Release Version: 20-July-2014

1. Purpose

Clinical research is undergoing a significant paradigm shift. The explosive growth of detailed clinical data captured during routine clinical care has enabled investigators to examine the impact of diagnostic and therapeutic interventions in real world settings. These new data sources, based on data that are not collected and maintained for the primary purpose of a specific research study, are commonly referred to as observational or secondary data. Large-scale, multi-institutional observational study designs are becoming an important complement the traditional small-scale controlled clinical trials for generating new insights and knowledge. National and international multi-institutional clinical research networks using data from electronic health records (EHRs), administrative/claims data from billing and reimbursement activities, patient reported observations, social media networks, real-time devices, and other novel data sources are expanding the scope and depth of available data to answer critical questions about care decisions and outcomes that matter to patients and families. The era of "big data" and large-scale data sharing networks has arrived, along with high expectations that these new capabilities will answer questions that cannot be examined within the traditional controlled clinical trial model.

At the same time, substantial literature suggests data collected in EHRs and other "operational" systems have significant data quality issues. Clinical and patient-centered systems usually are optimized for ease-of-use and efficiency, especially systems used in busy clinical care settings. While data quality concerns exist within systems that have been optimized for clinical research, these concerns are more significant in settings that are focused on clinical care delivery rather than research. As research data warehouses and large scale data networks become established sources of observational data, it is critical that consistent methods for assessing and reporting data quality are developed and adopted so that users of data and consumers of results understand the potential impact of data quality on research methods and findings interpretation.

This project seeks to develop a comprehensive data quality assessment and reporting framework that enables clinical investigators, patients & policy makers to understand the strengths, weaknesses and limitations of observational data used to generate new clinical knowledge. While applicable to prospective data collection, our main focus is to develop methods for evaluating the quality of data that is *not* collected and maintained for the primary purpose of a specific research study.

The current literature that describes data quality concepts, data quality assessment methods, and data quality reporting is fragmented and inconsistent. The inconsistent use of terms to describe different data quality features and ad-hoc approaches to data quality measurement make it difficult to understand data quality results generated by different groups. The lack of a consistent data quality reporting framework also make it difficult to compare data quality results across multiple data partners.

The underlying premise of this work is that by standardizing the concepts used to characterize data quality, the methods used to evaluate data quality, and the metrics and formats used to

report data quality findings, investigators, patients and policy-makers will better understand the limitations of the data, leading to improved statistical methods and an opportunity to provide feedback to data managers to improve data quality. Likewise, confidence and trust in the use of observational data in both single-site and multi-site studies will be improved. As a tangible side benefit, the existence of comprehensive data quality assessment methods can be used to highlight areas needing improvements in data collection and to track the impact of data quality improvement efforts over time.

2. Context & Scope:

Data quality assessment is a continuous process that can occur at multiple stages starting with the data collection steps, such as documentation of clinical findings in the EHR, and also during the extraction and use of data sets for secondary use such as clinical research. Figure 1 highlights at least three major "contexts" that are the focus of this work. Context 1 refers to the initial creation of a broad data resource *at a single institution* that is not tied to a specific study or intended use. Context 2 refers to the initial creation of a broad data resource *across multiple institutions*, also not tied to a specific study or intended use. Context 3 refers to the creation of a specific data resource that is intimately tied to a specific intended use. Context 1 may refer to data extraction from an EHR into a local data warehouse or from a clinical data warehouse to a local research data warehouse. Context 2 may refer to a multi-institutional data warehouse physically housed at a centralized data center or a virtual data warehouse where data are available for research queries but are not physically transferred to a centralized location (1). Context 3 may refer to data extraction directly from an EHR or data warehouse into a study-specific data mart or analytic data set. Context 3 differs from Contexts 1 and 2 where there exists a specific intended data use that may involve specific data quality requirements, such as statistical distribution assumptions for analytic variables. These use-specific data requirements are often labeled "fit for use" data quality features. Fit-for-use data quality requirements may impact the relevance or priority of specific data quality assessment and reporting needs.

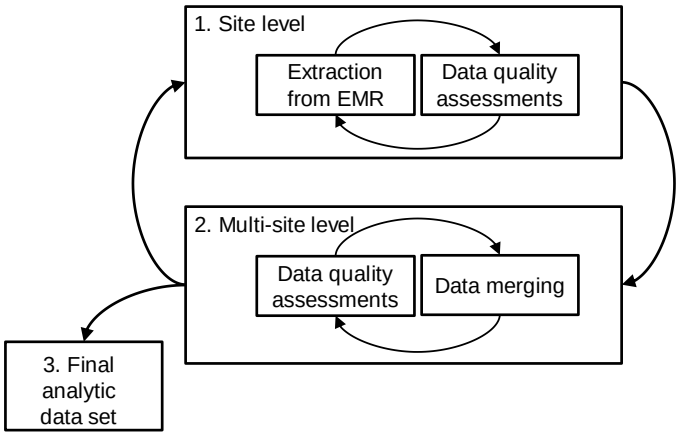


Figure 1: Three contexts of data quality assessment.

