Are Meaningful Use Stage 2 certified EHRs ready for interoperability? Findings from the SMART C-CDA Collaborative

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation

Published Version
doi:10.1136/amiajnl-2014-002883

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:13454846

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Are Meaningful Use Stage 2 certified EHRs ready for interoperability? Findings from the SMART C-CDA Collaborative

John D D’Amore,1,2 Joshua C Mandel,3,4 David A Kreda,5 Ashley Swain,1 George A Koromia,1 Sumesh Sundareswaran,1 Liora Alschuler,1 Robert H Dolin,1 Kenneth D Mandl,3,4,6 Isaac S Kohane,3,6 Rachel B Ramoni6,7

ABSTRACT
Background and objective Upgrades to electronic health record (EHR) systems scheduled to be introduced in the USA in 2014 will advance document interoperability between care providers. Specifically, the second stage of the federal incentive program for EHR adoption, known as Meaningful Use, requires use of the Consolidated Clinical Document Architecture (C-CDA) for document exchange. In an effort to examine and improve C-CDA based exchange, the SMART (Substitutable Medical Applications and Reusable Technology) C-CDA Collaborative brought together a group of certified EHR and other health information technology vendors.

Materials and methods We examined the machine-readable content of collected samples for semantic correctness and consistency. This included parsing with the open-source BlueButton.js tool, testing with a validator used in EHR certification, scoring with an automated open-source tool, and manual inspection. We also conducted group and individual review sessions with participating vendors to understand their interpretation of C-CDA specifications and requirements.

Results We contacted 107 health information technology organizations and collected 91 C-CDA sample documents from 21 distinct technologies. Manual and automated document inspection led to 615 observations of errors and data expression variation across represented technologies. Based upon our analysis and vendor discussions, we identified 11 specific areas that represent relevant barriers to the interoperability of C-CDA documents.

Conclusions We identified errors and permissible heterogeneity in C-CDA documents that will limit semantic interoperability. Our findings also point to several practical opportunities to improve C-CDA document quality and exchange in the coming years.

BACKGROUND AND SIGNIFICANCE
Health Level 7 (HL7), a leading standards development organization for electronic health information, defines interoperability as ‘the ability of two parties, either human or machine, to exchange data or information where this deterministic exchange preserves shared meaning.’12 In addition, semantic interoperability has been operationally defined to be ‘the ability to import utterances from another computer without prior negotiation, and have your decision support, data queries and business rules continue to work reliably against these utterances.’12

In our study, we apply the operational definition of semantic interoperability to assess structured data within Consolidated Clinical Document Architecture (C-CDA) documents, which certified electronic health record (EHR) systems must produce to satisfy federal regulation of EHR adoption. We study core variation in document samples to examine if reliable semantic interoperability is possible.

EHR adoption and Meaningful Use
EHR use in the USA has risen rapidly since 2009 with certified EHRs now used by 78% of office-based physicians and 85% of hospitals.1–4 Meaningful Use (MU), a staged federal incentive program enacted as part of the American Recovery and Reinvestment Act of 2009, has paid incentives of US$21 billion to hospitals and physicians for installing and using certified EHRs pursuant to specific objectives.5–8 Stage 1 of the program (MU1) commenced in 2011, Stage 2 (MU2) in 2014, and Stage 3 is expected by 2017.

While the term interoperability can refer to messages, documents, and services, MU provides several objectives that prioritize document interoperability. Although multiple document standards existed prior to MU1, providers with installed EHRs rarely had the capability to send structured patient care summaries to external providers or patients, as noted by the President’s Council of Advisors on Science and Technology and the Institute of Medicine.8,9 MU1 advanced document interoperability by requiring Continuity of Care Document (CCD) or Continuity of Care Record (CCR) implementation as part of EHR certification. Many vendors chose the CCD, which was created to harmonize the CCR with more widely implemented standards.10 11 In MU2, the C-CDA, an HL7 consolidation of the MU1 CCD with other clinical document types, became the primary standard for document-based exchange.12

C-CDA use in document interoperability
The C-CDA is a library of templates using extensible markup language (XML) to transmit patient-specific medical data in structured and unstructured formats.13 It builds upon the HL7’s Clinical Document Architecture release 2.0 (CDA) and the Reference Implementation Model (RIM), a consensus view of how information can be abstractly represented.14 The CDA constrains the RIM by...
applying principles of how to represent information in clinical documents. The C-CDA Implementation Guide 1.1 describes how to create nine CDA document types (table 1), each a combination of specific sections (eg, problems, allergies) and entries (eg, diagnosis of heart failure, medication allergy to penicillin). Moreover, different documents types (eg, a history and physical vs discharge summary) share common sections to achieve consistency in data representation.

