BRITISH COLUMBIA | Ministry of Health

Document No. DMS004

Healthideas Data Warehouse Standards for Third Party Users

Version 1.01 – May 6, 2019

Data Management and Stewardship Branch (DMS)

Health Sector Information, Analysis and Reporting Division

Document metadata

Revisions	**Date**	**Revisions**	**Author**
Version 1.0	April 2018	Version 1.0	DMS
Version 1.01	May 6, 2019	Reviewed; links updated; no other amendments	DMS
Contact/author	healthideas@gov.bc.ca		
Approved by	Andrew Elderfield, Executive Director, DMS		
Date approved	May 6, 2019		
Date effective	May 6, 2019		
Next review date	May 6, 2020		
ARCS/ORCS number	470-00		
Security Classification	Public		

Contents

I Document purpose....4
II Scope....4
III Responsibility....4
IV Introduction to Health*ideas*....5
V Definitions....5
VI Information classification....7
1.0 Standards....8
1.1 Confidentiality....8
1.2 Access control....8
1.3 Data usage and analysis....9
1.4 Transferring data or files into or out of the SAE....10
1.5 External data storage and transmission....10
2.0 Any questions / comments?....11
Appendix A: References....12

I Document purpose

The purpose of this document is to establish standards for Third Party users to ensure the protection of sensitive and personal information stored in, and accessed from, the Health*ideas* data warehouse environment.

II Scope

The standards in this document apply to Third Party users of the Health*ideas* data warehouse. The term "Third Party" is defined as:

- any person, group of persons, or organization, with which the Ministry of Health has an information-sharing agreement (regardless of the form of the agreement, for example, information-sharing agreement, information-sharing plan, research agreement, memorandum of understanding, common or integrated program agreement, etc.)

Contractors, subcontractors, or other agents of a Third Party are treated as part of that same Third Party.

Organizations that have been contracted to provide services to the Ministry of Health ("the Ministry") are out of scope for the purposes of this document.

III Responsibility

It is the responsibility of all Third Party users of Health*ideas* ("users") to read and ensure they comply with these standards. **Non-compliance with any of these standards may result in a user's access being suspended, either temporarily or on a permanent basis.**

The Ministry's Data Management and Stewardship (DMS) branch is responsible for the collection, storage, availability, use and access to Ministry of Health data sets for research and analysis (i.e. secondary use). This responsibility includes a requirement to provide reliable and secure services related to the operation of the Health*ideas* data warehouse.

Provincial legislation and policy require the Ministry to protect the confidentiality, integrity and availability of the Health*ideas* data sets and related information products. **In accordance with information-sharing agreement(s) with the Ministry, it is the responsibility of a Third Party to implement and maintain necessary security measures of their own to protect Health*ideas* data.**

This document builds upon the BC Government-wide Information Security Policy (ISP) and standards of the Office of the Chief Information Officer (OCIO), the Core Policy and Procedures Manual (CPPM), and relevant Ministry policies and standards, to meet the Ministry's goal of protecting its information and technology assets.

IV Introduction to Health*ideas*

Health*ideas* is the Ministry of Health's principal data warehouse for secondary use. It contains a range of data sets and derived information products that support analysis and research. These include information about health services provided to British Columbians, including hospital and physician services, as well as population and other reference data.

The Secure Analysis Environment (SAE) is a security service used to protect sensitive and personal information in Health*ideas*.

For further information on the SAE, see the *Secure Analysis Environment User Guide for Third Parties*.

V Definitions

"Aggregate Data" — data that has been compiled from record-level data to a level of aggregation that ensures that the identity of individuals cannot be determined by reasonably foreseeable methods.

"Anonymized-ID"— a de-identified type linkage key that is a replacement for a PHN type linkage key.

"Contact Information" — information to enable an individual or a place of business to be contacted and includes the name, position name or title, telephone number, address, email or fax number. This is a type of Direct Client Identifier.

"Crosswalk" — a file that enables data linking by providing a mapping of equivalent elements of two database tables. In the context of Health*ideas*, a crosswalk which maps PHNs and anonymized client labels allows a user to circumvent database security controls and invalidate the anonymization.

"De-identified Data"— personal information that has been modified using appropriate de-identification processes, so that the identity of the individual cannot be determined by a reasonably foreseeable method. De-identified Data does not contain Health*ideas* Direct Client Identifiers, and Anonymized-ID are provided instead. In Health*ideas* Level 1 and Level 2 the following Direct Client Identifiers have been removed:

- Client name, address or phone number
- Full postal codes
- Full birth dates (or Age in Days) or full death dates
- Registration and Premium Billing (R&PB) contract number
- Hospital chart number
- Pharmacare Family ID
- PHN

"Direct Client Identifier" — an identifier such as a full PHN, SIN, chart number or a high-risk client re-identification attribute, such as birth and death date, and postal code.

"Identifiable Data" — data that contains Direct Client Identifiers.

