Commentary

The long road to semantic interoperability in support of public health: Experiences from two states

Brian E. Dixon a,b,c,⇑, Daniel J. Vreeman d, Shaun J. Grannis d

a Department of BioHealth Informatics, Indiana University School of Informatics and Computing, Indianapolis, IN, USA
b Center for Biomedical Informatics, Regenstrief Institute, Inc., Indianapolis, IN, USA
c Center for Health Information and Communication, Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service CIN 13-416, Richard L. Roudebush VA Medical Center, 410 W. 10th St., Suite 2000, Indianapolis, IN 46202, USA
d Indiana University School of Medicine Indianapolis, IN, Regenstrief Institute, Inc., Indianapolis, IN, USA

A R T I C L E   I N F O

Article history:
Received 14 November 2013
Accepted 16 March 2014
Available online 25 March 2014

Keywords:
Medical informatics
Public health informatics
Infectious disease reporting
Public policy
Meaningful use

A B S T R A C T

Proliferation of health information technologies creates opportunities to improve clinical and public health, including high quality, safer care and lower costs. To maximize such potential benefits, health information technologies must readily and reliably exchange information with other systems. However, evidence from public health surveillance programs in two states suggests that operational clinical information systems often fail to use available standards, a barrier to semantic interoperability. Furthermore, analysis of existing policies incentivizing semantic interoperability suggests they have limited impact and are fragmented. In this essay, we discuss three approaches for increasing semantic interoperability to support national goals for using health information technologies. A clear, comprehensive strategy requiring collaborative efforts by clinical and public health stakeholders is suggested as a guide for the long road towards better population health data and outcomes.

Published by Elsevier Inc.

1. Introduction

Health information technology (health IT) is increasingly vital to the public’s health [1]. Health IT, including electronic health record (EHR) systems, telemedicine and clinical decision support, has the potential to support achievement of the triple aim: improving the quality of and satisfaction with patient care while improving the health of populations and reducing the per capita cost of health care [2,3]. For example, delivering the right information to the right person at the right time using health IT has the potential to reduce up to 18% of patient safety errors and as many as 70% of adverse drug events [4]. Health IT is further estimated to play a key role in health systems transformation by enabling care coordination initiatives, including patient-centered medical homes and accountable care organizations [5,6].

Recognizing known benefits and greater potential for improving health care, the Health Information Technology for Clinical and Economic Health (HITECH) legislation in the U.S. [7] incentivizes adoption and ‘meaningful use’ of health IT amongst hospitals and physician practices. The meaningful use program, administered by the Centers for Medicare and Medicaid Services (CMS), provides financial payments to hospitals and providers who adopt EHR systems that comply with criteria established by the federal government. The most recent published criteria from CMS [8] require eligible hospitals and providers to submit electronic health data to local and state health departments.

To maximize the value of health IT to clinical and public health, semantic interoperability is necessary. Semantic interoperability can be broadly defined as the ability for one IT system to receive information from another IT system and reliably apply its business rules against the information received [9]. This definition represents a well-established, consensus-based view from the international health information exchange community for shared messaging (syntax) and meaning (semantics) between health IT systems. The Center for IT Leadership estimates that among various health IT investments, introducing semantic interoperability would produce the greatest economic benefit to the U.S. health system [10]. To achieve semantic interoperability, the U.S. health system must adopt consistent clinical messaging and data standards that provide a framework and language for communicating shared meaning. While messaging (syntax) is critically important, we focus this essay on the semantic (meaning, data) aspects of interoperability.

⇑ Corresponding author at: Regenstrief Institute, 410 W. 10th St., Suite 2000, Indianapolis, IN 46202, USA. Fax: +1 317 274 9305.
E-mail address: bedixon@iupui.edu (B.E. Dixon).

http://dx.doi.org/10.1016/j.jbi.2014.03.011
1532-0464/Published by Elsevier Inc.
Despite being a requirement for the nation’s emerging health information infrastructure, a clear approach to achieving semantic interoperability remains elusive. Although clinical data standards are available, most hospitals, laboratories, and physician offices continue to rely on local, idiosyncratic and incompatible ways of identifying clinical observations (e.g., laboratory tests, clinical measurements) and their results. This may be due in part to the fact that translation of local terminology into available standards is a complex, costly and resource intensive process [11,12].