MU2 objectives include the use of C-CDA documents for both human display and machine-readable data exchange. Since C-CDA implementation guidance requires both data structured in XML and specific terminologies, healthcare providers can generate machine-readable documents for individual care transitions and across a practice to prevent EHR vendor lock-in. Previous research has cataloged issues associated with past interoperability standards, but research specific to C-CDA is still limited given the nascent utilization of the standard.

**OBJECTIVES**

Those of us (JCM, DAK, KDM, ISK, RBR) involved in the SMART (Substitutable Medical Applications and Reusable Technology) Platforms Project, an Office of the National Coordinator for Health Information Technology (ONC)-funded Special Health Advanced Research Project, have been exploring ways to integrate medical apps across diverse EHRs. To assess the current state of C-CDA interoperability and prepare recommendations to improve document quality, the SMART team engaged Lantana Consulting Group in April 2013 to form the SMART C-CDA Collaborative.

The Collaborative approached health information technology vendors for a study of C-CDA quality and variability. Vendors who participated in the Collaborative provided 2011 Certified EHR Technology for a majority of provider attestations for MU1 from 2011 to 2013. While several vendor applications received 2014 EHR certification before joining the Collaborative, most received it during the Collaborative’s term, which ended in December 2013. To identify both application-specific and general means to improve the quality of C-CDA documents, we engaged vendors in discussions and document reviews to refine our analysis, as well as to hear how and why vendors made certain implementation decisions. Our interaction with vendors may have influenced the quality of C-CDA documents used during certification, but many reported that the feedback from discussion would only be incorporated into future application releases.

**MATERIALS AND METHODS**

**Vendor outreach**

We e-mailed invitations to organizations listed on the Certified Health IT Product List (http://onchpl.force.com/ehrcert). In cases where SMART or Lantana had prior contact with individuals within vendor organizations, we sent personal invitations. We posted public announcements on the SMART Platforms website and on the HL7’s Structured Document Working Group mailing list. We provided further details to interested organizations by phone, informing them of the means for sample collection and group discussions.

**Collection of samples**

As a condition of participation, we required vendors to submit at least one C-CDA document that had been generated by exporting a fictional patient’s health record from their software application. To allay concerns, we allowed submitted documents to be kept private, but nonetheless encouraged vendor participants to select a sharing policy that included public posting to a GitHub repository managed by Boston Children’s Hospital (https://github.com/chb/sample_ccdas).

**Automated parsing of samples**

C-CDA samples were parsed using the open-source BlueButton.js tool v0.0.19, to which one of the authors (JCM) has contributed. We have previously used BlueButton.js to integrate C-CDA data into medical applications. Using node.js, we parsed each C-CDA twice: a first pass wrote all data to a text file, and a second pass only recorded parsing times to isolate file writing time artifacts. All processing was performed on a quad-core AMD 2.2 GHz workstation with 6 Gb of RAM running Windows 7 (Microsoft, Redmond, Washington, USA). We counted each non-null section and JavaScript Object Notation data elements returned from the parser.

**Manual analysis of samples**

While only vendor-supplied C-CDA samples were part of the formal analysis, C-CDA documents from HL7 and other non-vendor organizations were reviewed for comparison to collected samples. Two of the authors (JDD, AS) performed the manual

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Data domains required by each C-CDA document type</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-CDA document type</td>
<td>Data domains</td>
</tr>
<tr>
<td></td>
<td>Demographics</td>
</tr>
<tr>
<td>Continuity of care document</td>
<td>✓</td>
</tr>
<tr>
<td>Consultation note</td>
<td>✓</td>
</tr>
<tr>
<td>Diagnostic imaging report</td>
<td>✓</td>
</tr>
<tr>
<td>Discharge summary</td>
<td>✓</td>
</tr>
<tr>
<td>History and physical note</td>
<td>✓</td>
</tr>
<tr>
<td>Operative note</td>
<td>✓</td>
</tr>
<tr>
<td>Procedure note</td>
<td>✓</td>
</tr>
<tr>
<td>Progress note</td>
<td>✓</td>
</tr>
<tr>
<td>Unstructured document</td>
<td>✓</td>
</tr>
</tbody>
</table>

Only the required domains are shown for each C-CDA document type. Additional information required by MU2 (ie, care team, functional and cognitive status, plan of care goals and instructions, immunizations, and referral information) are also supported in C-CDA documents. Because C-CDA documents are open templates, vendors may add optional data domains in order to meet regulatory and business requirements.

