

CPCSSN Operating Procedures

SOP number:	CPCSSN-SOP-Data Access Processes_001
Version number:	2.00
Category	Data Access
Effective Date:	August 31 st , 2021

References

References
No References

Revision History

Version	Description of Change	Author	Date
NA	NA	NA	NA

Approvals

	Name, Surname, Title	Signature	Date
Authored by	Ayat Salman		
Reviewed by	Stephanie Garies		May 21 2021
Reviewed by	Sabrina Wong		May 20 2021
Reviewed by	Alex Singer		July 11 2021
Reviewed by	Michelle Griever		May 14 2021
Approved by	Steering Committee		August 31 st , 2021

Disclosure (Internal, External)	Internal, External
Confidentiality (Strictly, Confidential, Public)	Confidential
Status (Draft, Under review, Final)	Final

TABLE OF CONTENTS: CPCSSN Data Access Procedures

Objective	3
CPCSSN Data Repository	3
CPCSSN Access Policy	4
Objective of the Access Procedures	4
Communication	5
Access charges	5
No preferential or exclusive access	6
Data available in the CPCSSN CPCSSN Repository	8
Legal and ethics approval.....	8
Intellectual Property Rights	8
Application Review Stages	9
.....	11
Reconsideration of DARs by the Data Access Committee	11
SRE	12
SRE User Agreement.....	12
Provision of Data	13
Role of parties involved in the access review	13
<i>APPENDIX I</i>	<i>14</i>
<i>APPENDIX II.....</i>	<i>22</i>
<i>APPENDIX III</i>	<i>60</i>
<i>APPENDIX IV.....</i>	<i>61</i>

Objective

The overriding objective of these Access Procedures is to encourage the extensive and appropriate use of the CPCSSN Repository. In particular, the intention is that the process for applying to use the CPCSSN Repository is simple and streamlined, without undue delays in providing access to data and samples for research.

The Access Procedures reflect the Access Policy which is outlined in CPCSSN's Ethics & Governance Framework. The procedures and processes that have been applied to access to the CPCSSN Repository derive from the following key principles:

- The CPCSSN Repository is available to all researchers for all types of health-related research that is in the public interest, without preferential or exclusive access for any person. All researchers, whether in universities, charities, government agencies or commercial companies, and whether based in the Canada or abroad, will be subject to the same application process and approval criteria.
- Applications to use the CPCSSN Repository will be checked to ensure that research proposals are consistent with these Access Procedures, the Ethics & Governance Framework, and the consent that was provided by the participating providers (including having relevant scientific and ethics approval).
- Safeguards will be maintained to help ensure the anonymity and confidentiality of participants' data and samples. Researchers will enter a legal agreement with CPCSSN not to make any attempt to identify participants, and data provided to researchers from the CPCSSN Repository will not identify any particular participant (they will be "coded").
- Researchers will have to pay for access to the CPCSSN Repository on a cost-recovery basis for their proposed research, with a fixed charge for initiating the application review process and a variable charge depending on how many samples, tests and/or data are required for the research project.
- CPCSSN will remain the owner of the repository but will have no claim over any inventions that are developed by researchers using the CPCSSN Repository (unless they are used to restrict health-related research, access to healthcare unreasonably or are related to processes associated to CPCSSN operations).

CPCSSN Repository

Established in 2008 through a partnership between the College of Family Physicians of Canada and Public Health Agency of Canada, the Canadian Primary Care Sentinel Surveillance Network (CPCSSN) is a network of 13 practice-based research and learning networks in 8 provinces and 1 territory. Currently, over 2 million patients and 1500 primary care clinicians participate in CPCSSN.

CPCSSN receives de-identified patient data; these data are extracted by participating Networks (or groups of Networks) from the electronic medical records of consenting primary care practices across the country. The Networks apply CPCSSN cleaning, coding and standardization algorithms to transform raw EMR data into meaningful information and forward these data to CPCSSN to form a national CPCSSN Repository that can be used for quality improvement, chronic disease surveillance, and research.

The rationale for recruitment of a cohort of such large size was to allow reliable quantification of the relevance of a large number of risk factors (e.g. lifestyle, environment and genes), both separately and in combination, to a wide range of diseases developing during follow-up.

CPCSSN Access Policy

CPCSSN has established a detailed Ethics & Governance Framework (EGF) which has been drafted and adopted. The EGF sets out the governance structure for CPCSSN and provides the framework that CPCSSN operates within (including the Policy governing access to the CPCSSN Repository).

CPCSSN's purpose is described in the EGF. Researchers who apply to use the Repository will be required to explain explicitly how their research project supports this stated purpose.

Objective of the Access Procedures

The objective of these Access Procedures is to facilitate access to data so that they get the widest possible usage while ensuring that such access and usage is consistent with the undertaking of CPCSSN.

At all times, decisions to grant access should maintain the undertakings made by each participating Network to the physicians when they consented to take part: (i) to help ensure that any uses of the CPCSSN Repository are consistent with its stated aims; (ii) to protect participants' anonymity and confidentiality; (iii) to ensure that research projects have relevant scientific and ethics approval, and (iv) to make information publicly available about the uses of the CPCSSN Repository.

It is intended that these Access Procedures are clear and transparent and are implemented in a manner which is proportionate, accountable and fair. The Access Procedures provide a framework for addressing and determining access issues. They deliberately do not prescribe what will be done in each and every circumstance because CPCSSN cannot (and would be unwise to try to) predict the nature of access requests that it will receive over the long-term. Built into the Access Procedures is sufficient flexibility to address both the expected and the unexpected, including the ability to revise them in the light of practical experience.

The functioning of these Access Procedures will be kept under review by the CPCSSN Central and the Data Access Committee. Input from members of the Steering Committee, the participants, researchers, funders, EGC and other interested parties will be taken into account. CPCSSN's Board will amend the Access Procedures as required on a periodic basis.

Communication

The EGF sets out the basis of the relationships that CPCSSN aims to foster with participants, researchers and the public at large. CPCSSN is committed to maintaining a dialogue with these different communities and will keep them updated on the progress of CPCSSN and the research work which is carried out using the CPCSSN Repository.

Access charges

CPCSSN has carefully considered a variety of charging models and has decided to adopt the following approach:

CPCSSN will follow the price schedule (Figure1) below:

Secure Research Environment (SRE) Access Fee	Fee Amount
SRE initial access	\$2,500 (3 users/1 year)
Additional user	\$500 per person
Switch users	\$100 per switch
Backup files recovery	\$500 per request
Project Data Fee	
Academic	\$5,000
Non-Academic	\$20,000
Administrative Fee	
Application review processing	\$2,000
Analytical Services	
Additional customized data processing	\$150 per hour depending on the expertise of the analyst
Data Analysis	\$150 per hour depending on the expertise of the data analyst

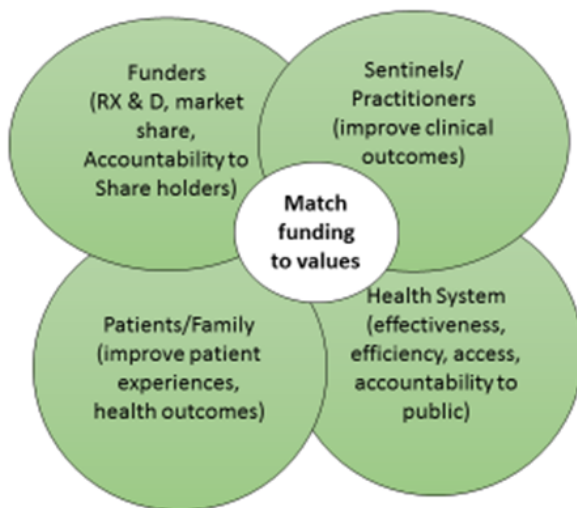
CPCSSN will keep this charging policy under review (including the possibility of different levels of charge for different types of researcher) to ensure that it continues to represent an equitable, balanced and pragmatic approach.

No preferential or exclusive access

CPCSSN data are a significant opportunity to generate new knowledge about the organization and delivery of primary healthcare with the overall goal to spur change that will make a difference in the lives of patients and their families across Canada. This decision-making framework will address the CPCSSN's values and goals.

Figure 1 shows the complex relationship between this network and its partners, particularly potential funders.

Improving the care and quality of life of people with multiple chronic conditions-- multimorbidity



CPCSSN's aim is to provide fair, consistent and transparent access to the CPCSSN Repository in order to promote health-related research by bona-fide researchers that is in the public interest. No distinction will be made between Applicants on the basis of whether they are:

- From the Canada or abroad; and/or
- CPCSSN members¹; and/or

¹ CPCSSN members are defined as: all CPCSSN staff, CPCSSN investigators, members of the Steering Committee, participating PBRNs and physician sentinel practices who are actively participating in CPCSSN

- Charitable, academic, governmental or commercial (industry) entities²

In addition, the following terms are defined as follow;

Primary Care Academics: Primary Care Academics include university academic researchers who are in or associated with Departments of Family Medicine.

Non-Profit Partners:

CPCSSN recognizes that there may be times where it is mutually beneficial to work with for-profit partners. There are three additional principles that the CPCSSN will use to guide working with for-profit partners:

- Alignment with the CPCSSN mission and values – CPCSSN private sector work must provide a public benefit;
- Transparency – full results reports, the identities of for-profit entities that use the CPCSSN data, and dataset creation plans will all be publicly available
- For-profit entity work undertaken with the CPCSSN must be in addition to, and in no way detracting from, CPCSSN research or quality improvement work – this includes the requirement that work with for-profit entities is financially viable

In the event of work with for-profit partners: (a) the research is conducted at arm's length from the partner; (b) no partner's employees, contractors, or agents are part of the project team; (c) the partner has no influence on study direction or analysis; and (d) the sponsor has no access to data, other than a final report. Apriori primary and secondary outcomes must be registered with the CPCSSN and that the CPCSSN will make the results of any for-profit analyses publicly available through such means as peer-reviewed publication, report or brief available on its website and upon request.