Given that semantic interoperability is necessary but lacking, we argue that the U.S. needs a clear strategy for achieving semantic interoperability among health IT systems. Similar to strategies published in recent years for the adoption of e-health [13] as well as clinical decision support [14–16], a strategy for semantic interoperability should outline principles and a roadmap that stakeholders can measurably apply to adopt standard vocabularies. An ideal strategy provides an equitable, acceptable pathway that is efficiently implemented at a reasonable cost. The dimensions of equity, acceptability, efficiency, and bureaucracy (e.g., implementation) are adopted from the field of policy analysis [17], and they represent criteria by which a government or public entity can weigh alternative strategies for achieving semantic interoperability.

In this essay, we illustrate the current state of semantic interoperability using a case example drawn from public health and discuss three policy strategies for strengthening interoperability. Public health is a representative subset of the health care system with information needs and challenges that are similar to the clinical care enterprise. Thus, the example is applicable to the health care system writ-large. We first present the current landscape for our public health example, then analyze alternative strategies for achieving semantic interoperability using the above criteria, and finally present recommendations for improving the use of clinical data and information for public health, including disease surveillance, community assessment and measurement of care quality. We conclude by discussing how enabling interoperability for public health and other clinical care contexts will support the broader aims of national health IT adoption to achieve the triple aim in the U.S. and other nations.

2. The current state of semantic interoperability

Electronic laboratory reporting (ELR) provides an illustrative use case. ELR involves the direct, electronic submission of communicable disease case information, following laboratory-confirmed diagnosis, to public health authorities. ELR can significantly improve disease surveillance by increasing the timeliness and completeness of case reports submitted to public health agencies [18,19]. We chose ELR as an exemplar use case because surveillance of disease remains the “foundation of public health practice” [20], and the concepts illustrated by ELR apply to the broad spectrum of information exchange. Furthermore, ELR is a core requirement of the meaningful use initiative driving the adoption and use of health IT in the U.S. [21]

2.1. Vocabulary standards for electronic laboratory reporting

Standard vocabularies for representing clinical data exist, and the current versions of available standards are mature [22]. Logical Observation Identifiers Names and Codes (LOINC) provides universal identifiers for laboratory tests and other clinical observations [23]. Systematized Nomenclature of Medicine–Clinical Terms (SNOMED CT) provides universal identifiers for organisms, substances, diseases, and other findings that may be recorded in the medical record or identified in test results [24]. Together, LOINC and SNOMED CT are internationally recognized vocabulary standards for communicating tests (LOINC) and results (SNOMED CT) for notifiable disease reporting using ELR [25]. Many commercial health IT systems claim to support these standards, and the meaningful use criteria for ELR require ‘certified’ EHR systems to support them.

2.2. Quantifying use of vocabulary standards in operational systems

Given a lack of published evidence on the use of current standards for ELR, we sought to gauge the use of LOINC and SNOMED CT using operational systems in two states: Indiana and Wisconsin. Regenstrief is the birthplace of the Indiana Network for Patient Care (INPC), one of the largest and longest tenured regional health information exchange (HIE) networks in the U.S. [26,27]. The INPC processes over 5 million ELR messages each month, and the HIE forwards ELR messages for reportable conditions to the Indiana State Department of Health on behalf of its members. Atlas Public Health, a division of Atlas Development Corporation, develops and deploys solutions for public agencies and health care providers. Atlas provides operational ELR services that receive data for the Wisconsin Department of Health Services (WDHS).

We examined over 7 million ELR messages between 2010 and 2011 for test results reported to the INPC or WDHS [28]. Analysis focused on the applicable data fields where LOINC and SNOMED CT values are expected, which according to Health Level 7 version 2 specifications are OBX-3 (identifies the test performed) and OBX-5 (identifies the result of the test performed). Using data from each state, we calculated the proportion of field values that appropriately contained either LOINC or SNOMED CT codes in cases where a semantically interoperable code was expected.

Less than 17% of incoming Indiana ELR messages contained a standardized LOINC code identifying the test performed, and none of the results contained a standardized SNOMED CT identifier. For the Wisconsin dataset, none of the ELR messages contain a standardized LOINC code for identifying the test performed, and less than 13% of the test results contain a standardized SNOMED CT identifier [29].