C-CDA, Consolidated Clinical Document Architecture; MU, Meaningful Use.
inspection, adapting techniques previously used for analysis of CDA documents.\(^5\)

Only a single C-CDA document submission was required to participate in the SMART C-CDA Collaborative. To give equal weighting in our analysis to each vendor application, when multiple samples were submitted from a single application, we selected the one with as many domains as possible and largest in kilobytes, together a proxy for the most data. The manual inspection identified errors and heterogeneity in the studied samples, but was confined to seven domains from the ‘Common Data Set’ defined in MU2: demographics, problems, allergies, medications, results, vital signs, and smoking status.\(^4\)

We defined an error in our study as any XML usage that conflicted with mandatory guidance from the HL7 C-CDA 1.1 Implementation Guide.\(^13\) Given this definition, any document with an error would not satisfy MU2 requirements for document interoperability. While many errors can be identified by automated software tools, some require human review (eg, where the dose of a structured medication entry contradicts dosing information in the narrative).

Identifying heterogeneity in structured data meant finding variations in data inclusion, omission, or expression across examined documents that did not qualify as errors defined above. Again, while some heterogeneity can be detected by automated software tools, human reviewers identified other types of heterogeneity which are currently not identifiable by software (eg, the omission of result interpretation as structured information when known from value and reference range).

Our inspection recorded only the first instance of any specific error or heterogeneity found in each domain of each sample. Recording repeated instances of the same issue in an individual C-CDA would document data frequency and not prevalence of error types.

We mapped observed errors and heterogeneity to one of six mutually exclusive categories: (1) incorrect data within XML elements; (2) terminology misuse or omission; (3) inappropriate or variable XML organization or identifiers; (4) inclusion versus omission of optional elements; (5) problematic reference to narrative text from structured body; and (6) inconsistent data representation.

Automated analysis of samples
Automated analysis of the samples made use of the Transport Testing Tool (TTT) release V175 (http://transport-testing.nist.gov/ttt) from the National Institute of Standard and Technology (NIST) and the SMART C-CDA Scorecard (http://ccda-scorecard.smartplatforms.org) from one of the authors (JCM).

TTT returns schema and schematron errors and warnings describing the conformance of a C-CDA document to the XML templates and conformance statements published by HL7.

The SMART C-CDA Scorecard performs a set of semantic checks that official validation tools omit. These checks include the validation of RxNorm, Systematized Nomenclature of Medicine (SNOMED), Logical Observation Identifiers Name and Codes (LOINC), and the Unified Code for Units of Measure (UCUM) use within a C-CDA document. The Scorecard computes a series of rubrics, each corresponding to a best practice for C-CDA implementation derived from discussion on an HL7 community mailing list. For example, two rubrics are: ‘Document uses official C-CDA templateIds whenever possible’ and ‘Vitals are expressed with UCUM units.’ The Scorecard assigns a score from zero to five for each rubric, allowing partial credit for documents with incomplete adherence to each rubric. No score is assigned for a rubric if no relevant data are available.

These scores are combined into section-wide scores by dividing the number of points earned by the total points possible. A composite score reported as a percentage (0–100%) is produced by summing the number of points earned across sections and dividing by the total points possible.

Group web conferences
From July through December 2013, SMART and Lantana conducted ten 60-min group meetings to discuss C-CDA implementation. The protocol consisted of a short review of issues identified through analysis of the collected samples and polling each health information technology vendor to respond to each issue (eg, ‘When do you include reference ranges as structured elements vs text strings?’). Written notes, compiled by one of us (JDD) for each meeting, were published weekly on a participant message board, which allowed for feedback between meetings (https://trello.com/b/CicwWIdW/smart-c-cda-collaborative).

One-on-one vendor reviews
From September through December 2013, SMART and Lantana scheduled sessions with individual vendors to review their respective C-CDA samples. Reviews covered specific observations about errors and explored variation in C-CDA data representations. Each vendor could request a second session and submit an additional C-CDA sample. Vendor feedback from these sessions was blinded when used in subsequent group discussions.