Requests for Data for Purposes Other Than Research: The Data Access Committee will examine the data being requested, will ensure the data request will not jeopardize the database, CPCSSN's principles of privacy, individual PBRN ethics approvals or CPCSSN's reputation.

CPCSSN will apply a standard set of criteria (subject to ongoing review and amendment by CPCSSN's Co-Chairs and Data Access Committee) to the assessment of all DARs, including the compatibility of the research project with the purposes of CPCSSN; the feasibility of the research project; and the facilities for managing data (Appendix I).

The level of scrutiny used to assess DARs will be proportionate to the nature and scale of the research project, and whether there are considered to be potentially contentious issues involved.

² Those partners with formal agreements in place with CPCSSN in the form of partnership agreements or memorandums of understanding will be exempt from this approval process and will be subject to the processes outlined in their signed agreements. Industry researchers and investigators can make requests for data from the CPCSSN central data repository. Data requests will be considered if the research will be made public when the research is concluded.

CPCSSN wishes to encourage collaboration between prospective researchers in order to increase the efficient use of the CPCSSN Repository.

There will be no restrictions on the number of researchers who can be provided with the same or overlapping data that are already contained within the CPCSSN Repository. This approach is intended to encourage rapid reporting of findings and different approaches to the analysis and interpretation of the data, as well as allowing researchers to confirm or refute published findings

Data available in the CPCSSN Repository

A CPCSSN Data Dictionary (Appendix II) will be available on the CPCSSN website for all data requestors. Requestors will have the opportunity to choose which variables to include within the data cut. They will have the opportunity to indicate the variable of interest in the DAR. Data within enriched fields may be requested but will need explanation in order to be reviewed by the Data Access Committee.

Legal and ethics approval

Research proposals accompanying the DAR must have received a Research Ethics Board approval prior for CPCSSN to provide or even review a DAR. In certain circumstance, CPCSSN can act as the science reviewer of the research proposal if it has not received Research Ethics Board Approval (REB) approval. At the request of the data request applicant, CPCSSN can provide a letter attesting scientific review.

CPCSSN requires REB approval for all DARs from their local Institutional Review Board (IRB).

CPCSSN requires equivalent commitments from all researchers in the SRE User Agreement.

Intellectual Property Rights

CPCSSN's approach to Intellectual Property Rights (IPR) is structured on the basis that it seeks to encourage use of the CPCSSN Repository for health-related purposes by researchers. To this end, CPCSSN will retain ownership of its rights in the CPCSSN Repository (so that it is available to all other approved researchers), while at the same time facilitating the development of clinical advances (e.g. diagnostics and treatments) arising from its use.

CPCSSN is the custodian of the database (which will be added to, and updated, throughout the life of the CPCSSN Repository) and retains all the intrinsic IPRs in the data in the CPCSSN Repository (notably database rights and copyright).

Researchers are granted limited licences (but not any ownership rights) to use the data to conduct the approved Research Project for a particular period of time. These rights are not assignable or transferable, and nor is there any ability to sub-license.

If a researcher creates separate datasets as a result of their use of the CPCSSN Repository, then IPRs in the researcher-generated datasets will be owned by the researchers and/or their institutions, subject to the requirement to grant a non-exclusive licence back to CPCSSN for its use on an irrevocable, perpetual, worldwide, fully paid-up, royalty-free, fully sub-licensable basis. These datasets will, therefore, be available for use by other researchers who are granted access to use the CPCSSN Repository.

CPCSSN will have no claim over inventions and associated IPRs that are developed by researchers as a result of using the CPCSSN Repository, unless such IPRs are used to restrict health-related research and/or access to healthcare unreasonably.

In the event that conduct is considered unreasonably restrictive by CPCSSN, it reserves the right to require that a licence of such rights is granted back to CPCSSN on an irrevocable, perpetual, global, royalty-free, fully sub-licensable basis so that other researchers who are granted access to use the CPCSSN Repository can exercise such rights to the extent necessary to conduct their research project.

CPCSSN will keep this policy on IPRs under review to ensure that it continues to represent an equitable, balanced and pragmatic approach.

Application Review Stages

The application process is entirely on-line via the CPCSSN website.

1. Data Request Application: To allow approved researchers to determine: (i) whether their proposed research use is likely to be approved; (ii) whether the CPCSSN Repository contains the data and/or samples required for their proposed research; and (iii) the indicative cost of obtaining such data and/or samples (e.g. in preparation for a funding application);
2. Review of the data request application: To allow CPCSSN to assess: (i) whether the proposed research use meets the required criteria for access (including having relevant scientific and ethics approval); (ii) whether the amount of depletable sample required is scientifically justified; and (iii) the cost of providing such data.
3. The DAR will be submitted to the Data Access Committee for review and approval.
4. REB approval: Applicants are asked to provide a local REB approval before research data is released to the researcher. This document must be sent with the DAR for review and approval.
5. SRE For approved applications, a User Agreement form will need to be executed and access charges paid before release of data to the approved researcher.

CPCSSN will produce a yearly report of all projects submitted, approved and not approved.

There are 4 stages to the application process (see schematic Figure 1):

Timelines: In order to facilitate access to the CPCSSN Repository, it is intended that each review stage will be conducted in accordance with an indicative timeline (although the need to seek further

information or guidance on particular applications may lead to a more prolonged process). The intended timelines for each stage are:

- Review DAR: 2 weeks to review and respond to requestor that the application is complete and will move to the review stage.
- Approval of the DAR: 4 weeks to review and respond
- Networks will continue to retain the ability to opt out or opt-in to projects. They will be notified by email that a project is currently under review and will have a two-week period to respond.
- if there is no identification of Network regions/sectors (general data across Canada), grants and publications may generally proceed after DAR Approval. There must be notification of all Networks involved a minimum of two weeks ahead of submission of grant or publication; they can opt-out if they so wish. Networks can be notified at the grant writing and publication writing time; this is encouraged to build capacity.
 - if there is identification of Network regions/sectors, opt-in by each Network involved is required for grants and publications
 - if there is industry funding, opt-in by each Network involved is required for grants and publications
- Access of Data through the SRE: Data will typically be available in the SRE within 2-3 weeks of approval.
- Once data analysis is complete, the analysis output will be reviewed by a CPCSSN member for data check before being transferred to the researcher outside the SRE.

The process from the start of the review DAR to release of data within the SRE may take up to about 2 months. There would then be a further period of time before data or samples (or related assays) would be ready for release.

The process of reviewing applications will be managed by the CPCSSN Central, which will conduct a set of standard checks and seek advice on particular applications as required. Recommendations and approvals on each application will be made by the Data Access Committee.

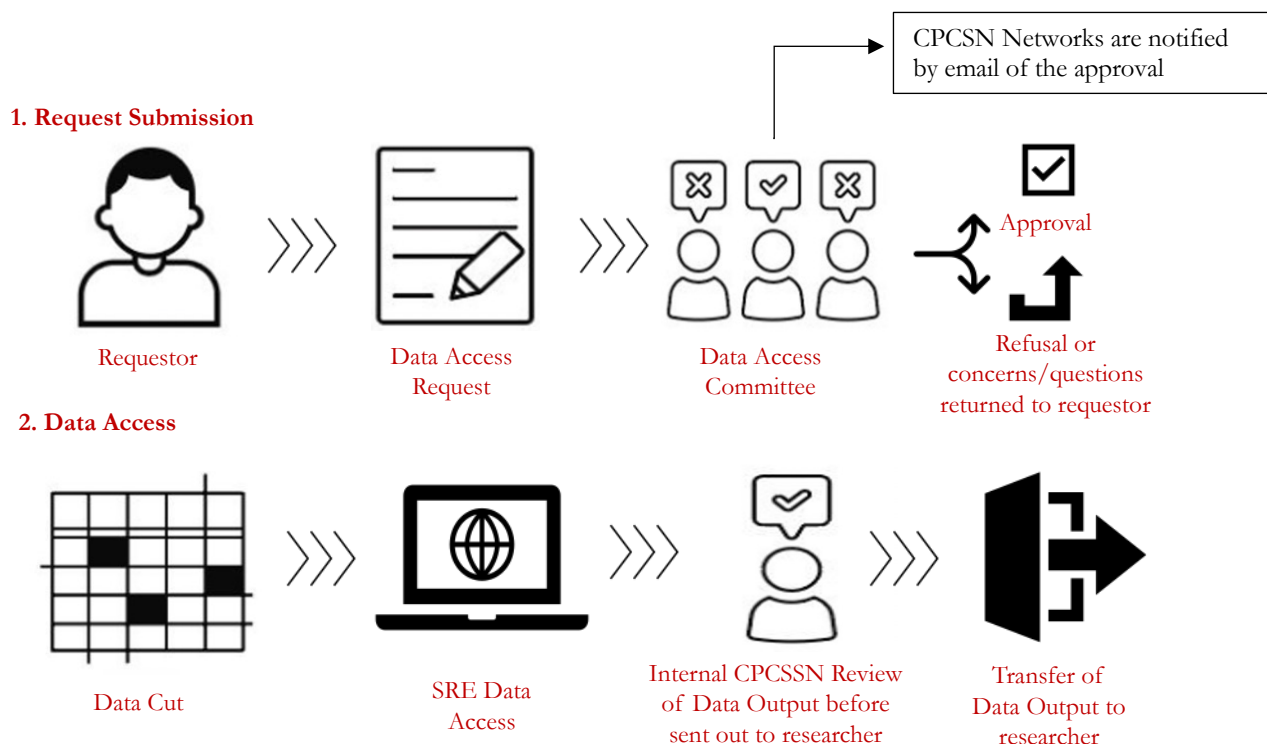


Figure 1: Data Request Review Schematic

Documents required at the DAR submission:

- Data Access Request Form
- SRE User Request Form
- User Agreement for each SRE User
- Research Protocol
- REB Approval
- Science Review (if available)

Incomplete applications will not be processed or reviewed.

