

12-2009

Using a Community of Practice to Enhance Data Quality in a Distributed Healthcare Information System

Biswadip Ghosh

Department of Information Systems, Metropolitan State College of Denver, bghosh@mscd.edu

Judy E. Scott

Business School, University of Colorado at Denver, judy.scott@ucdenver.edu

Follow this and additional works at: <http://aisel.aisnet.org/acis2009>

Recommended Citation

Ghosh, Biswadip and Scott, Judy E., "Using a Community of Practice to Enhance Data Quality in a Distributed Healthcare Information System" (2009). *ACIS 2009 Proceedings*. 71.

<http://aisel.aisnet.org/acis2009/71>

This material is brought to you by the Australasian (ACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ACIS 2009 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Using a Community of Practice to Enhance Data Quality in a Distributed Healthcare Information System

Biswadip Ghosh, Ph.D.
Department of Information Systems
Metropolitan State College of Denver
Denver, Colorado, USA
Email: bghosh@mscd.edu

Judy E. Scott, Ph.D.
Business School
University of Colorado at Denver
Denver, Colorado, USA
Email: judy.scott@ucdenver.edu

Abstract

Healthcare provider networks typically have severe budget constraints, which necessitate effective allocation of scarce resources and judicious adoption of expensive advances in medical procedures and treatment regimens. Elective surgery is particularly risky and resource intensive and require information systems that allow judicious decision making at multiple levels –healthcare policy makers, hospital administrators and clinicians. Given the large number of variables and the diversity in the facilities, underlying systems, procedures and personnel, quality of the input data can be threatened. Ensuring data quality for such an information system requires effective organizational practices coupled with technology. This paper studies the role of building a community of practice (CoP) among recording nurses for increasing knowledge sharing to enhance the quality of data collection from surgeries. The study finds support for a CoP in building social capital and increased knowledge sharing, leading to improvement in data quality and downstream impacts on patient care.

Keywords

Data quality, Community of practice, Social capital, Knowledge management. Grounded theory

INTRODUCTION

Given the high costs, healthcare provider networks typically have severe budget constraints, which necessitate effective allocation of scarce resources and judicious adoption of expensive advances in medical procedures and treatment regimens. Elective surgery is particularly resource intensive and risky and therefore benefits from information systems that facilitate decision-making at multiple levels – by healthcare policy makers, administrators and clinicians. Administrators need to track resource usage, outcomes of selected treatment and procedures and plan unit based resource allocation and standardized procedures (Epstein, 2006). Healthcare system policy makers need data across the entire healthcare network to standardize treatments and procedures. Clinicians need historical data to facilitate decisions on elective surgery and the suitability of treatments and procedures for a presenting patient. Although all of these decisions benefit from data across the healthcare network, given the diversity of facilities in a typically large distributed healthcare provider network and differences in their patient record systems, infrastructure, staff and the patient mix, ensuring data quality for an overall network wide information system is often problematic. Because of the differences in clinical personnel and their interpretation of patient variables and the lack of standardization of data recording practices, the values in the individual patient record systems (PRS) at different facilities are not suitable for a network wide information system without adopting data collection/cleaning processes.

Asserting data quality is a critical component of any information systems project (Brodie, 1980). Large distributed healthcare organizations have multiple administrative and clinical databases for patient medical records, insurance billing and benefits administration. As in other domains, data quality problems are increasing in such organizational databases (Wang, Strong and Firth, 1995). Prior studies have reported that between 50-80% of records in many such databases may be inaccurate, incomplete or ambiguous (Redman, 1998). The ramifications of using poor quality data in building an information system can lead to incorrect results that are flawed and unusable.

The research literature has established multiple dimensions in the definition of data quality (Wang and Strong, 1996; Strong, Lee and Wang, 1997a). Some definitions view data quality as an intrinsic property of the data and

include such characteristics as accuracy, completeness and objectivity of the data. The intrinsic category also has organizationally derived characteristics that include such factors as reputation and believability of the data as well. But other definitions provide a much broader concept of data quality that also includes an extrinsic category that stems from the organizational standpoint of data quality. The latter view accepts the fact that the quality of data cannot be established without the inclusion of additional factors that originate from (a) the management of the data and (b) the users of the data and information products produced from the data. Prior research has also identified several threats to data quality (Strong, Lee and Wang, 1997b). These threats can compromise data quality if not addressed with adequate organizational, managerial or technological solutions during the data lifecycle. These threats are also common in the health informatics domain and often lead to problems in producing quality clinical research data sets. Threats to data quality originate from factors such as getting data from multiple sources, when subjective judgements are needed in abstracting data, inadequate use of technology to support data collection and not understanding the definition and downstream uses for the data in the information system (Ghosh, 2009). These issues are particularly severe in many patient record systems (Moore, 2009).