RESULTS
Vendor outreach
Of the 107 individual organizations contacted, 44 (41%) responded to the invitation. Fourteen organizations submitted one or more samples from a single application and one organization provided multiple samples from three separate technologies. Several respondents did not submit a C-CDA sample. Supplemental samples came from four organizations who had openly published their C-CDAIs. In total, 91 C-CDA documents were collected with an average of 4.3 (range 1–20) documents per vendor application. Samples were categorized (table 2) by whether the vendor application had been certified for MU2 by study conclusion.\(^23\)

Automated parsing of samples
All 91 samples were parsed using BlueButton.js. Parsing results omitted smoking status because BlueButton.js does not support the C-CDA section of social history. Since not every C-CDA

| Table 2 SMART C-CDA Collaborative tallies: vendors, applications, and C-CDA samples |
|----------------------------------|--------|-------|
| MU2 certification status as of December 2013 | Vendors | Applications | C-CDA samples |
| Certified EHR | 12 | 14 | 55 |
| Certified modular EHR for C-CDA exchange (HIE) | 3 | 3 | 13 |
| Non-certified health IT | 4 | 4 | 23 |
| Total | 19 | 21 | 91 |

Results are categorized by the certification status of a vendor’s application as of December 2013 but the C-CDA samples submitted by a vendor may have been different from those submitted for EHR certification.

C-CDA, Consolidated Clinical Document Architecture; EHR, electronic health record; HIE, health information exchange; IT, information technology; MU, Meaningful Use; SMART, Substitutable Medical Applications and Reusable Technology.
We applied the TTT to each of the 21 samples that had been manually inspected and observed:

- Ten vendor applications returned no errors.
- The remaining 11 had an average of 71 errors (range of 2–297) with the higher values being observed among non-certified vendor applications.
- Warnings were issued for all samples, generally for omission of XML elements, with an average of 78 warnings (range of 7–381) per vendor application.

We submitted the same samples for scoring by the SMART C-CDA Scorecard, obtaining an average score of 63% (range 23–100%; figure 2). As expected, no correlation ($R^2<0.01$) was observed between TTT results and SMART C-CDA Scorecard scores because they examine wholly different aspects of C-CDA document correctness. De-identified group results were presented publically and identified results were shared during individual vendor sessions.

**Group and one-on-one vendor web conferences**

Of the 19 organizations represented by C-CDA samples, 12 attended at least one group call. Six organizations who did not submit a sample during the outreach also joined the group calls. On average, eight organizations participated in each group call. Eleven organizations discussed their samples in one-on-one sessions with the research team, and three requested a second session. Individual sessions averaged 66 min (range 30–90 min) for a total of 930 min. Five organizations submitted revised samples to the Collaborative.

**Summation: common trouble spots in C-CDA samples**

Based upon our analysis and discussions with Collaborative participants, we identified 11 specific areas (i.e., ‘trouble spots’) in examined C-CDA documents. Although not comprehensive, each trouble spot represents a relevant, common issue in C-CDA documents. Since not all vendors elected to publicize their participation in the Collaborative, de-identified results were presented in the last group call (figure 3). The severity and clinical relevance of these trouble spots vary according to the context of C-CDA document use. Data heterogeneity or omission may impose a minimal burden in cases where humans or computers can normalize or supplement information from other sources. In other cases, a missing or erroneous code (e.g., terminology misuse; table 4) could disrupt vital care activities, such as automated surveillance for drug–allergy interactions. Because the severity of trouble spots depends upon specific clinical

---

**Table 3** Categorized observations (N=615) across 21 C-CDA samples examined

<table>
<thead>
<tr>
<th>Examined domains from MU common data set</th>
<th>Demographics</th>
<th>Allergies</th>
<th>Medications</th>
<th>Problems</th>
<th>Results</th>
<th>Smoking status</th>
<th>Vital signs</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect data within XML elements</td>
<td>10</td>
<td>12</td>
<td>27</td>
<td>24</td>
<td>5</td>
<td>14</td>
<td>5</td>
<td>97</td>
</tr>
<tr>
<td>Terminology misuse or omission</td>
<td>9</td>
<td>40</td>
<td>29</td>
<td>12</td>
<td>31</td>
<td>2</td>
<td>19</td>
<td>142</td>
</tr>
<tr>
<td>Inappropriate or variable XML organization or identifiers</td>
<td>7</td>
<td>20</td>
<td>13</td>
<td>17</td>
<td>23</td>
<td>10</td>
<td>20</td>
<td>110</td>
</tr>
<tr>
<td>Element optionality through inclusion or omission</td>
<td>49</td>
<td>20</td>
<td>40</td>
<td>16</td>
<td>22</td>
<td>1</td>
<td>13</td>
<td>161</td>
</tr>
<tr>
<td>Problematic reference to narrative text from structured body</td>
<td>0</td>
<td>6</td>
<td>10</td>
<td>11</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Inconsistent data representation</td>
<td>23</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>12</td>
<td>0</td>
<td>2</td>
<td>52</td>
</tr>
<tr>
<td>Not elsewhere classified</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>99</td>
<td>108</td>
<td>125</td>
<td>85</td>
<td>97</td>
<td>33</td>
<td>68</td>
<td>615</td>
</tr>
</tbody>
</table>