Reconsideration of DARs by the Data Access Committee

If a requestor is advised that the DAC is minded to decline a Data Request Application, then they may request that the application be reconsidered by CPCSSN. See Annex III for evaluation criteria. The Data Access Committee, if an application is not approved, will provide feedback in the form of comments, suggestions and changes to the applicant. The applicant may resubmit a modified DRA for review to the CPCSSN Data Access Committee.

If a DAR is refused, the process for having a DAR reconsidered is as follows:

- Within 3 months of the relevant decision, the applicant PI should submit a written request, giving their reasons why they consider that the decision should be revised.
- Within 4-6 weeks of receipt of such a request, the Data Access Committee will aim to consider it along with the original application (and any other information that it considers pertinent) and CPCSSN will then respond to the applicant.
 - If considered necessary, the CPCSSN Data Access Committee may seek additional advice (e.g. from scientific or other experts, or the EGC), in which case the applicant PI will be advised by CPCSSN of any revision to the timetable for review.
 - If, following reconsideration, the application is declined then the applicant will not be able to submit the same proposal again within a 12-month period (unless the CPCSSN Data Access Committee has indicated specifically that it may be submitted subject to specific changes being made).

Note: Internal projects evaluating Data Quality Improvement (DQI) will not be reviewed at the Data Access Committee. These special projects will be review at the Quality Assurance and Privacy Committee. These projects, in certain circumstances, will not be accessed through the SRE and will be evaluate on a case per case basis.

SRE

The CPCSSN secure research environment (SRE) is a controlled and secured service environment for undertaking research using sensitive CPCSSN data. The CPCSSN SRE provides robust controls and safeguards to enable the secure transfer of sensitive data into a highly secure environment where it can be stored, manipulated and analysed by approved members of a research team.

Research teams, institutions, organisations can apply to use the SRE as a service. In order to do so, the following documents are required:

- CPCSSN SRE Request Form (CSR)
- SRE User Request Form
- User Agreement for each SRE User
- Research Protocol
- REB Approval (Via Queen's University or an approved IRB)

The package will be sent to CPCSSN Administrator who will process the request and submit it for review to CPCSSN Co-Chairs. An MOU must be executed before submitting an SRE application package.

SRE User Agreement

The Requestor will fill out a User Agreement that will be submitted with the DAR. In addition, an SRE User Request Form is also required during the DAR submission. Each SRE User will be required to sign an SRE User Agreement.

Provision of Data

Data from the CPCSSN Repository will be provided in the following manner:

- Data sets for analytical purposes will be made available solely through the CPCSSN Secure Research Environment (SRE) (Appendix III) so that the relevant data can be accessed only through the SRE. The data cannot be downloaded, copied or moved outside the SRE.
- The SRE provides researchers all the software and tools necessary to analyze the data.
- Analysis output developed in the SRE will be transferred outside the SRE when the researcher via email in the same format they have been built in the SRE.
- For SRE access and user Guide (Refer to CPCSSN_SOP_SRE_User Guide)

Role of parties involved in the access review

The roles of, and interactions between, the different parties involved in the DAR review and oversight process are described in Appendix IV.

APPENDIX I

CPCSSN Data Access Committee DAR Evaluation Criteria

When considering a data request the CPCSSN DATA ACCESS COMMITTEE will consider, among other things:

- The purpose of the research – will it advance the practice of primary health care?
- Has the researcher considered and planned for ethics, privacy, etc.
- Does this research request conflict or overlap research being done by other CPCSSN investigators? (priority will always be given to CPCSSN “members” as defined below)

Consistency

A DAR submission involves multiple documents which outline the details of the study. Project information in all documents must be consistent. We recommend using the exact wording on all documents (copy and paste is encouraged to reduce minor inconsistencies)

- Data Stewards are making decisions based on the study specifics in the DAR and supporting documents. If, for example, the DAR and Ethics application have different cohort definitions it is not possible for adjudication to proceed.

CPCSSN Framework for managing conflict of interest and bias regarding relationships with Funders (and Other) Partners

Experts in policy and ethics nationally and internationally are calling for more explicit and preplanned ethics evaluation in policy related to health and health care. Ethical analysis should not to be considered an ‘add-on’ but an integral part of the process of engaging in any partnership to conduct research. This document is meant to describe how the Canadian Primary Care Sentinel Surveillance Network (CPCSSN) intends to match its overall mission and priorities to those of potential partners and mitigate any possible misalignment or risks.

A. Objectives:

1. To describe the foundational ethical values and principles necessary for partnering with the CPCSSN;
2. To describe the decision-making framework that the CPCSSN will use in partnerships with entities such as governmental, non-governmental organizations and for-profit entities.

B. Background

The Canadian Primary Care Sentinel Surveillance Network (CPCSSN) is the only pan-Canadian network that collects electronic medical record (EMR) data for research, surveillance and

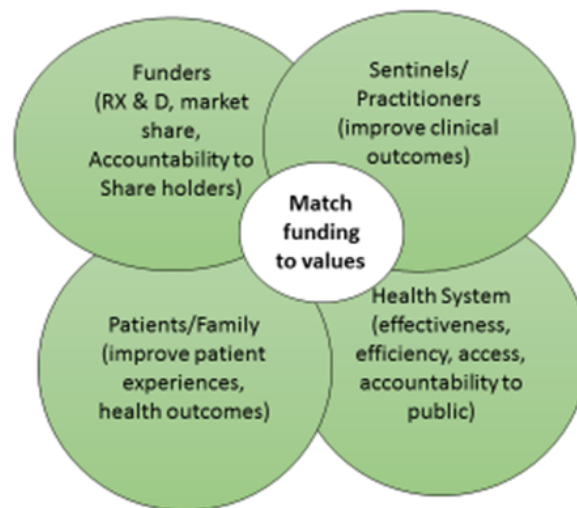
primary healthcare practice quality improvement. CPCSSN extracts, cleans, and maintains longitudinal de-identified patient data from more than 1,200,000 Canadians using the EMRs of more than 1,000 primary care practitioners. Examples of the extracted data includes health conditions, risk factors and behaviors, prescribed medications, laboratory tests, and referrals to other healthcare providers.

CPCSSN provides a well-established foundation with which to engage health researchers from a variety of backgrounds as well as information technology (IT) developers, and computer scientists. For the latter, CPCSSN's electronic medical record platform has great potential for use in the development and evaluation of patient-specific technology tools, including apps and patient access portals.

CPCSSN data are a significant opportunity to generate new knowledge about the organization and delivery of primary healthcare with the overall goal to spur change that will make a difference in the lives of patients and their families across Canada. This decision-making framework will address the CPCSSN's values and goals.

Figure 1 shows the complex relationship between this network and its partners, particularly potential funders.

Improving the care and quality of life of people with multiple chronic conditions-- multimorbidity



C. Definitions

Values are “concepts we use to explain how and why various realities matter. Where moral values are secure, we can cultivate a wide array of values. Where moral values are in danger, all values are at risk.” [ref: EBE Vol. 5, 2515-2520].

Principles are generally understood as guides to decision and action embodying sets of values. Ethical frameworks are not a theoretical task; they are directed to guide *practical* decisions and actions. We ask questions such as: “What are our criteria for decisions and actions regarding partnership funding?”

Our values based decision-making framework should first and foremost guide the overall priorities and activities for the CPCSSN. The values ought to be reflected in any approved project and in the subsequent knowledge translation activities generated by each project.

Conflict of interest can be described as ‘a set of conditions in which professional judgment concerning a primary interest (such as a patient's welfare or the validity of research) tends to be unduly influenced by a secondary interest (such as financial gain)’ (Thompson, 1993; Lemmens and Freedman, 2000; Krimsky, 2006). Conflict of interest has three components — a trust-based relationship, a conflicting interest and an exercise of judgment. The relationship between the parties in question is often trust based, such that one party (the trustor) is entitled to trust that the other (the trustee) will promote or protect their interests in relation to matters within that relationship. The second component is an influence that tends to make the trustee's judgment on a given decision less reliable for promoting or protecting the trustor's interests than it would normally be (Davis, 1998). Third, the trustee must be in a position to make a decision that affects the trustor's interests.

D. CPCSSN Mission

CPCSSN’s mission is to improve primary healthcare delivery, practice and policy through surveillance, research and quality improvement. CPCSSN’s achieves its mission by creating a central data repository of EMR data elements from it different network nodes across Canada. This repository has three main purposes: (1) surveillance of chronic conditions; (2) quality improvement feedback to participating primary care practitioners and their clinics/organizations; and (3) development of new knowledge through research.

CPCSSN, through its partnerships with participating sentinels, is a data steward of de-identified pan-Canadian EMR data which can serve as a bridge between primary care, public health, and other health sectors. The EMR data that makes up the repository consists of “coded information”, which is defined as information that has been stripped of identifiers and replaced with a CPCSSN ID number.

E. CPCSSN values

CPCSSN’s values include the following core concepts:

- Providing excellent and robust data and analyses, where bias and conflict of interests are mitigated as much as possible;
- Improving patient care and health service delivery with a particular focus on primary care

- Inclusiveness and collaborative relationships with researchers, clinicians, and stakeholder groups whose mission, values, and practices are in alignment with CPCSSN;
- Transparency and accountability for the surveillance, quality improvement, and research produced;
- Effective stewardship of the data repository created by participating sentinels ensuring data quality, security, privacy and appropriate which meet or exceed local and national regulatory standards;
- No identifiable data regarding CPCSSN practitioners or patients is released to third parties.