Prior studies have not identified the organizational practices that can enhance data quality and support data collection processes for a network wide healthcare information system. Identifying the organizational practices is important because data quality problems may result in poor decisions that affect patients and increase administration costs. For example, surgeons' decisions on patient treatment could be compromised if their risk models were based on inaccurate historical data. This research finds that supporting a community of practice (CoP) enhances data quality in a distributed healthcare information system by building social capital and increasing knowledge sharing among the nurses. The fundamental motivation of this research is to study how a CoP among the data collectors spread across multiple facilities (1) addressed the threats to data quality for a distributed healthcare information system, (2) increased knowledge sharing among nurse recorders in the healthcare network to improve data quality and patient outcomes, and (3) impacted social capital resources.

METHODOLOGY

This study uses qualitative research methods based on interviews of team members (Glaser and Strauss, 1999). Since there is a paucity of research literature on data quality issues in a large healthcare network, this study uses grounded theory methodology (Straus and Corbin, 1990). Grounded theory is appropriate in the early stages of research on a topic, because it is inductive and does not rely on previous literature or prior empirical evidence (Eisenhardt, 1989). Moreover, this method explains process, 'how' research questions, and context, and provides detailed data for deducing constructs for theory generation and elaboration. Grounded theory places emphasis on abstracting participants' accounts of experiences and events and relating those to existing literature to explain the phenomena (Suddaby, 2006).

Case Selection

The organization selected for this case study is a distributed national healthcare organization of hospitals, patient consultation offices, nursing homes and labs established to serve over 10 million patients. The healthcare system comprises over 150 acute care hospitals and 14,000 staff physicians. Additionally the system has over 350 outpatient clinics, which provide treatment on 24 million annual patient visits. There is wide variety in the facilities in the distributed healthcare system, ranging from "standard" rural centers with limited scope of medical practice to highly "advanced" urban hospitals with state of the art medical technology for patient care. Issues with data quality, such as incomplete records, lack of outcome variables, inconsistent classification of pre-Operative and post-Operative variables that needed judgement (e.g. does patient have "acute" renal failure?) prevented the organization wide use of the patient data.

These issues were alleviated by building and supporting knowledge sharing capabilities and processes among the recording nurses who are responsible for data collection and reporting. Approximately 150 nurse recorders were assigned to collect, review, update (if necessary) and transmit the data on surgeries. They brought consistency to the data collection practices and built an input data set for the information system that is homogeneous and comparable across the distributed healthcare network. Each nurse recorder was provided with detailed definitions of each of the preoperative, operative, and outcome variables of interest for the information system. To support, train and build uniformity in the data and collection practices, the nurse recorders were brought together in a "community of practice" (CoP) that was coordinated by a national administrator and ten regional leaders who were available on a daily basis to support the data collection activities, help mitigate conflicts and answer questions (Wenger, et.al., 2002). The CoP became the means to establish consistency in the data collection process by the recording nurses and contributed to improving the data quality.

Interview Protocol

The objective of this research is to generate theory to explain how data quality was enhanced in an information system across a large healthcare network. To accomplish this goal, multiple participants in the data collection, cleaning and transmittal processes in the organization were interviewed. The interviews were open ended – however the following questions were posed to guide the theory building:

- What types of challenges do you face in the data collection practices?
- What types of knowledge need to be shared?
- How is knowledge sharing facilitated among data collectors?
- How do you improve the “community of practice” among nurses?
- What are the outcomes of this knowledge sharing?

Three nurse recorders, two regional leaders and the national nurse executive along with the CoP administrator were interviewed over several weeks. At the same time, we searched and analyzed the relevant literature. Including more than one participant’s perspective, and incorporating multiple theoretical perspectives at multiple levels of analysis into our discussion, strengthens the generalizability of the findings.

The data collection and analysis were iterative and guided by literature sources on communities of practice, knowledge management and social capital that functioned as secondary data, increasing theoretical sensitivity. Grounded theory analyzes phenomena by examining causal and intervening conditions and consequences of an action/interaction strategy. The information collection and analysis were iterative and guided by literature sources on information quality, communities of practice, knowledge management and social capital that functioned as secondary information, increasing theoretical sensitivity. Grounded theory analyzes phenomena by examining causal and intervening conditions and consequences of an action/interaction strategy.

Table 1 below lists the causal and intervening conditions, and consequences of the two processes - the community development and data recording and transmittal processes. The six constructs (community of practice, social capital, knowledge sharing, technology features and usage, data quality and patient outcomes) emerged from open coding, a technique of categorizing data; the processes emerged from axial coding, a technique that links the constructs. Analysis continued until no further concepts emerge - the point at which theoretical saturation was reached.

