary research and policy use. SAGE provides technical infrastructure and governance processes that protect participant/client privacy and ensure ethical re-use of data. Our goal is to improve outcomes for Alberta's children, youth, families and communities by enabling new uses of data to inform social, educational and health policy, practice and service planning.

Why access data through SAGE?

Millions of dollars are being invested in data collection in both the academic and public domain. According to a report from the Canadian Research Data Summit, many valuable research data are being “under-utilized and [...] important publicly funded asset[s] are being wasted”. While true for research data, this under-utilization can also be seen for data collected by service delivery organizations and within government sectors. Within these sectors, data is collected primarily for the provision of services – however, there is a wealth of knowledge to be discovered through its re-use.

Re-use of existing data beyond its original purpose has numerous benefits. These include:

- Replication and verification of results, enabling greater transparency
- Increased collaborative opportunities with the original data collector
- In instances where data linkage is possible, datasets from multiple sources can be combined to answer more complex question and potentially engage more interdisciplinary expertise
- Reduced burden on client/participant by avoiding data collection duplication
- Increased return on investment

SAGE provides a number of value added services that optimizes the use of existing data. For one, SAGE evaluates and enhances data quality in partnership with the original data collector to increase re-use potential, which can be beneficial for secondary data users. SAGE also ensures that datasets have comprehensive documentation and metadata (data about data), so that the studies and their variables are easy to understand, thus allowing for easier and appropriate re-use. Second, SAGE is leveraging PolicyWise' expertise in secondary research, data linkage, and analytics, lending advice to secondary data users where possible. Most importantly, SAGE does not only build bridges to new research, but facilitates the building of partnerships – be it amongst researchers, policy makers, or community practitioners.

Section II: Data Access Process

SAGE works with Secondary Data Users to ensure easy and feasible data re-use. This document describes the processes and procedures required for accessing data from SAGE.

Access to SAGE data is dependent on, but separate from, funding and ethics. Prior to data release, applicants are required to have ethical approval and encouraged to have funding in place.

Figure 1- Procedure overview



Step 1: Secondary data user submits Data Inquiry Form

Secondary Data Users complete the online [Data Inquiry form](#) stating the purpose and potential impact of their research.

1.1. Eligibility

Secondary Data Users should:

- Demonstrate that they or their research team (if applicable) have the content and analytic skills, knowledge, and resources to carry out the proposed project.
- Be able to obtain ethics approval from a recognized ethical review board.
- Propose a project that is for non-commercial purposes.
- Must abide by the “Terms of Use” for the dataset(s) of interest as found on [SAGE’s Dataverse](#)

Metadata about SAGE data assets is available on [SAGE’s Dataverse](#) to help guide proposal/hypothesis/aim development. Within the each dataset’s metadata, Secondary Data Users are encouraged to pay particular attention to the “Embargo” and “Terms of Use” which outlines availability of variables and access procedures, respectively.

Step 2: SAGE staff consults with Secondary data user

After review of the inquiry, SAGE staff will contact the Secondary Data User to discuss the feasibility of the project. Where requested, SAGE staff can support datasets exploration, identification of relevant variables, ethic submissions, as well as provide methodological and statistical advice.

An important goal of SAGE is to facilitate collaboration and to support research and data capacity building. During the consultation stage, SAGE will discuss potential collaborative opportunities between the Secondary Data User and the original data producer (i.e. researcher or service delivery organization). In some instances, collaboration with the original research team is required at various stages of the project.

SAGE will also discuss any dataset specific terms of use, as stipulated by the data producer (for example, form of data producer acknowledgement). SAGE staff will also ensure that the Secondary Data User understands that access to the data will be via the secure Analytic Environment.

Step 3: Secondary user data submits Data Access Request form

After consultation, SAGE staff will send the Secondary Data Users a Data Access Request form. The Secondary Data User will complete and submit the Data Access Request online. The Data Access Request is akin to completing a grant funding application. This request clearly states research aims and hypotheses, as well as describes the relevant exposure, outcome, and confounders to justify the data being requested. A list of study variables will also be submitted as part of the Data Access Request. A budget, as well as applicant and co-applicants academic CVs will be

submitted. We recommend that the Secondary Data User work with their co-applicants (if appropriate) to review the Data Access Request before submitting it. Note that if the scope of your research changes, you need more variables or if your research team changes from that stated on the Data Access Request form, we may need to review your request again and sign a new Data Access Agreement.

Step 4: Data Access Request Review

All Data Access Requests are reviewed by SAGE staff and the original Data Producer (or representative) if they have indicated that they wish to participate in Data Access Request reviews. In some cases, one or more external peer reviewer will also be asked to review the request.

SAGE staff review the request for feasibility and appropriate use of data. The Data Producer and the external reviewer(s) (if appropriate) ensure that the proposed research is relevant, sound, and consistent with the data being requested. Note that postgraduate research students need to submit a copy of their committee's thesis approval as part of their Data Access Request.

4.1 External Review Criteria

If an external review is required, each reviewer evaluates and provides comments for the research proposal based on the following criteria:

- Overall Impression
- Background, objectives, potential contribution to knowledge
- Approach and Methods
- Impact and Knowledge Translation
- Project Management, Research Team, and Partnerships

Additional criteria (does not count towards overall score)

- Budget

The detailed reviewed criteria can be found in the *SAGE Access Review Criteria* document.

4.2 Review timeline

Review processes take approximately 4-6 weeks.

