Conformance Testing of Healthcare Data Exchange Standards for EHR Certification

R. Snelick[1]
[1]National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA

Abstract – *Seamlessly sharing and using healthcare data as intended among distributed healthcare information systems is difficult. The adoption and adherence to clear and unambiguous standards can help manage this complexity. Well-defined standards, and conformance to those standards, provide the foundation for reliable, functioning, usable, and interoperable healthcare information systems. Recent federal programs (in the US) have included incentive payments for healthcare providers who adopt and "meaningfully use" certified electronic health record (EHR) technologies; however, unless these products are developed using clearly-defined standards, the adoption rate will increase, but the promise of improved quality of healthcare will not be realized. The proliferation of healthcare information systems designed without compliance to standards will likely exacerbate, not lessen, current patient care challenges by creating a landscape saturated with systems lacking usefulness, usability, and interoperability that will be rejected by the end-user community. Additionally, the standards must be used and deployed as intended, and conformance testing is the process that helps ensure adherence to the standards. In this paper, we explore conformance testing and the tools that are used to perform HL7 (Health Level Seven) v2-based conformance testing for certification of EHR technologies.*

Keywords: Conformance; Conformance Testing; Data Exchange Standards; Healthcare Information Systems; Interoperability.

1 Introduction

As described in [1], use of electronic health records (EHRs), especially systems with clinical decision support capabilities, has been shown to enable quality improvement in healthcare as well as to help reduce the cost of that care when used regularly in the practice of medicine. Recognition of these findings led to the enactment of the Health Information Technology for Economic and Clinical Health (HITECH) Act, which provides funding for incentive payments to physicians and hospitals that adopt health information technology (HIT). Initially focusing on adoption of EHRs, approximately $17 billion in the Center of Medicare and Medicaid Services (CMS) incentive payments were made available through CMS's HITECH-based EHR Meaningful Use (MU) Program, to be paid to providers that attest to or demonstrate "meaningful use" of "certified" EHR technology (CEHRT) [1]. In-line with the CMS program, the Office of the National Coordinator (ONC) published EHR certification criteria [2] and established a program for certifying EHR technologies [2]. ONC, in collaboration with the National Institute of Standards and Technology (NIST), developed test procedures and conformance test tools [3] based on the ONC's EHR certification criteria. EHR technologies are tested for compliance to the criteria by ONC-Accredited Testing Laboratories (ATL). This paper provides an overview of the testing approach and test tools used by the ATLs to verify that vendors' EHR technologies meet the certification criteria that specify HL7 v2 data exchange standards.

2 HL7 V2 Data Exchange Standard

There are numerous healthcare data exchange standards in the US and internationally for communicating administrative and clinical data. The most widely used standard is the Health Level Seven (HL7) Version 2.x *Application Protocol for Electronic Data Exchange* (hereafter HL7) [4,5]. This standard is designed to support application-to-application message exchange. An HL7 *message* contains data for real-world events such as admitting a patient or sending a laboratory result (e.g., a CBC—complete blood count). For each event, HL7 defines an *abstract message definition* that is composed of a collection of segments (data units, e.g., Patient Demographics) in a predetermined sequence. Rules for building an abstract message definition are specified in the HL7 message framework, which is hierarchical in nature and consists of building blocks generically called *elements* [4,5]. These elements are *segment groups*, *segments*, *fields*, *components*, and *sub-components*. The requirements for a message are defined by the message definition and the constraints placed on each message element. The constraint mechanisms are defined by the HL7 conformance constructs which include usage, cardinality, value set, length, and data type. Additionally, explicit conformance statements are used to specify other requirements that can't be addressed by the conformance constructs.