We acknowledge that data quality concerns actually begin prior to the data extraction activities that we have defined in Figure 1. For example, relevant data quality assessment planning should begin at the point of clinical observation and data recording into an electronic data collection system such as the EHR. Other relevant efforts are focused on improving the quality of data capture for research purposes at the point of care which, if effective, should improve the data quality measures obtained in all three contexts in Figure 1 (2).

3. Goals of a comprehensive data quality assessment & reporting methodology

A formal methodology provides value to those who use it when the methodology addresses widely accepted critical needs. A formal methodology for data quality assessment and reporting should be able to address critical questions or concerns about data quality. A representative list of data quality concerns, expressed as questions about data quality, that have been described by data users include:

1. **Transparency**: What do we know about the data we are receiving or using?
2. **Provenance**: Do we know how the provided data were created, transformed, combined, cleaned, and imputed from source to final data?
3. **Guidance**: Do we have a clear understanding of the processes/procedures/metrics were used to implement data quality assessments?
4. **Reproducibility**: Are the processes/procedures/metrics used to implement data quality assessments common across groups such that two groups applying the same processes or the same group applying the processes over time would yield the same results?
5. **Comprehension:** Do we have insight into all relevant features of the data we intend to use?
6. **Explanation**: Can we use the results to determine the root-cause of any unexpected data characteristics?
7. **Improvement**: Can we determine which aspects of the data can be improved?

Additional goals that are mostly relevant to Context 3 in Figure 1 are listed below. Context 3 differs from Contexts 1 & 2 because there exists an explicit specific intended data use, such as a specific research hypothesis with an associated statistical method. In that setting, a comprehensive data quality methodology would need to address more specific questions regarding the fitness of the data to be used in the specific intended context/problem. We append these questions with the term "**Fit for use**" to highlight that in Context 3 there is a known intended use of the data:

8. **Fit for use: Assurance**: What evidence do we have that we can use the data for our intended purpose?
9. **Fit for use: Single-site comparability**: Is it safe to combine/merge/integrate data from within an organization together to create a more expansive data set?
10. **Fit for use: Multi-site comparability**: Is it safe to combine similar data sets across multiple organizations to create a larger data set?
11. **Fit for use: Impact**: How do the data quality findings impact the intended study analytic methods and results? Can this impact be quantified?

The questions listed above are useful to guide the development and evaluation of a proposed data quality methodology. As new requirements are added to these lists, the existing data quality methodology needs to be re-evaluated and extended if incomplete.

4. Data quality categories: A harmonized framework

The current data quality literature is inconsistent in the definition of concepts that describe the complex multi-dimensional aspects of data quality. Without a common set of definitions and a consistent use of terms tied to these defitions, it is difficult to compare different data quality assessment methods or to understand the scope and completeness of data quality assessment results. An important contribution of this work is to develop community-based consensus definitions of the key categories of data quality and operational definitions for measuring these categories. While many publications provide definitions for data quality, we started with the glossary of terms provided in the Canadian Institute for Health Information Data Quality Framework (CIHI DQF) (2009), available at no cost at http://www.cihi.ca/CIHI-ext-portal/pdf/internet/data_quality_framework_2009_en.

5. Levels of reporting for Data Quality metrics

Reports of data quality may be requested for different levels of analysis. At each of these levels, different metrics may apply or be available in different settings.

- Data quality metrics may be considered at the level of a data value, a data element, a data vector, a data set, a database, or a collection of databases.
    - Data elements are collections of individual data values (a "column" of a data set)
    - Data vectors are collections of data values for different data elements (a "row" of a data set)
    - Data sets are collections of different data vectors
    - Databases are collections of named data elements and their values, typically capturing important relationships among values.
    - Multi-institutional databases are databases generated by independent organizations that attempt to represent the same data concepts across institutions.
- A different categorization of data quality metrics may be the data generation context, such as during routine clinical care, research data collection, or extraction from a primary or secondary data source. In this setting, a primary data source might be the electronic medical record or patient-entered diary. A research data collection source might be an electronic case report form or research survey. A secondary data source might be a clinical data warehouse.
- Metrics might also be reported based on unique features of the data, such as patient populations, time periods, data sources, or contributing institutions.

