Who We Are
Established in 1994, CIHI is an independent, not-for-profit corporation that provides essential information on Canada’s health system and the health of Canadians. Funded by federal, provincial and territorial governments, we are guided by a Board of Directors made up of health leaders across the country.

Our Vision
To help improve Canada’s health system and the well-being of Canadians by being a leading source of unbiased, credible and comparable information that will enable health leaders to make better-informed decisions.
# Table of Contents

1 Introduction ........................................................................................................................... 1  
  1.1 Background .................................................................................................................. 1  
  1.2 Organization of This Specification .............................................................................. 2  
  1.3 Scope ............................................................................................................................ 3  
    1.3.1 In Scope ................................................................................................................ 3  
    1.3.2 Out of Scope ......................................................................................................... 3  
  1.4 Stability ....................................................................................................................... 4  
  1.5 EMR System Conformance ......................................................................................... 4  
2 Assumptions ........................................................................................................................ 5  
3 Strategy and Approach ...................................................................................................... 6  
  3.1 Conceptual-to-Physical EMR Implementation ........................................................... 6  
    3.1.1 Conceptual Data Model ....................................................................................... 6  
4 EMR System PHC Data Submission Considerations ....................................................... 8  
  4.1 Example Data Flow Scenarios .................................................................................... 8  
  4.2 Example EMR System Data Extract Submission Business Process ......................... 8  
  4.3 Treatment of Sensitive Data ....................................................................................... 12  
  4.4 Security Controls ........................................................................................................ 12  
    4.4.1 Privacy .................................................................................................................. 13  
  4.5 Transport Layer Interoperability ............................................................................... 13  
  4.6 Error Handling ........................................................................................................... 13  
5 PHC EMR Data Extract XML Schema Definition .......................................................... 14
1 Introduction

1.1 Background

Primary health care (PHC) is the most common health care service provided to Canadians. Across Canada, electronic medical records (EMRs) are increasingly being used, with the goal of improving patient care and the health of Canadians. However, it is evident that enhancements to EMRs are essential if we are to achieve the benefits of improved patient care and health system management.

As described elsewhere, the Canadian Institute for Health Information (CIHI) collaborated with the jurisdictions and Canada Health Infoway to develop the Pan-Canadian Primary Health Care Electronic Medical Record Content Standard (PHC EMR CS) to ensure that standardized information is available in EMRs to improve access, quality, outcomes and chronic disease prevention and management. This standard contains specifications for key concepts and value sets that describe a subset of important data elements in EMRs. The development and implementation of this standard coincided with the rapid evolution and increased use of EMRs in PHC settings across all Canadian jurisdictions. Consequently, it is critical that these EMRs be upgraded to provide the technical and functional capabilities required to realize their intended benefits.

For more information on CIHI’s PHC Information program and the PHC EMR CS, please refer to the PHC EMR CS Frequently Asked Questions and PHC EMR CS Information Sheet, both of which are available by sending a request to phc@cihi.ca or at www.cihi.ca/phc.

This PHC EMR Data Extract Specification Business View is intended to provide background information to EMR vendors and a broad range of stakeholders about the requirements associated with collecting and extracting standardized PHC EMR data.
This specification is designed to

- Be a resource for representing the PHC EMR CS in an extract;
- Provide the technical requirements for all PHC data elements that support PHC data reporting, including those that will be identified in future releases of the PHC EMR CS; and
- Define the clinical document structure required for extracting and submitting PHC data for health system analysis and use by authorized organizations.

In accordance with Canada Health Infoway’s Standards Collaborative license agreement for HL7 specifications, the technical content is available on the EMR and Integration (EMRi) Wiki. Information may be accessed at https://emri.infoway-inforoute.ca/References/Standards/CIHI.

### 1.2 Organization of This Specification

This specification was designed to document the extraction of the PHC EMR CS and to provide both detailed requirements and guidelines (or considerations) for stakeholders.

**Section 2** defines key high-level assumptions that informed the development of this specification.

**Section 3** identifies the overall strategy and approach for collecting and extracting a subset of PHC EMR data for health system use (HSU) purposes.

**Section 4** identifies considerations for EMR system vendors that are related to the implementation requirements supporting PHC data submission using the PHC EMR CS and PHC EMR Data Extract Specification.