CPCSSN recognizes that there may be times where it is mutually beneficial to work with for-profit partners. There are three additional principles that the CPCSSN will use to guide working with for-profit partners:

- Alignment with the CPCSSN mission and values – CPCSSN private sector work must provide a public benefit;
- Transparency – full results reports, the identities of for-profit entities that use the CPCSSN data, and dataset creation plans will all be publicly available
- For-profit entity work undertaken with the CPCSSN must be in addition to, and in no way detracting from, CPCSSN research or quality improvement work – this includes the requirement that work with for-profit entities is financially viable

In the event of work with for-profit partners: (a) the research is conducted at arm's length from the partner; (b) no partner's employees, contractors, or agents are part of the project team; (c) the partner has no influence on study direction or analysis; and (d) the sponsor has no access to data, other than a final report. Apriori primary and secondary outcomes must be registered with the CPCSSN and that the CPCSSN will make the results of any for-profit analyses publicly available through such means as peer-reviewed publication, report or brief available on its website and upon request.

F. Safeguarding CPCSSN Data for secondary use (includes Research, Surveillance or Quality Improvement)

The CPCSSN platform adheres to a privacy and information system security framework (Williamson et al., 2013, p. 5) using:

- ▶ Privacy by Design Principles, the standard principles adopted by governments and privacy commissioner offices throughout the world
- ▶ Information Standards Organization (ISO) 27001/2, governing information system security
- ▶ Tri-Council Policy Statement, Ethical Conduct of Research Involving Humans (TCPS2) 2014, governing the use of health information in research

The CPCSSN has completed and will regularly review privacy impact assessments (PIA) and threat risk assessments (TRA). The scope of the PIAs include every stage of the information flow between the network nodes and the CPCSSN central repository. The PIAs are intended to pay close attention to CPCSSN's technological, organizational, and physical safeguards for

health information. After the baseline PIAs were completed, each network node continues to use the CPCSSN Privacy Compliance Checklist to undertake their own compliance monitoring on an annual basis. A TRA was also conducted and included CPCSSN's management of personal information and personal health information, and all end user connection points (access by regional data managers, researchers, research associates, etc.). The CPCSSN data are housed at the Centre for Advance Computing (CAC) at Queen's University in Kingston, Ontario. The decision to use the CAC was made with security in mind. The regional server requires that all remote connections to it be encrypted (e.g. through a VPN or SSH connection). This ensures that even if someone gains unauthorized access to the regional server, patient data will remain secure. Other benefits of the central management with the CAC are: 1) Physical access to the CAC server room is controlled, logged, and monitored (e.g. with video cameras), 2) Access to the regional server can be revoked in the event that an employee is terminated or leaves the project, 3) A strong, consisted firewall policy can be applied for all CPCSSN servers by the CAC, and 4) Software security updates can be uniformly applied to all of the servers.

Another key ethical safeguard is the continuous renewals of approvals from Research Ethics Boards (REB) in Canada and yearly updates for investigators to declare any potential conflicts of interest. Each PCRN had to obtain REB approval before its collected health information could be used for research purposes. The CPCSSN executive committee ensures that annual renewal will be obtained from all local PCRN University affiliated.

Safeguards are taken to prevent identifiable patient data from being stored in the CPCSSN repository such as (e.g. data management tools to identify and remove names, health care numbers, credit card numbers, etc). Safeguards will be taken to prevent identifiable patient data from being released from the CPCSSN repository.

Standard Operating Procedures (SOP) in Privacy, Ethics and Information Security the CPCSSN have been developed. These address security incidents and reporting. CPCSSN has a protocol for reporting incidents and in the event of a real or perceived incident there is a step-by-step process of what to do in the event of an incident (Appendix A).

All research utilizing CPCSSN repository data must be approved by relevant Research Ethics Boards and the CPCSSN Standing Research Committee in accordance with CPCSSN's Standard Operating Procedures (SOP) in Privacy, Ethics and Information Security and the CPCSSN Information and Data Sharing Agreement. In the case of research, the researcher will be allowed limited use of the CPCSSN data according to what is agreed upon in the CPCSSN Information and Data Sharing agreement. The research will not be allowed to use the data obtained for other purposes outside that agreement.

Any decisions regarding publication and presentation of findings related to CPCSSN data rest entirely with investigators. No partner organizations or individuals will be given authority to limit sharing of findings with the public assuming the information falls within the previously stated privacy limitations and partnership agreements.

Consent of Participating Primary Care Practitioners

All participating CPCSSN sentinels are offered the opportunity to opt out of participating in research projects at two broad levels:

- 1) Full participation with all studies and partnerships with organizations and individuals deemed acceptable by the CPCSSN Steering Committee (this includes charities and for-profit entities).
- 2) Participation restricted to studies and partnerships only with publicly funded agencies and granting agencies.

G. Guide to Decision Making and Approval of studies involving Partnerships *(eg. funders sentinels, patients, researchers, decision-makers)*: *(Informed by Canadian Institutes of Health Research Ethics Framework for Partnerships with the Private Sector, 2013)*

Who are the CPCSSN's Right Partner(s)? The CPCSSN should ideally partner only with organizations and individuals whose mission, values and practices are in alignment with CPCSSN. There will be situations where CPCSSN's mission, values and practices do not fully align with those of a partner. In these cases, CPCSSN and its partner would need to ensure we have sufficient alignment of mission, values and practices.

For-profit entities or charities whose mission, values or practices conflict with our values (e.g., companies that manufacture or sell products that are known to be harmful to human health, without counter-balancing benefits, or that knowingly engage in practices that violate human rights) should be excluded from partnerships CPCSSN.

If conflicts arise or are discovered, the CPCSSN will reserve the right to terminate or reassess the project scope. In accordance with local and national REB standards all projects will be reviewed yearly to ensure compliance with all CPCSSN policies, and the values stated in this document.

Are we better together? Partnerships should promote initiatives that each partner, on its own, could not address as comprehensively or as effectively.

Is CPCSSN interested in the following:

1. Open to funding from partners to create new disease algorithms and contributions to our code library?
2. Is the timeframes for the work feasible?
3. Is there interest in providing educational content via CPCSSN?
4. Is more EMR data extraction realistic in the next 3-5 years?

Do we have shared objectives? Partners should agree explicitly on the objectives of the partnership, even though they may have different motivations and values. To achieve success, partners should establish a common understanding regarding the objectives of the partnership and should negotiate the management of the partnership in ways that can best achieve them by building trust and maintaining open-mindedness. In some instances, such as

when the stakes in the outcome of a partnership are high, greater harmonization between partners, their visions or their operations may be necessary to ensure the success of a partnership. All partnership agreements will be ratified by the CPCSSN Steering committee upon entering in agreement and yearly in coordination with REB approvals

Are there any conflicting objectives? If partner objectives are not aligned with the CPCSSN mission and values, use of the CPCSSN data is likely not possible. Examples of the types of projects not acceptable to the CPCSSN are those whose primary goal is marketing, the data are skewed for purposes such as disease awareness, the answer to the question being asked has already been completed, or evidence can be found using other data or methods (e.g. meta analyses).

Do we have fair selection? The selection of partners should be transparent. To help select right partners among potential organizations, the CPCSSN Steering Committee should develop objective criteria and transparent procedures that are applied across the organization to ensure fairness. Entering into partnerships will be discussed and agreed upon with a majority vote by the advisory board.

Do we have an appropriate proportion of for-profit partners? The CPCSSN is at risk of obligation and conflict of interest where for-profit entities represent the majority proportion of infrastructure or project funding. The CPCSSN will have a diversified funding scheme made up of tri-council funding, governmental and non-governmental organizations and for-profit entities. No governmental or non-governmental organization or the for-profit entities shall make up more than 50% of the CPCSSN's infrastructure and project costs at any one time.

Appendix A –CPCSSN values expanded

Empowerment of practitioners and patients

CPCSSN recognizes that demonstrating integrity is essential for building trust among the organizations and people with whom it deals. It is critical for establishing and maintaining an excellent reputation with the Canadian public. We operate with professional integrity to maintain confidence in the organization itself, in its relationships, its data gathering, management and analysis services, its information and knowledge products.

Transparency

The CPCSSN practices being transparent about what it is doing, how and when it is doing it, with whom it is engaged, and what impact it believes the activity or product may or will have. Information is conveyed in timely and appropriate ways to relevant stakeholders and audiences in the spirit of intentional openness. Transparency is created by communicating regularly with participating sentinels, making information available on our website and reporting to our funders. Transparency is a pre-condition for any form of accountability.

Accountability

We appreciate the importance of establishing a highly accountable organizational culture. Attaining and maintaining organizational accountability necessitates our willingness to accept full personal accountability for the assigned roles and responsibilities related to governance, management practices, the implementation of the CPCSSN's strategic plan, and to the delivery of value to internal and external stakeholders. It also requires risk and liability management which CPCSSN mitigates by complying with all ethical, legal, privacy and regulatory requirements and procedures.

Collaboration

Success is supported by our commitment to be a trusted and respected partner and collaborator. When appropriate, we share our skills, talents, knowledge and CPCSSN Repositories with others who hold a shared purpose and direction. As a pan-Canadian 'network of networks' we nurture our unique capacity to contribute to the strengthening and advancement of primary healthcare (primary care and public health) research, practice and policy at regional, provincial and national levels through partnership and collaborative initiatives. In the spirit of collaboration, we may serve as a bridge or connector between information technology, healthcare and academia.

Stewardship

CPCSSN is a steward of a data C containing clinical electronic medical record data from primary care. Referring to data and knowledge product accessibility, we support processes and agreements that enable timely access to accurate and quality data, data sharing, analysis, translation and interpretation, and the information exchange of such. We work to improve ease of use of the above.

