Regional Variation in Medical Data Quality Management

Daniel P. Lorence and Amanda Spink
The Pennsylvania State University, dpl10@psu.edu, spink@ist.psu.edu

ABSTRACT

This paper reports findings from a U.S. national survey on the adoption of data quality management. Findings suggest that on a nationwide basis, regional variation was found in the adoption of the data quality management function within healthcare organizations.

DATA QUALITY MANAGEMENT

How widespread is the problem of poor data quality in the U.S. healthcare system? According to a recent audit of Medicare-based reimbursement, the Office of Inspector General (OIG) of the Department of Health and Human Services (HHS) found a 5 percent increase in erroneous payments to physicians attributed to databased flaws or lack of data documentation compared to the previous year. Inadequate or inappropriate data documentation also was the top reason for erroneous payments to physicians (Gibbs-Brown, 2000). More troubling is the fact that incorrect coding accounted for almost 15 percent of the total improper payment made by Medicare; more than 90 percent of those payments were made to physicians. Although the reimbursement aspects of incorrect coding are apparent, systematically identifying the causes and remedies remains a problem for medical systems developers (Gibbs-Brown, 2000).

Data warehouse managers must be able to facilitate accuracy, determine an application's purpose, and focus on the rationale for collecting certain data elements (Mendelsohn & Kraemer, 1998). Information managers are further charged with ensuring accuracy, requiring appropriate education and training, and communicating timely and appropriate data definitions to those who collect data. Data warehouse managers are charged with ensuring that appropriate edits are in place to guarantee accuracy. Adherence to approved coding principles that generate coded data of the highest quality remains important, but that is just a part of the overall data quality picture (Gorla & Krehbiel, 1999). Ensuring the accuracy of coded data is a responsibility shared between information systems professionals and clinicians.

Most significantly, the accreditation standards of the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) as well as the Medicare Conditions of Participation also require that final diagnoses and procedures are recorded in the health record and authenticated by healthcare practitioners. Congress has mandated a series of regulations, under the Health Insurance Portability and Accountability Act of 1996 (HIPAA), intended to reduce the administrative costs and burden associated with health care by standardizing data and facilitating electronic transmission of many administrative and financial transactions. Because of the belief that the electronic movement of health information creates patient privacy and security concerns, Congress also directed the Secretary of HHS to develop standards to protect the privacy and security of individually identifiable health information (P.L. 104-191, 1996). Such requirements for standardized information management practices suggest that, while the implementation and impact of such mandates remain uncertain, the goal of HIPPA is clear: organizations need to devise some type of formal data management strategy to ensure the standardized collection of information within their structural bounds (HHS, 2000).

RELATED STUDIES

The provision of information management services itself has suffered, to an extent, from a lack of centralized strategic development, with local standards and practices often emerging from initiatives in academic institutions, resulting in a service that has shown a large regional variation in implementation (Leape & Park, 1990). Data quality improvement typically has accounted for a small fraction of operational expenditures, from delivery organizations as well as from national funding sources. Little research has been done to assess the implications of quality of data in an evidence-based environment for health services delivery. Although primarily concerned with U.S. healthcare, the interregional and international perspectives must also be considered. The scope of evidence-based medicine is such that interregional co-operation will be required, and the potential for access to advanced information services via the global network will break down traditional boundaries to health care systems (Park, 1993).

This study is one attempt to determine, on a national level, do healthcare organizations uniformly adopt prescribed data quality management as part of their overall organizational structure? If not, how does adoption vary?

DATA COLLECTION

Data from a nationwide survey of healthcare information managers, the AHA Annual Survey, the U.S. census, Interstudy publications, state and regional health service departments, and *The Market Statistics Report* were used to examine the organizational and environmental characteristics of data quality in a variety of healthcare settings. A comparison of selected data quality characteristics was made across practice settings, geographic areas, and selected healthcare demographic characteristics. The survey data include measures of the quality of data obtained via automated encoders, the impact of organizational mergers and acquisitions on data quality, and the existence of an organizational data quality committee or team. The survey further identifies the existence of a data quality manager, the most common methods used for identifying data quality problems, and the uses of edit checks in automated data fields. Finally, the data provide measures of the existence of organizational master data dictionaries, the existence of organizational policies and procedures related to the timeliness of data capture, and the existence of an organizational process to identify and prioritize data acquisition.