An HL7 conformance profile (also referred to as a message profile) is a constraint on the base standard that defines specified requirements for a given use case (event), a common practice in HL7 specifications [5,6]. The message profile can be represented as an XML document, which forms the basis for conformance testing. Figure 1 shows an example snippet (condensed) of a lab result profile [7]. Each element in the message profile is listed along with its associated attributes. For a more detailed description of a message profile refer to the HL7 standard [4] and [5,6]. It is important to note that the attributes and constraints a profile applies to a message provide a clear and unambiguous definition, thereby facilitating the design, implementation, and testing of interfaces [5,6]. The NIST EHR conformance test tools use the XML message profile as the basis for validation [3,5,8,9].

Fig. 1. Snippet from a Message Profile

```
<Segment Name="OBX" Max="1" Min="1" Usage="R"
LongName="Observation/Result">
  <Field Name="Observation Identifier" Max="1" Min="1"
  Usage="R" Datatype="LRI_CWE_CR" Table="VS_LOINC">
    <Component Name="Identifier" Usage="R" MinLength="1"
    MaxLength="20" Datatype="LRI_ST">
    <ConformanceStatement id="LN-001"><EnglishDescription>If
    CWE.3 (Name of Coding System) is valued "LN" then CWE.1
    SHALL be a valid LOINC code identifier format.
    </EnglishDescription><Assertion><Custom id="1"
    className="gov.nist.healthcare.mu.lri.custom.Loinc"/></Asse
    rtion></ConformanceStatement>
    </Component>
    …
  <Field Name="Observation Value" Max="1" Min="0" Usage="RE"
  Datatype="varies"></Field>
  <Field Name="Abnormal Flags" Max="*" Min="0" Usage="RE"
  Datatype="LRI_IS" Table="0078">
```

The message structure defines a *template* to which the message must comply; it explicitly defines the elements and the sequencing of the elements in a message instance. Conformance constructs are used to define and constrain requirements on message elements. Figure 1 provides a representative lab result observation that is defined by the OBX segment [7]. The OBX segment has multiple fields such as Observation Identifier (e.g., Cholesterol), Observation Value (e.g, 196), Units (e.g., mg/dL), and Abnormal Flags (e.g., N). There are other segments in the message definition such as PID—Patient Identification. Fields can also contain structure, i.e., components and sub-components. For every element, constraints are defined, e.g., Usage (indicates if the element is required, conditional, etc.), Cardinality (indicates the number of times the element may occur), or Value Set (indicates a defined vocabulary). The message structure and the element constraints define the requirements and are used for message validation.

The 2014 Edition/Stage 2 ONC EHR certification standards and criteria specify four HL7 v2 implementation guides that apply to five certification criteria: (1) transmission to immunization registries, (2) syndromic surveillance to public health agencies, (3) transmission of reportable lab results to public health agencies, (4) transmission of electronic lab results to ambulatory providers, and (5) incorporation of lab tests and results [2]. A conformance profile is defined for each interaction (event) covering the specific use, for example, sending an immunization record from the EHR system to the Immunization Information System (IIS). The conformance test tools described below are used to ensure that EHR systems correctly implement this interface standard (profile).

3 Conformance Testing

Conformance testing is a process that determines if an entity (message, document, application, system, etc.) adheres to the requirements stated in a specification. Conformance testing is a multi-faceted operation that can range from a simple assessment of the validity of a message value to a nuanced determination of a system's reaction to a complex sequence of events. Conformance testing strives to establish a degree of confidence in the conformity of a given entity (implementation) based on the quantity and the quality of the tests performed. Interoperability testing assesses whether applications (or software systems) can communicate with one another effectively and correctly, and whether they can provide the expected services in accordance with defined requirements (i.e., have a common understanding and use of the data exchanged). Such testing is critical, since many modern system architectures are designed as distributed systems and rely on seamless operations.

NIST developed the HL7 v2 validation tools for ONC 2014 Edition/Stage 2 HIT certification testing. These tools covered the standards and criteria described in section 2. Although, the initial focus of this effort targeted ONC HIT certification, the tools are equally applicable and valuable for use at site installations. In fact, a number of local public health registries have incorporated the tools into their operational environments or, as in the case of the Arkansas Department of Health for instance, require them for on-boarding [10].