The interviews were conducted in two phases –first phase identifying the two processes and the constructs in those two processes and the second phase serving as the vehicle for building the relationships among those constructs. The research literature was consulted to define the constructs in the first phase and to support the relationships for the second phase.

Table 1. Constructs and Processes for Data Quality

Grounded Theory	Community Development	Data Recording & Transmittal
Context	Diverse physicians, hospital administrators and policy, multiple patient record systems, policy makers, diverse clinician practices	
Causal Conditions	(1) Community of Practice	(3) Knowledge Sharing
Intervening Conditions	(2) Social Capital	(4) Technology Features & Usage
Consequences	(3) Knowledge Sharing	(5) Data Quality (6) Patient Outcomes

ANALYSIS

In the healthcare organization that we studied, typical problems range from misunderstanding data definitions, errors in data entry, incomplete data sets, conflicts between multiple data sources, and inaccessibility of data from remote facilities. These problems were addressed with a community development process.

Community Development Process

The group of recording nurses are provided ongoing training on the descriptions of the data variables being collected. National educational conferences are held annually. Further, national and regional team teleconference calls provide the nurse recorders with ongoing education. This training process allows the nurses to attend review session about all the policies, procedures, and data collection methods being advocated. According to one of the nurses:

“We have annual conference to meet and build relationships. We have multiple teleconferences where we discuss questions and concerns. If a nurse has a question, then they can pose that to the regional leaders. The nurses also pose questions to the CoP facilitator and she not being clinical has to refer back to a regional leader to get the answer. The Q/A is made widely available to all the nurses using the Web.”

The nurse recorders also periodically complete a training module and examination by electronic means. As a nurse said:

“We are doing chart review prospectively, so any knowledge shared to help us interpret the patient charts in a uniform way can help.”

Additionally, the nurse recorders that collect variable data are required to take inter-rater reliability (IRR) tests. A standardized case study exam is developed by the IRR Committee and administered electronically to all nurse recorders simultaneously on a pre-determined examination date. Nurses are supported by the community in developing a deeper understanding of variable definitions or honing their judgement of patient conditions, all leading to better outcomes on the IRR test. Variable agreement has historically been excellent on most variables. A nurse explains:

“[We] need to do inter-rater reliability tests and post the results, so that nurses can use the results as a bargaining tool with physicians as to when they rate some variable differently than the physician.”

The conferences and frequent interactions among the nurses allows them to build rapport and a body of cooperation. Since the nurses participate in these activities to improve their capabilities in data collection, the body can be considered a community of practice (CoP) (Wenger, et.al., 2002).

Community of Practice

A CoP is loosely defined in the research literature as a group of individuals bound together by shared expertise with a passion for joint enterprise towards a shared goal (Wenger and Snyder, 1999). CoPs can be an effective vehicle for organizations to promote the spread of best practices, develop employee skills and support knowledge sharing in the support of business goals and processes (Brown and Duguid, 1999; Duguid, 2005). CoP's typically also have a coordinator, who is the main sponsor of the CoP and who facilitates activities and supports the members to build a stronger practice. The core outcome for the CoP is the effective discovery, diffusion and application of knowledge around their area of interest and/or the business needs.

Through the participation in the community, the nurses support each other by identifying ways to use technology to more effectively collect, process and transmit data. The interactions also build trust, identify and norms within the team, which allows the nurses to share knowledge to more effectively carry on data collection activities, which can span multiple facilities and systems. A nurse noted the importance of trust across facilities as requests for remote data are frequent:

“It takes a lot of extra time to get the charts and review data from a completely different system and coding/definitions. We use email and calls. Calls are more since patient identifying data cannot be emailed.”

The core outcome for the CoP is the effective discovery, diffusion and application of knowledge around their area of interest and/or the business needs. The collection, cleaning and reporting of data on surgeries in the healthcare network, require the nurse recorders to work within the community. The CoP among these nurse recorders allows them to share their knowledge among the group to support each other in their training and data collection activities, which spans multiple locations in the healthcare network. One of the nurses compares the CoP to a family:

“Openness of the community – “family” – good, bad and the ugly is said – open communications is desired.”