**Section 5** provides detailed technical specifications for extracting PHC EMR data that can be used for HSU, including

- The required structure to implement a PHC EMR data extract (instantiated as an XML document);
- The approach used in the design of the PHC EMR Data Extract Specification schema; and
- The PHC EMR CS data elements–to-schema mapping.

---

i. Health system use (HSU) is defined as the use of health data to improve the health of Canadians and the health care system. It is commonly used synonymously with the term “secondary use.”
The appendices provide supporting information relevant to the PHC EMR Data Extract Specification, including XML schematics and examples. The technical content outlined may be accessed on Canada Health Infoway’s EMRi Wiki at https://emri.infoway-inforoute.ca/References/Standards/CIHI.

1.3 Scope

1.3.1 In Scope
The goals of the PHC EMR Data Extract Specification are to

- Provide detailed pan-Canadian specifications that are required for collecting and extracting data required for HSU;
- Provide context for submissions that support data flows to organizations that use the data for HSU purposes (henceforth referred to as HSU organizations); and
- Provide the logical information model and physical XML schema required to map data for PHC EMR data extracts that are required to support HSU.

1.3.2 Out of Scope
The PHC EMR Data Extract Specification does not include

- Jurisdiction-specific schemas;
- Information related to vendor product certification and criteria for acceptance;
- Algorithms for data encryption, de-identification and anonymization;
- Data transformation rules for HSU repositories;
- Extract, transformation and load scripts and functions; and
- Specifications of functionality required by an EMR system to gain access to data that may be sourced from another system (for example, any regional, jurisdictional or national electronic health record/health information access layer component, such as laboratory, drug, referral or immunization repositories).
1.4 Stability

This specification depends on a number of other specifications and factors that may affect its stability over time.

The PHC EMR Data Extract Specification relies heavily on the PHC EMR CS (described specifically in Section 5) and the terminology reference sets to support the standardized use of the PHC EMR data elements. Each artifact will have its own sustainment and maintenance life cycle. Changes to the PHC EMR CS or terminology reference sets may affect this specification; thus it is expected that related artifacts will need to evolve together. Please note that at the time this specification was published, not all terminology reference sets were defined.

The HL7 v3 Clinical Document Architecture Release 2 (CDA R2) specification was used as a starting point for many constructs in this specification, but the PHC EMR Data Extract Specification is not expected to conform to CDA R2 due to some intentional constraints. This deviation from CDA R2 is described in more detail in Section 5. Future or yet-to-be-identified changes to CDA R2 may suggest changes to this specification and thus may influence its evolution. As a result of aligning with CDA R2, the PHC EMR Data Extract Specification does not directly align with the Pan-Canadian Messaging Specifications (a localization of HL7 v3).

For questions or clarification on the PHC EMR Data Extract Specification, please contact CIHI by email at phc@cihi.ca.

1.5 EMR System Conformance

The PHC EMR Data Extract Specification does not have a formal conformance profile. An EMR system’s conformance will be based on the implementation requirements of any organization (for example, EMR vendors or jurisdictions) that implements the PHC EMR Data Extract Specification.

At a minimum, conformance is the proper implementation of the XML schema in vendor products. This would be tested by an EMR system’s ability to produce a PHC EMR data extract (instantiated as an XML document) that can be validated against the PHC EMR data extract schema.

It is also assumed that the recipient systems, including CIHI’s Primary Health Care Voluntary Reporting System (PHC VRS), will test any submission against the XML schema as a key criterion for accepting a data submission and during the parsing of each submission.
2 Assumptions

Key assumptions include the following:

1. Where relevant and appropriate, mechanisms to support patient and provider consent and the sharing of PHC data are handled through a combination of the EMR software and physical processes, policies and governance in support of jurisdictional legislation.

2. Non-coded data will evolve to a coded format as part of the planned EMR enhancements; thus the PHC EMR Data Extract Specification will need to support both coded and free-text (non-coded) representation of data.

3. Data collected through the EMR software must be edited at source (in the system it is entered in) or be transformed before extraction to adhere to the PHC EMR CS and the XML schema criteria specified as part of the PHC EMR Data Extract Specification.

