2003-04

Building Performance Models from Expert Knowledge

Abernethy, Margaret A.

Building Performance Models from Expert Knowledge

Margaret A. Abernethy, *The University of Melbourne*

Malcolm Horne, *Monash University* and *Southern Health*

Anne M. Lillis, *The University of Melbourne*

Mary A. Malina, *Naval Postgraduate School*

Frank H. Selto, *University of Colorado at Boulder* and *The University of Melbourne*

April 2003

Corresponding author:

Frank H. Selto

Leeds School of Business

University of Colorado at Boulder

Boulder, CO USA 80309-0419

frank.selto@colorado.edu

Tel: 1-303-492-1549

Fax: 1-303-492-5962

This research is supported by an Australian Research Council Linkage Grant, a Faculty of Economics and Commerce (University of Melbourne) Research Grant, and financial assistance from Southern Health and the University of Colorado at Boulder. We gratefully acknowledge research assistance from Yanhua Yang, Jialin Sun, Cameron Houston, and Louise Corben and helpful comments from workshop participants at the University of Melbourne, Griffith University, the 8th Biennial Management Accounting Research Conference, Sydney, February 2003, and the European Accounting Association Annual Congress, Sevilla, April 2003.

Building Performance Models from Expert Knowledge

Abstract

Improving management control of knowledge-based organizations motivates building performance management models (PMM) of causally related, key success factors (KSF) . This study elicits knowledge maps of KSF from field experts. These knowledge maps are layered to create the foundation of the organization's PMM.

The study elicits causal knowledge from experts who through their experience, training, etc. have encoded relational or causal knowledge about complex systems; that is, they understand how things fit and work together, although they might not have articulated that knowledge. Converting experts' tacit causal knowledge into organizational capability or explicit knowledge should a) perpetuate that knowledge in the organization, b) enable improved training of less experienced employees, and c) lead to deployment of improved systems (eg, PMM). Because no single method for eliciting mental models or knowledge maps dominates the literature, the study uses multiple methods and overlays their results to build a comprehensive causal model.

This study reports the results of a field study to build the foundation of a PMM in a clinical department of a large hospital. The study uses three qualitative methods to elicit mental models of KSF and their interactions from key clinical program administrators, physicians, and nurses. The motivation of the present study is to report the results of (1) tapping the causal knowledge of individual experts in the field and (2) triangulating multiple methods of qualitative data analysis. The alternative method of building a PMM by using archival data-mining is rejected for reasons of (1) limited archival time-series data, (2) limited scope of archival data, (3) myopic focus on conveniently available data, and (4) inability to screen out spurious relations. Because these limitations are generally present in knowledge-intensive organizations, this study's approach can have general application.

Key words: causal knowledge, qualitative method, knowledge map, performance measurement

Building Performance Models from Expert Knowledge

Introduction

During the last decade management accounting research has devoted increasing attention to the design of more effective management control systems by measuring and including multiple indicators of performance. An allegedly improved management control system is composed of a performance management model (PMM), which contains related financial and non-financial performance measures, complementary controls, and incentives to manage performance. Recent research on organizations' PMM includes prescribing forms of PMM, inferring PMM from related KSF, describing PMM in practice, and testing existing PMM for causal and other desirable properties.

Some researchers have prescribed comprehensive PMM (e.g. Otley's (1999) performance management model, Kaplan and Norton's (2001) balanced scorecard (BSC); Epstein et al's 2000 APL model). Others have studied the types of PMM being implemented in firms (e.g., Ittner and Larcker's (2001) valued-based management review). These models not only identify the importance of diverse performance measures but also emphasize the importance of the relations among diverse performance measures in PMM. Apart from reports of BSC adoptions, little empirical evidence exists to evaluate impacts of these models.

Empirical researchers have used archival data to assess relations between limited pairs of non-financial and financial performance measures that could be elements of a PMM (e.g. Sedatole, 2003; Abernethy and Lillis, 2001; Amir and Lev, 1996; Banker et al. 1993, 1995, 1996, and 2000; Behn and Riley, 1999; Foster and Gupta, 1990, 1999; Ittner and Larcker, 1997, 1998; Perera, et al. 1997). A few reports have assessed the causal properties of expected links among performance measures in firms that have implemented PMM (Malina and Selto, 2001, 2003; Rucci et al., 1998).