"Indirect Identifiers" – data that may identify an individual when they are connected with other pieces of information to single out an individual (e.g. age, gender, date of visit, demographics). While indirect

identifiers on their own may not be personal information, they are considered personal information if they can be combined together to identify an individual. This is commonly referred to as the mosaic effect.

"Information Product" — a data set that has been derived from source data domains.

"Level 0", "Level 1", "Level 2", "Level 3" and "Level 4" — security levels in Health*ideas* [see section VI for detailed definitions]

"Personal Information" — recorded information about an identifiable individual other than Contact Information.

"Practitioner" — a "health care practitioner" as defined in section 1 of the *Medicare Protection Act*.

"Secure Analysis Environment" ("SAE") — a virtual desktop which it allows analysis to be performed in a controlled secure environment. Data movement is restricted in and out.

"Sensitive Information" — information that if compromised could result in serious consequences for individuals, organizations or government.

"Team Schema" — a database structure which allow the sharing of data within the data warehouse between database users. Team schemas can be created using three different sources: secure LAN, Oracle database and SAS.

Note that specialized roles must be set up for accessing and creating team schemas. When a user has access to a team schema, that user has access to any data placed within the team schema. As a consequence, access approval to team schemas is strictly controlled. For more information please contact healthideas@gov.bc.ca.

"Third Party" — any person, group of persons or organization with which the public body (i.e. Ministry of Health) has an information-sharing agreement (regardless of the form of the agreement, for example, information-sharing agreement, information-sharing plan, research agreement, memorandum of understanding, common or integrated program agreement, etc.)

VI Information classification

Information in Health*ideas* is classified as follows:

Health*ideas* data warehouse security level	Definition	Description	Data type	OCIO Classification
4	▪ Contact information direct client identifiers[1] are visible; ▪ Non-contact direct client identifiers[2] are visible; and ▪ Practitioner names are visible.	Row-level client-identifiable data	Sensitive information and/or personal information (SI/PI)	HIGH
3	▪ Contact information direct client identifiers are NOT visible; ▪ Non-contact direct client identifiers are visible; and ▪ Practitioner names are visible.			
2	▪ Contact information is NOT visible; ▪ Direct client identifiers are NOT visible, dates and postal code de-identified, anonymized-IDs assigned; and ▪ Practitioner names are visible	Row-level client de-identified data		
1	▪ Contact information is NOT visible; ▪ Direct client identifiers are NOT visible, dates and postal codes are de-identified and anonymized IDs are assigned; and ▪ Practitioner names are NOT visible			
0	▪ Aggregate patient data, inspected and approved for distribution	Aggregate data	Anonymized information or Open Data	PUBLIC

[1] "Contact information direct client identifiers" include, for example, the following: person's name, full address, phone number, and e-mail address.

[2] "Non-contact direct and indirect client identifiers" include, for example, the following: full PHNs; SINs; chart numbers; and high-risk client re-identification attributes, such as birth and death date or postal code.

1.0 Standards

1.1 Confidentiality

1.1.1 All information sourced from Health*ideas* must be treated as strictly confidential.

Users may only access, use, transfer, disclose, publish, or release data sourced from Health*ideas* (or permit information sourced from Health*ideas* to be accessed, used, transferred, disclosed, published, or released) as authorized by an information-sharing agreement with the Ministry and/or as expressly authorized or required by law.

1.2 Access control

1.2.1 Access by a Third Party to data in Health*ideas* must be authorized by an information-sharing agreement.

An information-sharing agreement (regardless of the form of the agreement, for example, information-sharing agreement, information sharing plan, research agreement, memorandum of understanding, common or integrated program agreement, etc.) defines the purpose of data access and usage by the Third Party, and must be signed and authorized by all relevant parties prior to access being granted.

1.2.2 All Third Party users may only access sensitive and/or personal information in Health*ideas* through the Secure Analysis Environment (SAE).

This standard exists to ensure data analysis is performed in a secure and protected environment, and that all activities are controlled and monitored. Note this applies to all levels of data access, from 1 to 4.

1.2.3 Unless there is a proven need to access Identifiable Data, which is authorized in an information-sharing agreement, Third Party users may only access De-identified Data.

The Ministry standard for providing access is by default De-identified Data (Level 1 and/or Level 2), unless otherwise explicitly authorized in the relevant information-sharing agreement. This standard exists to limit access to Identifiable Data to only those users that have a legitimate business requirement to access such data.

1.2.4 Users requiring access to both De-identified and Identifiable Data will have separate user ID accounts to access De-identified and Identifiable Data.

If a user requires access to both De-identified Data (Level 1 or 2) and Identifiable Data (Level 3 or 4), he/she must use two different user IDs, one approved for access to De-identified Data (Level 1 or 2), and the other approved for access to Identifiable Data (Level 3 or 4). This standard exists to prevent the creation of unauthorized crosswalks. (See the related standard 1.3.2).