4. Implementers of EMR systems are responsible for mapping local terminologies and vocabularies to the terminology reference sets of the PHC EMR CS. Transformations that will be required to convert local vocabularies to the specific standards in the PHC EMR CS will be part of the EMR system or of the submitting organization’s tooled process.

5. The PHC EMR CS and the PHC EMR Data Extract Specification will continue to evolve over time as implementations continue to drive requirements; thus the EMR software will need to support this evolution through regular maintenance and sustainment cycles.

6. EMR systems will need to align with clearly defined concepts that are intended for the purposes of HSU of PHC EMR data. In lieu of an accepted pan-Canadian model, the PHC EMR Data Extract Specification used CIHI’s reference data model as a basis for building relationships between the PHC EMR CS data entities.
3 Strategy and Approach

3.1 Conceptual-to-Physical EMR Implementation

A number of tools provided in this specification are meant to help EMR vendors implement the specification in EMR systems. These tools include a set of conceptual, logical and physical data models that can be used to map and align the PHC EMR CS and the PHC EMR Data Extract XML Schema to an EMR system’s existing models and schematics.

3.1.1 Conceptual Data Model

The following represents the PHC EMR Data Extract Specification’s conceptual data model. The model is used as a high-level representation of the relationships between the data concepts found in the PHC EMR CS. Each entity is further defined in the PHC EMR CS.

Figure 1: Conceptual Data Model

Note
Although not explicitly identified in this model, each subtype of the Observation entity is a Coded Observation, except for the single instance of Measured Observation.
To allow for HSU of data, the extraction of PHC EMR data must be representative of a patient-centric, longitudinal model to support data analysis based on an episode of care. It is expected that when comparing information in an existing EMR system with the PHC EMR CS for the purposes of designing an extract, the conceptual model can be used as an alignment foundation.

An approach for ensuring that the appropriate data is consistently aligned with a schema (as outlined in Figure 2) is the following:

1. Map the EMR system’s conceptual data model to the PHC EMR Data Extract Specification’s conceptual model, giving an overall understanding of the data relationship alignment.

2. Map the EMR system’s logical data model to the PHC EMR Data Extract Specification’s logical data model (found in Section 5), resulting in an understanding of data element alignment.

3. Use the PHC EMR CS–to–PHC EMR Data Extract Specification’s schema mapping (found in Section 5) to identify aligned PHC EMR data extract segments, providing the information to transform data.

4. Use the PHC EMR Data Extract XML Schema to produce an instantiated XML document.
4 EMR System PHC Data Submission Considerations

The following examples of data flows and an EMR system data extract submission business process are provided for consideration and to provide context for a jurisdiction or EMR system vendor. These considerations do not specifically influence the PHC EMR Data Extract Specification but would influence many implementation decisions.

4.1 Example Data Flow Scenarios

PHC EMR data may flow from PHC organizations to an HSU organization directly or through a series of HSU organizations. The PHC EMR Data Extract XML Schema will support each of the following examples and can maintain the representation of Custodian and Author:

- From the provider organization to an HSU organization;
- From the provider organization to a jurisdiction (as an HSU organization) and then to another HSU organization;
- From the provider organization to a collaborating group of HSU organizations (henceforth referred to as HSU networks) to an HSU organization; or
- From the provider organization (whose data is maintained by a vendor organization) to an HSU organization.

4.2 Example EMR System Data Extract Submission Business Process

The following model provides an example process for producing a PHC EMR data extract. It describes key considerations for EMR vendors when developing functionality to produce a data extract in accordance with the PHC EMR CS and the PHC EMR Data Extract Specification. The model also provides an example process for receiving data, which helps to complete the model but is out of context for the production of a data extract using the PHC EMR Data Extract Specification.
Figure 3: Example EMR System Data Extract Process

<table>
<thead>
<tr>
<th>Process</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DataSubmitter</strong></td>
<td><strong>DataReceiver</strong></td>
</tr>
<tr>
<td><strong>Start</strong></td>
<td><strong>Start</strong></td>
</tr>
<tr>
<td>1. EMR System Data Entry</td>
<td>7. Receiving System</td>
</tr>
<tr>
<td>2. EMR System Gets External Data</td>
<td>8. XML Extract Validation</td>
</tr>
<tr>
<td>4. Transformation and, if Required, De-identification</td>
<td>End</td>
</tr>
<tr>
<td>5. Generate XML Extract</td>
<td>Out of Context</td>
</tr>
<tr>
<td>6. Send</td>
<td></td>
</tr>
</tbody>
</table>
Each step in the data extract process describes a function that an EMR system would be required to handle to produce a PHC EMR data extract based on the PHC EMR CS.