While this literature is beginning to improve our understanding of PMM functions, it is mostly silent on how PMM components (KSF and their interrelations) are identified. Interesting questions include: Where do PMM come from? Where should they come from? How are key success factors (KSFs), interrelations among these factors, and their links to organizational outcomes identified? It appears from descriptions of PMM in practice, that top management usually identifies KSFs and their interrelations (e.g. Kaplan and Norton, 2001; Malina and Selto,

2001) and/or existing archival data are "mined" to extract proxies that best represent the KSFs (Rucci, et al, 1998). There is some doubt as to the efficacy of top-down management initiatives in knowledge-based organizations (e.g., Hellstrom et al, 2001; Brown and Duguid, 2000; Davenport et al, 1996). In these organizations, knowledge is widely dispersed, and information asymmetry can exist between top management and the managers that control core operating activities. Although top-down PMM might not be always effective in improving performance, the topic of PMM in knowledge-intensive organizations has received limited attention in the accounting literature (Widener, 2003).

Knowledge-based organizations are dependent on the efficient management of human resources as this resource is the prime source of the organization's knowledge, capabilities, and systems. Because human resources are mobile and governed by self-interest, the theory of learning organizations predicts that organizations seek to convert individuals' tacit or unobserved knowledge to explicit or structural knowledge in order to build organizational capabilities (e.g. Nonaka, 1994; Nonaka and Takeuchi, 1995). Included in this tacit knowledge is knowledge that is relevant to the design of an effective PMM (Forrester, 1994). Through their experience and training, experts within organizations encode relational or causal knowledge about complex systems; that is, they understand how things fit and work together, although they might not have articulated that knowledge. Converting individuals' tacit knowledge to the organization's explicit knowledge is important to the effective management of intellectual resources (e.g. Huff and Jenkins, 2002; Forrester, 1994).

The purpose of this research report is to describe a general method for extracting tacit knowledge from experts in knowledge-intensive organizations. The study uses revealed tacit knowledge to create a knowledge map to form the conceptual, knowledge-based foundation of a PMM. This study draws on the psychology, management, and systems literatures to support the method used here to extract tacit knowledge and develop knowledge maps that form the foundation for a PMM. Because no single method for developing knowledge maps dominates the literature, the study uses multiple methods and overlays their results to build an inclusive performance model.

The study is undertaken in a clinical program within in a large public, teaching hospital. This setting was selected for several reasons. First, teaching hospitals represent the archetypical knowledge-based firm. Knowledge is dispersed and impacted in clinical programs where core operating tasks are performed and controlled by medical experts. Often a lack of goal congruence exists between the medical experts controlling the core production processes and top hospital management. Similar to other knowledge-based organizations, the medical-care production process is not well understood and organizational outcomes are difficult to measure quantitatively. These conditions create a management control environment where the monitoring and measuring performance of core operating activities is particularly problematic. It is also a setting where tacit knowledge is at risk. Converting tacit knowledge to explicit knowledge provides a means of developing a PMM for activities controlled by medical experts. Conversion of tacit knowledge is also the primary step before this type of organization can build its distinctive capabilities (Lorino and Tarondeau, 2002; Morecroft et al, 2002; Sanchez, 2001). Second, public teaching hospitals have a weak performance-management history, which presents an opportunity to demonstrate the value of documenting the tacit knowledge of expert clinical managers into KSFs and their interactions. Some researchers have suggested extending management control innovations to the health care environment (e.g., Capettini, et al, 1998; Forgione, 1999; Handler, et al., 2001; Steward an Lockamy, 2001). However, surprisingly relatively little is known about management control of healthcare, despite its being one of the largest and fastest growing sectors in most developed countries (Evans, 1998; Abernethy and Lillis, 2001). This study also contributes to management control in this specific domain.

The paper is structured as follows. The following section synthesizes the theoretical literature that forms the basis of the study. Next is a description of the research method used to build the KM and the method's three approaches to qualitative data analysis. The succeeding section presents a synthesis of the three approaches to create a composite KM. Finally, the paper presents conclusions and extensions to future research, including construct and statistical validity tests.

Literature Review

Prior Research

This study synthesizes several literature streams to build its theoretical and methodological foundation. The study draws on the accounting and management control literature as the basis for a performance management model, the psychology literature that has focused on methods of extracting experts' tacit knowledge, the strategic management literature on creating knowledge maps, and the systems dynamics literature to determine how mental data are used to construct and validate models of a complex process.

Performance management models

Organizational success depends on the effective management of the three sequential components of the production function – inputs (i.e. labour and capital), processes, and outcomes. Modern management control texts (e.g., Merchant, 1998; Simons, 2000) reflect a 40-year history of concern with systematically managing financial outcomes, inputs, and processes, consistently with an organization's strategy. Until recently, however, little attention has been devoted to understanding the *relations* among the three productive components or how to effectively manage the leading drivers or *causes* of financial outcomes (e.g., Otley, 1999; Ittner and Larcker, 2001) to achieve strategic goals. Reliance on financial performance is managing *through the rearview mirror* and leads to obvious dangers in a rapidly changing environment. Hence both researchers and managers see the potential usefulness of a performance management model that links the leading indicators associated with inputs and processes to outcomes. The balanced scorecard of Kaplan and Norton (e.g., 1996) was the first widely disseminated approach to articulate the links between leading inputs, processes, and lagging outcomes believed necessary to achieve strategies. While the balanced scorecard (BSC) has stimulated much related research in recent years, the literature contains no analysis of the foundations of such a comprehensive performance management model.