6. Intrinsic and extrinsic sources for data quality evaluations

Intrinsic attributes of data quality do not require application of knowledge to assess. These are typically context-independent measures of reliability and completeness of data elements over time or across instruments.

Extrinsic sources are frequently required to generate context-dependent validity and "fitness for use" metrics that depend on one or more [relative] gold standards, external data, or heuristics. In implementation, these can be made more scalable and independent of context by leveraging

resources from well-regarded databases of national statistics or well-curated knowledge sources such as the Unified Medical Language System at the National Library of Medicine.

7. The Role of Metadata

Metadata is information about data (3). Metadata is critical to ensuring a common understanding of the intended meaning, strengths, weaknesses, and assumptions for interpreting a data element and its value. The Data Documentation Initiative (DDI) is a metadata standard developed for describing social science data sets (4) but this standard has not been widely applied to clinical data. Metadata is frequently required to assess fitness for use in any given case.

Key Data Element (DE) Descriptions (metadata):

- Well-documented conceptual and operational definitions, including any changes in conceptual or operational definitions over time.
- Enumeration of valid values, including use of syntactic data standards such as standardized coding systems, units and significant digits, and any changes in coding standards or valid values over time.
- A list of business rules applied to its data values.
- A description of measurement/observation method(s) with relevant citations.
- A statement of known sources of variation that influence measure/observation method/source
- A statement that the data element is (or is not) part of a validated instrument
- Data showing the data element performance characteristics in terms of Sensitivity, Specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV)

General principles for meta-data standards are under development in ISO11179. While developing a comprehensive model for clinical metadata is out of scope for the proposed work we do propose a growing list of fundamental metadata elements that are driven by CER/PCOR use cases. These might include, for example, the settings in which data elements have been collected, the particular instruments or EHRs, the nature of the health systems, a history of data processing rules used to generate the final values, etc.

Table 1 provides an integrated and harmonized set of categories to describe the various aspects of data quality. Some categories contain subcategories that address specific features of that data quality category. As a starting point, we used terms found in data quality publications and standard operating procedures (SOP) manuals from multiple well-established national data coordinating centers (5–15).

In reviewing the various data quality assessment approaches, the categories described in Table 1 are consistently mentioned. While the published definitions for these features vary somewhat, the key data quality concepts are similar, indicating sufficient commonality that these are critical features to include in a combined data quality category. High-level definitions for each category are provided. Detailed operational and measurement definitions will be filled in a later version of this document.

Table 1: Harmonized data quality categories, subcategories, high-level definitions and examples.

DQ Category	General description	Operational definition	Examples
Validity	The degree to which the data succeed in quantifying the intended concept. Validity has subcategories:		
	• Concurrence (Criterion Validity): The extent to which data are in agreement with **external "gold" standards** and/or known relationships	• Data values are consistent with well-regarded external data/knowledge resources in similar context (e.g. populations) . See diagram of hierarchy of data accuracy assessments by Zozus (Table 2). • Distributions of data values within matched sub-groups are not significantly different from gold standards	• The observed distribution of ethnicity from a clinical data source is not significantly different from the [gold standard] census data.
	• Plausibility (Face validity): The extent to which data are accepted or regarded as true, real, and credible based on **heuristics and/or expert judgment**	• Data fit expected constraints, such as dependencies with related variables, extremes boundaries, and reality/external knowledge.	• Sex agreement with sex-specific contexts (pregnancy, prostate cancer) • Collections of related measures are physiologically consistent (AST, ALT, bilirubin, PTT in liver failure)
	• Concurrence (Construct validity): Values or distribution of data values for comparable and independently collected data are **similar** (concur with each other)	• Concurrence of independent measures of data values known to measure the same quantity in two different data sources. Analogous to with inter-rater reliability but focused on independent measures rather than the repeated observations of the same measure • Similarity of distributions in two different data sources that are expected to measure the same data elements in the same sub-population	• HbA1c results from two laboratory measurements are statistically not different. For example, HbA1c values measured at the same time but performed at a local lab and a central lab are similar. • Regional prevalence of diabetes are similar between two independent claims databases from different national payers that serve the same general population.
Completeness	The extent to which the quantity or volume of available data is appropriate. This category has a temporal and temporal subcategories.	• Measures of "data density" against a relevant denominator dimension, such as per patient, per time frame, between date ranges. • The temporal sequence of data is sufficient to capture the dynamic aspects of an event or process. Includes measures of trends, transitions and retention. Serial measurements over time are sufficient to detect clinical state. • Missingness measures: Missing completely at random (MCAR), missing at random (MAR), missing not at random (MNAR) -- see http://en.wikipedia.org/wiki/Missing_data	• The number of patients enrolled per year does not show extreme spikes or dips not explained by external events such as mergers and separations. • Serial hematocrits -- annual measurements may be sufficient in stable patients but insufficient in patients with GI bleeding. • Missingness by indication (MNAR)-- liver function tests are only obtained when the patient has a liver disease.
Non-redundancy	The extent to which records represented as independent observations are not duplicates of the same source information		Encounters obtained via electronic health records and via billing claims records are reconciled and not double counted.