Note that if a user has been granted multiple SAE accounts, use of each account must be in accordance with its specified purpose.

1.2.5 Access to a Team Schema must be approved by the owner of that Team Schema.

The Ministry's approval process for accessing a Team Schema ensures that authorization is granted by the owner of the Team Schema, who ensures the validity of the request. Once granted access, users may only use the data within a Team Schema for the explicit purpose(s) as outlined in the relevant information-sharing agreement.

1.2.6 It is the responsibility of a Third Party to notify the Ministry when a Third Party user from their organization no longer requires access to Health*ideas.*

If a Third Party user no longer requires access to Health*ideas*, the Third Party must, without delay, submit a 7076 form for such removal of access to the Ministry's DMS Branch.

1.3 Data usage and analysis

1.3.1 Third Party SAE accounts must only be used for the purpose(s) for which they are specifically granted.

Third Party access to Health*ideas* is limited to what is specified in the relevant information-sharing agreement(s); users must perform their analysis in accordance with the relevant agreement(s). This standard exists is to prevent the creation of unauthorized work products.

1.3.2 Creation and usage of crosswalks is prohibited.

When Information Products containing Direct Client Identifiers are linked with data that has an Anonymized-ID linkage key, such as in the creation of tables including both Anonymized-IDs and PHNs, this creates a crosswalk and introduces the risk of re-identification. In order to protect personal and sensitive information, such linkage tables are prohibited.

For assistance or questions regarding table linkages, please contact healthideas@gov.bc.ca.

1.3.3 Third Party users must perform all analysis work within the SAE.

In order to prevent unnecessary transfer and/or outside storage of sensitive and/or personal information, all analysis must be completed within the SAE using the available tools.

1.4 Transferring data or files into or out of the SAE

1.4.1	Extracting, importing and/or reproducing any data from the SAE by methods other than the Managed Transfer Process (MTP) is prohibited.

The MTP is the only allowed method for data import and export from the SAE. Any other method, such as e-mail, screen-scraping, photography, or manually writing down data for future use, is prohibited. Such methods increase the risk of data misuse and do not meet an acceptable standard of data protection.

1.4.2	Third Party use of the Managed Transfer Process (MTP) to transfer files into or out of the SAE must be authorized by an information-sharing agreement and must also be defined and authorized in an MTP Plan.

MTP Plans are drafted in accordance with the associated information-sharing agreement, and identify staff at the Third Party organization or the Ministry who will be responsible for authorizing and completing transfers of files into and out of the SAE.

Third Parties should take note that use of the MTP is limited to the least number of staff possible. All transfers of files into and out of the SAE are overseen by the relevant MTP approver(s) and the Ministry.

For further information or assistance with the MTP process, please write to healthideas@gov.bc.ca.

1.4.3	Only final work products may be exported from the SAE.

Unless explicitly authorized within the Third Party's information-sharing agreement with the Ministry, and described in the Third Party's MTP Plan, only files which are final work products may be exported.

1.5 External data storage and transmission

Note: this standard applies to data transferred to a Third Party's infrastructure.

1.5.1	Third Parties must have established data retention and disposal processes in line with the relevant information-sharing agreement(s) and BC Government policies and standards.

If a Third Party has a requirement to store sensitive and/or personal information it has exported from the SAE (i.e. outside of the SAE Secure LAN) on its own infrastructure, it must have established processes in place to ensure the data is protected and to minimize privacy and security risks.

For the BC Government's Information Security Policy, see: Office of the Chief Information Officer IM/IT Standards and the Information Security Policy. (See hyperlinks in Appendix A).

1.5.2	After data is transferred outside the SAE, any subsequent data transfers containing sensitive and/or personal information from Health*ideas* must use an encryption method.

Transmission of unencrypted sensitive and/or personal information increases privacy and security risks and is therefore prohibited. For further information, see *Cryptographic Standards for Information Protection*.

Users should also not use USB flash drives to transfer sensitive and/or personal information exported from the SAE. This is in accordance with recommendation 1 of the Office of the Information and Privacy Commissioner for BC (OIPC) report F13-02 which states: "The Ministry should develop and implement additions to the BC Government policy on the use of portable storage devices to require the use of other, more secure, forms of information transfer. Portable storage devices should only be used as a last resort and must always be encrypted."[1]

2.0 Any questions / comments?

This is the first version of this document, and it will be updated annually and as required. Comments and suggestions are welcome, and may be used to inform future versions. For clarification and assistance with this document, or to provide feedback, please email: healthideas@gov.bc.ca.

Appendix A: References

Core Policy and Procedures Manual - Chapter 12 IM/IT Management

Cryptographic Standards for Information Protection

Information Security Policy

Office of the Chief Information Officer IM/IT Standards

Office of the Information and Privacy Commissioner for BC (OIPC), *Investigation Report F13-02*

Secure Analysis Environment User Guide for Third Parties is available at the following link: *https://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/data-access-health-data-central/ministry-of-health*