CoP's have different characteristics and roles and may require varied management attention from the CoP coordinator to sustain (Wenger, et.al., 2002). CoP's can differ in their membership structure – with core and/or one or more peripheral groups and in their stage of development ranging from starting/coalescing to sustaining to transforming. They also differ in the “rhythm” that guides the pace of their activities, the balance between

openness and ownership of issues and their goals, which can consist of a mixture of exploration verses exploitation (Brown and Duguid, 2001). CoPs need to be sponsored by supporting relationship building among the CoP's members. It is important to root out potential problems such as multiple agendas and craft a sense of intimacy among the often diverse members to derive optimal benefits (Cross, et.al., 2006). A nurse mentioned:

"We have annual conference to meet and build relationships. We have multiple teleconferences where we discuss questions and concerns. If a nurse has a question, then they can pose that to the regional leaders. The nurses also pose questions to the CoP facilitator and she not being clinical has to refer back to a regional leader to get the answer. The Q&A is made widely available to all the nurses using the Web."

The CoP also provides an engine to develop an important organizational resource – social capital.

Social Capital

Prior research has also identified an important organizational resource—social capital (Nahapiet and Ghoshal, 1998). Social capital theory (SCT) argues that effective knowledge exchange and transfer occur when sufficient opportunities for interactions and exchange are present between individual knowledge workers (Nahapiet and Ghoshal, 1998). These interactions help build social capital in the group. The three dimensions of social capital – (a) trust, (b) common identity and (c) common norms play a large part in the effectiveness/success of the CoP and how much cooperation is achieved and knowledge sharing is furthered. The regional leader mentioned the impact of group identity among the nurse recorders:

"Nurses feel they are part of a bigger picture – they want to help the next "Joe" not get pneumonia after surgery. So this is their incentive to go the extra mile and collect the data – feed it to the system so that the model can be generated and used by clinicians."

Trust enables the members of the CoP to work more cooperatively, limiting the power and positional rivalries. A nurse explains:

"Trust is very important when we call each other and request patient data. The nurse on the other end needs to have a sense of who the person is that is making the request. This is particularly true for the tertiary facilities."

Social capital comprises a set of values, norms and informal sanctions that are shared among members through interactions that permit them to cooperate with each other. Social capital can provide the lubricant for workers to get knowledge and advice well beyond the organizational hierarchy, to enable them "to get things done more effectively." A nurse said:

"Lots of sites are so small that the nurse may not need to assess a particular variable or face a complicated case in years. This is when community and conference calls helps in sharing knowledge."

Several research studies have found a high correlation between social capital and knowledge sharing. IS researchers have found that the functioning of a CoP can be an engine for the creation of social capital (Lesser and Storck, 2001). Several IS researchers have utilized SCT to study organizational knowledge management systems and found significant relationships between social capital and knowledge sharing (Kankanhali, Tan and Wie, 2005; Wasko and Faraj, 2005).

Knowledge Sharing

Knowledge is information combined with experience, context, interpretation and reflection. Since knowledge has a personalization aspect to it, hence it is more difficult to manage and manipulate than mere information. Most literature categorizes knowledge as either tacit or explicit. Tacit knowledge is that which lies within a person, which can be difficult to code and share with others. Whereas explicit knowledge is often considered skills, which can be externally coded, acquired and interchanged among people. Knowledge may also be procedural or process oriented in describing how something is done or knowledge may be causal describing why something happens. Knowledge cannot be easily decoupled from the social environment in which it is developed, therefore its sharing must account for the socially situated nature of knowledge (Brown and Duguid, 2001). As a nurse said:

"We are doing chart review prospectively, so any knowledge shared to help us interpret the patient charts in a uniform way can help."

Published research has reported the benefits of sharing knowledge in business processes, various activities related to building an information system and supporting the data lifecycle (Poston, et.al., 2007). In the business context, knowledge is defined as any information that is relevant, actionable and is based on a person's

experience (Davenport and Prusak, 1998). For example, interpreting values and patient notes in the system needs to be supported by knowledge sharing among the nurses involved with the patient. A nurse mentioned

“Blood pressure meds are used for a lot of different reasons. – one might be small vessel disease. So if blood pressure meds are in the system, and the doctor does not document why, it is difficult to judge the charts. Similarly, a lot of nursing judgment goes into determining the “functional status”. We have to read into the charts to determine this variable – where does the patient live, wheel chair yes/no, etc.”

Most literature categorizes knowledge as either tacit or explicit. Tacit knowledge is that which lies within a person, which can be difficult to code and share with others. Whereas explicit knowledge is often considered skills, which can be externally coded, acquired and interchanged among people. Such skills may be the judgement to abstract data from the patient records, as one recording nurse said:

“Judgment is built up among the nurses and helps better data reporting. How to say YES/No on peripheral vascular disease? There is not one variable that will give this information, rather several things (in Vista) need to be consulted to say yes or no. All nurses need to agree on the definition.”

Knowledge may also be procedural or process oriented in describing how something is done or knowledge may be causal describing why something happens. Grover and Davenport (2001) state that knowledge stems partly from organizational artifacts like processes, structures and technologies, however, the dynamic context to knowledge is provided by the people – knowledge workers – their culture and the flow of interactions. A nurse explains how the flow of knowledge can be problematic in some data collection scenarios:

“Coding is delayed causing delays in data collection and case completion. [We] need to speed up the coding to improve knowledge flow.”