DQ Category	General description	Operational definition	Examples
			Medication records obtained from drug fulfillment and from EHR prescriptions are reconciled and not double counted.
Reliability	The extent to which values are the same across repeated trials, measurements or observations. Reliability has multiple subcategories:		
	• Inter-rater reliability – the extent to which two observers or instruments agree on a value of a data element	• Inter-rater reliability comparison of two observers'/systems' data values intended to represent a single quantity. [This may be redundant with convergent validity in our case]	HbA1c measurement values from hospital and outpatient labs are statistically similar under the same conditions: same person, same sample time.
	• Temporal reliability (Test-retest reliability) – the extent to which a data element retains stable values over time.	• Changes in value of a data element • Changes in the distribution of values of data elements	No significant or unexpected changes in frequency of diabetes diagnoses between adjacent time intervals.
	• Data processing reliability – the extent to which derived data values are faithful to the quantity of interest (e.g. bugs have not been introduced during programming)	• Convergent programming • Comparison to upstream data sources	Two programmers develop extracts and get the same summary statistics A data value does not "go missing" between data processing/transmission steps
Quality of documentation and metadata	• In absence of a standard metadata model, can a **standard set** of accuracy/fitness for use questions be answered? • In the *presence* of a standard metadata model, what is the missingness and completeness of the metadata.	• Are the "instruments" for collecting and encoding each data element well documented? • Are the source values and data processing rules well-documented? • Is the context of data collection for each data element well-documented?	
Additional features of data quality that are can be significant issues <u>in specific settings</u> but are deemed not in scope include: systems-related features such as security, user rights, system availability, and documentation; provenance-related issues such as non-repudiation and fraud; governance-related issues such as availability (e.g., willingness to share data)			

Table 2: Hierarchy of accuracy of data *values* assessment methods. By Zozus [under review].

Comparison to a source of truth Comparison to an independent measurement	↑ Accuracy
Comparison to independently managed data Comparison to an upstream data source	↓ Partial accuracy
Comparison to a known standard Comparison to valid values	↑ Discrepancy detection
Comparison to validated indicators Comparison to aggregate statistics	↓ Gestalt

A data value might be evaluated at different levels of confidence. From strongest to weakest assessment of accuracy:

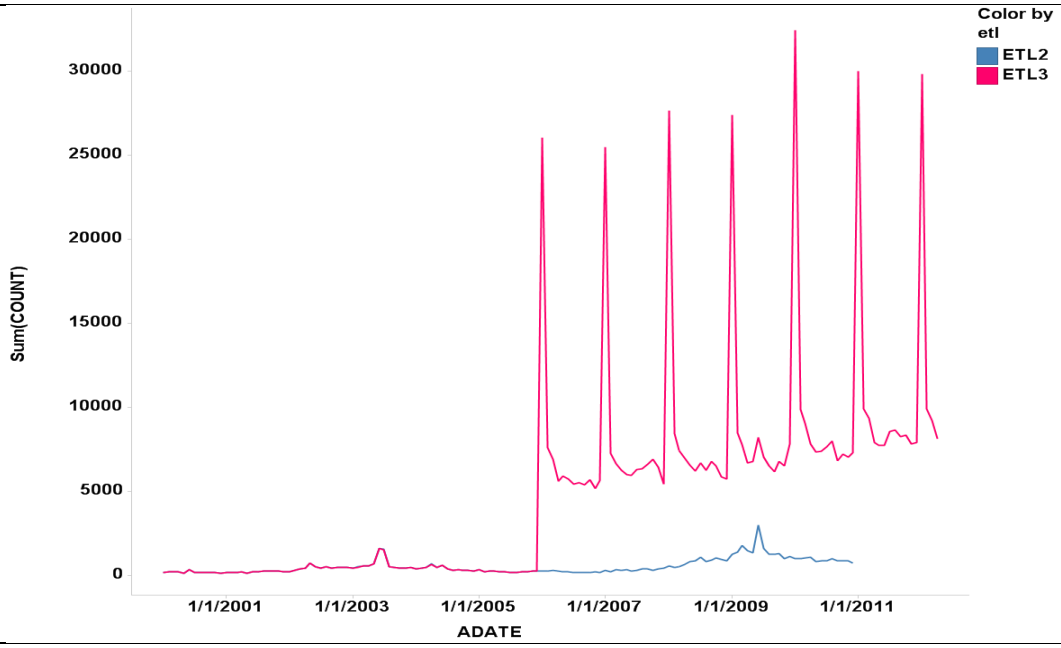
- "Source of truth" is comparison to a gold standard (criterion validity)
- Independent measurement of the same data element
- Comparison to independently managed data - convergent evidence (construct validity)
- Known standard (e.g. range check)
- Validated values (e.g. nothing negative in age)
- Validated indicators (face validity heuristics with demonstrated correlations with other data quality)
- Comparison to aggregate statistics (e.g. outlier values within a dataset)