CPCSSN remains flexible to enable positive, creative, organized and timely reaction to the needs of the organization, our partners, funders, investors, stakeholders, and the broader community. We endeavour to engage in regular knowledge exchanges at multiple levels to facilitate response to 'environmental' changes and emergent opportunities or concerns.

APPENDIX II

CPCSSN Data Dictionary 2020-Q4

Table: Network

- A master list of all the practice-based research networks within CPCSSN whose data has been sent to the central repository
- Linking table that contains the names and locations of each practice-based research network within CPCSSN

Variable Name	Variable Type	Form at	Definition	Allowed Values	Notes
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	
GeographicArea	Categorical		Geographical area of the Network.	'Southern Alberta', 'Northern Alberta', 'Greater Toronto Area', 'Eastern Ontario', 'Quebec', 'Newfoundland & Labrador', 'Manitoba', 'Maritime Provinces', 'British Columbia', 'Southern Ontario', 'National Capital Region',	

				'Northern Ontario'	
--	--	--	--	-----------------------	--

Table: Provider

- Contains limited demographic information on providers (physicians and nurse practitioners) who are currently participating in CPCSSN

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
Provider_ID	Provider table key	4 digits	Unique integer identifying each CPCSSN provider.	Positive Integers	<ul style="list-style-type: none"> Unique number assigned by each network to the providers in its sites
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> Allows this table to be linked to the Network table and all other tables that include the Network table key Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
BirthYear	Date	yyyy	Provider's year of birth.	BirthYear + 25 <= extraction year <= BirthYear + 85	
Sex	Categorical		Provider's sex.	Male, Female	<ul style="list-style-type: none"> Contains provider-reported sex

ProviderType	Categorical		The type of health care practitioner that the provider is.		<ul style="list-style-type: none"> • Not currently standardised
StartDate	Date	yyyy-mm-dd	Date that the provider joined CPCSSN at their current practice site.		<ul style="list-style-type: none"> • If a provider changes sites but remains part of CPCSSN, StartDate will be the date they started at the new site

Table: Patient

- List of EMR patients whose primary provider is a consenting physician in the CPCSSN project (has an entry in the Provider table)
- Contains limited demographic information on each patient

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
Patient_ID	Patient Table Key	15 or 16-digits	Unique integer randomly assigned to each patient within CPCSSN.	Positive integers	<ul style="list-style-type: none"> • The Patient_ID can be used to link records for each patient that are stored in other tables
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> • Allows this table to be linked to the Network table and all other tables that include the Network table key • Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Provider_ID	Provider Table Key	15-16 digits	Unique integer identifying each CPCSSN provider.	Positive Integers	<ul style="list-style-type: none"> • Only a Provider_ID that exists in the Provider table can be referenced here
StartDate	Date	yyyy-mm-dd	Date that the provider starts providing care to the patient at this site.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> • If this information is not recorded in the EMR, the date on the earliest record in the EMR may be used
Sex	Categorical		The sex (not gender) of the patient.	'Male' or 'Female'	

Age	Integer	1-2 digits	The age of the patient as of 2020-12-31.	0 - 125	<ul style="list-style-type: none"> • Must pick one of the following options: <ul style="list-style-type: none"> • Option 1: Age (to nearest year) and Location = 'Urban' or 'Rural' • Option 2: Age (to nearest 5 years) and Location = full postal code (if available)
Location	Categorical		Urban/rural or postal code.	'Urban', 'Rural',	<ul style="list-style-type: none"> • Must pick one of the following options: <ul style="list-style-type: none"> • Option 1: Age (to nearest year) and Location = 'Urban' or 'Rural' • Option 2: Age (to nearest 5 years) and Location = full postal code (if available)
DateCreated	Date	yyyy-mm-dd	The date on which the record was input into the EMR.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> • If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used

Table: AllergyIntolerance

- All allergy and intolerance data for the patient

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
AllergyIntolerance_ID	AllergyIntolerance Table Key	15-16 digits	Unique integer identifying an allergy/intolerance record.	Positive Integers	
Network_ID	Network Table Key	1 or 2- digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> • Allows this table to be linked to the Network table and all other tables that include the Network table key • Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16- digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> • Allows this table to be linked to the patient table and all tables that contain the patient table key • Only a Patient_ID that exists in the Patient table can be referenced here
Encounter_ID	Encounter Table Key	15 or 16- digits	Unique integer ID that links a record in this table to a specific Encounter record (i.e., to a single interaction with a primary care provider).	See Encounter table	<ul style="list-style-type: none"> • Allows this table to be linked to the Encounter table and all other tables that include the Encounter_ID • Only an Encounter_ID that exists in the Encounter table can be referenced here • Multiple records in this table may be linked to a single Encounter record • May be empty if the record cannot be linked to a specific encounter

StartDate	Date	yyyy-mm-dd	Date on which the allergy was first identified.	1902-01-01 to CutOffDate (see the Cycle table)	
StopDate	Date	yyyy-mm-dd	Date on which the allergy noted as inactive.	1902-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> • If an allergy is inactive and StopDate cannot be found, Status field must still be set to 'Inactive'
DIN	Categorical	8 digits	Drug Identification Number for the medication.	00000000 - 99999999	<ul style="list-style-type: none"> • DIN look-up and downloadable files can be found on Health Canada's Drug Product Database website
Name_calc	Categorical		Standardised text description that accompanies the specific ATC code recorded in Code_calc.	Descriptions from Health Canada's Drug Product Database	<ul style="list-style-type: none"> • Populated by the CPCSSN ATC coding algorithm, which uses the medication descriptions from Health Canada's Drug Product Database
CodeType_cal c	Categorical		Code set from which the code in Code_calc was drawn. Currently, can only be 'ATC'.	ATC	<ul style="list-style-type: none"> • Must be 'ATC' or empty
Code_calc	Categorical		An ATC code.	Valid ATC code only	<ul style="list-style-type: none"> • Populated by the CPCSSN ATC coding algorithm, which uses the ATC codes from Health Canada's Drug Product Database
Severity	Categorical		The severity of the allergy or intolerance reaction.	'Mild', 'Moderate', 'Severe'	
AllergyStatus	Categorical		Current status of the allergy or intolerance.	'Active', 'Inactive'	
DateCreated	Date	yyyy-mm-dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> • If the date of record creation cannot be found within the EMR, the date on which the

					record was last modified may be used
--	--	--	--	--	---

Table: Billing

- Contains patient-level EMR billing records that have been processed by CPCSSN text processing algorithms
- Each record is for a single patient, labelled by Patient_ID
- Each patient may have more than one record for a single encounter (i.e., physician billed for multiple services). Linking to other records corresponding to single encounters is accomplished through the Encounter_ID variable

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
Billing_ID	Billing Table Key	15 or 16-digits	Unique integer identifying a Billing record.	Positive Integers	
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> • Allows this table to be linked to the Network table and all other tables that include the Network table key • Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16-digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> • Allows this table to be linked to the the patient table and all tables that contain the patient table key • Only a Patient_ID that exists in the Patient table can be referenced here
Encounter_ID	Encounter Table Key	15 or 16-digits	Unique integer ID that links a Billing record to a specific Encounter record (i.e., to a single interaction with a primary care provider).	See Encounter table	<ul style="list-style-type: none"> • Allows this table to be linked to the Encounter table and all other tables that include the Encounter_ID • Only an Encounter_ID that exists in the Encounter table can be referenced here • Multiple records in this table may be

					linked to a single Encounter record • May be empty if the record cannot be linked to a specific encounter
ServiceDate	Date	yyyy-mm-dd	Date on which the service was provided.	1980-01-01 to CutOffDate (see the Cycle table)	
ServiceCode	Categorical	Varies by province	Province-specific code for the service being billed for.	Varies by province	• Allowed service code sets vary by province and may change with time. • We currently have data for five provinces. Links to documentation are below: BC: MSC Payment Schedule AB: Health Service Codes MB: Manitoba Physician's Manual ON: OHIP Schedule of Benefits & Fees NS: MSI Billing
DiagnosisText_calc	Categorical		Standardised text description that accompanies the specific ICD-9-CM code recorded in DiagnosisCode_calc.	Descriptions from ICD-9-CM only	• Populated by the CPCSSN ICD-9-CM coding algorithm, which uses the descriptions of disease entities from ICD-9-CM • Text terms can be found through an ICD-9 look-up website
DiagnosisCodeType_calc	Categorical		The diagnosis code set from which the code in DiagnosisCode_calc was drawn. Currently, can only be 'ICD9'.	'ICD9'	• Must be 'ICD9' or empty
DiagnosisCode_calc	Categorical		An ICD-9-CM code.	Valid ICD-9-CM code	• Populated by the CPCSSN ICD-9-CM coding algorithm, which uses the descriptions of disease

					<p>entities from ICD-9-CM</p> <ul style="list-style-type: none"> • ICD-9 codes can be found through an ICD-9 look-up website
DateCreated	Date	yyyy-mm-dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> • If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used

Table: Encounter

- Encounters (i.e., interactions) between patient and provider
- Not all interactions are captured in this table
- The interaction provider does not need to be participating in CPCSSN, although if they are not, the provider's information is suppressed

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
Encounter_ID	Encounter Table Key	15-16 digits	Unique integer identifying an Encounter record.	Positive Integers	<ul style="list-style-type: none"> • Encounter_IDs can be used to link records that are located in separate CPCSSN tables, but are from a single encounter between patient and provider • Not all Encounter_IDs link to records in other CPCSSN tables • The majority of records in other CPCSSN tables have no Encounter_ID
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 13	<ul style="list-style-type: none"> • Allows this table to be linked to the Network table and all other tables that include the Network table key • Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16-digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient Table	<ul style="list-style-type: none"> • Allows this table to be linked to the patient table and all tables that contain the patient table key • Only a Patient_ID that exists in the Patient table can be referenced here