The research on knowledge management shows that knowledge sharing, specifically the sharing of tacit knowledge is particularly difficult (Burgess, 2005). Factors that impact knowledge sharing are the characteristics of the organizations, relationships, the type of knowledge and the transfer process (Argote, 1999; Ko, Kirsch and King, 2005; Inkpen and Tsang, 2005). Hansen et.al. (1999) described two knowledge sharing strategies – personalization, which is facilitated by personal relationships and codification, which is facilitated by technology usage. A nurse recorder mentioned the need to share and capture patient knowledge in order to build a complete data set. They use both strategies – the effective use of systems and personalization with other clinicians. This is established in the nurse’s quote below:

“Surgeons are interested in cutting. They do not like documentation as it takes away from their work. This can inhibit the knowledge sharing. We need to collaborate among the nurses to understand what is going on with the patient. Need to call back the surgeons to get a lot of additional information.”

Data Collection and Transmittal Process

The data collection process utilizes the training and community development established. The nurse recorders report on over 140 variables per surgery. The data can be recorded in multiple screens in the patient record system(s). The nurse needs to build a data case that can be considered a “parallel” patient record. While some data fields come in automatically (such as labs) others must be collected manually by the nurse through informal interactions with clinicians. If the patient has been treated at affiliated facilities, then the data must be collected from that facility. This necessitates phone calls or visits to check the charts. Charting by clinicians can be erratic and do not follow standards owing to the large diversity in staff and recording practices. As a nurse mentioned:

“I also have to double check all the automated data from the system as to what I report on the data form. This data form is like a “new” patient documentation (case) for me and so I have to document why I am reporting some variable in a certain way. I don’t and can’t go back and update the patient record, so the reported data case is a second form of documentation and it can differ from the original case. So I have to maintain documentation to support the values I am reporting.”

Moreover some of the reporting variables need to be coded as either having a condition or not. This requires judgement by the nurse. Hence data abstraction is necessary by the nurse, who needs to do a chart review and interpret progress notes to understand the patient’s symptoms and outcomes. These decisions need to be defended by the recording nurse upon scrutiny. Quoting a nurse recorder:

“I also have to double check all the automated data from the system as to what I report on the data form. The data transmittal form is like a “new” patient documentation (case) for me and so I have to document why I am reporting some variable in a certain way. I can report different values, which can differ from the original patient case. So I have to maintain documentation to support the values I am reporting.”

Technology Features and Usage

Technological approaches to improving data quality include taking steps to utilize the technology to assist in data collection steps such as locating, capturing and transmitting the data. Technology is particularly well suited for working with large volumes of data by partitioning them into manageable pieces. Building software code to check data elements as they are entered into the repository and the execution of consistency checks can also be done with the aid of technology. The knowledge about particular data collection strategies using the technology are shared among nurses:

“Some hospitals have really good systems and that help the nurses to collect data. Other centers do not have all the subsystems in the HIT (Health Information Technology) functional. Often the conference call between nurses can convey this information and then the nurse can go back and request that IT implement the particular feature at their location.”

Technology also can be used in the data collection stage by providing forms and templates to the data collectors to use during that process (Chen, et.al., 2007). This allows for standardization of the data collection process across departments thorough the use of software. A nurse relates:

“The anesthesia sheets are scanned into the imaging package. We have to get the ASA classification (risk of anesthesia) and weight from these scanned images. Often the attending’s classification is different from the anesthesiologist for ASA. So we have to go back to the operative notes to determine what value to use.”

Technology performs format conversions and allows data from multiple sources to be aggregated. Statistical tools compare the concordance of two data sources, when there are common fields between those sources. A nurse explains drug conflicts:

“There are often conflicts with drugs in the system. Discontinued prescriptions are not marked as “discontinued”, so the same drug could exist with multiple dosage in the system causing conflicts. Which one is correct? Manual verification is needed. Hypertension and diabetic meds are problematic as they can conflict a lot.”

Wireless handheld devices put in to the hands of the data collectors allow for a more efficient collection process. The end result of the data collection, the data product itself can be placed in the users hands to also increase the adoption and use of the data product.