The survey, completed in May of 1999, was a response to the need of the healthcare industry for more timely and frequent practice information in the field of health information management The initial survey was initially fielded in June 1998, with follow-up assessments accomplished through May 1999. It was designed to provide representative information on the population of health information managers. The sample included respondents from a variety of practice settings and job titles and excluded students. The survey obtained data from 16,591 health information managers, for a 50.4 percent gross response rate. Preliminary findings from the survey were summarized for analysis and are reported elsewhere (Lorence, 1999).

RESULTS

Data Quality Management

Data Quality Manager
Adoption of the dedicated data quality management role, though widespread, was still found in only a minority of organizations. Adop-

tion of specified person for the data quality management role was also found in only a minority of organizations. Overall, 43.7% of respondents reported having adopted a specified data quality manager. Variation was prevalent across practice settings regarding adoption of the data quality manager role. Hospital practice setting (50%) respondents were more likely than clinic (30%) and "Other" practice setting (37%) respondents to report that their organization has a data quality manager.

Preferences also varied across organizational roles, regarding adoption of the data quality manager role. Nearly one-half (48%) of the respondents with an "Other" job title report their organization has a data quality manager, which is significantly different from Managers (38%).

Significant variation also existed across geographic boundaries (Table 1).

Mid Atlantic organizations (44%) were more likely to have adopted the data quality manager role, a significant difference compared to the West North Central (34%) and West South Central (38%) regions.

Merger activity was only marginally associated with variation in adoption of data quality managers. Respondents whose organizations have merged and have a data quality manager (50%) reported a significantly higher adoption rate compared to organizations that have not merged (40%). The dedicated data quality management role, while growing in acceptance, was still found in only a minority of organizations.

Data Quality Committee or Team

Overall, 48.5% of respondents reported having adopted a data quality team or committee. Slightly less have adopted dedicated data quality managers within their organizations. Despite federal mandates to standardize data practices, and professional protocols suggesting more stringent, uniform data quality measures, why does there still exist only partial adoption of data quality functions within organizations, and like variation in adoption patterns of such positions?

Variation was prevalent across practice settings regarding adoption preferences. More than one-half (55%) of the hospital setting respondents' organizations have a data quality committee or team, a significant percentage compared to clinic (31%) and "Other" (41%) practice setting respondents' organizations. Mean response from hospital practice settings (4.19) respondents was significantly higher when compared to clinic (3.60) and "Other" (3.65) setting respondents,

when rating the quality of coding with encoders compared to manual coding.

Preferences also varied across organizational roles, regarding adoption of data quality teams. Respondents with a job title of "Other" (52%) were more likely to report for their organizations having a data quality committee or team. This was significantly higher than responses from managers, where only (44%) responded that they had adopted a data quality team.

Significant variation also existed across geographic boundaries (Table 2).

Fifty-four percent (54%) of New England region respondents report their organizations have a data quality committee or team, a significant finding compared to the Mid Atlantic (44%), West North Central (43%) and West South Central (45%) regions.

Merger activity was also somewhat associated with variation in adoption of data quality teams. Organizations that have gone through a merger or acquisition (54%) and have a data quality committee or team are significantly different compared to organizations that have not merged (46%).

DISCUSSION

Why should data quality be a dedicated organizational function? Primarily, because there now exists a need to incorporate a multidimensional data quality model into day-to-day management practices. A comprehensive data quality model goes far beyond simple data accuracy audits. It requires continuous assessment and adjustment of a multi-factored model (Williams & Yang, 1996). It should, for example, include some measure of relevancy, or assurance that the data collected are meaningful to the performance of the process or application for which they are collected. In addition, accessibility requires that data items are easily obtainable and legally proper to collect.