4 Testing Sending Applications

When testing the ability of the System Under Test (SUT) to create messages[1], the focus of the conformance testing is on validating the message produced by the sending

[1] The concepts apply equally to documents and other message protocols.

system (e.g., an EHR). The *sender* SUT is treated as a "black box" — only the content of the message is of interest not how the message is created or transmitted.

The NIST conformance test tools used for validating sending systems have two operational modes: (1) Context-free and (2) Context-based. The Context-free mode validates any HL7 v2-based message created by the SUT for the given subject (e.g., immunization messaging, lab result messaging). It is not dependent on a specific use case instance, Test Case, or specific test data content. The Context-based mode validates messages created by the SUT that are associated with a given use case instance and a Test Case that includes specific test data that are entered into the SUT. The validation assesses the technical requirements and content-specific requirements specified in the Test Case. Context-based validation expands the test space, enabling more comprehensive testing (e.g., testing of conformance usage constructs such as "conditional" and "required, but may be empty").

Context-based conformance testing of the technical requirements and capabilities of the EHR technology is central to certification testing. Through collaboration with subject matter experts from the ONC Standards and Interoperability (S&I) Framework, the Centers for Disease Control and Prevention (CDC), the Association of Public Health Laboratories (APHL), and the International Society for Disease Surveillance (ISDS), NIST developed the Test Cases that targeted the most important use cases and capabilities specified in the referenced standards.

The test data are provided to assist the Tester in verifying that the vendors' EHR technologies are capable of supporting the required functions. Verifying the ability to support the specific test data content is a secondary aspect to the certification testing. Testing and verification related to specific content usually are more appropriate for local installations of the EHR technologies; however, for certain aspects of a certification Test Case, examining exact content is necessary to verify that a capability exists in the EHR technology. An added benefit of providing realistic test data for common use cases is reinforcement of the expected interpretation and use of the referenced standards.

Both Context-free and Context-based modes are useful for message validation. Since Context-free testing is not tied directly to Test Case data, any message instance can be validated. This method suits site installations well, enabling in-house testing on messages that are tailored to local requirements. Context-based testing is driven by Test Cases, targeting specific Test Scenarios that enable more precise testing. Context-based testing is the method used in the ONC HIT Certification Program, however, the NIST HL7 v2 conformance test tools [3,9] support both modes of validation.

4.1 Case Study: Transmitting Lab Results

In this section, a case study based on the ONC Edition 2014 certification criterion for the transmission of laboratory results is used to explain the principles of Context-based testing for a sending application.

The focus of conformance testing for a sending system is on validating the message. The SUT is treated as a "black box"—how the message is created or transformed is not in scope. If we consider a laboratory information system (LIS) or a laboratory module that is integrated with an EHR system (hereafter called "lab component"), testing is not concerned with the detailed architecture of the lab component, but rather with what it produces (a lab results message) based on a given set of inputs (i.e., a lab results interface Test Case). The "black box" can consist of a self-contained lab system or multiple interrelated modules. The Use Case for the transmission of lab results could consist of the following steps:

1. A lab test is ordered for a patient
2. The specimen is collected (if applicable), and is received and processed by the lab
3. The lab result is produced, imported, and stored by the LIS
4. **The lab result message is created**
5. The lab result is transmitted to an ambulatory electronic health record (EHR) system
6. **The lab result is incorporated into the ambulatory EHR system**

The scope covered by the ONC transmission of lab results criterion is step 4 above – the lab result message is created. Step 6 – the lab result is incorporated into the ambulatory EHR – is covered by the incorporate lab results criterion; see section 5 for a case study for testing receiving systems.

