North American Association of Central Cancer Registries, Inc.

Interoperability Ad Hoc Committee
Clinical Data Work Group

Protocol for the

Pilot Project of Hospital Cancer Registry Data Transmission to the Central Cancer Registry Using Health Level Seven (HL7) Clinical Document Architecture (CDA)

April 25, 2007

Liora Alschuler
Sanjeev Baral
Larry Derrick
Steve Fuschlin
Ken Gerlach
Barry Gordon
Lori Havener
Carol Kosary
Jim Martin
Marsha Reichman

Table of Content

1   Background

NAACCR Standards for Cancer Registries, Volumes I and II, have provided detailed information for the NAACCR data exchange record layout for over 10 years. As health information technology has evolved over the years NAACCR has continued to utilize the same flat-file format to exchange data. Although the current format has served its purpose over the years there are limitations to the flat-file format (e.g., adding new data items, transmitting text fields that are longer than the allowable values, changing the character length of an existing data item, etc.) and the NAACCR date data items are not consistent with national standards (e.g., NAACCR date fields include codes [0s, 8s, and 9s] other than dates).

In January 2007 during the NAACCR Leadership Retreat, it was determined that NAACCRs number one priority is to achieve syntactic and semantic interoperability of cancer registration standards with national standards. The Interoperability Ad Hoc Committee was initiated by the Board in February. During this time the IT Committee was restructured to elevate current activities related to the interoperability project. The Cancer Abstract Transmission Work Group has been renamed the Clinical Data Work Group and moved under the Interoperability Ad Hoc Committee.

One of the NAACCR Information and Technology (IT) Committee objectives is to explore alternate mechanisms to transmit and receive the cancer abstract. In November 2005 the IT committee convened the Cancer Abstract Transmission (CAT) Work Group (WG) to explore alternative mechanisms or messages to transmit and receive the cancer abstract. The CAT WG (from this point on referred to as the Clinical Data WG) developed a matrix to review the core criteria and software specific criteria of selected data transmission tools. The tools evaluated: Health Level Seven (HL7) version 2.X SML, HL7 version 3, HL7 Clinical Document Architecture (CDA), and Comma Separated Value (CSV). Criteria were scored (see appendix A) for each tool to determine which would be assessed as an alternative mechanism or message to transmit data and CDA was selected.

The development of this protocol is to establish a method to pilot test the utility of HL7 CDA as a data transmission mechanism to transmit data from a hospital cancer registry to the central cancer registry. Future phases of this project will review the impact on software tools and continue to assess the impact on hospital cancer registries, central cancer registries, standard-setting organizations, etc.

2   Overview

The initial phase of this project will be to pilot test transmission of a cancer abstract using HL7 CDA from a hospital cancer registry to the central cancer registry. This project will evaluate: the transmission of specific data fields and text fields; the impact of a new transmission on software systems (hospital and central); the ability to visually review data without special software; ease of adding and revising data items; compatibility with software tools (e.g., Transform Tool); data security; consistency with national health information technology standards; and, costs. The transmission of data from central cancer registry to central cancer registry and central cancer registry to national programs will not be tested in this phase of the pilot project.

An issue with the current NAACCR data exchange layout is the limited character length of text fields. This study will also assess the ability to transmit all text that is received (i.e., text fields will not be truncated) using the HL7 CDA transmission format.

The American College of Surgeons (ACoS) Commission on Cancer (CoC), the Centers for Disease Control and Prevention (CDC) National Program of Cancer Registries (NPCR), and the National Cancer Institute (NCI) Surveillance Epidemiology and End Results Program (SEER) establish data collection and transmission requirements recorded in the NAACCR Standards for Cancer Registries, Volume II, *Data Standards and Data Dictionary*. This project includes the requirements of each of these standard-setting organizations.

3 Participant Selection

Abstract Plus and C/NExT, hospital cancer registry software vendors, and Rocky Mountain Cancer Data System (RMCDS) and Eureka, central cancer registry vendors, have agreed to pilot test HL7 CDA messaging transmission from a hospital cancer registry to the central cancer registry.

Eureka will be working with the California Cancer Registry and two hospital cancer registries, using CNExT, to participate in the pilot project (see section 4.2 for transmission method). RMCDS will be working with the Virginia Cancer Registry (VCR) and one hospital cancer registry, using Abstract Plus, to participate in the pilot project (see section 4.3 for transmission method).

4 Study Design and Method

The pilot test will include two central cancer registry vendors, two hospital cancer registry vendors and three hospitals. Hospital registries will transmit data to the central cancer registries in the NAACCR record type A (full case abstract = incidence and confidential data plus text summaries) as stated in the Standards for Cancer Registries, Volume II, *Data Standards and Data Dictionary* Version 11.1.

All required (R) and required when available (R*) data items as noted in the Standards Volume II, Version 11.1, Chapter VIII, Required Status Table will be transmitted. This will include CoC, NPCR and SEER requirements. Dummy data will be used in the initial phase of the project. The table in appendix B includes the different transmission forms the vendors will utilize to transmit data.