Our analysis demonstrates that very few real-world ELR messages emanating from operational laboratory systems contain standardized codes, even post-meaningful use regulations. To effectively use the data in ELR messages, Indiana and Wisconsin state health departments employ software and personnel to translate the incoming data into standardized LOINC and SNOMED CT concepts that can be utilized by the states’ surveillance systems. In Indiana, the INPC performs the translation of the lab-provided test identifier to LOINC code prior to forwarding the ELR data to the state health department. In Wisconsin, the WDHS implemented commercial software from Atlas Public Health to translate the lab-provided test identifiers and test results to LOINC and SNOMED CT respectively. Similar terminology mediation strategies have been used for data exchange between the U.S. Veterans Administration and the U.S. Department of Defense [30].

2.3. Mapping local codes to standards is challenging and costly

Mapping local terms to concepts in standard vocabularies is challenging because it is a complex and resource intensive process [11,12]. Identifying the correct concept from the standard vocabulary requires both specific domain knowledge and knowledge of the target vocabulary standards. In practice, even physicians and laboratory personnel with a good understanding of the tests at their institution often lack the resolution of knowledge required to successfully map all of their local concepts to standard vocabularies [12].
Unfortunately, local test names often lack information needed to appropriately identify the correct standard concept [12,31]. For example, test names may lack an indication of the specimen type or whether the result returned is quantitative or ordinal (e.g. positive/negative). Similarly, the units of measure associated with the result value may not be available during mapping. Several studies have evaluated different automated tools to assist with the process of mapping local laboratory tests to LOINC [32–36]. Yet, even with the best available automated tools, expert human review is still needed to resolve computer-generated candidate mappings.

In addition, because local and standard vocabularies evolve, the burden of maintaining the mappings is significant, ongoing, and easy to underestimate [37]. We have found that mapping local terms to vocabulary standards can be a rate-limiting step in creating semantic interoperability between systems.

3. Potential strategies for improving semantic interoperability

At present, the United States lacks a comprehensive strategy for broadly leveraging standardized vocabularies necessary for full semantic interoperability of health IT systems. As meaningful use drives more clinical systems to exchange data with public health agencies, both senders and receivers must increase their capacity for semantic interoperability. Informed by our experiences in managing and supporting standardized vocabularies and health IT systems used in public health settings, we describe and discuss three potential strategies for increasing semantic interoperability capacity for surveillance and other public health functions. For our examples, senders are clinical care systems, and receivers represent public health stakeholders.

First, the data sender could be solely charged with the responsibility for translating local codes to standardized concepts. In the case of ELR, this would mean that CMS or public health agencies by fiat would require labs, hospitals, and physician practices to translate outbound ELR message content to specified standards. Second, the onus for translation could be placed on the data receiver. In this case, public health agencies receiving data would accept local or standard codes and perform the necessary translation. A third option would involve a collaborative strategy in which data senders and receivers together work towards semantic integration.

3.1. Option 1: Require Data Senders to Use Standards

If policy mandates that data senders translate their local codes into standard codes by fiat, then the cost and burden solely falls to labs, hospitals, and physician offices. This is consistent with the approach established by the Stage 2 meaningful use regulations [8], in which CMS requires hospitals to submit ELR data to public health agencies. The requirements for this objective specify that laboratory tests must be identified using a LOINC code and non-numeric result values be coded with SNOMED CT (where such codes exist for the items being reported). Therefore this policy is not equitable in the sense burden falls to just one stakeholder group. Furthermore, its acceptance by the affected stakeholder group may be limited. As observed with the final rule for the implementation of ICD-10 in the U.S. [38], health care providers are resistant to one-sided policy solutions that require significant cost and burden [39].

Under this scenario, labs, hospitals and physician offices will continue to primarily use idiosyncratic terms in local concept dictionaries created locally or provided by EHR system vendors. When data from the EHR is to be sent to public health for reporting purposes, the locally used terms will be translated into the required standard terminologies. Translating existing concept dictionaries and maintaining them over time will require dedicated and sustained resources. This would include employing highly trained human resources to perform translation services, or the outsourcing of translation services to a third-party vendor. Allocation of ongoing financial and human resources to maintain translation tables and services will require hospitals, laboratories, and physician practices to rationalize these expenses to their stakeholders.