Appendix (added 20 July 2014)

Critiques of earlier versions of this document requested examples of data quality issues that are tied back to the data quality categories in Table 1. Members of the EDM Forum and PCORI data quality collaborative provided recent examples of data quality issues. As more examples are added to the library, they will be included here and linked to the proposed data quality categories.

The examples below have been altered to prevent association of the example to any organization or data-sharing network.

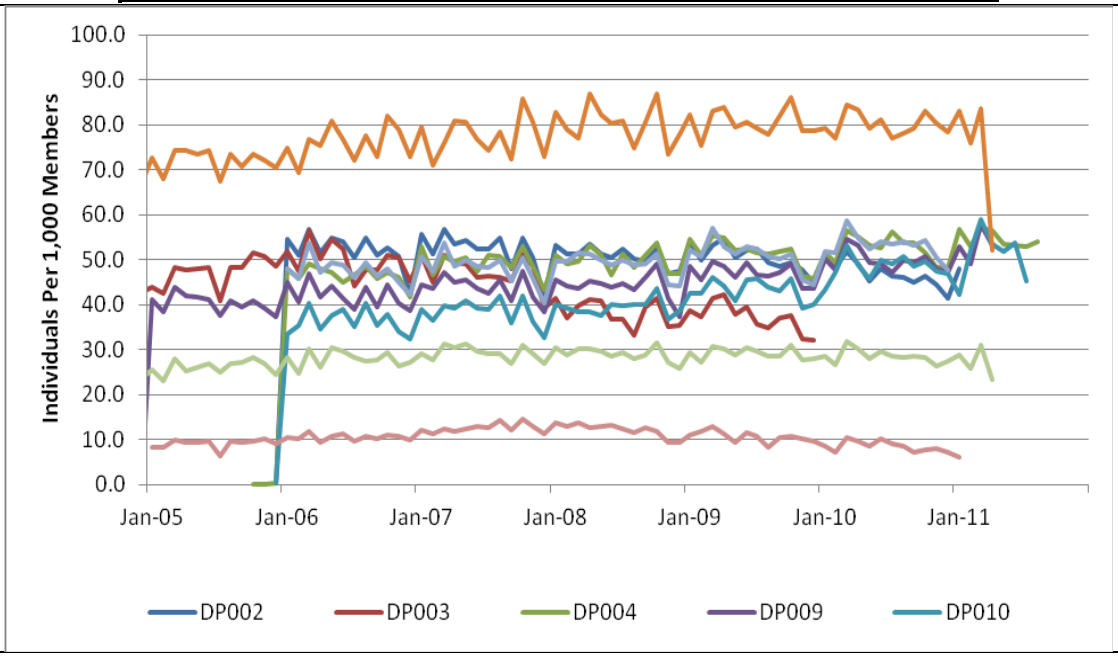
A recent review of diagnosis codes (ICD9-CM) and medication codes (NDC) in data extracted from four different electronic health record systems showed that approximately 1-4% of entered codes did not comply with published standard coding practices. This is an example of concurrence (criterion validity) issues.



An example of a plausibility (face validity) issue. The number of visits per month over time is not expected to show large spikes during the month of December.