Test Cases are provided for specific laboratory tests for which a lab results message will be imported into the conformance test tool (e.g., a CBC—Complete Blood Count). A Test Case consists primarily of a narrative Test Story (one possible path described by the Use Case) and a Test Data Specification. The Test Story describes a real world situation and provides the context for the Test Case. The Test Story also provides details associated with the Test Case such as pre-conditions, post-conditions, test objectives, and notes to testers. The Test Data Specification provides the data associated with the Test Story and consists of typically available information in the clinical setting. Together the Test Story and the Test Data Specification provide sufficient information to

be entered into the SUT for a particular Test Case, e.g., creating a lab results message. A Message Content Data Sheet is provided to show a conformant message instance for the Test Case. It also lists the category for each message element, indicating the kind of data and the expected source, which are based on the test case and test case objectives. How the data are categorized is directly related to how the message content is validated by the Test Tool. In some cases the validator is examining a message element for the presence or absence of data, and in other cases it is examining the message element for both the presence of data and exact content. The Message Content Data Sheet provides the evaluation criteria (expectations) and can be thought of as the "answer" to the "question" given in the Test Story and the Test Data Specification.

Fig. 2. Context-based Validation Test Flow

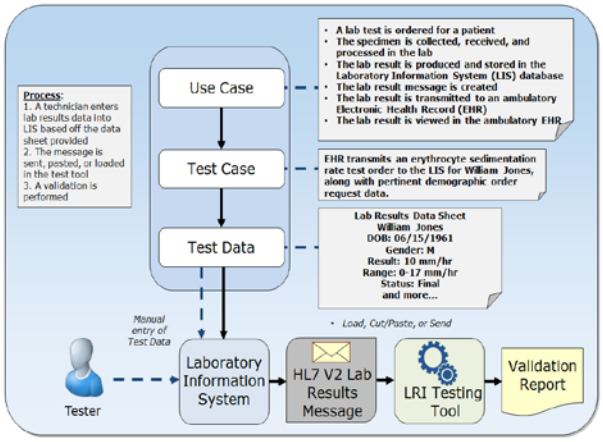
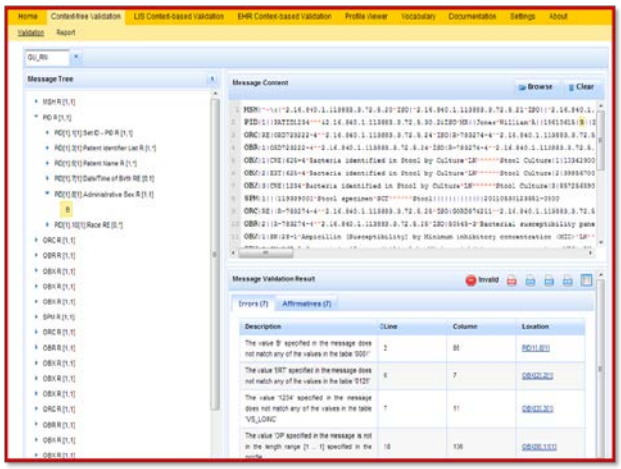


Figure 2 summarizes the Context-based testing flow for a representative Test Case. The Test Case artifacts are accessed by the Tester, the test data are loaded into the LIS, and a lab result message is generated. The Tester selects the corresponding Test Case in the Test Tool, imports the generated message, and the Test Tool validates the message based on the requirements in the Lab Results Interface (LRI) specification [7]. The web-based Test Tool provides interactive validation results for each message along with validation report documents.