Further, a measure of comprehensiveness is needed, that requires assurance that all necessary data items are included in the database. Consistency, likewise, requires that the value of the data is reliable and remains the same across applications; that is, data are consistent when the value of the data (such as the patient's health record number) is the same across applications and systems. Timeliness of data is also important, assuring that a given value is up-to-date when it is current for a specific point in time. Clear data definitions likewise make it possible for current and future data users to know what the data mean. Precision assessments also require that data values are just large enough to support the application or process. Finally, data relevancy, or assurance that data collected are meaningful to the performance of the process or application for which they are collected, is also a key component of the data quality dimension (Fletcher, 1998).

While it may be that the quality of information, such as local costs, is less a reflection of actual truth than the need to report data that maximizes reimbursement, both researchers and clinical providers often rely on this information for evaluating clinical performance, even more so in an evidence-based environment (Chassin et al., 1997).

Geographic variation in data quality has implications for both clinical quality as well as reimbursement, since information on costs for a specific procedure can vary greatly from one city to another. The key question: is this due to actual variation in cost or charges, or simply inaccurate data being reported? In many cases, the total amount paid for medical services per beneficiary is more than twice as much in some markets as it is in others, even

Table 1: Regional variation in the appointment of a data quality manager

	New England	Mid. Atlantic	E. No. Central	W. No. Central	So. Atlantic	E. So. Central	W. So. Central	Mountain	Pacific
Yes	44.3%	44.5%	44.2%	33.9%	42.6%	43.5%	37.5%	41.7%	40.7%
No	55.7%	55.5%	55.8%	66.1%	57.4%	56.5%	62.5%	58.3%	59.3%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%

Table 2: Regional variation in the appointment of a data quality committee or team

	New England	Mid. Atlantic	E. No. Central	W. No. Central	So. Atlantic	E. So. Central	W. So. Central	Mountain	Pacific
Yes	53.7%	44%	48.5%	42.6%	47.8%	48.6%	44.6%	53.1%	46.4%
No	46.3%	56%	51.5%	57.4%	52.2%	51.4%	55.4%	46.9%	53.6%
Total	1005	100%	100%	100%	100%	100%	100%	100%	100%

after adjusting for differences in the local prices of services. This is often based upon cost data collected and maintained with questionable information management practices. Payments for medical services per beneficiary in the Southeast, for example, have been found to less than half as much as in the Midwest, even after adjusting for price. In home-care services, regional Medicare reimbursement adjusted for price may cost as much as 12 times more per person in Tennessee, compared to areas of Minnesota (Wennberg, 1997).

Regions with relatively high reimbursement for outpatient and home-care services provided to Medicare patients also provide high inpatient hospital services. Often, analysts wrongfully assume that increased spending for ambulatory care automatically results in lower inpatient costs. Frequently, however, spending in these sectors is completely independent. For example, the rate of coronary-bypass surgery per 1,000 Medicare beneficiaries in Northern Illinois has been found to be more than four times the rate in southern Colorado. Similarly, rates of balloon angioplasty procedures for opening diseased heart vessels ranged from 1.6 per thousand in Buffalo, New York, to 12.8 per thousand in Stockton, California (Wennberg, 1997).

CONCLUSION

There is an increasing reliance on health care data that is accurate, reliable, and of the highest quality. This comes, however with an increasing reliance on information system performance and how it ultimately affects the quality of data maintained. This is the challenge information information systems developers face in an evidence-based healthcare system. Evidence-based measurement is at the heart of efforts to promote more consistent, and effective, data quality management, as the US healthcare system grows increasingly reliant on data representations in lieu of extensive patient-provider interaction. This exists in an environment of concurrent utilization of information systems in research and day-to-day clinical practice. Variation in data errors within the Medicare system that is based largely on traditional fee-for-service insurance reimbursement is especially timely as Congress seeks to curtail spending growth in that program without cutting necessary services to beneficiaries.