Although no common definition of the balanced scorecard exists, its approach is conceptually consistent with an organization's economic production process, albeit with modifications to the labels used to define the components.The balanced scorecard defines labour and capital inputs as "investments in people and technology." The BSC also separates outcomes by customer and financial perspectives. A BSC-type model is appealing because it captures the core elements of any organization's production process and recognizes causal links among them. Figure 1 illustrates a simple BSC-type model with outcomes classified as either effectiveness or efficiency outcomes. Effectiveness outcomes in hospitals, for example, include patient outcomes such as quality of care, and improvements in health status. Efficiency outcomes might include patient throughput and financial budget performance (profit is not commonly used as an efficiency outcome measure in hospitals). These labels are conceptually consistent with the "financial perspective" and "customer perspective" of the BSC.

Figure 1
Preliminary Conceptual Model

The relations among the core elements of an organization's production function in figure 1 become a PMM when the model reflects the organization's explicit knowledge of KSFs associated with each element, their interrelations, and the organization's outcomes. Such a model could convincingly communicate strategic intent, support operational decisions, model alternative outcomes given both decisions and environmental conditions, and provide reliable feedback for learning, communication, and improvement (Kaplan and Norton, 2001; de Geus, 1994). However, how to identify the KSFs and the relations among them remains unclear. The BSCs reported in practice appear to be the result of a) top-down imposition of desired KSF and interrelations (e.g., Malina and Selto, 2001), b) data-mining of existing archival sources (e.g., Porac, et al., 2002; Rucci, et al., 1998), and c) interviews of top or divisional managers (e.g., Ambrosini and Bowman, 2002). Clearly, all are feasible methods to gather PMM-relevant data, but all are somewhat flawed unless performance-relevant knowledge has been made explicit. Top-down models might not reflect tacit knowledge, routines, and capabilities that really drive performance (e.g., Huff and Jenkins, 2002). Data-mining relies on conveniently available data that might be unrelated to what should be measured, but has not. Top managers might understand the organization's intended strategy and policies but might be ignorant of or unwilling to discuss actual observed system behavior (e.g., Morecroft and Sterman, 1994; Forrester, 1994). These problems can be especially acute in knowledge-based organizations, where knowledge of the relation between causes and effects among organization key success factors depends on eliciting tacit knowledge from those closest to the KSF (Forrester, 1994).

Capturing tacit knowledge to build performance management models

The psychology, systems dynamics, and strategic management literatures use "mental data" from experts to build formal knowledge maps. Knowledge mapping is a means of converting key individuals' tacit knowledge to a model of explicit KSF and their interrelations. Eden (1992) agrees that cognitive or knowledge maps visualize knowledge and can communicate the visualization to individuals, groups or organizations. The benefits of mapping experts' tacit knowledge also identified by Huff and Jenkins (2002) are particularly relevant for providing the foundation of a performance management model:

1. Knowledge maps (KM) or mental models convert dispersed (and at-risk) tacit knowledge to explicit organizational knowledge.
2. KM organize and express the rationale of complex systems, thus aiding planning and evaluation activities.
3. KM represent micro- or macro-levels of knowledge of activities, processes, and systems, thus aiding individuals at all levels of the organization.
4. KM reveal and allow testing of "common knowledge" that is normally taken for granted but rarely articulated and exposed to refinement.
5. KM facilitate communication, learning, and creation of new knowledge and can be the key tool to building a learning organization.

This study relies on performance management literature and the literatures that have used tacit knowledge to develop knowledge maps of organizational practices. The systems dynamics literature provides the basis for the research method used to extract the mental data.

Research Method

This section describes the research method used to elicit mental data from experts to identify KSFs and the interrelations among them. These data are then used to prepare the KM foundation. The study uses a knowledge modeling approach similar to that reported by Vennix and Gubbels (1994), which includes the following major steps:

1. Identify the management problem.
2. Develop a preliminary conceptual model.
3. Involve experts in a structured, interactive modeling task.
4. Prepare the final conceptual KM
5. Evaluate the quality of the final KM

The first step, identifying the management problem, has been described earlier – Hospitals have a weak performance management history and socially costly inefficiencies. Part of the management problem is the difficulty of identifying relevant KSFs for clinical services. This study begins to address that problem by developing a performance management model for clinical services in a large, public teaching hospital. In the second step to build a preliminary conceptual model, the study expands the naïve, BSC-type model illustrated in figure 1 through researchers' use of complementary qualitative methods. The third step allows the experts themselves to directly articulate their causal knowledge. The fourth step overlays the researchers indirect and experts direct KMs to build a final model. The fifth step will rely on establishing face validity and preliminary statistical analyses of the final KM, as permitted by the limited measures currently available (see the appendix).

