MEDINFO 2023 — The Future Is Accessible
J. Bichel-Findlay et al. (Eds.)

doi:10.3233/SHTI230924

Interoperability in the Wild: Comparison of Real-World Electronic C-CDA Documents from Two Sources

Brian E. DIXON[a,b,1] and Nate C. APATHY[b,c]
[a] *Fairbanks School of Public Health, Indiana University, Indianapolis, Indiana, USA*
[b] *Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, Indiana, USA*
[c] *National Center for Human Factors in Healthcare, MedStar Health Research Institute, Washington, DC, USA*
ORCiD ID: Brian E. Dixon https://orcid.org/0000-0002-1121-0607,
Nate C. Apathy https://orcid.org/0000-0002-7775-8544

Abstract. Although health information exchange (HIE) networks exist in multiple nations, providers still require access multiple sources to obtain medical records. We sought to measure and compare differences in data presence and concordance across regional HIE and EHR vendor-based networks. Using 1,054 randomly selected patients from a large health system in the US, we generated consolidated clinical document architecture (C-CDA) documents from each network. 778 (74%) patients had at least one C-CDA document present from either source. Among these patients, two-thirds had information in only one source. All documents contained demographics, but less than half of patients had data in clinical data domains. Moreover, data across HIE networks were not concordant. Results suggest that HIE networks have different, likely complementary, data available for the same patient, suggesting the need for better integration and deduplication for national HIE efforts.

Keywords. Health information interoperability, health information exchange, information sources, medical records

1. Introduction

Health information exchange (HIE) networks exist in multiple nations, and they provide a variety of methods for electronic sharing of clinical documents and data [1]. Two dominant forms of HIE networks include: community HIE networks at the local, regional, or state level; and vendor-mediated exchange networks [1]. Community HIE networks often centralize data storage, and participation is generally available to any health care provider organization in the relevant region that is able to contribute data. EHR vendor-mediated networks are brokered by an EHR vendor or a third-party consortium of EHR vendors and are limited in their participation to organizations that use a single EHR (in the case of Epic Systems' CareEverywhere) or one of a set of participating EHR vendors (for example, CommonWell Health Alliance).

---

[1] Corresponding Author: Brian E. Dixon, Regenstrief Institute, 1101 W. 10th St., Indianapolis, IN, 46202, USA, Contact details: +1 317-278-3072 bedixon@regenstrief.org

As of 2019, more than 60% of hospitals reported participation in community HIE networks, and half of hospitals reported use of EHR vendor alliance or single EHR vendor methods of querying patient information from outside organizations [2]. On average, hospitals reported frequent use of two HIE systems, suggesting complementarity across HIE system types [3]. Yet the literature lacks studies directly comparing HIE networks in terms of data accessibility, consistency, or other measures of data quality, which have important implications for the value of HIE networks to organizations taking on the cost and implementation effort to participate in these networks. Given this gap, we aimed to measure differences in data presence and concordance across the two most widely adopted HIE networks.

2. Methods

2.1. Setting

A large academic health system with multiple acute, primary, and specialty care locations in a large metropolitan area (MSA) with a population of 2.1 million. Most residents report their race as white (71%) followed by black (15%), Hispanic (7%), and Asian (4%). The median age is 36.6 years, and the median household income is USD 62,502.

2.2. Participants & Data Sources

The study sample included 1,054 randomly selected active patients. Patients were attributed if they possessed a current address in the MSA and had at least two encounters in 2018 and one in 2017, to illustrate regular and recent receipt of care in the health system. We used Consolidated Clinical Document Architecture (C-CDA) documents from two HIE networks to compare data accessibility and consistency. The C-CDA standard is a base standard, providing a common structure, coding approach, and semantic framework for electronic clinical documents.