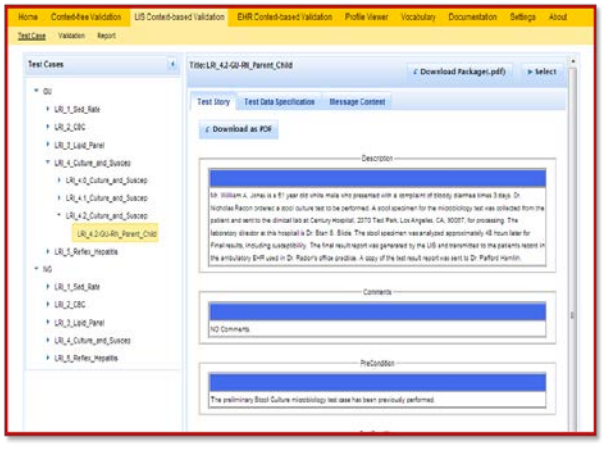
The Context-free and Context-based testing modes in the NIST Test Tools have certain basic features in common. When the user selects the Context-free tab, they then select the desired message profile and import the test message into the Test Tool. Once the message is imported, the message validation is performed automatically (Figure 3). The left-hand panel of the Validation screen shows the user a tree structure view of the message where individual data-content can be examined. The upper-right panel is the message content window. The validation results are displayed in the lower-right panel of the tool and include a description of the error and its location.

Fig. 3. Message Validation Panel



In the Context-based mode the user first selects a specific Test Case that provides a particular scenario and test data. These data are entered into the LIS which creates a message that corresponds to the test data. The validation process then proceeds much like the Context-free mode, except the validation is bound to specific data requirements defined by the Test Case. Figure 4 shows a screenshot of the Test Case panel that includes the Test Story (shown), the Test Data Specification, and the Message Content. The tool also includes a Profile Viewer and Vocabulary tab that allow browsing of the requirements specified in the implementation guide. The Test Cases and the Context-based validation are linked; i.e., in addition to validating the technical requirements specified in the implementation guide, the Test Tool performs selective content validation based on the provided Test Data Specification with the associated data categorization.

Fig. 4. Test Case Panel



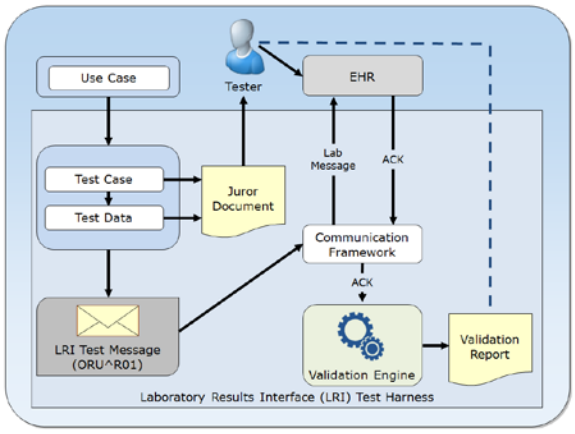
5 Testing Receiving Applications

Testing a *receiving* system is a challenge, because, typically, the system does not produce a *tangible* object (e.g., a message) for the Tester to assess directly. From

the available testing approaches used for testing receiving systems, the Inspection Testing approach has proven most suitable for the ONC HIT certification testing environment.

Inspection Testing relies on human validation (a visual inspection) of the SUT in order to collect evidence for the conformity assessment. Usually, the Inspection Testing process involves priming or knowing the state of the receiving system; providing a known and documented stimulus to the system; and evaluating the system's response to the stimulus against expected results based on the input and requirements.

