

AN EVALUATION OF DATA QUALITY RELATED PROBLEM PATTERNS IN HEALTHCARE INFORMATION SYSTEMS

-RESEARCH IN PROGRESS-

Cathal Ó Riain, Markus Helfert
School of Computing
Dublin City University, Ireland

ABSTRACT

Governments increasingly recognize the importance of healthcare information systems (HIS) as a tool to decrease costs, streamline workflow and enhance patient care. Many approaches have been taken to improve HIS in the international community. Our research aims to categorise problems related to these HIS using the concept of data quality and the information product approach. This paper illustrates how the implementation of total data quality management (TDQM) methodologies could help to solve problems currently evident in HIS in Europe. Information management strategies employed by healthcare systems in Europe are discussed and examined. It is our hypothesis that data quality-based management methodologies provide an efficient vehicle for reform in healthcare systems. We are undertaking further research in this area to examine the requirements for, and effects of data quality management in healthcare settings through empirical evaluation of healthcare information management strategies in Europe.

KEYWORDS

Data quality, data quality management, information quality, information management, information product

1. INTRODUCTION

Health is one of the largest consumers of public funds in the world, accounting for an average of 8.4% of GDP in OECD countries (OECD Health Data, 2003). It exerts significant influence over the economic policy of governments, which tend towards two main objectives in the health sector: to improve the quality of care, while reducing the cost of its provision. Expenditure on health systems is growing rapidly worldwide (Sein and Waheed, 2003). But strategies to improve health information systems (HIS) are generally not well designed and articulated (WHO, 2001), resulting in inefficiency and failure to deliver optimal health services.

The introduction of modern information systems has improved efficiency and effectiveness in many information intensive industries, healthcare among them. The consensus is that data quality is a serious problem for most organisations, and one that deserves more managerial attention. Healthcare organisations are no exception in this regard (Wang et al. 1995; Abate et al. 1998; Eckerson 2002; Gendron and D'Onofrio 2000; Leitheiser 2001). In this paper, we will classify important issues in HIS in Europe using the information product approach to data quality management (DQM), showing that these issues follow established DQM problem patterns. Based on this observation, we conclude that DQM provides an effective vehicle through which healthcare strategies can be formulated to enable HIS to achieve their goals.

This literature review paper is organized as follows: Sections 2 discusses the concept of data quality in information systems. Sections 3 and 4 provide an evaluation of HIS strategies in Europe, and the categorization of issues encountered in European HIS using the information product approach to DQM. Section 5 highlights the impacts of poor data quality in healthcare. We conclude with our plan for continued research in the area.

2. DATA QUALITY IN INFORMATION SYSTEMS

DQM has developed beyond the traditional view of data quality as a synonym for data accuracy. Wang and Strong (1996) note that in order to improve data quality in an organization, a multi-dimensional view of the concept must be taken. Indeed, accuracy is just one of four dimensions of data quality identified by researchers, where completeness, consistency and timeliness are seen as the other main facets of the concept (Ballou and Pazer 1985; Wand and Wang 1996). Tayi and Ballou (1998) state that the term “data quality” can be defined as “fitness for use”. This implies that data used in one area may not be suitable for the needs of another. Accurate data are not “fit for use” if they are not provided in a timely, easily understood manner.

3. EVALUATION OF HEALTHCARE SYSTEMS IN EUROPE

Gnesotto and DeVogli (2003) note many weaknesses inherent in HIS in Europe. The mandates of organizations managing systems are too vague, with standard operating procedures frequently missing. Information system related strategies often lack focus, particularly in southern European countries (e.g. Italy and Portugal). Assignment of tasks among organizations for collection, analysis and communication of information is regularly unclear. Some organizational processes are designed as isolated elements, instead of parts of an overall system.

These weaknesses manifest themselves in the form of inefficiency. Tasks are undertaken by many organizations, instead of each one focusing on what it does best. HIS outputs are often late, unseen, unused or plainly irrelevant. The poor quality of information provided means that decision makers have access to diverse or contradictory information, compromising data credibility and impairing strategy formulation.

In response to these problems, many countries have initiated strategies to improve their health information capacity. Increased demand for evidence on which to base health decisions and the advance of technological capabilities further motivate the development of these strategies. However, these strategies are often prone to failure when implementation is attempted (Hackney and McBride, 2002). We observe below that many problems in HIS in Europe are related to poor data quality. Thus, we hypothesize that data quality-aligned management strategies provide a comprehensive framework to overcome these problems. Further study to evaluate European health information strategies and how they propose to rectify these issues is in progress.