Data Quality

Data quality is defined by its dimensions. For this study, data quality consists of the following four categories– (1) intrinsic, (2) accessibility, (3) contextual and (4) representational (Strong, Lee and Wang, 1997a). The last three categories are referred to as the extrinsic categories of data quality. Each dimension of data quality is subject to threats from multiple sources. The intrinsic dimension is threatened by actual problems in the data – wrong or missing values and conflicting values from multiple sources. Quoting a nurse recorder:

“Wound infection – sometimes the chief of staff wants to lower the severity of the infection in the documentation. We have to report the correct value otherwise the risk model will not be correct. We usually come to a true assessment of the wound infection, by consulting the full chart, the progress notes. If something like – “organ space repair” is noted then the wound is a severe infection”.

Accessibility to data can be problematic when multiple facilities are involved and data is not electronically accessible. A nurse explains the effort to access data from facilities associated with universities:

“Some locations farm their surgeries to the university hospitals. It takes a lot of extra time to get the charts and review data from a completely different system and coding/definitions”

Contextual issues originate from data that does not match the needs of the systems – such as cases where the coding required does not match the stored format. A nurse recorder talks about data conflicts:

“The physicians tend to copy and paste a lot of stuff from different screens to fill out the patient history screens. This causes data conflicts as a lot of old resolved symptoms come “back to life” in the system.”

Representational issues originate from when the interpretability of the data is problematic. One of the nurses gives an example of how they need to interpret data:

“A lot of nursing judgment goes into determining the “functional status”. The nurse has to read into the charts to determine this variable – where does the patient live, wheel chair yes/no, etc.”

Improving data quality is possible by attention to the above issues. The knowledge sharing among the nurse recorders and their adoption and use of technology has improved the data quality.

Patient Outcomes

The data collection and transmittal process consists of the creation of a “parallel” patient chart. The recording nurse collects, interprets and assembles information from multiple systems using information available in technology based solutions and interactions with clinicians to assemble the surgical case to transmit. This extensive review of the patients’ charts often leads to discoveries that can impact the quality of care. Examples of this are found in the quotes below:

“We do chart review and find serious quality of care issues – no followup visits, wrong site wounds, positive pathology results for cancer and that the physician has not initiated care. We point these out and it results in improved quality of care.”

“We are a watch dog that are reading the charts with a lot of clinical acumen to improve patient care.”

DISCUSSION

This study discussed the threats to data quality in an information system for a large healthcare network. These stem from the use of multiple patient record systems in multiple facilities and the diversity in the processes and operating practices among the diverse set of clinicians. The study finds that by increasing knowledge sharing among nurse recorders in the healthcare network the above threats to data quality can be addressed. Our innovative model explains how a CoP can foster social capital for enhancing knowledge sharing in supporting data collection activities. It posits that a CoP provides potential control of data quality problems in a distributed system with multiple sources of data. The CoP was supported remotely by a CoP administrator and regional leaders who facilitated interactions and with technologies such as a FAQ and web pages.

The grounded theory approach culminated in a conceptual model (Figure 1) that sheds light on a fresh theoretical perspective of improving the data quality in distributed information systems. The conceptual model is a synthesis of two process models that provide explanations for necessary but not sufficient conditions; for example the training and community development process is necessary but not a sufficient condition for knowledge sharing during data collection and transmittal. Furthermore, process models provide ‘the story, ..., and sequence of events that occur over time’ and focus ‘on the dynamics of social change, explaining how and why the results occur’ (Mohr, 1982). Although process models can become analytically complex, they do have a specific form and should not be discounted as unscientific or less rigorous than the more commonly used factor models. The conceptual model represents the interaction of context with the two processes. The process models illustrate relationships among constructs developed with grounded theory.

The conceptual model developed from the theory building is presented in Figure 1.