Fig. 5. Testing Incorporation of Lab Results

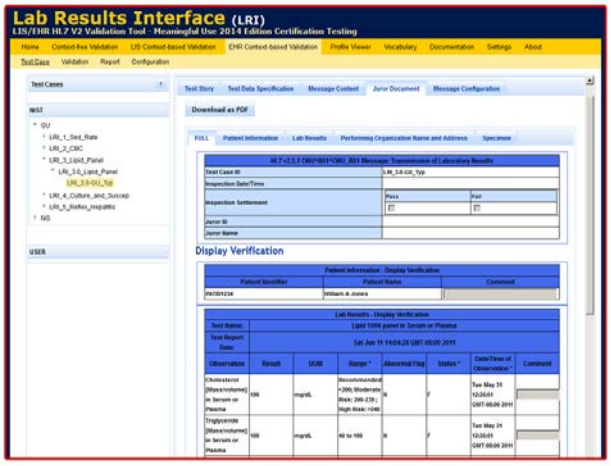


A "test harness" contains the Test Case and associated test material, such as the test message and a Juror Document. The Juror Document (Test Case-specific inspection check list) is used by the Tester to inspect the receiving SUT for conformance to the specification. The information contained in the Juror Document is based on the data provided in the test message, the known state of the system, and requirements listed in the given test criterion. During the inspection process, evidence of the SUT's conformance can be obtained through a variety of methods, including viewing the system's display screens, browsing the system's data base, viewing the configuration files, or other mechanisms supported by the SUT.

5.1 Case Study: Lab Results Incorporation

Testing for the incorporation of laboratory results provides a good example of the challenges faced when testing a receiving system. No output artifact is produced that can be assessed directly by the Tester during this test [1]. For this ONC criterion, the ambulatory EHR, as the receiving SUT, is examined for evidence of the *incorporation* of laboratory results information from the received message and also for the ability to display seven types of information that are part of a laboratory results report (per requirements adopted from the Clinical Laboratory Improvement Amendments (CLIA), which are regulatory standards for clinical laboratory operations in the US). The ONC criterion for incorporation of laboratory results specifies the Laboratory Results Interface (LRI) Implementation Guide [7] for generating the laboratory results message, and the related conformance testing involves a Juror Document and a human inspector. The content of the Juror Document is derived mostly from the Test Case and test message. Figure 5 illustrates the testing flow when using the NIST conformance test tool for the incorporation of laboratory results test procedure.

Fig. 6. Juror Document Panel



The Test Tool [9] provides a test harness that interacts with the EHR SUT, simulating the function of an LIS (or a laboratory component) that would create the LRI message. The use case described in the LRI implementation guide for creating lab result messages is the counterpart to the use case described for incorporating these messages; therefore, the same Test Cases developed for creating lab result messages can be used for incorporation of these messages, which allows for reuse of certain testing artifacts. The EHR SUT is primed with data (i.e., patient demographic information) to enable incorporation of the lab result message data elements into a specific patient's record; and the Test Tool, EHR, and test message are configured to enable communication between the systems.

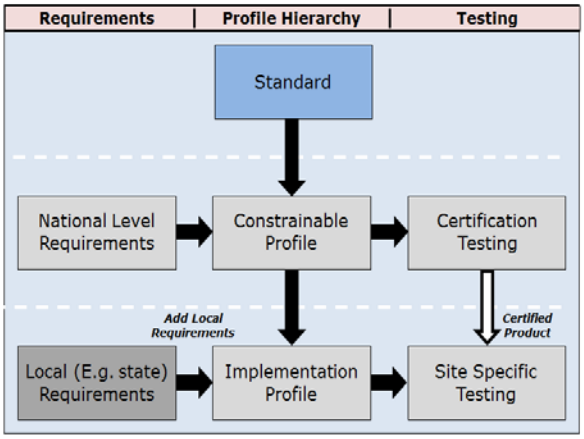
The lab result test message is sent from the LRI test harness to the EHR SUT, and the Juror Document is generated automatically by the test tool infrastructure based on the Test Case and test data. The Tester obtains the Juror Document and uses it to examine the EHR and to verify and document the presence or absence of the data elements transmitted in the test message. The data elements are categorized in the Juror Document according to how they are to be verified. For example, some data elements must be displayed to the clinical user on the EHR screen as well as stored in the EHR, while

other data elements are required to be stored or derivable only. The results gathered by the inspector are used in combination with the Validation Report from the ACK message to determine if the SUT passed or failed the test.