4. DQM RELATED PROBLEM PATTERNS IN HEALTHCARE

Wang et al. (1998) introduced the information product approach to DQM. The authors note that managers observed across many industries do not convincingly act on the belief that DQM is key to their success, and formulate a framework for the categorization of an organization’s data quality problems where DQM is not implemented. Below, we use this framework to examine and categorize HIS deficiencies in Europe.

4.1 Failure to Understand Consumers’ Needs

A need exists to recognize the information requirements of two different types of information consumer – both the external and the internal customer. For a patient healthcare episode, many stakeholders each have their own information requirements. Doctors need timely data to assess patient condition and evidence-based information on best practice for treatment. Nurses need information on test results and what interventions are required for the patient’s condition. The patient needs to know how to aid the recovery process. The forgotten consumer in healthcare is hospital management, who need to be able to plan, resource and finance the entire process in an efficient manner. Treatment cost information is frequently lost in the system so value for money cannot be determined; beds lie empty while patients remain on waiting lists. The Irish Department of Health and Children (2004) recognizes that its current infrastructure is inadequate to meet the complex information needs of a modern health service, including core areas of performance management and value for money. To provide an optimal service, all information consumers’ needs must be understood and addressed.

4.2 Poorly Defined Information Production Process

Information tends to be produced by organizations within national health industries in non-standard ways. The importance of standards and clearly defined industry-wide information production methods are essential in healthcare. Data records for a patient are of little use in one hospital if the patient is being treated in a different setting, where the records cannot be accessed and have no meaning. Dermot Smith (2004) recognized the fragmented nature of the technical architecture of Irish Healthcare systems, and the lack of standards and modeling techniques to define healthcare processes in Ireland, calling for systems integration in the sector to address the issue.

The patient electronic healthcare record (EHR) should be an interoperable set of data, usable by systems in any healthcare setting. European countries vary in their development of the EHR to date. In less developed countries such as Greece, a lack of technological preparedness has limited development of an EHR in health systems (Orfanidis et al, 2004). Countries with higher living standards are more advanced; Denmark, the UK and Sweden have developed 10 digit numeric codes. An identifier based on the social security number is being developed in Ireland.

4.3 No Product Life Cycle

In the absence of the EHR in most healthcare settings, paper based processes have evolved to satisfy requirements. One of these processes dictates that at one studied hospital, healthcare episode files are shipped off site 2 months after the episode occurs due to on-site storage constraints (Ó Riain and Helfert, 2004). Effectively, once the immediate information requirement has been satisfied in the system and the patient is discharged, the tendency is to mark the issue as 'solved' and neglect to maintain and improve it. However, this information continues to be useful after it is rendered more difficult to access, especially if the patient presents again with similar symptoms after the files have been moved. The life cycle of this data should be re-examined and should endure as long as the organization has use for it. A DQM program would highlight the importance of the information life cycle and rectify this deficiency in the system.

4.4 Lack of an Information Product Manager (IPM)

IT departments in hospitals regularly find themselves reacting to ad hoc requests focusing on the short term, more immediate goal of care for the patient and provision of the best medical care rather than formulating a clear strategy to mitigate risk and improve customer service. Without an IPM to formulate clear information strategy, administrative procedures, such as admissions, have evolved that lead to inefficiency in the system. English (1999) asserts that the IPM is accountable for implementing processes to assure and improve information quality. Beds lie empty while patients remain untreated. Controls to remedy the inefficient admissions processes are not in place. Metrics to measure performance of the process or the system as a whole are not available. These are fundamental points that would come under the remit of an IPM in a hospital. In the absence of such a role, inefficiency can flourish. A DQM program would highlight these IPM responsibilities and allow hospital management to develop longer-term strategy and focus.

5. THE IMPACT OF POOR DATA QUALITY

Having categorized the problems in HIS and recognized the linkage with data quality, we now turn our attention to the impacts of poor data quality. Data quality is a notoriously difficult concept to measure. Redman (1996) estimates that the typical industrial data quality error rate of 1-5% can constitute a 10% loss in revenue. The Institute of Medicine (2000) believes that poor data quality is responsible for up to 98,000 deaths in U.S. hospitals each year. Eckerson (2002) estimates that data quality problems cost U.S. businesses more than \$600 billion per year. Health providers and insurers are not immune to this. Empirical research is not readily available to show the effect of poor data quality in the European healthcare arena specifically, but these figures are compelling nonetheless. We are creating a research plan to investigate these costs in detail and develop a cost-benefit analysis for data quality program in a healthcare setting.