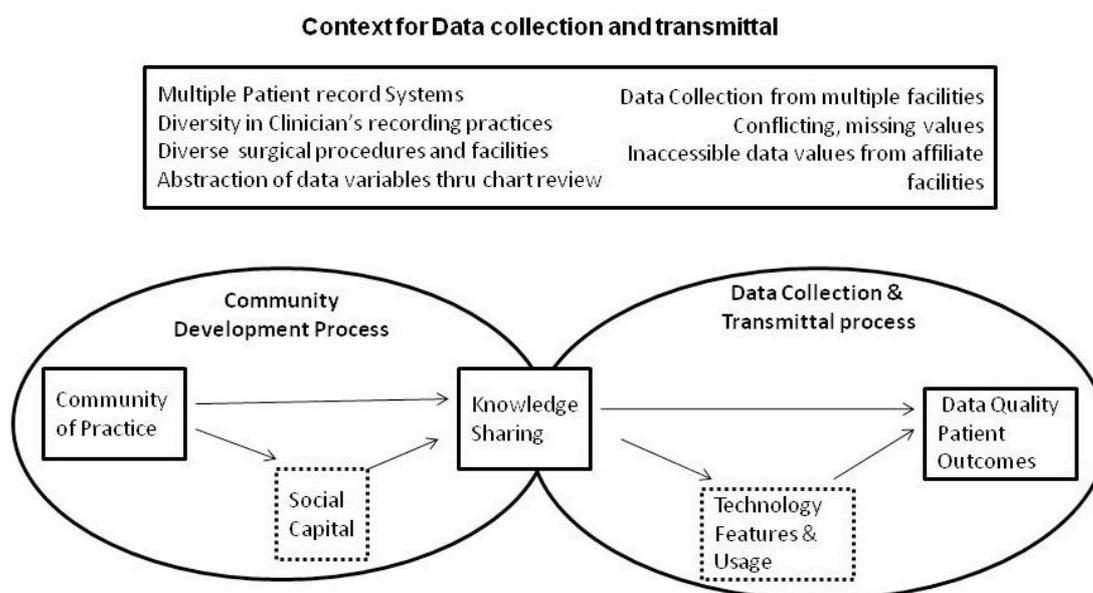


Figure 1: Conceptual Model

The role of the CoP was found to be far reaching – even impacting the technology feature adoption and usage among the nurses. The CoP also moderated the data collection practices adopted by the nurses. One of the

secondary observations from our interviews was the identification of the vast differences that exist across electronic patient record systems and the associated recording practices of clinicians. Without adequate interventions to “standardize” the input data set (as in the information system in our study) it is difficult to aggregate and utilize data from multiple facilities. Indeed the facilities and the clinical practitioners vary greatly in their health information system capabilities.

The grounded theory developed in this study meets the criteria of applicability. First, it fits the substantive area of study – a distributed system. Second, it is understandable to the practitioner. Third, it has generality to other healthcare systems and other industries that collect data from multiple sources. Finally, it provides potential control for the action and conditions to which it applies.

Implications for professionals

The importance of information systems that support distributed healthcare networks is quite apparent. This research contributes by addressing data quality concerns in such a diverse environment through understanding the process of enhancing data collection processes by a team of nurses supported by a community of practice. Our study advances the existing body of IS research on facilitating electronic healthcare delivery systems. The importance of information systems that support distributed healthcare networks is evident. This research contributes by addressing information quality concerns in such a diverse environment. This study deepens our understanding of enhancing information collection processes by a team of nurses supported by a CoP. Researchers can use the model for further research on the impacts of improving the quality of patient information in a HIS. There are currently many databases holding medical information, however, without asserting the quality of the data they cannot be used in any clinical setting. This paper contributes by describing a model for how data quality can be improved in such data sets and used in clinical systems to improve care.

Practitioners can benefit from this study in several ways. First, the study shows the importance of focusing on the quality of input data collection to build trust in the reliability of system outputs. Second, adequate training and support helps ensure that the data is suitable for aggregation and comparison across a distributed organization. A similar process may also be adopted to improve data quality in other information systems that span diverse organizations. Specifically, the importance of the CoP highlights that there is a possible way to build social capital that supports critical knowledge sharing.

REFERENCES

- Argote, L. 1999. *Organizational Learning: Creating, Retaining and Transferring Knowledge*. Boston: Kluwer Academic.
- Broodie, M. L. 1980. “Data quality in information systems”. *Information and Management*, 3, pp. 245-258.
- Brown, J., & Duguid, P. 1991. “Organizational learning and communities of practice: Toward a unified view of working, learning, and innovation”, *Organization Science*, 2, pp. 40-57.
- Brown, J.S. and Duguid, P. 2001. “Knowledge and Organizations: A Social-Practice Perspective”, *Organization Science*, 12(2), pp. 50-66.
- Burgess, D. 2005. “What Motivates Employees to Transfer Knowledge outside their work unit?”, *Journal of Business Communications*, 42(4), pp. 324-348.
- Cross, R., Leseter, T., Parker, A. and Velasquez, G. 2006. “Using Social Network Analysis to Improve Communities of Practice”, *California Management Review* 49(1), pp. 32-60.
- Duguid, P. 2005. “The Art of Knowing”: Social and Tacit Dimensions of Knowledge and the Limits of the Community of Practice”, *The Information Society*, 21, 109-118.
- Davenport, T.E., Prusak, L. 1998. *Working Knowledge: How organizations manage what they know*. Boston: Harvard Business School Press.
- Eisenhardt, K.M. 1989, “Building Theories from case study research,” *Academy of Management Review* (14:4), pp 532-550.
- Epstein, A.J. 2006. “Do Cardiac Surgery Report Cards Reduce Mortality? Assessing the Evidence”, *Medical Care Research and Review* (63:4), pp. 403-426.
- Ghosh, B. 2009. “Improving Data Quality in Healthcare Research Data Sets”, *Handbook of Research on IT Management and Clinical Data Administration in Healthcare* (ed. A. Dwivedi), IGI Press.
- Ghosh, B. and Scott, J.E. 2007a. “Effective Knowledge Management Systems for Nurses”, *Information Systems Management*, 24 (1), pp. 73-84.