Both errors and heterogeneity observations were recorded in each category with the exception of ‘Inconsistent data representation’ which only included heterogeneity.

C-CDA, Consolidated Clinical Document Architecture; MU, Meaningful Use; XML, extensible markup language.

---

**Figure 1** Parsing times for C-CDA document samples (N=91). C-CDA, Consolidated Clinical Document Architecture.
<table>
<thead>
<tr>
<th>Category</th>
<th>C-CDA XML code</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect data</td>
<td>Incorrect data</td>
<td>Error</td>
</tr>
<tr>
<td></td>
<td>doseQuantity is ‘40 mg’ but should be ‘1’ to correspond to the RxNorm code that specifies tablet dosing</td>
<td></td>
</tr>
<tr>
<td>Terminology misuse</td>
<td>RxNorm code 7982 is ‘penicillin’, while the display and narrative state ‘codeine’</td>
<td>Error</td>
</tr>
<tr>
<td>Inappropriate organization</td>
<td>Code for a vaccine recorded in the diagnostic results section whereas it should be in immunizations</td>
<td>Error</td>
</tr>
<tr>
<td>Element optionality</td>
<td>Method code is optional and included on only one sample (eg, patient position as seated for blood pressure).</td>
<td>Heterogeneity</td>
</tr>
<tr>
<td>Interpretation code is optional for results and often omitted or left blank. In this example normal can be inferred from reference range</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference to narrative text</td>
<td>Reference to allergic reaction (cough) has no reference to allergen (aspirin)</td>
<td>Heterogeneity</td>
</tr>
<tr>
<td>Inconsistent representation</td>
<td>Two samples showing a medication to be administered ‘every day’ but units vary from hours to days</td>
<td>Heterogeneity</td>
</tr>
</tbody>
</table>
flows, we confine our discussion to the knowable barriers they create to semantic interoperability.

DISCUSSION
We demonstrated that aggregated, structured data covering a range of clinical information from MU2 C-CDA samples can be parsed with the open-source BlueButton.js library. This allowed us to inspect manually and programmatically the structured content of vendor-supplied documents to answer this question: will the exchange of C-CDA documents generated by 2014 Certified EHR Technology be capable of achieving semantic interoperability? Analyzing these documents, we identified many barriers to such interoperability. This leads to recommendations on how to improve C-CDA document quality, and with such improvements, advance automated document exchange.

Barriers to semantic interoperability
Our observations identify several barriers that will challenge reliable import and interpretation of parsed C-CDA documents. It can be helpful to categorize these issues based on how erroneous data can be detected.

Present in automated detection
Some observed violations of the C-CDA specification are already detected by NIST’s TTT validator. For example, the validator flagged inappropriate null values of ‘UNC,’ which were likely intended to be ‘UNK,’ meaning unknown. Such errors were unusual among EHR applications since certification requires the production of TTT-validated documents.

Potential for automated detection
Some barriers to semantic interoperability could be detected with additions to the TTT validator. One observed area was internal C-CDA consistency, which could be evaluated using logical correlations of structured entries. For example, if a C-CDA problem has an observation status asserting that the problem is biologically active, it would be incorrect for the concern status code to be ‘completed’ or for the patient’s timing information to include a problem resolution date.