The program for change in the Irish Healthcare system (2001) recognizes the importance of the data quality question. The strategy, lauded as a benchmark by Gnesotto and DeVogli (2003), broadly outlines four principle goals for the Irish Healthcare System; Better Health for Everyone, Fair Access, Responsive and Appropriate Care Delivery, and High Performance. Each goal can only be planned, supported and evaluated through the effective use of information. In order to enable strategic management of the system, quality information is required to evaluate each of the four goals. Without this information, managers are doomed to tactical, short-term management, which breeds the inefficiency currently prevalent in the Irish health system.

We believe that the healthcare issues outlined in this paper can be rectified through the implementation of DQM programmes. Due to the importance of DQM and the lack of valid data in this area, we conclude that further empirical research is required to assess the impact of poor data quality in HIS in Europe.

6. CONCLUSIONS

This paper has illustrated that DQM is central to major issues in healthcare. Examples of this have been highlighted through the targeted categorization of healthcare issues, using the information product approach to DQM. The implementation of DQM is not a matter of choice for the healthcare sector. It is a necessity for the development of the efficient health service desired by the national governments. Poor data quality is detrimental to efficiency and cost-effectiveness, the two principle aims of health services around the world.

The study of DQM in healthcare, its effects and its use to formulate and implement reference models are central themes of our future research agenda. We will aim to examine cost-benefit analysis in the implementation of DQM programs in HIS, while empirically validating the conceptual framework established in this paper. We propose to study the effect of the IS configuration on the provision of quality data, to test the impact of providing sub-optimal DQ to HIS users, and to examine the theoretical impact of their subsequent interventions on virtual patients and the subsequent cost and benefit incurred to the organization. This information can be used to form a graph (see Figure 1, below) charting the cost of poor DQ against different levels of DQ in HIS. Simultaneously, the cost of DQ maintenance will be charted against different levels of DQ in NIS. This will enable the development of an overall DQ cost model for NIS and the establishment of an optimal level of DQ investment for NIS through cost/benefit analysis.

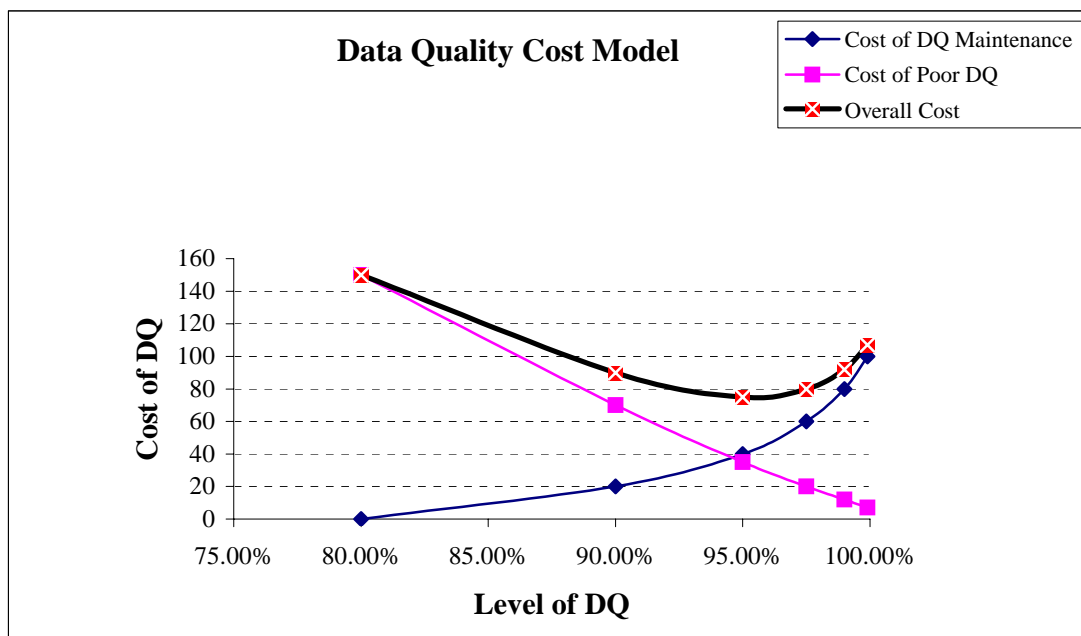


Figure 1: Theoretical cost/benefit analysis for investment in DQ assurance in HIS

ACKNOWLEDGEMENT

This research is partly funded by the Irish Research Council for Science, Engineering and Technology (<http://www.ircset.ie>) under the Post Graduate Research Scholarship scheme.