4.1 HL7 CDA Implementation Guide

The Clinical Data WG is developing an HL7 CDA draft implementation guide to pilot test HL7 CDA transmitting data from a hospital cancer registry to a central cancer registry. The CDC-NPCR has a contractual agreement with Alschuler Associates, LLC to provide HL7 CDA (Release 2) subject matter expertise and technical assistance for the development of the implementation guide for the transmission of the cancer registry abstract report.

The HL7 CDA implementation guide describes how to code CDA documents for exchanging cancer abstracts (CDA elements with conformance rules). It is intended for implementers who are familiar with the NAACCR record layout and are now creating support for CDA. The implementation guide does not replace the NAACCR Standards for Cancer Registries, Volume II: Data Standards and Data Dictionary.

4.2   CNExT and Eureka (California Cancer Registry) Data Transmission Method

The participants in California plan to use the targeted transmission form (hospital registry HL7 CDA -> central registry HL7 CDA) as noted in #6 of appendix B. Two hospitals using CNExT software will directly export test case reports in the HL7 CDA format and transmit them to the California Cancer Registry using Eureka software to load them in to a test system.

4.3   Abstract Plus and RMCDS (Virginia Cancer Registry) Data Transmission Method

The primary test that will be performed is the transmission from the hospital system (which will be using Abstract Plus) to the VCR. Batch files will be loaded into the VCR system centrally using RMCDS. The procedure will start at the hospital where Abstract Plus will export abstracts in the NAACCR flat-file. The file will be converted at the hospital using the transformation software into the HL7 CDA format. Depending on time, resources, and other factors, the approach of having abstract plus directly generate an HL7 CDA format file to be transmitted to the VCR may be tested.

The method of transporting the file to VCR has not been fully decided, but should not affect the testing goal. If Virginia department of health information technology can establish a secure web site by the time of the test, this means will be used for data transmission. Alternate approaches for the secure transfer of data between the hospital and the VCR are available should the secure web site not be available. The transmission protocol is being worked on at the hospital. The transmission will be very likely batch and not single records.

When the file(s) arrive at the VCR the HL7 CDA data will be transformed and loaded into the VCR database. Two approaches will be tested.

The first approach will be to use transformation software to convert the HL7 CDA format file back into a NAACCR flat-file, which will be loaded into the VCR database. The original hospital abstracts will be compared to the data that ends up on the VCR database. Any differences that appear will be traced through the process to determine how and why they occurred.

The second approach that will be tested is to have RMCDS software directly transform the HL7 CDA format file and load it into the VCR. This will give RMCDS an opportunity to evaluate the feasibility of capturing additional information from the file. This information could be additional text in the text fields or repeated information in certain fields. Like the procedure outlined above, abstracts originating in the hospital would be compared to those ending up on the VCR database.

RMCDS will also be investigating the feasibility and difficulty of putting hospital abstracts directly into the HL7 CDA format. They will investigate what the approximate time, effort, and cost might be to implement going directly to HL7 CDA format and reading directly in HL7 CDA format files.

4.4   Description of Data Being Transmitted

All required (R) and required when available (R*) data items as noted in the Standards Volume II, Version 11.1, Chapter VIII, Required Status Table will be transmitted. This will include CoC, NPCR and SEER requirements.

California will transmit the full contents of the NAACCR Standards Volume II, Version 11.1 new case abstract as CDA. We will extend this in two ways: 1) include California required extra fields and 2) transmit longer text blocks than the required NAACCR maximum. Registrars will know in advance which text will be truncated when received by systems that haven't implemented long

text yet. We will carry out test transmits using both single cases and large numbers of cases and will compare the results with parallel transmits using the flat file standards. We will compare data between systems after transmit to make sure our mapping rules do not mistranslate data.

4.5   Data Security

Central cancer registries are required to protect the privacy of the individual patient and the reporting facilities. Measures will be taken to ensure the security of confidential data.

4.5.1   CNExT and Eureka (California Cancer Registry) Data Security

California data will be transmitted using an existing https-based encrypted application-to-application method between CNExT in the hospitals and the California Cancer Registry Eureka servers.

4.5.2   Abstract Plus and RMCDS (Virginia Cancer Registry) Data Security

If the Virginia department of health information technology can establish a secure web site by the time of the test, this means will be used for data transmission. Alternate approaches for the secure transfer of data between the hospital and the VCR are available should the secure web site not be available.

4.6   Software Tools

4.6.1   Transformation Software Tool

The initial implementation of a new data transmission mechanism will be phased-in by hospital registries and central registries at different time intervals. It will be important to run parallel systems (flat-file format and CDA) until all registries have implemented the new data transmission mechanism. Data conversion is necessary to allow this phase-in process.

As part of the contractual agreement with Alschuler Associates, LLC, a transformation software tool will be developed which will convert a NAACCR flat-file format (Standards Volume II Version 11.1) into the CDA format specified in the implementation guide and vice versa. This pilot will test this tool in Virginia where the NAACCR flat-file will be generated and then converted to the CDA format using the transformation software tool. The transformed CDA format will be transmitted from the hospital registry to the central registry. RMCDS will receive the CDA formatted data and will transform the CDA format to the NAACCR flat-file format using the transformation software tool. The NAACCR flat-file format sent can be compared to the NAACCR flat-file converted in the central cancer registry. See appendix B, Pilot Verification.