1. PHC data is entered into the EMR system by clinicians. The structure of this data should align with the high-level data model described in Section 3 (see Figure 1). Please refer to the PHC EMR CS for data concept descriptions and vocabulary requirements. Section 5 further describes the potential data concept relationships that the extract supports.

2. PHC data may also include data that is transferred into the EMR system from other external sources, such as a lab system (for example, lab results), diagnostic imaging repositories (for example, diagnostic imaging reports), provincial electronic health record (EHR) repositories (clinical documents, allergies, etc.) or other EMRs (clinical summaries, referrals, etc.). Depending on jurisdictional policies or requirements, this information may or may not be included in the PHC extract. The schema does support the appropriate identification of the originating source of such information while maintaining the submitter as a custodian organization and should be used when representing external sources of data.

3. At a regular interval (such as monthly, quarterly or yearly), the EMR user (or system operator) will be able to extract the PHC EMR data. Jurisdictional implementation requirements may dictate that EMR systems automate a scheduled extract. A manually or electronically scheduled extract requires an EMR system to determine
   a) Which patients should be included—all patients that have had any changes to their PHC data since the last extract; and
   b) What PHC data should be included—all PHC data for that patient according to the PHC EMR CS.

4. Prepare data according to the PHC extract format and terminology as described in the PHC EMR CS and PHC EMR Data Extract XML Schema Definition (see Section 5). If the EMR system does not use the prescribed PHC EMR CS terminology reference sets, the EMR system will have to perform the appropriate data transformations.

The manner in which this transformation occurs is outside the scope of this specification. This will usually require the EMR system to use a pre-mapped, locally entered system terminology—to—standardized terminology as a basis for coded or configured transformations. For each of the PHC EMR Data Extract Specification data elements that requires a coded term, a corresponding original text can be included with the extract, thus supporting appropriate quality control on the transformation into a coded data element.
Any jurisdictionally or organizationally required de-identification or consent directive–based masking would also take place during the transformation step of the extract process. Please refer to Section 4.3 for further explanation of applying de-identification and consent directives. The user of the personal health information should collect only what is necessary for the purpose identified by the HSU organization.

5. Create the PHC EMR data extract (XML instantiation) according to Section 5. It can be persisted locally in preparation for its submission to an HSU organization. Depending on jurisdictional or organizational policy, the data may require encryption (above and beyond any de-identification or consent directive–based masking applied).

Other jurisdictional or organizational requirements may include very specific formatting for a file or web service–based submission. Below is an example set of requirements based on CIHI’s PHC VRS file submission process:

a) Example File Format: The PHC VRS will accept data in XML format conformant to the XML schema defined in Section 5 of the PHC EMR Data Extract XML Schema definition.

b) Example File Naming Conventions: File names are standardized to facilitate the receipt and processing of PHC VRS data. This file naming convention must be adhered to or the file will not be accepted. If the file name does not meet the specifications outlined in the eData Submission Services (eDSS) agreement, the submitting organization will be notified either through eDSS notification or by CIHI staff.

The file name will be in the following format:

PHCyyyypphhhhhsss.zip

Where

PHC = Primary Health Care (static title in the file name)

yyyy = Fiscal Year

pp = Submission Period (01 to 12, referencing a month of the fiscal year, April to March)

hhhhh = Submission Organization Identifier

sss = Submission Sequence Number (001, 002, 003, etc.)
c) Example Submission Sequence Number: This sequence number is used by CIHI to process the files in order.

It is recommended that at the beginning of each submission period the Submission Sequence Number used in the file names be set to 001.

It is important that within each Submission Period a Submission Sequence Number be used only once. If two files are submitted with the same file name, the most recent file will supersede all previously submitted files.