Provider_ID	Provider Table Key	15-16 digits	Unique integer identifying each CPCSSN provider.	See Provider Table	<ul style="list-style-type: none"> Only a Provider_ID that exists in the Provider table can be referenced here If the provider for the visit is not part of CPCSSN, the Provider_ID will be empty
Encounter Date	Date	yyyy-mm-dd	Date on which the encounter occurred.	1980-01-01 to cut-off date	
Encounter Type_calc	Categorical		Standard coding of the Encounter_orig text.	'Email', 'Office Visit', 'Out of Office Visit', 'Phone Call'	
DateCreated	Date	yyyy-mm-dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used

Table: EncounterDiagnosis

- Diagnoses resulting from each encounter between patient and provider

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
EncounterDiagnosis_ID	EncounterDiagnosis Table Key	15 or 16-digits	Unique integer identifying an EncounterDiagnosis record.	Positive Integers	
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> Allows this table to be linked to the Network table and all other tables that include the Network table key Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16-digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> Allows this table to be linked to the patient table and all tables that contain the patient table key Only a Patient_ID that exists in the Patient table can be referenced here
Encounter_ID	Encounter Table Key	15 or 16-digits	Unique integer ID that links an EncounterDiagnosis record to a specific Encounter record (i.e., to a single interaction with a primary care provider).	See Encounter table	<ul style="list-style-type: none"> Allows this table to be linked to the Encounter table and all other tables that include the Encounter_ID Only an Encounter_ID that exists in the Encounter table can be referenced here Multiple records in this table may be linked to a single Encounter record May be empty if the record cannot be linked to a specific encounter
DiagnosisText_calc	Categorical		Standardised text description that accompanies the specific ICD-9-CM code recorded in DiagnosisCode_calc.	Descriptions from ICD-9-CM only	<ul style="list-style-type: none"> Populated by the CPCSSN ICD-9-CM coding algorithm, which uses the descriptions of disease entities from ICD-9-CM Text terms can be found through an ICD-9 look-up website
DiagnosisCodeType_calc	Categorical		The diagnosis code set from which the code in	'ICD9'	<ul style="list-style-type: none"> Must be 'ICD9' or empty

			DiagnosisCode_calc was drawn. Currently, can only be 'ICD9'.		
DiagnosisCode_cal c	Categori cal		An ICD-9-CM code.	Valid ICD-9- CM code	<ul style="list-style-type: none"> • Populated by the CPCSSN ICD-9-CM coding algorithm, which uses the descriptions of disease entities from ICD-9-CM • ICD-9 codes can be found through an ICD-9 look-up website
DateCreated	Date	yyyy-mm- dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> • If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used

Table: Exam

- Results of physical exams performed on the patient
- Eight exams are currently extracted by CPCSSN (allowed ranges and units in brackets)
 - Body mass index (BMI) [5-200 kg/m²]
 - Diabetic foot exam
 - Height [30-230 cm]
 - Peak expiratory flow rate (PEFR) [0-200 L/min]
 - Systolic and diastolic blood pressure (sBP and dBP) [sBP: 50-300 mmHg; dBP: 20-200 mmHg]
 - Waist circumference [10-300 cm]
 - Waist-to-hip ratio [0.1-10]
 - Weight [1-500 kg]

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
Exam_ID	Exam Table Key	15 or 16-digits	Unique integer identifying an Exam record.	Positive Integers	
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> • Allows this table to be linked to the Network table and all other tables that include the Network table key • Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16-digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> • Allows this table to be linked to the the patient table and all tables that contain the patient table key • Only a Patient_ID that exists in the Patient table can be referenced here
Encounter_ID	Encounter Table Key	15 or 16-digits	Unique integer ID that links an exam record to a specific Encounter record (i.e., to a single interaction with a primary care provider).	See Encounter table	<ul style="list-style-type: none"> • Allows this table to be linked to the Encounter table and all other tables that include the Encounter_ID • Only an Encounter_ID that exists in the Encounter table can be referenced here • Multiple records in this table may be linked to a single Encounter record • May be empty if the record cannot be linked to a specific encounter

Exam1	Categorical		Name of the physical exam.	BMI, Foot Exam, Height, Joint Pain, PEFR, sBP, Waist Circumference, Waist Hip Ratio, Weight	
Result1_calc	Numeric		Result1_orig converted to CPCSSN standard units.		<ul style="list-style-type: none"> Only populated if Result1_orig (and Result2_orig, if applicable) can be converted from the UnitOfMeasure_orig to the CPCSSN standard units, and the converted value is within allowed bounds CPCSSN does not yet code 'Foot Exam results to standard text, so all Result1_calc entries for this exam type will be blank
Exam2	Categorical		Name of the physical exam that is paired with Exam1	'dBP' or empty	<ul style="list-style-type: none"> Currently the only paired exam is blood pressure When Exam 1 is 'sBP (mmHg)' then Exam2 must be 'dBP (mmHg)'
Result2_calc	Integer		Result2_orig converted to CPCSSN standard units.		<ul style="list-style-type: none"> Only populated if Result1_orig (and Result2_orig, if applicable) can be converted from the UnitOfMeasure_orig to the CPCSSN standard units, and the converted value is within allowed bounds
UnitOfMeasure_calc	Categorical		The CPCSSN standard units for each exam.	cm, kg, kg/m ² , L/min, mmHg	<ul style="list-style-type: none"> Only populated if Result1_orig (and Result2_orig, if applicable) can be converted from the UnitOfMeasure_orig to the CPCSSN standard units, and the converted value is within allowed bounds

DateCreated	Date	yyyy-mm-dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used
-------------	------	------------	-------------------------------	--	---

Table: FamilyHistory

- Family history of the patient

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
FamilyHistory_ID	FamilyHistory Table Key	15 or 16-digits	Unique integer identifying a FamilyHistory record.	Positive Integers	
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> Allows this table to be linked to the Network table and all other tables that include the Network table key Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16-digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> Allows this table to be linked to the the patient table and all tables that contain the patient table key Only a Patient_ID that exists in the Patient table can be referenced here
Encounter_ID	Encounter Table Key	15 or 16-digits	Unique integer ID that links an family history record to a specific Encounter record (i.e., to a single interaction with a primary care provider).	See Encounter table	<ul style="list-style-type: none"> Allows this table to be linked to the Encounter table and all other tables that include the Encounter_ID Only an Encounter_ID that exists in the Encounter table can be referenced here Multiple records in this table may be linked to a single Encounter record May be empty if the record cannot be linked to a specific encounter
DiagnosisText_calc	Categorical		Standardised text description that accompanies the specific ICD-9-CM code recorded	Descriptions from ICD-9-CM only	<ul style="list-style-type: none"> Populated by the CPCSSN ICD-9-CM coding algorithm, which uses the descriptions of disease entities from ICD-9-CM Text terms can be found through an ICD-9 look-up website

			in DiagnosisCod e_calc.		
DiagnosisCodeT ype_calc	Categorical		The diagnosis code set from which the code in DiagnosisCod e_calc was drawn. Currently, can only be 'ICD9'.	'ICD9'	<ul style="list-style-type: none"> Must be 'ICD9' or empty
DiagnosisCode_ calc	Categorical		An ICD-9-CM code.	Valid ICD- 9-CM code	<ul style="list-style-type: none"> Populated by the CPCSSN ICD-9-CM coding algorithm, which uses the descriptions of disease entities from ICD-9-CM ICD-9 codes can be found through an ICD-9 look-up website
Relationship_cal c	Categorical		Standardised relationship type, determined from the CPCSSN relationship coding algorithm.	'Aunt' 'Brother' 'Cousin' 'Daughter' 'Father' 'Granddau ghter' 'Grandfath er' 'Grandmot her' 'Grandson' ' 'Great Aunt' 'Great Granddau ghter' 'Great Grandfath er' 'Great Grandmot her' 'Great Grandson' 'Great Uncle' 'Half Brother'	<ul style="list-style-type: none"> Both Relationship_orig and DiagnosisText_orig entries factor into the coding; however, only when the Relationship_calcs determined from both fields are consistent (or only one of the field is empty), will this variable be populated The allowed values may have different capitalisation or hyphenation than indicated

				'Half Sister' 'Mother' 'Nephew' 'Niece' 'Sister' 'Son' 'Uncle'	
RelationshipSide_calc	Categorical		Was the relationship on the mother's or father's side?	'Maternal', 'Paternal'	
RelationshipDegree_calc	Categorical		Network degree of the relationship.	1, 2	<ul style="list-style-type: none"> Relationship degree defined here: http://www.cdc.gov/genomics/CPCSSN/Repositories/diseases/breast_ovarian_cancer/risk_categories.htm
AgeAtOnset	Integer	1-3 digits	Age of onset for the condition.	0-125	
VitalStatus	Categorical		Whether the relative was alive or deceased at the time when the family history was recorded.	'Alive', 'Deceased'	
WasCauseOfDeath	Integer	1 digit	Was this condition the cause of death?	0, 1	<ul style="list-style-type: none"> 0 = not the cause of death 1 = was the cause of death
AgeAtDeath	Integer	1-3 digits	Relation's age at death.	0-125	
DateCreated	Date	yyyy-mm-dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the	<ul style="list-style-type: none"> If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used

				Cycle table)	
--	--	--	--	-----------------	--

Table: HealthCondition

- The problem list and past medical history of the patient.