The first source was the Indiana Network for Patient Care (INPC) – a mature regional HIE network. INPC contains clinical data for more than 15 million patients from 117 hospitals and 17,000 outpatient practices statewide, covering 80% of the health care provider organizations in the MSA, and 95% of the MSA population. Our second source was the CommonWell Health Alliance data sharing network made available through the health system's Cerner EHR. Hospitals using EHRs equipped with CommonWell's interoperability services can query directly from a patient's chart to obtain records from outside providers. As of January 2021, 248 provider organizations in Indiana were connected to CommonWell, including the academic health system.

Data from both sources were extracted between October 2019 and February 2020, and all data from academic health system visits were filtered out of the C-CDAs, as this information was available to providers directly in the Cerner EHR. C-CDAs from both sources were deposited into a staging database for post-processing into relational tables, which were then provided to the research team for analysis.

2.3. Measures – Data Presence and Concordance

We sought to descriptively measure two aspects of data quality at the patient level: presence and concordance (Figure 1).

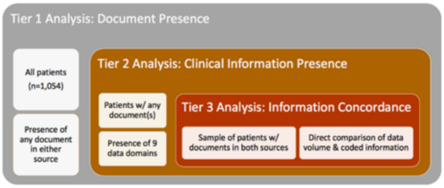

Figure 1. Methods used to examine presence and concordance in C-CDA documents from two sources.

Data presence: a binary measure capturing the ability to obtain patient information. We measured this for each patient for each HIE network and across clinical information domains. At the HIE network level, if a patient had any data in either HIE network, we denoted that data as "present." Importantly, a lack of data in both HIE networks does not necessarily indicate missing information, as these patients may only seek care within the academic health system. Due to this, we limited our clinical information domain presence measures to patients for whom data was available from at least one HIE network, which implies that it could be present in both. We measured the ability to obtain information across the most reliably populated clinical information domains across both sources.

Data concordance: the extent to which both the volume and content of clinical information was identical across sources. This analysis was only possible for patients with information available from both sources. Because a sample of 10 patients with data across all domains would bias the findings, we randomly sampled 10 patients for each data domain separately. We also limited to the following five data domains: encounters, laboratory results, problems, immunizations, and vital signs. We measured concordance by examining individual data elements in each source, measuring the proportion of all patient data elements that were available in both HIE networks, only INPC, or only CommonWell. This step comprised of manual review of patient information and was also done for the sample of 10 patients within each of the 5 data domains listed above.

2.3.1. Data Analysis

First, we calculated the proportion of patients for whom C-CDA documents were present in either or both sources as our measure of overall HIE network data presence. Second, for patients with any information from either source, we calculated for each data domain the proportion of patients for whom data was present from both sources. For volume concordance, we calculated the ratio of INPC to CommonWell data elements for a sample of 10 patients from each of the five data domains noted above. We averaged these for an overall volume concordance measure as well as reporting measures for each data domain. Finally, we manually compared the clinical data present in each source for the same 10 patients across these 5 data domains and calculated the proportion of all data elements that matched in both sources. Data was considered concordant across sources if the observation matched in content and date (e.g., LOINC 4544-3, Hematocrit by automated count, on Jan 1, 2018; ICD-10-CM E11.9, Type 2 diabetes mellitus without complications, on July 12, 2019). Standards are used by providers and Regenstrief to normalize data in INPC. To calculate overall content concordance, we computed a weighted average of the proportion of concordant records, weighted by total number of unique records, such that the proportion of records that were concordant for patients with a greater volume of overall records were weighted more than those of patients for whom, for example, only one or two records were present.

3. Results

3.1. Data Presence

In analyzing data presence at the HIE network level, 778 (74%) of the full sample had at least one C-CDA document present from either source. Two-thirds (N=513; 66%) of those patients had information in only one source. These patients' documents were not evenly distributed across sources: 438 (85% of single-source patients) had information only in INPC, while the remaining 75 (15%) only appeared in CommonWell.