Data QC1: Demographics

	Site A		Site B		Site C		Site D		Total	
Number of Records	31,518		54,252		170,365		124,705		380,840	
Number of Patients	31,518		54,252		170,365		124,705		380,840	
Sex										
Female	20,403	65%	34,094	63%	106,282	62%	77,974	63%	238,753	63%
Male	11,111	35%	20,158	37%	64,068	38%	46,717	37%	142,054	37%
Other/Missing	4	0%	0	0%	15	0%	14	0%	33	0%
Total	31,518		54,252		170,365		124,705		380,840	
Age 2009										
0-4	34	0%	55	0%	317	0%	255	0%	661	0%
5-9	101	0%	321	1%	716	0%	734	1%	1,872	0%
10-19	1,132	4%	2,714	5%	6,035	4%	5,661	5%	15,542	4%
20-39	5,873	19%	12,165	22%	30,335	18%	21,526	17%	69,899	18%
40-64	15,952	51%	29,860	55%	81,804	48%	58,139	47%	185,755	49%
65-74	3,695	12%	5,219	10%	23,454	14%	18,840	15%	51,208	13%
75-99	4,688	15%	3,891	7%	27,547	16%	19,476	16%	55,602	15%
100-119	41	0%	27	0%	157	0%	74	0%	299	0%
Other/ Missing	2	0%	0	0%	0	0%	0	0%	2	0%
Total	31,518		54,252		170,365		124,705		380,840	
Range of DOB										
Min DOB	1892.05		1905.01		1896.04		1898.12			
Max DOB	2008.10		2008.03		2008.11		2008.07			
mean	53.7		48.9		54.3		54.0		48.9 - 54.3	



Two examples of concurrent (construct) validity checks. In the top panel, the highlighted rows are consistent across sites as expected. A discrepancy could be acceptable if there was a known difference in the included populations (e.g. a prostate cancer population versus a pediatric population). The bottom panel shows the number of test results per 1,000 members over time. It illustrates a concurrence validity concern. Each data partner (DP) appears to have an internally consistent definition of a test and a member (relatively consistent test per member rates within a data partner) but markedly different definitions of either a test or a member (markedly different test per member rates across data partners).