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
HealthCondition_ID	HealthCondition Table Key	15 or 16-digits	Unique integer identifying a HealthCondition record.	Positive Integers	
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> Allows this table to be linked to the Network table and all other tables that include the Network table key Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16-digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> Allows this table to be linked to the patient table and all tables that contain the patient table key Only a Patient_ID that exists in the Patient table can be referenced here
Encounter_ID	Encounter Table Key	15 or 16-digits	Unique integer ID that links an family history record to a specific Encounter record (i.e., to a single interaction with a primary care provider).	See Encounter table	<ul style="list-style-type: none"> Allows this table to be linked to the Encounter table and all other tables that include the Encounter_ID Only an Encounter_ID that exists in the Encounter table can be referenced here Multiple records in this table may be linked to a single Encounter record May be empty if the record cannot be linked to a specific encounter
DiagnosisText_calc	Categorical		Standardised text description that accompanies the specific ICD-9-CM code recorded in DiagnosisCode_calc.	Descriptions from ICD-9-CM only	<ul style="list-style-type: none"> Populated by the CPCSSN ICD-9-CM coding algorithm, which uses the descriptions of disease entities from ICD-9-CM Text terms can be found through an ICD-9 look-up website
DiagnosisCodeType_calc	Categorical		The diagnosis code set from which the code in	'ICD9'	<ul style="list-style-type: none"> Must be 'ICD9' or empty

			DiagnosisCode_calc was drawn. Currently, can only be 'ICD9'.		
DiagnosisCode_calc	Categorical		An ICD-9-CM code.	Valid ICD-9-CM code	<ul style="list-style-type: none"> • Populated by the CPCSSN ICD-9-CM coding algorithm, which uses the descriptions of disease entities from ICD-9-CM • ICD-9 codes can be found through an ICD-9 look-up website
DateOfOnset	Date	yyyy-mm-dd	Date on which the health condition began.	1902-01-01 to CutOffDate (see the Cycle table)	
Status	Categorical		An indicator that the condition is active at the time of data extraction.	'Active', 'Inactive'	
DateCreated	Date	yyyy-mm-dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> • If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used

Table: Lab

- Results of lab tests

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
Lab_ID	Lab Table Key	15-16 digits	Unique integer identifying a lab record.	Positive Integers	
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> • Allows this table to be linked to the Network table and all other tables that include the Network table key • Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16-digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> • Allows this table to be linked to the patient table and all tables that contain the patient table key • Only a Patient_ID that exists in the Patient table can be referenced here
Encounter_ID	Encounter Table Key	15 or 16-digits	Unique integer ID that links a record in this table to a specific Encounter record (i.e., to a single interaction with a primary care provider).	See Encounter table	<ul style="list-style-type: none"> • Allows this table to be linked to the Encounter table and all other tables that include the Encounter_ID • Only an Encounter_ID that exists in the Encounter table can be referenced here • Multiple records in this table may be linked to a single Encounter record • May be empty if the record cannot be linked to a specific encounter

PerformedDate	Date	yyyy-mm-dd	Date on which the lab test was performed.	1902-01-01 to CutOffDate (see the Cycle table)	
Name_calc	Categorical		CPCSSN or LOINC standardised text description for the LOINC code recorded in Code_calc.		<ul style="list-style-type: none"> • Populated by the CPCSSN lab coding algorithm • There are standardised CPCSSN names for 50 lab tests, with remaining labs named as per the LOINC standard
CodeType_calc	Categorical		Code set from which the code in Code_calc was drawn. Currently, can only be 'LOINC'.	'LOINC'	<ul style="list-style-type: none"> • Must be 'LOINC' or empty
Code_calc	Categorical		A LOINC code.	Valid LOINC code only	<ul style="list-style-type: none"> • Populated by the CPCSSN lab coding algorithm • All LOINC codes, including for the 50 CPCSSN lab tests, are drawn from the LOINC data set
TestResult_calc	Open text		TestResult_orig converted into the standard units used for the 50 CPCSSN lab tests or standardised text.		<ul style="list-style-type: none"> • Populated by the CPCSSN lab test cleaning algorithm, and only applied to the 50 CPCSSN lab tests • There are both numeric and text lab results • Numeric lab results may include comparison operators (<, <=, >, >=)
UnitOfMeasure_calc	Categorical		The standardised unit of measure for the 50 CPCSSN lab tests.		<ul style="list-style-type: none"> • Populated by the CPCSSN lab test cleaning algorithm, and only applied to the 50 CPCSSN lab tests
DateCreated	Date	yyyy-mm-dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> • If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used

Table: Medication

- Medications prescribed for the patient

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
Medication_ID	Medication Table Key	15-16 digits	Unique integer identifying a medication record.	Positive Integers	
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> • Allows this table to be linked to the Network table and all other tables that include the Network table key • Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16-digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> • Allows this table to be linked to the patient table and all tables that contain the patient table key • Only a Patient_ID that exists in the Patient table can be referenced here
Encounter_ID	Encounter Table Key	15 or 16-digits	Unique integer ID that links a record in this table to a specific Encounter record (i.e., to a single interaction with a primary care provider).	See Encounter table	<ul style="list-style-type: none"> • Allows this table to be linked to the Encounter table and all other tables that include the Encounter_ID • Only an Encounter_ID that exists in the Encounter table can be referenced here • Multiple records in this table may be linked to a single Encounter record

					<ul style="list-style-type: none"> May be empty if the record cannot be linked to a specific encounter
StartDate	Date	yyyy-mm-dd	Date on which the medication was first prescribed.	1902-01-01 to CutOffDate (see the Cycle table)	
StopDate	Date	yyyy-mm-dd	Date on which the patient stopped taking the medication.	1902-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> If an allergy is inactive and StopDate cannot be found, Status field must still be set to 'Inactive'
DIN	Categorical	8 digits	Drug Identification Number for the medication.	00000000 - 99999999	<ul style="list-style-type: none"> DIN look-up and downloadable files can be found on Health Canada's Drug Product Database website
Name_calc	Categorical		Standardised text description that accompanies the specific ATC code recorded in Code_calc.	Descriptions from Health Canada's Drug Product Database	<ul style="list-style-type: none"> Populated by the CPCSSN ATC coding algorithm, which uses the medication descriptions from Health Canada's Drug Product Database
CodeType_calc	Categorical		Code set from which the code in Code_calc was drawn. Currently, can only be 'ATC'.	ATC	<ul style="list-style-type: none"> Must be 'ATC' or empty
Code_calc	Categorical		An ATC code.	Valid ATC code only	<ul style="list-style-type: none"> Populated by the CPCSSN ATC coding algorithm, which uses the ATC codes from Health Canada's Drug Product Database
DispensedForm_calc	Categorical		DispensedForm_orig converted to standardised text.	'Bottle', 'Capsule', 'Cream', 'Drops', 'Gel', 'Inhaler',	

				'Kit', 'Lotion', 'MDI', 'Ointment', 'Spray', 'Syringe', 'Tablet', 'Vial'	
DateCreated	Date	yyyy- mm-dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> ● If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used

Table: Referral

- All referrals (to specialists, physiotherapy, etc.) made by their assigned provider or clinic
- Excludes referrals made by specialists

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
Referral_ID	Referral Table Key	15-16 digits	Unique integer identifying a referral record.	Positive Integers	
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> • Allows this table to be linked to the Network table and all other tables that include the Network table key • Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16-digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> • Allows this table to be linked to the patient table and all tables that contain the patient table key • Only a Patient_ID that exists in the Patient table can be referenced here
Encounter_ID	Encounter Table Key	15 or 16-digits	Unique integer ID that links a record in this table to a specific Encounter record (i.e., to a single interaction with a primary care provider).	See Encounter table	<ul style="list-style-type: none"> • Allows this table to be linked to the Encounter table and all other tables that include the Encounter_ID • Only an Encounter_ID that exists in the Encounter table can be referenced here • Multiple records in this table may be linked to a single Encounter record • May be empty if the record cannot be linked to a specific encounter

CompletedDate	Date	yyyy-mm-dd	Date of the patient's visit to the referred to provider's clinic.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> Is not the date on which the patient's provider created the referral or the referral letter was written
Name_calc	Categorical		Standardised text description that accompanies the specific SNOMED-CT code recorded in ConceptCode_calc.	Name of a valid SNOMED-CT codes	<ul style="list-style-type: none"> Populated by the CPCSSN Referral Cleaning algorithm, which is based on SNOMED-CT coding
ConceptCode_calc	Integer		SNOMED concept code.	Valid SNOMED-CT codes	<ul style="list-style-type: none"> Populated by the CPCSSN Referral Cleaning algorithm, which is based on SNOMED-CT coding
DateCreated	Date	yyyy-mm-dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used

Table: RiskFactor

- Risk factors recorded for the patient
- CPCSSN currently codes for six risk factors

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
RiskFactor_ID	RiskFactor Table Key	15-16 digits	Unique integer identifying a risk factor record.	Positive Integers	
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> • Allows this table to be linked to the Network table and all other tables that include the Network table key • Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16-digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> • Allows this table to be linked to the patient table and all tables that contain the patient table key • Only a Patient_ID that exists in the Patient table can be referenced here
Encounter_ID	Encounter Table Key	15 or 16-digits	Unique integer ID that links a record in this table to a specific Encounter record (i.e., to a single interaction with a primary care provider).	See Encounter table	<ul style="list-style-type: none"> • Allows this table to be linked to the Encounter table and all other tables that include the Encounter_ID • Only an Encounter_ID that exists in the Encounter table can be referenced here • Multiple records in this table may be linked to a single Encounter record

					<ul style="list-style-type: none"> May be empty if the record cannot be linked to a specific encounter
StartDate	Date	yyyy-mm-dd	Date on which the risk factor began.	1902-01-01 to CutOffDate (see the Cycle table)	
EndDate	Date	yyyy-mm-dd	Date on which the risk factor ended.	1902-01-01 to CutOffDate (see the Cycle table)	
Name_calc	Categorical		Standardised risk factor name coded from the text in Name_orig.	'Alcohol', 'Diet', 'Exercise', 'Obesity', 'Psychosocial Stress', 'Smoking'	<ul style="list-style-type: none"> Populated by the CPCSSN RiskFactor name coding algorithm
Status_calc	Categorical		Standardised risk factor status coded from the text in Name_orig.	'Current', 'Never', 'Not Current', 'Past'	<ul style="list-style-type: none"> Populated by the CPCSSN RiskFactor status coding algorithm 'Not Current' indicates that it is unclear whether the patient never had or does not currently have the risk factor (e.g., 'non-smoker' could indicate has stopped smoking or has never smoked)
DateCreated	Date	yyyy-mm-dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used