- Ghosh, B. and Scott, J.E. 2007b. "Social Capital in Knowledge based Business Process Outsourcing". Proceedings of the Thirteenth Americas Conference on Information Systems, Keystone, Colorado.
- Glaser, B.G. and Strauss, A.L. 1999. *The Discovery of Grounded Theory: Strategies for qualitative research*. New York, NY, Aldine de Gruyter.
- Grover, V. and Davenport, T.H. 2001. "General perspectives on knowledge management: Fostering a research agenda", *Journal of Management Information Systems*, 18(1), pp. 5-22.
- Hansen, T., Nohria, N. and Tierney, T. 1999. "What's your strategy for Managing Knowledge?", *Harvard Business Review*, March-April, pp. 106-116.
- Inkpen, A.C. and Tsang, E.W.K. 2005. "Social Capital, Networks and Knowledge Transfer", *Academy of Management Review*, 30(1), pp. 146-165.
- Kanhanhali, A., Tan, B.C.Y. and Wei, K.K. 2005. "Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation", *MIS Quarterly*, 29(1), pp. 113-143.
- Ko, D.G., Kirsch, L.J. and King, W.R. 2005. "Antecedents of Knowledge Transfer from Consultants to Clients in Enterprise System Implementations", *MIS Quarterly*, 29(1), pp. 59-85.
- Lee, Y. W., & Strong, D. M. 2003. "Knowing-why about data processes and data quality", *Journal of Management Information Systems*, 20(3), pp. 13-39.
- Lesser, E.L. and Storck, J. 2001. "Communities of Practice and Organizational Performance", *IBM Systems Journal* 40(4), pp. 831-841.
- Mohr, L.B. 1982. *Explaining Organizational Behavior*, San Francisco, CA, Jossey-Bass
- Moore, J. (2009), "Bad Data and PHR Adoption", Chilmark Research, Retrieved 30 September, 2009 from <http://chilmarkresearch.com/2009/04/13/bad-data-amp-phr-adoption>
- Nahapiet, J. and Ghoshal, S. 1998. "Social Capital, Intellectual Capital and the Organizational Advantage", *Academy of Management Review*, 23(2), pp. 242-268.
- Poston, R.S., Reynolds, R.B. and Gillenson, M.L. 2007. "Technology Solutions for Data Accuracy and Availability Problems in Healthcare Records", *Information Systems Management*, 24 (1), pp. 59-70.
- Redman, T. C. 1992. *Data quality: Management and technology*. New York: Bantam Books.
- Suddaby, R., 2006, "From the Editors: What Grounded Theory is Not," *Academy of Management Journal* (49:4), pp 633-642.
- Strauss, A. and Corbin J. 1990. *Basics of Qualitative Research: Grounded Theory, Procedures and Techniques*, Newbury Park, CA, Sage.
- Strong, D. M., Lee, Y. W., & Wang, R. Y. 1997a. "Data quality in context", *Communications of the ACM*, 40(3), pp. 103-110.
- Strong, D. M., Lee, Y. W., & Wang, R. Y. 1997b. "Ten potholes in the road to information quality", *IEEE Computer*, 30(8), pp. 38-46.
- Wang, R. Y., & Strong, D. M. 1996. "Beyond accuracy: What data quality means to data consumers", *Journal of Management Information Systems*, 12(4), pp. 5-34.
- Wang, R. Y., Storey, V. C., & Firth, C. P. 1995. "A framework for analysis of data quality research", *IEEE Transactions on Knowledge Data Engineering*, 7(4), pp. 623-640.
- Wasko, M.M. and Faraj, S. 2005. "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice", *MIS Quarterly*, 29(1), pp. 35-57.
- Wenger, E. And Snyder, W.M. 1999. "Communities of Practice: The Organizational Frontier", *Harvard Business Review*, January-February, pp. 139-145.
- Wenger, E., McDermott, R. And Snyder, W.M. 2002. *Cultivating Communities of Practice: A Guide to Managing Knowledge*. Boston, MA., Harvard Business School Press.

COPYRIGHT

Ghosh & Scott © 2009. The authors assign to ACIS and educational and non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full

and this copyright statement is reproduced. The authors also grant a non-exclusive licence to ACIS to publish this document in full in the Conference Papers and Proceedings. Those documents may be published on the World Wide Web, CD-ROM, in printed form, and on mirror sites on the World Wide Web. Any other usage is prohibited without the express permission of the authors.