Terminology issues were prevalent and also amendable to automated detection. In several samples we observed the use of non-existent, deprecated, or misleading codes, and non-adherence to required value sets. For example, one sample used the deprecated LOINC code ‘41909-3’ which has been superseded by LOINC code ‘39156-5’ to represent body mass index. There were also more complex concerns. For example, medication allergies should be encoded at the ingredient level (eg, ‘aspirin’) or drug class level (eg, ‘sulfonamides’), but some samples reported allergens at the semantic clinical drug level (eg, ‘aspirin 81 mg oral tablet’). While the latter representation is syntactically correct, it is clinically questionable to say that someone is allergic to a specific form and dose of aspirin. To reconcile such terminology issues, receivers of C-CDA documents would need to perform substantial manual reconciliation or apply intricate normalizing logic to the hierarchy of potential RxNorm codes.

Issues difficult to detect automatically
Heterogeneity in data representation imposes interoperability barriers that are difficult to detect automatically without clear guidance. We frequently observed variations where the C-CDA specification does not provide uniform guidance. Telephone
numbers illustrate this, where examples in the C-CDA show multiple ways to encode and no testable conformance is provided. In collected samples, we found 12 distinct patterns for recording a telephone number through combination of dashes, parentheses, periods, and leading characters. These representations are straightforward for humans to interpret but automated tools require specificity on permissible representations. Many variations in data representation may be addressed through lightweight data consumption normalization algorithms (eg, regular expressions for a telephone number).

Data optionality introduces two large challenges for semantic interoperability. First, the data are not present for certain downstream clinical workflows and applications. For example, the absence of medication administration timing (eg, ‘take every 8 hours’) prevents generation of automated reminders to promote medication adherence. Second, the absence of data may only reflect that the certified technology never populates and does not convey whether the data were known, unknown, or not structured in a vendor’s application. Such heterogeneity creates instances where the receiver cannot disambiguate data context.

Many of these observations may have a straightforward explanation. Several vendors explained that they focused development efforts on C-CDA generation to pass TTT validation and less on provider demands for semantic interoperability. Almost all vendors commented that they had too few implementation examples to guide them in expressing common clinical data and ambiguous guidance from regulatory and standards development organizations.

### Improving C-CDA document quality

We identify four areas—spanning standards development, implementation, and policy—that can lead to improved C-CDA document quality. Each of the recommendations we make in these areas can be weighed for its potential benefit against burden for implementation.

#### Provide richer samples in publically accessible format

Vendors commented in the Collaborative that they did not always know how to represent data within the C-CDA. While the ONC created a website to assist in C-CDA implementation and testing (http://www.sitenv.org/) and HL7 increased its help desk content, vendors suggested these were inadequate and sometimes unclear. There is need for a site where public samples and common clinical scenarios of C-CDA documents, sections, and entries can be queried. We posted samples to the Boston Children’s Hospital’s public C-CDA repository, when permitted by vendors. HL7 also supports this goal through the commission of a CDA Example Task Force (http://wiki.hl7.org/index.php?title=CDA_Example_Task_Force). A simple and powerful solution would be to require every technology to validate codes and value set membership. Because the C-CDA includes dozens of reference vocabularies in its implementation, testing for appropriate conformance to common vocabularies, such as SNOMED, LOINC, RxNorm, and UCUM, should be part of certification. Although many of the large value sets referenced by C-CDA are dynamic and are subject to change, this is reasonably addressable if reference terminology systems were maintained and hosted by an authority, such as the National Library of Medicine. Because there is no such authority today, the SMART C-CDA Scorecard hosted unofficial vocabulary sources for its C-CDA scoring.

#### Reduce data optionality

MU2 regulations have taken steps to reduce optionality by requiring a Common Data Set. The Common Data Set, however, does not constrain C-CDA optionality at a granular level. In effect, this permits vendors to omit data. For example, vendors must include an appropriate RxNorm code for each medication in C-CDA documents but otherwise may populate its dose, timing, and route of administration with null values. Moreover, MU2 data requirements imperfectly correspond to HL7 C-CDA specifications. For example, no single document type in the C-CDA library requires all components of the Common Data Set, so optional sections must be used to meet MU2 requirements. We therefore recommend that regulations require EHR vendors to populate known data in C-CDA documents.

#### Monitor and track real-world document quality

In real-world clinical environments, a multitude of C-CDA documents will be generated to satisfy clinical workflows. To quantify and improve document quality, metrics could be calculated using a C-CDA quality surveillance service running within the firewall or at a trusted cloud provider. Such a service could use existing tools, such as the TTT validator and SMART C-CDA Scorecard. These services could also be offered through health information exchanges that transmit C-CDA documents between organizations.