Table: Vaccine

- Vaccinations given (or attempted to be given) to the patient

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
Vaccine_ID	Vaccine Table Key	15-16 digits	Unique integer identifying a vaccine record.	Positive Integers	
Network_ID	Network Table Key	1 or 2-digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> • Allows this table to be linked to the Network table and all other tables that include the Network table key • Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16-digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> • Allows this table to be linked to the patient table and all tables that contain the patient table key • Only a Patient_ID that exists in the Patient table can be referenced here
Encounter_ID	Encounter Table Key	15 or 16-digits	Unique integer ID that links a record in this table to a specific Encounter record (i.e., to a single interaction with a primary care provider).	See Encounter table	<ul style="list-style-type: none"> • Allows this table to be linked to the Encounter table and all other tables that include the Encounter_ID • Only an Encounter_ID that exists in the Encounter table can be referenced here • Multiple records in this table may be linked to a single Encounter record • May be empty if the record cannot be linked to a specific encounter

GivenDate	Date	yyyy-mm-dd	Date of vaccine administration.	1902-01-01 to CutOffDate (see the Cycle table)	
ExpiryDate	Date	yyyy-mm-dd	Vaccine expiry date.	1902-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> The vaccine-batch expiry date, not the end-date of vaccine efficacy
Name_calc	Categorical		Standardised text description that accompanies the specific ATC code recorded in Code_calc.	Descriptions from Health Canada's Drug Product Database	<ul style="list-style-type: none"> Populated by the CPCSSN ATC coding algorithm, which uses the medication descriptions from Health Canada's Drug Product Database
CodeType_cal c	Categorical		Code set from which the code in Code_calc was drawn. Currently, can only be 'ATC'.	ATC	<ul style="list-style-type: none"> Must be 'ATC' or empty
Code_calc	Categorical		An ATC code.	Valid ATC code only	<ul style="list-style-type: none"> Populated by the CPCSSN ATC coding algorithm, which uses the ATC codes from Health Canada's Drug Product Database
DIN	Categorical	8 digits	Drug Identification Number for the medication.	00000000 - 99999999	<ul style="list-style-type: none"> DIN look-up and downloadable files can be found on Health Canada's Drug Product Database website
NotGiven	Integer		Whether or not the vaccine was administered.	0 or 1	<ul style="list-style-type: none"> 'No' = 0 (vaccine administered) 'Yes' = 1 (vaccine NOT administered)
DateCreated	Date	yyyy-mm-dd	EMR date stamp of the record.	1980-01-01 to CutOffDate (see the Cycle table)	<ul style="list-style-type: none"> If the date of record creation cannot be found within the EMR, the date on which the record was last modified may be used

Table: DiseaseCase

- Patients who have been identified by the CPCSSN disease case detection algorithms as having one or more diseases

Variable Name	Variable Type	Format	Definition	Allowed Values	Notes
Disease_ID	Disease Table Key	15-16 digits	Unique integer identifying a disease record.	Positive Integers	
Network_ID	Network Table Key	1 or 2- digits	Unique integer assigned to each regional practice-based research network within CPCSSN.	1 - 14	<ul style="list-style-type: none"> Allows this table to be linked to the Network table and all other tables that include the Network table key Refer to the Network table for the specific regional practice-based research network covered by a given Network_ID
Patient_ID	Patient Table Key	15 or 16- digits	Unique and randomly assigned integer CPCSSN patient ID.	See Patient table	<ul style="list-style-type: none"> Allows this table to be linked to the patient table and all tables that contain the patient table key Only a Patient_ID that exists in the Patient table can be referenced here
Disease	Categorical		The disease name.	See the CPCSSN disease case documentation	
Version	Integer		The version number for the disease case definition.	See the CPCSSN disease case documentation	
IsDefaultDefin ition	Integer		Whether the disease is the most current or is the CPCSSN standard definition, OR is an older or less accurate definition.	0 or 1	<ul style="list-style-type: none"> 'No' = 0 'Yes' = 1

DateCaseDetected	Date	yyyy-mm-dd	The date on which the first indication of disease, for a given case definition, appears in the CPCSSN record.	1980-01-01 to CutOffDate (see the Cycle table)	
------------------	------	------------	---	--	--

APPENDIX III

CPSSN Secure Research Environment (SRE) Overview

WHAT IS IT: The CPCSSN secure research environment is a controlled and secured service environment for undertaking research using sensitive CPCSSN data. The CPCSSN SRE provides robust controls and safeguards to enable the secure transfer of sensitive data into a highly secure environment where it can be stored, manipulated and analysed by approved members of a research team.

WHERE IS IT LOCATED: The CPCSSN secure research environment (SRE) is located at the Queen's Centre for Advanced Computing (CAC) and benefits from the high levels of security already in place. The security is based on a Secure Research Environment (SRE) service provided by the CAC. SRE provides remote user access through a Virtual Private Network (VPN) and allows the CPCSSN administrator to enforce different levels of restrictions on data access. For example, data movement into and out of the SRE can be limited to a set of privileged users. Users login to the SRE through a VPN into a remote desktop that offers the available analytics software.

Architecture for SRE

The data SRE runs on multiple VMs within a SRE. A backend VM hosts the databases and analytic software and a frontend VM hosts the user interfaces. The data SRE is built using open source and commercial software. The commercial software licenses (Microsoft, SAS and IBM) are all obtained through academic/research agreements between the companies and Queen's University. The current data SRE prototype provides users with a secure login to a remote desktop on Windows Server and gives them the choice of using SAS, IBM Cognos and Python to perform analytics. Additional software can be easily included. A federated database is implemented in the backend using IBM Infosphere that can provide access to a number of database management systems. The CPCSSN data are currently stored in Microsoft SQL Server.

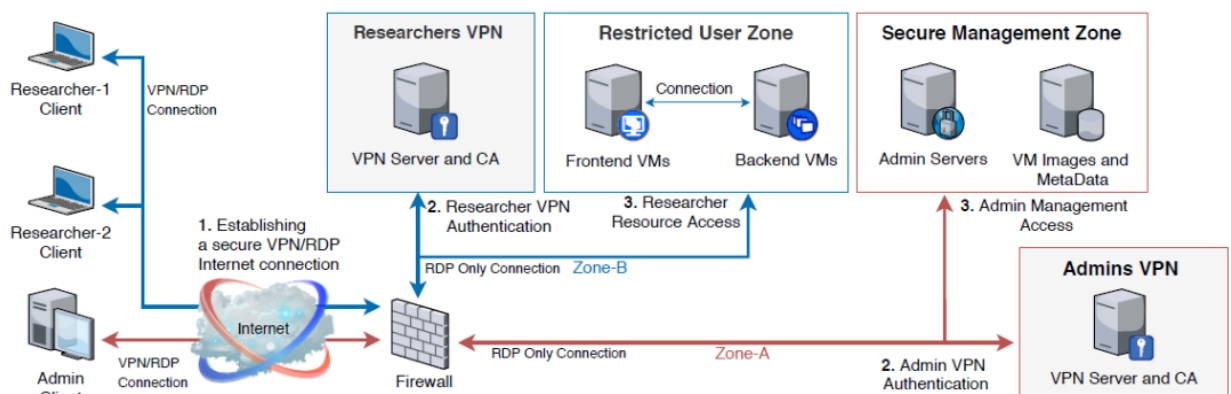


Figure 2 - SRE Design Overview

For further information on setting up the SRE please refer to CPCSSN_SOP001_SRE_User Guide

APPENDIX IV

Data Access Committee (DAC)

Purpose/Role

The CPCSSN Data Access Committee (DAC) has overall responsibility for the access procedures and all access decisions. The Committee includes individuals with expertise in the use of epidemiological research CPCSSN Repositories.

The DAC will inform the Steering Committee of all surveillance and research projects and make recommendations on areas of study/data collection. The DAC acts as the coordinating body for the vetting of projects put forward by CPCSSN investigators or those proposed from outside CPCSSN.

Review and approval of requests for access to CPCSSN data for research purposes will primarily rest with the DAC, although the DAC maintains the discretion to seek input from the CPCSSN's Steering Committee where appropriate.

It is constituted as follows:

- Chair DAC (non-science membre)
- Scientist/Researcher Academic x2
- Ethicist or member with ethics background
- Data Scientist
- CPCSSN Data Analyst
- Physician
- Patient/Community Representative (may invite 2 members)
- Other Ad-Hoc may be invited to the meeting such as a member of Correctional Services Canada

It is chaired by a “non-scientist” member (with a legal and/or ethics background).

Committee Meetings

- Meetings will be held monthly or at the call of CPCSSN Co-Chairs
- There will be a minimum of one meeting per month.
- Minutes will be taken at each meeting by a designated recording secretary. Minutes will be approved by the Chair and circulated to committee members.
- A brief status report, with minutes attached, will be provided to the Steering Committee at least every quarter, more often if required.
- Special guests can be invited to meetings by the Chair as needed for the purpose of expert advice or information sharing.

Research Proposal and Request for Data Access Approval System

All requests for access to CPCSSN data must be made through the CPCSSN website.

The DAC Chair will call a meeting of the Committee who will consider the data request. The Committee also has the option of asking the researcher to attend a meeting to answer questions (in person or via conference call) prior to making a decision on the disposition of the request. For requests where the Committee is not in unanimous agreement or where issues or concerns are present, the Committee will make a formal recommendation to the Steering Committee who will provide final approval of the request.