6 Perspectives of Testing

Profiling and the application of a profile hierarchy for specifying requirements of data exchange standards are critical for achieving interoperability [6]. A conformance profile is a refinement to either the underlying standard or another conformance profile, and it normally specifies constraints on messages or documents. A relationship can be drawn between the profile level and the type of testing that can be performed, as illustrated in Figure 7. The relationships shown are not the only relationships possible, but they are the typical ones. SDOs (Standards Development Organizations) that create implementation guides often do so at the national (realm) level. These constrainable profiles defined within an implementation guide express a minimum core set of capabilities that each implementer must meet. Beyond this core definition, a certain amount of flexibility is allowed for data elements not fully qualified in the implementation guide; for example, the "optional" usage defined for some elements could be redefined by implementers as "required".

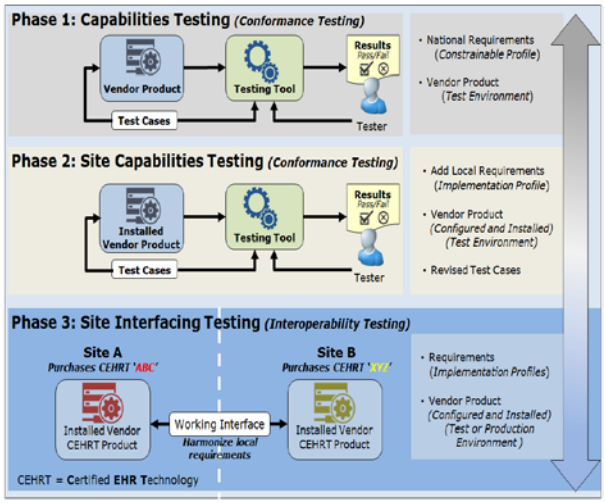
Fig. 7. Profile Level and Testing Relationship



The established *baseline* ensures a consensus level of functionality that satisfies the targeted use cases. Vendor and local implementations further constrain and define requirements that are compliant with the national level requirements. Local, that is site[2] specific, variations are denoted explicitly.

National certification programs (such as the ONC HIT Certification Program in the US) develop and/or reference national level profiles for their certification criteria. One of the objectives of certification testing is to assess whether the capabilities in each vendor's product meet the requirements defined by the specified profile. Capability Testing is the type of testing used for ONC HIT certification testing and is based on constrainable profiles as depicted in Figure 8. A key point to bear in mind for Capability Testing is that its purpose is to verify that a product has the required capabilities, not to verify how the product might be used when installed in a production environment. Once any needed additional local requirements are established and documented, site testing is performed using an implementation profile (i.e., a completely defined specification). Site specific testing focuses on the ability of a product to support its intended use at an actual installation, which may be based on partner agreements.

Fig. 8. Levels of Testing



Certification testing seeks to ensure that every product that is certified supports the capabilities defined by the national level standard. National level certification testing brings a set of stakeholders one step (or phase) closer to achieving interoperability, but it is only the first step (see Figure 8). It is incorrect to assume that installing certified products will lead to "*out of the box interoperability*" when interfacing two or more of these products. The scope of ONC HIT certification testing is phase one. After Capability Testing of vendor products is performed in a test setting, a second round of Capability Testing that includes testing based on local requirements should be performed; we refer to this level of testing as Site Capability Testing (See Figure 8).

At the national standard profile level, local requirements and variations have not been taken into account. Once local agreements are defined and the profiles have been documented, site specific testing can occur. The distinction between the different profiles and the associated levels of testing is important. Capability

[2] The term "site" in this context can mean a single site, multiple sites, or a group of sites that define the same (implementation) profile requirements.

Testing occurs in phase one of the process and focuses on conformance testing. For site installations, the baseline requirements are customized to meet local requirements, and additional conformance testing needs to occur. This local conformance testing is required to ensure that the local requirements are implemented and that the national requirements have not been compromised (think of this as a form of "regression" testing). Once all parties participating in the site installation have completed this second round of conformance testing, then interoperability testing can proceed.