#### Advancing C-CDA document exchange

MU2 requires providers to exchange C-CDA documents for 10% of care transitions and for certified EHR technology to be capable of ingesting select data upon receipt. This is a significant advance from MU1, where only data display and testing of exchange were required. According to MU2 regulations, however, the intake of clinical data for certified systems need not be fully automated. This is entirely appropriate given the issues identified in this research.

To advance automated document exchange, we suggest vendors and policy makers consider Postel’s Law. This law, also known in computing as the robustness principle, states ‘be conservative in what you send, be liberal in what you accept from others.’ Our recommendations to provide more robust examples, validate codes, and reduce data optionality will reduce variability in the export of C-CDA documents, addressing the first half of this principle.

To improve the liberal consumption of C-CDA documents, intelligent parsing could normalize some aspects of heterogeneity. Such software could detect common variations in units, terminology, and data expression to return a more consistent data set from C-CDA documents. Our recommendation to monitor actual C-CDA document exchange would serve vendors well as they write normalizing algorithms. However, even the best engineering cannot reliably populate missing data or resolve conflicting statements, so there will be an upper bound as to what can be normalized. While Postel’s Law cannot be directly enforced through regulation, a combination of policy changes with community support of our recommendations would move real-world C-CDA exchange closer to realizing this principle.
A further challenge for real-world document exchange also emerged in this study. Latency of C-CDA document production and consumption, while not a barrier to semantic interoperability, may limit application responsiveness. Automated parsing using BlueButton.js, which provides minimal C-CDA normalizing logic, requires up to several seconds for larger documents. Intelligent normalization, network latency, and C-CDA generation time, none of which we measured in our research, would add to such computational time. Together these considerations suggest a limited role for C-CDA usage in low latency services, such as concurrent clinical decision support or other third party applications.

Caveats
When providers adopt certified software for MU2, we expect C-CDA exports will exhibit many of the challenges we observed in our samples. Nonetheless, our findings about the readiness of C-CDA documents for interoperability have several limitations. First, our samples represented technologies that voluntarily submitted or publicized their C-CDA documents. While participating vendors represent a majority of the certified EHR market, we examined only fictional patient records from vendor test or development environments. These fictional records represented varying clinical scenarios individually created by vendors. Using a standard data set based on consistent clinical conditions from all technologies would have yielded greater comparative analysis. Second, our findings do not capture real-world implementation by hospitals and physicians, since MU2 had not yet been implemented during the research. We anticipate additional issues will surface in the thousands of upcoming C-CDA deployments. Third, we focused exclusively on seven clinical domains from the Common Data Set. Had we scrutinized other C-CDA domains, we would likely have recorded additional errors and heterogeneity. Finally, while we examined C-CDA documents and discussed their production rationale with vendors as external observers, we were unable to examine any vendor’s C-CDA consumption and reconciliation algorithms. This would have provided further insight into the challenges of semantic interoperability but was beyond the scope of our research.

These limitations in aggregate have likely caused us to understate the frequency and types of observed errors and heterogeneity that will be observed in real C-CDA exchange. Our findings, however, materially capture the problems facing C-CDA document exchange for MU2.

CONCLUSION
Although progress has been made since Stage 1 of MU, we expect C-CDA exports will exhibit many of the challenges we observed in our samples. Nonetheless, our findings about the readiness of C-CDA documents for interoperability have several limitations. First, our samples represented technologies that voluntarily submitted or publicized their C-CDA documents. While participating vendors represent a majority of the certified EHR market, we examined only fictional patient records from vendor test or development environments. These fictional records represented varying clinical scenarios individually created by vendors. Using a standard data set based on consistent clinical conditions from all technologies would have yielded greater comparative analysis. Second, our findings do not capture real-world implementation by hospitals and physicians, since MU2 had not yet been implemented during the research. We anticipate additional issues will surface in the thousands of upcoming C-CDA deployments. Third, we focused exclusively on seven clinical domains from the Common Data Set. Had we scrutinized other C-CDA domains, we would likely have recorded additional errors and heterogeneity. Finally, while we examined C-CDA documents and discussed their production rationale with vendors as external observers, we were unable to examine any vendor’s C-CDA consumption and reconciliation algorithms. This would have provided further insight into the challenges of semantic interoperability but was beyond the scope of our research.

These limitations in aggregate have likely caused us to understate the frequency and types of observed errors and heterogeneity that will be observed in real C-CDA exchange. Our findings, however, materially capture the problems facing C-CDA document exchange for MU2.


