



Published in final edited form as:

*Stud Health Technol Inform.* 2013 ; 192: 1208.

## Data Standards Data Quality

**Meredith Nahm and W. Ed Hammond**

Duke Center for Health Informatics, Duke University, Durham, NC, USA

### Abstract

The relationship between data quality and data standards has not been clearly articulated. While some directly state that data standards increase data quality, others claim the opposite. Depending on the type of data standard and the aspects of data quality considered, both arguments may in fact be correct. We deconstruct a typology of data standards and apply a dimensional definition of data quality to clearly articulate the relationship between the two, providing a framework for data quality planning.

### Keywords

Data standards; data quality; information quality; Clinical informatics; Clinical research informatics

### Introduction

Depending on the type of data standard and the aspects of data quality considered, data standards may or may not impact data quality. Unfortunately, the language we use lacks specificity both about data standards and data quality. Quality likewise is an imprecise term. There are at least two world-views from which quality is approached: 1) quality as conformance to specifications vs. 2) quality as fitness for use. [1] While these need not be different, in practice, there may be divergence. In the absence of context dependent definition of data quality, significant expectation differences between suppliers and customers can result. Likewise, there are different perspectives from which data quality is approached. For example, some view data quality as those things pertaining to the data values themselves while others also include contextual features such as timeliness, and ease of access. Lack of specificity about data standards and data quality enables diverging expectations of the value of data standards and impedes their adoption and progress. We hope instead that clarity in language and intent will hasten the promise of standardization that has benefitted so many other industries.

---

© 2013 IMIA and IOS Press.

This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License.

Address for correspondence: Meredith Nahm, Meredith.nahm@duke.edu, 300 W. Morgan St. suite 1100 Durham, NC, 27701, USA.

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

## Types of Data Standards

We have yet to see a classification of data standards. Such a classification might be based on the aspects of reality (or data about it) that are represented. Such a classification system would surely increase the explicitness with which we talk about data standards, and would aid in clarifying the relationship between data standards and data quality.

## Types of Data Quality

Data quality has been defined as a multidimensional concept. [2] Some dimensions such as accuracy are properties of individual data values and are context independent while others are dependent on the context of use. Data quality dimensions are used to define and articulate expectations of data. Data quality dimensions have been articulated for regulatory decision-making in clinical trials for marketed therapeutics. [3] Such dimensions have not yet been articulated for other healthcare contexts. It is clear, however, that use of data by those other than those who originally collected them, necessitates additional dimensions. Until definitions of data quality for healthcare have been articulated, the term, “data quality” with respect to health care data has little meaning.

## Data Standards and Data Quality

By merely deconstructing 1) the types of data standards and 2) the dimensions of data quality, the inequality of data standards and data quality becomes apparent. Different data uses, necessitate different quality dimensions. Different types of standards impact data quality in different dimensions. With this approach, we have a way to examine a data standard and assess its impact on data quality.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported by the National Library of Medicine through grant number 1K99LM011128-01A1.

## References

1. Dictionary entry for “quality”, Quality Glossary. American Society for Quality (ASQ); available from <http://asq.org/glossary/q.html> [Accessed December 10, 2012]
2. Wang R, Strong D. Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*. 1996; 12(4):30.
3. Department of Health and Human Services, Food and Drug Administration. *Guidance for Industry Computer Systems Used in Clinical Trials*. Apr.1999