4.6.2   Tools to be Evaluated Later

The initial phase of this pilot project is to evaluate data transmission from the hospital cancer registry to the central cancer registry. Once this is completed and data transmission issues have been resolved another phase of the project will be to evaluate other software tools.

4.6.2.1   SEER*Prep

Work Group participants from IMS indicate that SEER*Prep can be configured to ready any format and that they do not anticipate significant challenges in this area. However, this pilot project will focus on the transmission from a hospital cancer registry to a central cancer registry and will not test the capability of SEER*Prep to read a HL7/CDA formatted file.

4.6.2.2 EDITS

The developers of EDITS plan to adjust the tool to read an XML formatted file, but do not plan to initiate development until an XML prototype is in place. CDA is an XML formatted file and could serve as a prototype in this development. However, this pilot project will not test the capabilities of EDITS to read a HL7/CDA formatted file.

4.7 National Health Information Technology Standards

The NAACCR Program Manager of Standards will participate on the HITSP, Systematized Nomenclature of Medicine (SNOMED) Surgical Pathology working group, Public Health Data Standards Consortium (PHDSC), and the HL7 Anatomic Pathology Special Interest Group (SIG) providing updates pertinent to this project.

HL7 vocabularies will be reviewed through the development of the HL7 CDA implementation guide for this pilot project. Although this is a pilot project, discrepancies between HL7 and NAACCR standards that have been identified from the HL7 vocabulary review will be forwarded to the Semantic WG for their review.

4.8 Advantages and Disadvantages

The Clinical Data WG will create a report on implementation issues identified during the pilot project. Implementation issues that will impact registry operations will be addressed through a work group including representatives from the Registry Operations Committee.

Success will be determined by the ability to produce, transmit, receive and process records, obtaining the same results, in terms of information content, as under legacy systems. Barriers to accomplishment of this goal will be documented. The report will include determination of the time involved to accomplish transmission (including needed translations) and a comparison to the legacy system. Modifications to the legacy systems needed to accomplish goals will be discussed and these will be ranked in terms of how major they are considering resources required.

4.9 Financial Costs and Non-Financial Benefits Assessment

The Clinical Data WG will create a report on the assessment of the financial costs and non-financial benefits to implementing a new data transmission mechanism. This assessment will consider personnel (FTEs), software and hardware costs as well as look at non-financial benefits (e.g., the ability to transmit expanded data sets).

5   **Appendix A Core Criteria and Software Specific Criteria Review**

NAACCR Clinical Data Work Group Comparison Table to Assess the Different Formats				
Criteria	**HL7 2.x XML**	**HL7 V.3**	**HL7 CDA**	**CSV**
Compatibility with statistical software	2	2	2	2
Ability to handle repetitions and other structures	3	3	3	1
Ability to transmit large text data	3	3	3	2
Ease of visually reviewing the data without special software	3	3	3	1
Consistency with national (USA) healthcare industry formats	3	3	3	2
Ease of adding and revising data items (cost)	3	3	3	2
Self-identifying	3	2	3	0
Internet friendly or compatibility with Internet systems	3	3	3	3
Availability of development tools e.g. NeoTools.	3	3	3	1

0 – Not Possible
1 – Possible/Minimal Support
2 – Supported
3 – Strongly Supported

6 Appendix B Transmission Forms from Flat-File to HL7 CDA

NAACCR Cancer Abstract Report (Full, Record Type 'A'): Transition From Flat-File to HL7 CDA

Revision date: 12-08-2006

| Hospital Registry | | Central Registry |

Current

| 1 : Abstract |
| + Format = flat |

→ Send → | 1 : Abstract |
| + Format = flat |
→ Add to DB

Pilot – Verification

| 2a : Abstract |
| + Format = flat |
→ Convert → | 2b : Abstract |
| + Format = HL7/CDA |
→ Send → | 2b : Abstract |
| + Format = HL7/CDA |
→ Convert → | 2a : Abstract |
| + Format = flat |
→ Add to DB

Intermediate

| 3a : Abstract |
| + Format = flat |
→ Convert → | 3b : Abstract |
| + Format = HL7/CDA |
→ Send → | 3b : Abstract |
| + Format = HL7/CDA |
→ Add to DB

| 4a : Abstract |
| + Format = flat |
→ Send → | 4a : Abstract |
| + Format = flat |
→ Convert → | 4b : Abstract |
| + Format = HL7/CDA |
→ Add to DB

| 5 : Abstract |
| + Format = HL7/CDA |
→ Send → | 5 : Abstract |
| + Format = HL7/CDA |
→ Convert → | 5a : Abstract |
| + Format = flat |
→ Add to DB

Targeted

| 6 : Abstract |
| + Format = HL7/CDA |
→ Send → | 6 : Abstract |
| + Format = HL7/CDA |
→ Add to DB