Several existing health IT initiatives suggest this option to be the default choice or status quo. While hospitals eligible for stage 1 meaningful use incentives could opt to choose ELR as one of several menu set options, Stage 2 meaningful use criteria mandate ELR as a core objective [8]. The standard vocabularies for the Stage 2 criteria specify use of LOINC and SNOMED CT, so hospitals seeking to qualify for this objective must map to these standards. Also through HITECH, the U.S. Centers for Disease Control and Prevention (CDC) has funded the Lab Interoperability Cooperative (LIC), a consortium that includes the American Hospital Association, the College of American Pathologists, and Surescripts [40,41], to support achievement of the meaningful use criteria by enabling 500 hospitals in the U.S. to participate in ELR over a two-year period. To accelerate ELR, the LIC offers education, training, and technical support services to assist hospitals in their mapping efforts.

Expanding the existing policy of placing burden on providers makes it easy to implement from the perspective of the government as it requires expansion of operational programs such as meaningful use. Yet it presents a daunting, under-funded challenge for providers, and it would likely result in fragmented silos of translated terms. While the LIC has been successful in recruiting hospitals to participate in its pilot work, the proportion of pilot participants which will be operational in ELR by then end of 2014 remains unknown. Furthermore, the LIC’s target enrollment of 500 hospitals represents only 8.7% of the nation’s 5724 registered hospitals [42]. This means that the vast majority of hospitals will need to hire, train, or outsource resources for translating their local dictionaries to the standards required for the meaningful use program. It is unlikely that EHR adoption incentives will provide sufficient resources to cover these costs, which are well beyond the costs associated with EHR system implementation. In addition, both the LIC and meaningful use program currently only support translation of local terms for hospitals. There are more than 5800 clinical laboratories that are independent of hospitals but support non-affiliated physician practices [43]. We are not aware of any current initiatives focused on vocabulary standardization for these kinds of laboratories or physician practice EHR systems. Thus current policies fragment semantic interoperability, excluding the systems used in the settings where the majority of Americans receive their care [44].

3.2. Option 2: Public Health Solely Responsible for Translation

In the second scenario, the responsibility to ensure use of standardized codes lies with the “receiver”. Public health agencies would be responsible for hiring, training, and maintaining human and technical resources for translating concepts from the various data senders. While many health departments currently allocate staff to support ELR, significant additional resources would be required for public health to assume full responsibility for translating local concepts for all providers within a jurisdiction.

However, public health funding has been dramatically reduced in recent years with more than half of all and 73% of large (population >500,000) local health departments reporting core funding cuts, necessitating reductions in staff and program resources [45]. Increased resources needed to support semantic interoperability are therefore unlikely due to the associated economic and political challenges. If our nation is to invest more resources into public health, those dollars will more likely be spent on disease...
prevention rather than translating local codes into standardized vocabulary concepts. Moreover, given existing momentum towards requiring data senders to translate local codes it is impractical to reverse course towards standardization at the terminal end.

Therefore this policy option is neither equitable in its distribution of responsibility, nor is it likely to be acceptable as an unfunded mandate to state government who individually fund public health activities in the U.S. Bureaucracy would be redundant and fragmented given that each state would manage its own implementation plan, and the strategy is inefficient as costs would not scale across states for providers who operate regionally or cross-nationally.

### 3.3. Option Three: A Strategic, Cooperative Approach

Under a third scenario, public health would collaboratively develop a strategic plan with data sharing partners whereby all stakeholders that generate and report clinical data would partner to improve semantic interoperability. The onus of translation would not fall disproportionately to any one group, making it equitable. Instead each stakeholder group would invest time and resources into the process of translation to enable full semantic interoperability across the myriad health IT systems and scenarios for public health reporting. So while implementation might be somewhat more complex in this scenario, it is likely to be more acceptable to all stakeholders and incur the lowest cost.

Public health organizations would seek to cooperatively develop minimum requirements for data senders that align with a core set of business processes for streamlining the reporting of clinical data. It is not uncommon in the U.S. for local, state, and federal public health organizations to ask clinical providers to report very similar data sets using disparate methods and formats. For example, infection preventionists in Indiana hospitals were asked to fax information on communicable diseases to the local health department in addition to reporting the same information to the state health department using an Internet-based application that requires manual data entry [46].