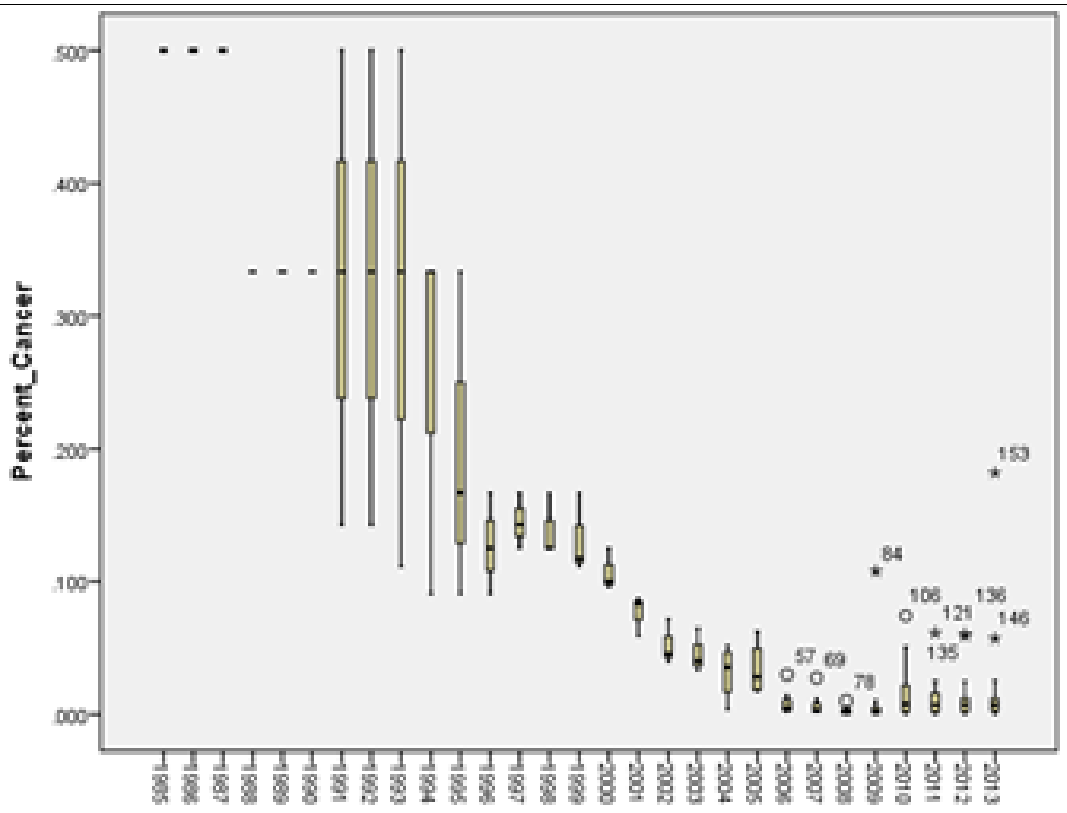
Data QC1: Procedure

	Site A		Site B		Site C		Site D		Total	
Number of Records	5,354,897		13,108,382		18,762,527		6,198,190		43,423,996	
Number of Patients	30,895		53,991		403,946		120,019		608,851	
Avg Px per Patient	173		243		46		52		71	
Encounter Type: Records										
AV	3,844,945	72%	6,786,218	52%	16,829,425	90%	5,100,796	82%	32,561,386	75%
ED	186,139	3%	935,130	7%	#N/A		602,739	10%	1,724,008	4%
IP	428,795	8%	717,638	5%	1,933,102	10%	494,655	8%	3,574,190	8%
IS	46,501	1%	258,436	2%	#N/A		#N/A		304,937	1%
LO	497,996	9%	1,040,657	8%	#N/A		#N/A		1,538,653	4%
OE	164,980	3%	3,160,090	24%	#N/A		#N/A		3,325,070	8%
RO	185,541	3%	210,104	2%	#N/A		#N/A		395,645	1%
TE	#N/A		109	0%	#N/A		#N/A		109	0%
Total	5,354,897		13,108,382		18,762,527		6,198,190		43,423,998	
Encounter Type: Patients w/										
AV	30,787	100%	53,677	99%	167,661	42%	118,090	98%	370,217	61%
ED	17,770	58%	36,743	68%	#N/A		58,227	49%	112,740	19%
IP	15,257	49%	24,005	44%	343,315	85%	61,725	51%	444,304	73%
IS	3,041	10%	4,913	9%	#N/A		#N/A		7,954	1%
LO	19,386	63%	43,115	80%	#N/A		#N/A		62,501	10%
OE	11,139	36%	50,664	94%	#N/A		#N/A		61,803	10%
RO	22,149	72%	33,094	61%	#N/A		#N/A		55,243	9%
TE	#N/A		62	0%	#N/A		#N/A		62	0%
Date Range									range	
Min	2002.01		2000.03		1985.08		2002.01		1985.08 - 2002.01	
Max	2008.12		2008.12		2009.03		2008.12		2008.12 - 2009.03	

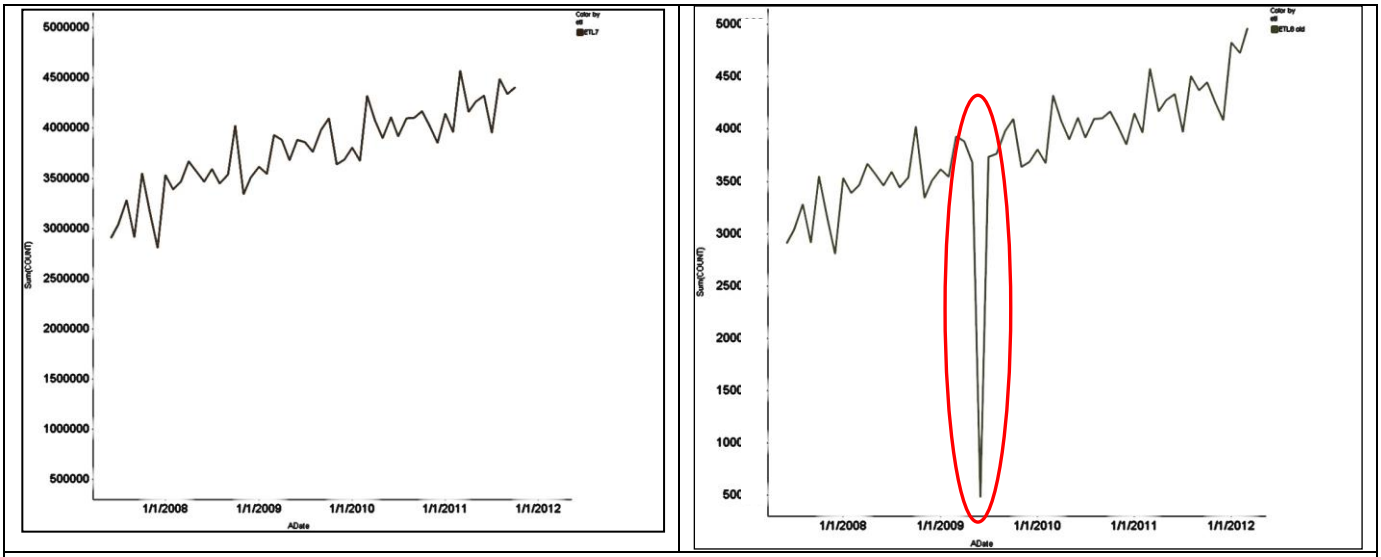
An example of missingness completeness issue in the distribution of Encounter Types. Sites C & D have substantial missing data is some categories but not in other categories. This is an example of missing not at random (MNAR).