REFERENCES

- Abate, M.L. et al, 1998. A Hierarchical Approach to Improving Data Quality. *Data Quality Journal*, Vol. 4, No. 1.
- Ballou, D.P., Pazer, H.L., 1985. Modelling data and process quality in multi-input multi-output information systems. *Management Science*, Vol. 31, No. 2, pp.150-162.
- Department of Health and Children, 2001. *Quality and Fairness: A Health System for You*. Department of Health and Children, Dublin, Ireland.
- Department of Health and Children, 2004. *Health Information: A National Strategy*. Department of Health and Children, Dublin, Ireland.
- Eckerson, W.W., 2002. *Data Quality and the Bottom Line: Achieving Business Success Through a Commitment to High Quality Data*. The Data Warehousing Institute Report Series, No.101, Chatsworth, USA.
- English, L.P., 1999. *Improving Data Warehouse and Business Information Quality*. John Wiley & Sons Inc., New York, USA.
- Gendron, M.S., D'Onofrio, M.J., 2000. An Exploratory Study Investigating Data Quality in the Healthcare Industry: What are the Implications for Data Warehousing? *Proceedings of the Americas Conference on Information Systems*. Long Beach, USA.
- Gnesotto, R., DeVogli, R., 2003. *Health Monitoring Systems in Europe: Structures and Processes*. EC Health Monitoring Programme. http://europa.eu.int/comm/health/ph_projects/2001/monitoring/fp_monitoring_2001_frep_13_en.pdf, last accessed 22 Nov, 2004.
- Hackney, R., McBride, N., 2002. Non-Implementation of an IS Strategy Within a UK Hospital: Observations From a Longitudinal Case Analysis. *Communications of the Association for Information Systems*, Vol.8, pp.130-140.
- Institute of Medicine, 2000. *To Err is Human: Building a Safer Health System*. National Academy Press, Washington D.C., USA.
- Leitheiser, R.L., 2001. Data Quality in Healthcare Data Warehouse Environments. *Proceedings of the 34th Hawaii International Conference on Systems Sciences*, Hawaii, USA, pp 1-2.
- Ó Riain, C., Helfert, M., 2004. *A Case Study for Data Quality Management in an Irish Hospital*. Working paper, School of Computing, Dublin City University, Dublin, Ireland.
- Orfanidis, L. et al, 2004. Data quality issues in electronic health records: An adaptation framework for the Greek health system. *Health Informatics Journal*, Vol.10, No.1, pp 23-36.
- Organization for Economic Cooperation and Development, 2003. OECD Health Data 2003 <http://www.oecd.org/health/healthdata>, last accessed 22 Nov, 2004.
- Redman, T.C. (1996) *Data Quality for the Information Age*. Artech House, Boston, USA.
- Smyth, D., 2004. *The Health Information Strategy*. Presentation to the Irish Forum for Health Informatics, Dublin, Oct 7, 2004. <http://www.ifhi.ie>, last accessed 22 Nov, 2004.
- Tayi, G.K., Ballou, D.P., 1998. Examining Data Quality. *Communications of ACM*, Vol. 41, No. 2, pp.54-57.
- Than Sein, U., Waheed, A., 2003. *National Health Accounts: Policy Brief on Concepts and Approaches*, WHOSEA Regional Health Forum, Vol. 7, No. 2, <http://w3.whosea.org/rhf/rhf7-2/HealthPolicy.htm>, last accessed 22 Nov, 2004.
- Wand, Y., Wang, R.Y., 1996. Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*, Vol. 39, No. 11, pp. 86-95.
- Wang, R.Y. et al, 1998. Manage Your Information as a Product. *Sloan Management Review*, Vol. 39, No. 4, pp.95-105.
- Wang, R.Y., Storey, V.C., Firth, C.P., 1995. A Framework for Analysis of Data Quality Research. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 7, No.4, pp.623-639.
- Wang, R.Y., Strong, D.M., 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, Vol. 12, No. 4, pp.5-34.
- WHO Regional Committee for the Western Mediterranean, 2001. *Technical Discussions: Health Systems Development*. <http://www.emro.who.int/RC48/Documents/EMRC48TechDisc1.doc>, last accessed 22 Nov, 2004.