To reduce the scope and redundancy of concepts that require translation into standard vocabularies, public health organizations need to strategically develop core ‘value sets’ into which common concepts are consolidated. An example of a value set is the Reportable Conditions Mapping Table, a consolidated list of all possible tests for diseases of interest to CDC and state health departments [47,48]. If a laboratory were to map its local test dictionary to the CDC-maintained value set, then reportable results would more easily be electronically identified and transmitted to public health authorities. Yet the value set contains over 200 tests for Lyme disease, many of which are not performed on a regular basis. Therefore such value sets need to be regularly updated with new concepts and aligned with best practice [49]. An alternate to “all inclusive” value set approach is to create sets representing the most commonly used concepts [50]. Focusing the mapping effort on use case driven value sets will save time and energy with respect to translation, improving the strategy's efficiency. Policies that require data reporting, including future efforts aimed at achieving meaningful use and accountable care, should require use of available, appropriate value sets to encourage adoption of standard vocabularies.

Providers should, in parallel, agree to incrementally transition from idiosyncratic concepts in favor of standard concepts. While there will always be a need for local codes to support special cases in clinical practice, standard terms should become the default. Further, providers must insist that their health IT systems not only “bolt-on” standards—e.g., allow standards to be mapped to local terms—but utilize standards a priori wherever possible as the native identifier of a test or result. Thus when a routine hemoglobin test result is reported for a patient with diabetes, a LOINC code representing the hemoglobin test is used natively, rather than a local code that later needs to be translated to LOINC. The same should be true for other commonly performed tests such as chlamydia screens and blood pressure measurements. Local policy in the form of governance by a Chief Medical Informatics Officer may be more effective than federal policy to implement this part of the strategic plan.

Cooperative action by public health organizations and providers will require support from upstream suppliers of electronic data—other stakeholders that may require policy-driven encouragement. For example, physician offices receive test result data from clinical laboratories. If laboratories utilize LOINC and SNOMED CT codes in their electronic messages, clinical decision support systems and public health IT systems could leverage the standardized codes without translation by the EHR or a third-party health IT system. If standardized codes could not be directly inserted into electronic clinical messages, then the reference laboratory could publish the mappings of its proprietary test codes to LOINC on its website as several do today [51–53]. When standard codes or translation tables are available, downstream health IT systems can more efficiently deliver or act upon clinical information as it is exchanged.

Revisions to the Clinical Laboratory Improvement Amendments (CLIA) program could include requirements for using standards in all electronic messages to health IT systems or the publication and maintenance of translation tables in the public domain that could be easily accessed by health IT systems.

The growing interest in vocabulary standards observed among other key stakeholders including laboratory instrument and test kit vendors is encouraging. Many of the large in vitro diagnostic (IVD) vendors have joined the IVD Industry Connectivity Consortium (IICC, http://www.ivdconnectivity.org), which is focused on modernizing instrument interface and interoperability standards. IICC members have committed to or are migrating toward the Health Level 7 messaging standard and LOINC observation identifier standard for their products. In addition, Regenstrief has had direct interactions with more than 15 laboratory testing vendors as they seek to identify the LOINC codes appropriate for their tests. Policies that promote and incentivize coordination, as well as leadership from public health, would enable downstream data recipients to benefit both from standardization at the point of testing and access to translation tables for historical codes.

### 4. Discussion

While this essay has primarily focused on ELR and public health reporting processes, the notion of semantic interoperability and the strategies described are applicable to a range of use cases in clinical and population health. For example, community health assessments, historically performed by health departments and now required of accountable care organizations (ACOs), may be easier if ACOs can quickly extract data from their information systems using vocabulary standards. Additionally, measuring quality of care is also supported by the use of standards. Several initiatives aim to create ‘e-Measures’ where quality indicators are automatically extracted from EHR systems [54]. Such initiatives will be hindered if they are unable to compare blood pressure values or hemoglobin tests when calculating numerators and denominators as outlined by the National Quality Forum. Native support for clinical data standards enables these and similar efforts that increasingly desire to use the growing volumes of clinical data captured by health IT systems.