REFERENCES

Chassin MR, Kosecoff JB, & Park RE. (1987). Does inappropriate use explain geographic variations in the use of health services? A study of three procedures. *JAMA. 258*:2533-2537.

Fletcher, D. (1998). *Working paper on A Data Quality Model*. HIM Data Quality Task Force, Chicago, IL.

Gibbs-Brown, June. 2000. *Improper Fiscal Year 1999 Medicare Fee-for-Service Payments, (A-17-99-01999). Department of Health and Human Services, Office of Inspector General – AUDIT. (February 17, 2000).*

Gorla, Narasimhaiah, Krehbiel, Steve. 1999. Summers Rubber Company Designs Its Data Warehouse. *Interfaces. (March):* 104-117.

HHS. (2000). *ANSI ASC X12N HIPAA, Implementation Guide*. U.S. Dept. of Health and Human Services, Washington, DC.

Leape LL, Park RE, Solomon DH, Chassin MR, Kosecoff JB, & Brook RH. (1990) Does inappropriate use explain small-area variations in the use of health services? *JAMA,. 263*: 669-672.

Lorence, D. (1999). National Implementation Patterns of Computer-based Patient Records, *Proceedings of the National Conference on Health Statistics, NCHS, Washington, DC, August 1999.*

Mendelson, Haim, & Kraemer, & Kenneth L. 1998. Clockspeed and Information Response: Evidence From the Information Technology Industry. *Information Systems Research (December):* 415-433.

Park, RE. (1993). Does inappropriate use explain small-area variations in health care services? A reply. *Health Serv Res.28*: 401-410

P.L. 104-191. (1996). *Health Insurance Portability And Accountability Act Of 1996,* Public Law 104-191, 104th Congress. 1996.

Wennberg, J.E. (1997). *Dartmouth Atlas of Health Care, Center for the Evaluative Clinical Sciences, Dartmouth Medical School, Hanover, NH.*

Williams JI & Young W. (1996). *A summary of studies on the quality of health care administrative databases in Canada.* In: Goel V, Williams JI, Anderson GM, Blackstien-Hirsch P, Fooks C, Naylor CD, editors. *Patterns of Health Care in Ontario.* The lCES Practice Atlas. 2nd ed. Ottawa: Canadian Medical Association.

Related Content

Science Mapping
Chaomei Chen, Rachael Dubinand Timothy Schultz (2015). *Encyclopedia of Information Science and Technology, Third Edition (pp. 4171-4184).*
www.irma-international.org/chapter/science-mapping/112859

An Approach to Clustering of Text Documents Using Graph Mining Techniques
Bapuji Raoand Brojo Kishore Mishra (2017). *International Journal of Rough Sets and Data Analysis (pp. 38-55).*
www.irma-international.org/article/an-approach-to-clustering-of-text-documents-using-graph-mining-techniques/169173

A Study of Knowledge Discovery and Pattern Recognition Based on Large-Scale Sentiment Data in Online Education for College Students
Guoliang Li, Bing Wangand Maoyin You (2023). *International Journal of Information Technologies and Systems Approach (pp. 1-13).*
www.irma-international.org/article/a-study-of-knowledge-discovery-and-pattern-recognition-based-on-large-scale-sentiment-data-in-online-education-for-college-students/323194

Estimation and Convergence Analysis of Traffic Structure Efficiency Based on an Undesirable Epsilon-Based Measure Model
Xudong Cao, Chenchen Chen, Lejia Zhangand Li Pan (2024). *International Journal of Information Technologies and Systems Approach (pp. 1-25).*
www.irma-international.org/article/estimation-and-convergence-analysis-of-traffic-structure-efficiency-based-on-an-undesirable-epsilon-based-measure-model/332798

WSN Management Self-Silence Design and Data Analysis for Neural Network Based Infrastructure
Nilayam Kumar Kamilaand Sunil Dhal (2017). *International Journal of Rough Sets and Data Analysis (pp. 82-100).*
www.irma-international.org/article/wsn-management-self-silence-design-and-data-analysis-for-neural-network-based-infrastructure/186860