4.3 Review outcome

After evaluation, SAGE staff will inform the Secondary Data User of the review outcome, as well as provide a copy of comments/suggestions offered by the reviewers. Unsuccessful applicants may revise and resubmit proposals. These Secondary Data Users are advised to speak to SAGE staff before resubmission.

Successful applicants and the title of their proposed research may be posted on PolicyWise' website.

Step 5: SAGE staff provides a preliminary cost estimate

Successful applicants may be charged an access fee based on a cost-recovery model (see *SAGE cost recovery guidelines*). These fees are determined on a project-to-project basis, and are required to reimburse operational costs such as processing of data requests and use, as well as, data preparation and extraction. This fee may be waived in certain situations such as undergraduate research project or for service delivery organizations.

Step 6: Secondary data user applies for ethical approval from a recognized research ethics board

Prior to data release, secondary research proposals need to have ethical approval from a recognized ethics board (i.e. institutional or community). Approval can be sought during or after the review process.

Step 7: SAGE staff and Secondary Data User sign Data Access Agreement

SAGE and Secondary Data Users sign a Data Access Agreement that outlines the terms and conditions for provision of services and data.

Some of the terms and conditions outlined in Data Access Agreement include:

7.1 Intellectual property

SAGE does not claim any intellectual property rights with regards to the results, discoveries, invention, or work that could stem from research where the data was used.

7.2 Repatriation of derived variables

Derived variables are variables constructed by the Secondary Data User during the course of a proposed secondary research project. These variables may include new data (where applicable) or derived data created from existing data.

To enrich the SAGE research platform, Secondary Data Users are required to return the results of the secondary analysis (i.e. derived data or new data), along with their appropriate documentation, back to SAGE. The exact nature of the results to be returned will be determined in collaboration between the Secondary Data User and SAGE. The derived variables will be incorporated into SAGE data assets for use by other Secondary Data Users. This will allow researchers to increase the value of data assets and build upon previous research.

7.3 Privacy of Participants

Approved Secondary Data Users agree to use and manage data from SAGE with strict confidentiality. Secondary Data Users agree that data will not be used to establish the individual identities of any of the participants.

The following documents should also be included with the Data Access Agreement: secondary data research proposal, list of requested variables, list of research team, and a copy of the ethic approval certificate. Once the data access agreement is

complete a request for new users is submitted to the SAGE IT manager. Also appropriate software is requested and deployed on the SAGE environment.

Step 8: All members of research team attend a SAGE orientation session and sign Code of Conduct Acknowledgement.

All new users attend a SAGE orientation session, before signing a Code of Conduct Acknowledgement and obtaining access to the SAGE Analytic Environment.

Step 9: Secondary Data User completes analysis in the SAGE Analytic Environment

The Secondary Data User will access their customized, de-identified dataset through the SAGE Analytic Environment. This Analytic Environment is a secure remote computer system that acts as a personalised access point to data and analytic tools (SPSS, STATA, SAS, Tableau, etc.) so that Secondary Data Users can store and work on their datasets. Since the environment is moated, there is no internet access and data cannot be removed from the environment. This type of environment limits the risk of privacy breaches, while facilitating multidisciplinary, multi-organizational collaboration.

In rare instances, Secondary Data Users may be required to access data via a physical enclave or secure file transfer protocol (SFTP). These methods will depend on the sensitivity of the data, and are up to the discretion of SAGE.

8.1 Alternate methods of data access

- Physical enclave
Very sensitive data will require access through a physical enclave. Enclave spaces are located in PolicyWise offices in Edmonton and Calgary.
- Secure file transfer protocol
A data file can be securely and electronically transferred to the Secondary Data User. Secondary data users will be assigned a username and password when they first access the dataset.

8.2 Research Output

As part of the process to safeguard the privacy of research participants during secondary use of data assets, SAGE staff review (or “vet”) all research outputs to be shared externally, including the following permitted outputs:

- Syntax documents (programs and/or logs of programs);
- Descriptive or Results tables;
- Figures

Please note that SAGE does not vet the accuracy of programs or outputs.

Step 10: Evaluation of research impact and tracking of publications

SAGE staff should be informed when research is complete, and access to data is no longer required. The period of data access is study specific and described in the Data

Access Agreement. However, this period may be extended by the Secondary Data User through written request.

All research output (i.e. manuscripts, presentations, abstracts, reports) should be submitted to SAGE prior to dissemination for information purposes only. Information obtained from these documents will not be used or disclosed without the written consent of the authors.

In efforts to highlight the benefits of data sharing, SAGE encourages publication and impact tracking. Upon project completion, Secondary Data Users are asked to provide a Final Research Report bi-annually, for up to three years. This report aims to summarize the research output and knowledge mobilization resulting from the secondary research project. A template of the Final Report can be found on the PolicyWise website.

Contact Information

If you have any questions regarding SAGE's Access Manual, please contact SAGE staff at data@policywise.com.

TEMPLATE

Exhibit F –

Special Considerations for Data

A. Data Considerations

Considerations	Details/Descriptions
Embargoes	<i>i.e. 6 months after first publication</i>
Accessor conditions	<i>i.e. Researchers, non-profits only</i>
Acknowledgement	
Security conditions	<i>i.e. physical enclave only</i>
Special Vetting requirements for research outputs	
Other (please specify)	

B. Level of Data Producer Involvement

Task	Description of Data Producer Involvement
Proposal Development	
Proposal Review (i.e. Access Review Roster)	
Analysis and/or Interpretation	
Research output (manuscripts, reports, etc.) development	
Dissemination	
Other (please specify)	