Data presence further varied across clinical data domains. Patient demographics was the only domain present for all patients. This is likely due to the requirement of demographic information like name, date of birth, sex, and race in the C-CDA data standard. Other clinical information domains were less reliably present. Of the 778 patients with documents, 65% had encounter information present, and 52% of patients had problem lists. All other domains were present for <50% of patients. Procedures (16%), immunizations (27%), and medications (29%) were the least present categories.

We found additional variation in a stratified analysis of data domain presence by HIE network. For example, among the 401 patients for whom problem list information was present, 23% of these patients had problem list information in both sources, 57% had this information present in INPC only, and the remaining 20% only had problem lists present in CommonWell. Across all other domains, <25% of patients had information present in both sources. The INPC excelled at providing data on encounters (82%), lab results (81%), and procedures (33%) for the roughly half of patients for whom these data were present. CommonWell excelled at providing medication (97%) and social history (100%) data among roughly one-third of patients with such data.

3.2. Data Concordance

The HIE networks did not illustrate data concordance in volume or content. With respect to volume, documents from the two HIE sources rarely contained the same volume of data. CommonWell generally contained a much larger volume of data, potentially due to duplication of data across C-CDAs for the same patient. For example, CommonWell documents often included multiple records for what appeared to be the same encounter. We found this pattern in 6 of 10 randomly selected patients.

With respect to content concordance, less than one-third of the content matched between documents for the same patient, on average. Immunizations had the highest content concordance, with 31.9% of records available in both sources. For laboratory results, 14.2% of results were available in both sources. Concordance also varied *within* each data domain across data elements. Within demographic data, addresses matched for only 34% of patients, largely due to semantic mismatches. Date of birth data matched for 100% of patients, while gender matched for 98.1% of patients. Marital status matched for 72.5% of patients, and race matched for only 33.2% of patients.

4. Discussion

In a comparison of two distinct HIE networks operating in the US, we found high presence but limited concordance among documents for the same patient population, suggesting that the two HIE networks have different data available for the same patient, likely complementing one another with respect to current and historical data.

Limited data presence suggests that HIE efforts in the US have a long road ahead. Although most patients had data present, few patients had data in both networks, which suggest health systems likely need to participate in multiple HIE networks to access the full breadth of data necessary for clinical decision-making. Furthermore, both HIE networks lacked complete coverage of data domains suggesting that some HIE participants may be submitting or providing incomplete clinical information to the HIE. These are important considerations for the Trusted Exchange Framework and Common Agreement (TEFCA) policy in the US [4].

5. Conclusions

Findings suggest these HIE networks are complementary sources of information about patients from outside providers, yet without a gold standard of comparison it is difficult to assess the relative clinical value of each information source. Given the findings, full integration of data from both sources would likely yield a clinically relevant set of patient data with limited duplication. This hypothesis could be tested in future research.

Acknowledgements

The authors thank the following individuals at the Regenstrief Institute for their support of this work: Ashley Wiensch, MPH, PMP, for project coordination; John Price, for data extraction; and Ian Hays, for data extraction. We further thank Adam Fair, PMP, of the Indiana Health Information Exchange for support with data extraction.

References

[1] Dixon BE. Introduction to health information exchange. In: Dixon BE, editor. Health information exchange: navigating and managing a network of health information systems. San Diego: Academic Press; 2023, p. 3-20, doi: 10.1016/B978-0-323-90802-3.00013-7.

[2] Pylypchuk Y, Johnson C, Patel V. State of interoperability among U.S. non-federal acute care hospitals in 2018. ONC Data Brief. 2020 Mar; (51):1-20.

[3] Everson J, Patel V. Hospital's adoption of multiple methods of obtaining outside information and use of that information. J Am Med Inform Assoc. 2022 Aug;29(9):1489-96, doi: 10.1093/jamia/ocac079.

[4] Adler-Milstein J, Worzala C, Dixon BE. Chapter 21 - Future directions for health information exchange. In: Dixon BE, editor. Health Information Exchange (Second Edition): Academic Press; 2023. p. 447-68, doi: 10.1016/B978-0-323-90802-3.00005-8