Because both clinical and public health organizations benefit from semantic interoperability, a case can be made that they should work together to prevent data pollutants from entering
the pipes feeding the informational waters of the health care system. Collaborative action to achieve common goals further shares costs and risks, making the third option the most equitable, acceptable, and efficient. A summary of all three options with their relative ratings on each of the criteria for policy analysis is presented in Table 1.

Given that many hospitals, laboratories, and public health agencies are already engaged in translation activities at some level, now is the time to create a comprehensive set of policies to create an infrastructure that reduces redundant efforts and maximizes available financial and human resources. Moreover, a clear roadmap with thoughtful, strategic investment in the necessary human and technical resources will maximize the likelihood for a sustainable coordinated pathway to semantic interoperability.

Providers and public health organizations are likely to rely on third-party organizations for assistance under any policy scenario for data exchange in the short term. In Indiana, providers rely on the INPC to assist with terminology mapping. In Wisconsin, the state health department relies on Atlas Public Health for terminology services. Foreign ministries of health are looking to large multi-national organizations to support national initiatives. These observations beg the question whether national efforts in the U.S. could benefit from a publically funded and managed terminology service operated by an entity such as the U.S. National Library of Medicine (NLM). Economies of scale could be created if large-scale mapping efforts were supported nationwide, and created mappings could be shared across EHR vendor platforms. Furthermore, concentrated public efforts could prevent further creation of idiosyncratic intermediary mappings created for specific efforts like meaningful use. While new funding at the national level is unlikely to be popular, it is a policy option worth considering.

Finally, while this analysis highlights priorities for the U.S., the evaluation framework we have outlined can inform strategy formation for non-U.S. entities seeking to develop and implement national eHealth strategies [55–57]. While varying country priorities may result in the selection of different specific strategies, this framework can assist a spectrum of stakeholders by providing a rational approach for evaluating relationships among various semantic interoperability approaches and their corresponding policy dimensions. For example, if a strong central government drives adoption of e-health, it may be preferable to choose the second option where standardization is performed centrally by the ministry of health. Strategy alternatives and criteria must be analyzed through a policy lens consistent with the vision and priorities of a given nation.

5. Conclusions

The adoption and use of health IT to improve population health is a journey, not a destination. There are a number of initiatives and health agencies advancing the goal of interoperable health IT. As an important stakeholder in the U.S. healthcare system, the public health community is well positioned to inform policy related to achieving semantic interoperability. Collaborating with clinical providers and laboratories to incrementally improve the adoption and use of standardized vocabularies will not only support public health reporting and surveillance activities, but also impact broader public health activities and costs associated with using electronic information systems.

Acknowledgments

The authors thank Jason Siegel, Tanya Oemig, and Keith Michelson of Atlas Public Health for their assistance with extraction and analysis of ELR data from the Wisconsin Department of Health Services. The research reported here was supported, in part, by a grant (R01HS020909) from the U.S. Agency for Healthcare Research and Quality (BED and SJJ), a contract (HHSN276200800006C) from the National Library of Medicine (DJV), and by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service (CIN 13-416 (BED). Dr. Dixon is a Health Research Scientist at the Richard L. Roudebush Veterans Affairs Medical Center in Indianapolis, Indiana. Drs. Dixon and Grannis also receive funding from a Robert Wood Johnson Foundation (RWJF) Public Health Services and Systems Research (PHSSR) Mentored Research Scientist Development Award (71596) to study the impact of health information exchange on public health reporting processes. The views expressed in this article are those of the authors and do not necessarily represent the views of AHRQ, NLM, RWJF or the Department of Veterans Affairs.
References


[38] Administrative simplification: adoption of a standard for a unique health plan identifier; addition to the National Provider Identifier requirements; and a change to the compliance date for the International Classification of Diseases. 10th Edition (ICD-10-CM and ICD-10-PCS) medical code data sets. Final Rule Fed Regist 2012;77:54663–726.

[39] Chute CG, Huff SM, Ferguson JA, Walker JM, Halamka JD. There are important reasons for delaying implementation of the new ICD-10 coding system. Health Affairs (Project Hope) 2012;31:836–42.

[40] Conn J. CDC awards $5 million grant for ‘lab interoperability’. Med Healthc Online 2011.


[52] Laboratory Corporation of America. LabCorp LOINC® Map; 2014.

[53] Mayo Medical Laboratories. Logical Observation Identifiers Names and Codes (LOINC®); 2014.