6. The PHC extract persistence is then ready to be submitted either to the corresponding jurisdictional PHC HSU organization or directly to an HSU organization like CIHI. For further information on transport mechanisms, refer to Section 4.5: Transport Layer Interoperability.

4.3 Treatment of Sensitive Data

The PHC EMR Data Extract Specification does not specify or require any de-identification or consent directive–based masking for its use.

Jurisdictional legislation and organizational policy will dictate the implementation requirement to apply appropriate de-identification and consent directive–based masking techniques to sensitive data, such as personal health information or associated personal details. For example, provider name and other personal details should be extracted only where appropriate and with the proper approvals. EMR vendors should plan to account for these requirements in their systems. Furthermore, the user of the personal health information should collect only what is necessary for the purpose identified by the HSU organization.

4.4 Security Controls

The PHC EMR Data Extract Specification does not specify or require any encryption or data transfer protocols for its use.

Secure file encryption and data transfer protocols are required by jurisdictional and organizational implementations to ensure that the privacy and security of PHC data is not compromised in any way. Vendors will be required to implement technical security safeguarding mechanisms based on a jurisdiction’s or organization’s implementation needs; these mechanisms would likely be a part of a jurisdictional or organizational implementation guide. Depending on the implementation model, administrative, physical and technical security safeguarding mechanisms will be a jurisdiction’s, organization’s or vendor’s responsibility (or any combination thereof).
4.4.1 Privacy

Users of the PHC EMR Content Standard, Version 2.0, including the Data Extract Specification and the Implementation Guide, should comply with the 10 privacy principles established in the Canadian Standards Association’s Model Code for the Protection of Personal Information, as well as the relevant jurisdictional privacy legislation and guidance provided by privacy oversight bodies.

Users of the PHC EMR Data Extract Specification need to ensure that they respect their organizations’ privacy policies and practices and that they meet the required standards in safeguarding the important and sensitive information they are trusted with.

4.5 Transport Layer Interoperability

The PHC EMR Data Extract Specification does not pre-determine or prescribe any specific transport layer requirements.

It is expected that the method used for moving or transmitting a PHC EMR data extract XML document instantiation (either in file or message) will be decided at implementation; the method would likely be a part of a jurisdictional or organizational implementation guide. An XML document instantiation of a PHC EMR Data Extract XML Schema may be transported as an encrypted file, for example, as required by CIHI’s PHC VRS (over a encrypted tunnel [HTTP/SSL] in a compressed [zipped] form) or via an electronic message using web services for another receiving HSU organization.

4.6 Error Handling

The PHC EMR Data Extract Specification does not prescribe any specific automated error handling capabilities.

Implementation-specific requirements for automated error handling are expected; these would likely be part of a jurisdictional or organizational implementation guide. These would likely include the following:

1. Schema conformance—A PHC EMR Data Extract Specification instantiated as an XML document needs to strictly conform to the PHC EMR Data Extract Specification XML schema as specified in Section 5. Any non-conformant XML document cannot be processed by a receiving HSU organization. For example, CIHI’s PHC VRS, using the PHC EMR Data Extract Specification, would require a level of conformance to the PHC EMR CS and to the PHC EMR Data Extract XML Schema. This conformance check would include data type, element size and relationship cardinality.
2. Transmission errors—Any interruption or other error with a transmission of a properly formed PHC EMR data extract XML document will be handled by an EMR system in accordance with the implementation specification of the transport layer. Depending on the implementation, EMR system support of an automated process for handling transport layer interoperability may be required to deliver a single PHC EMR data extract, including notification support for human intervention.

3. Constrained vocabulary errors—The submitting and receiving organizations will need to negotiate error resolution for non-conformant vocabulary within an implementation context.

5 PHC EMR Data Extract XML Schema Definition

Refer to Canada Health Infoway’s EMRi wiki for the content of this section, which includes the following:

- Overview and Approach
- Data Elements
- Logical Data Model
- PHC EMR Data Extract Schema Model
- XML Schema
- XML Examples

The technical content outlined may be accessed on Canada Health Infoway’s EMRi Wiki at https://emri.infoway-inforoute.ca/References/Standards/CIHI.