The "Avg Px per Patient" and "Date Range Min" rows highlights concerns over concurrence (construct) validity. For "Avg Px per Patient", further investigations might reveal significant differences in business rules being used to assign procedure (PX) records to patients in Sites C & D, raising concerns about the validity of the measure across sites.

In a health care organization in the United States, a significant number of patients noted to be English speaking were also noted to require English-language interpreters. Similarly, at a pediatric facility, a small number of six-year old patients had an assigned marital status of married, separated, divorced, or unknown. These are examples of possible plausibility (face validity) data quality issues.



Annual box plots of percent of patients assigned a diagnosis of depression across multiple clinics each year. Over time, the size of the box and extreme values narrowed significantly, resulting in a markedly improved diagnostic reliability over time.



An example of data processing reliability. The sudden appearance of a large reduction in counts at one point in time (right panel) that was not present in a previous extract (left panel) suggests a programming error.

An example of a data quality issue in the context of "fitness for use" is the use of billing diagnosis codes to detect clinical conditions in patients for cohort selection. Bill codes are generated using specific rules that are imposed on an organization by outside payer policies and payment regulations. Within these constraints, billing practices attempt to optimize financial reimbursement rather than maintain clinical accuracy. In addition, deficiencies in billing codes, such as the inability to distinguish between "rule-out myocardial infarction (MI)" from a proven MI, also contribute to data quality issues -- both concurrence (criterion validity) and plausibility (face validity) data quality concerns. Of note, in this example, "fitness for use" suggests that the same data may have markedly different data quality features. From the perspective of a billing department, billing data that may be unfit for defining clinical cohorts may be entirely acceptable and deemed of high quality for billing purposes

References

1. Brown JS, Holmes JH, Shah K, Hall K, Lazarus R, Platt R. Distributed health data networks: a practical and preferred approach to multi-institutional evaluations of comparative effectiveness, safety, and quality of care. Medical care. 2010 Jun;48:S45–51.

2. Kush R, Alschuler L, Ruggeri R, Cassells S, Gupta N, Bain L, et al. Implementing Single Source: the STARBRITE proof-of-concept study. J Am Med Inform Assoc. 2007 Sep;14:662–73.

3. National Information Standards Organization (U.S.). Understanding metadata. Bethesda, MD: NISO Press; 2004.

4. Vardigan M, Heus P, Thomas W. Data Documentation Initiative: Toward a Standard for the Social Sciences. International Journal of Digital Curation. 2008 Aug 6;3(1):107–13.

5. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. Journal of the American Medical Informatics Association : JAMIA. 2013 Jan 1;20:144–51.

6. Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. J Biomed Inform. 2013 Oct;46(5):830–6.

7. Arts DGT, De Keizer NF, Scheffer G-J. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. J Am Med Inform Assoc. 2002 Dec;9(6):600–11.

8. Brennan PF, Stead WW. Assessing data quality: from concordance, through correctness and completeness, to valid manipulatable representations. J Am Med Inform Assoc. 2000 Jan;7:106–7.

9. Kahn MG, Raebel MA, Glanz JM, Riedlinger K, Steiner JF. A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research. Medical care. 2012 Jul;50 Suppl:S21–9.

10. Mini-Sentinel Coordinating Center. Mini-Sentinel Standard Operating Procedure: Data Quality Checking and Profiling. 20; Available from: http://www.mini-sentinel.org/work_products/About_Us/Mini-Sentinel_SOP_Data-Quality-Checking-and-Profiling.pdf

11. Nahm M. Data quality in clinical research. Clinical Research Informatics. London: Springer-Verlag; 2012. p. 175–201.

12. Nahm ML, Pieper CF, Cunningham MM. Quantifying data quality for clinical trials using electronic data capture. PLoS ONE. 2008;3(8):e3049.

13. Wand Y, Wang R. Anchoring data quality dimensions in ontological foundations. Comm ACM. 1996;39:86–95.

14. Wang R, Strong D. Beyond accuracy: What data quality means to data consumers. J Management Information Systems. 1996;12:5–34.

15. Liaw ST, Rahimi A, Ray P, Taggart J, Dennis S, de Lusignan S, et al. Towards an ontology for data quality in integrated chronic disease management: A realist review of the literature. International Journal of Medical Informatics. 2013 Jan;82(1):10–24.