Step (or phase) three focuses on interoperability testing, and, ideally, conformance testing should continue to be included in the process. The need to include conformance testing here is especially critical if the implementations are being modified to achieve interoperability. It is important that conformance is not compromised to obtain interoperability. The sites wishing to interoperate likely have purchased certified products that have been customized to meet site requirements and have tested those implementations accordingly. Site Interface Testing is employed to determine that both data exchange and data use meet the business requirements. Such testing addresses the question: does the interface work for the intended use case? Site Interface Testing can be performed in a test or production environment.

Although we have presented the testing steps as a group, the concepts of conformance and interoperability testing are orthogonal. Conformance testing is performed on the various profile levels in the hierarchy, and the product is tested in isolation. Interoperability testing is performed among a set of products, be it in a test environment (such as the IHE Connect-a-thons [11]) or at a production site. Although orthogonal in nature, the sequence in which testing should occur is progressive and must take into consideration the realities of the production setting in which the HIT technologies are to be used. There is limited value in performing interoperability testing without prior agreements and conformance testing.

7 Summary

Improved outcomes, clinical decision support, and patient safety are a few of the many benefits provided by interoperable healthcare information systems such as EHRs [12]. To achieve interoperable systems, products must be developed to a set of well-defined standards that are universally adopted. To help ensure that the standards are implemented correctly, conformance testing is necessary. In the US, ONC has established a program to certify EHR technologies that uses NIST conformance testing tools. Certified products ensure a level of capabilities, which is the critical first step towards achieving interoperable systems. Beyond this step, refinement of standards within the framework of the established base standard is often necessary to accommodate site specific requirements. Subsequent conformance and interoperability testing also is necessary. Following this course of action will drive the industry closer to the goal of interoperable healthcare information technologies.

8 References

[1] *Understanding Meaningful Use with a Focus on Testing the HL7 V2 Messaging Standards*. R. Snelick, S. Taylor, HL7 May 2013 Newsletter. http://hl7.org.

[2] ONC Certification Programs; http://www.healthit.gov/providers-professionals

[3] *NIST Resources and Tools in Support of Health IT Standards and the ONC Meaningful Use certification program.* http://healthcare.nist.gov/

[4] Health Level 7 (HL7) Standard Version 2.7, ANSI/HL7, January, 2011, http://www.hl7.org.

[5] *Towards Interoperable Healthcare Information Systems: The HL7 Conformance Profile Approach*. R. Snelick, P. Rontey, L. Gebase, L. Carnahan. Enterprise Interoperability II: New Challenges and Approaches. Springer-Verlag, *London Limited 2007 pp*. 659-670.

[6] *Principles for Profiling Healthcare Data Communication Standards.* R. Snelick, F. Oemig. 2013 Software Engineering Research and Practice (SERP13), WORLDCOMP'13 July 22-25, 2013, Las Vegas, NV.

[7] *HL7 Version 2.5.1 Laboratory Results Interface (LRI) Implementation Guide.* Draft Standard for Trail Use. July 2012. http://www.hl7.org.

[8] *A Framework for testing Distributed Healthcare Applications*. R. Snelick, L. Gebase, G. O'Brien. 2009 Software Engineering Research and Practice (SERP09), WORLDCOMP'09 July 13-16, 2009, Las Vegas, NV.

[9] *NIST Laboratory Results Interface (LRI) EHR Meaningful Use Conformance Testing Tool*; http://hl7v2-lab-testing.nist.gov/mu-lab/

[10] Arkansas Department of Health Onboarding. http://www.healthy.arkansas.gov/programsServices/MeaningfulUse/Immunizations/Pages/OnBoarding.aspx

[11] *IHE Connect-a-thon* http://www.ihe.net/Connectathon

[12] Benefits of Electronic Health Records (EHRs) http://www.healthit.gov/providers-professionals/benefits-electronic-health-records-ehrs