Nature and collection of data

The research method elicits mental data from field experts as the basis for developing the KM. Although all forms of data can be useful in modeling efforts, Forrester (1994) encourages using mental data to build KM because it is far vaster than either written or numerical data, particularly at early modeling stages. He classifies mental data into three categories:

1. Observations about structure and policies, which are generally reliable because of repeated experience
2. Observed system behavior, which also can be reliable
3. Expectations about system behavior, which can be least reliable because humans are imperfect dynamic modelers

The study first obtained primarily the first two types of mental data via semi-structured interviews of highly experienced, mid-level employees. The first contact in the field was the medical director of the hospital, who approved the research project and the dedication of hospital resources to the study. The hospital medical director provided the research team with background information concerning the history and the internal structure of the hospital and provided access to relevant archival data (e.g., budget reports and routine operating statistics).

The researchers requested access to multiple key participants as the "experts" in this step of the study. The hospital medical director identified four key participants for the study within one of the major clinical programs in the hospital,[2] The participants includes the nurse managers of the two major wards of the clinical program, the medical director of a unit within the program,

[2] The hospital is divided into clinical (e.g., obstetrics), clinical support (e.g. laboratory services), and non-clinical programs (e.g. administrative departments). The.larger study includes other clinical programs as well as clinical support programs with quite diverse missions, resources, and personnel. Subsequent research will compare and contrast knowledge mapping and PMM development in the diverse programs.

and the director of surgery within the program. All of the participants are directly involved in treating patients and have significant roles in resource management in the program. They all have extensive training and expertise in their fields, and perform different roles. Table 1 includes personal statements from each of the participants concerning their roles and responsibilities within the clinical program.

Table 1

The researchers then developed a semi-structured interview protocol (table 2) following Yin (1994), which was used to elicit KSF and causal knowledge from experts in the field. The structure of the interview protocol reflects the objective of the study to generate reliable mental data and to identify interrelated KSF. The preliminary model illustrated in Figure 1 motivated questions within the interview protocol, but the questions do not constrain responses to fit this naïve model. The form of the questions also is designed to elicit "stories of performance," because stories are vivid, contextual devices for relating personal knowledge and experience. By telling stories of how the organization functions and can succeed, participants can make explicit what might have remained tacit knowledge about goals, processes, performance, and outcomes (Ambrosini and Bowman, 2002; Boje, 1991).

Table 2

Preliminary and experts' conceptual models

The research team conducted two rounds of interviews. The objective of the first round of interviews was to elicit knowledge from the participants regarding desired performance outcomes, the drivers of those outcomes, the perceived causal interdependencies among performance drivers, and the causal time lags between enhancements in performance drivers and their effects. These interviews closely followed the interview protocol with the researchers using follow-up questions where necessary. The second round of interviews asked the experts to participate directly in building a KM (described as method 3 below). All interviews lasted between one and two hours and were tape-recorded.[3]

[3] Researchers adhered to strict guidelines (required by their universities) to ensure that each participant was aware of the nature of the study. Each participant formally agreed to participate in the study and agreed that the interview could be tape-recorded. Disinterested contract typists transcribed the interviews verbatim. Participants had the opportunity to review the transcribed interviews. Researchers corrected any factual errors.

This study uses three complementary methods to analyze qualitative data collected from the interviews. The first method relies on a relatively objective, computerized analysis of coded first interview transcripts. The second method reflects a traditional, ethnographic interpretation of first interviews and interview context. Methods one and two extend the naïve model in figure 1 to a more complex conceptual KM. The third method used the experts themselves to visually build KM with cards containing KSF that have been extracted from their coded interview transcripts.

The goal of each method is to elicit a hypothesized KM for the clinical program. All three methods start with mental data collected in the first round of interviews. Within each method, composite KM are created across individuals by using common constructs as "glue points" in combining the individual models (Clarke et al., 2000). Similarly, a composite, final KM is created across methods. This approach insures that all elicited constructs and linkages are retained in the final model. Although the resultant final model is complex, no prior reasons exist to trim constructs or linkages. Trimming is the proper task for later validation and testing.

The following subsection describes the use of computer software to facilitate data coding and retrieval, as the basis for the model-building methods that follow. Subsequent subsections describe the three model-building methods, giving background on each of the three methods, stating advantages and disadvantages of each method, and describing the resulting KM from each of the three methods.

Computer-assisted coding

A major development for qualitative-method researchers is efficient and flexible qualitative database software.[4] This software can serve two purposes. First, researchers use this software to create a database through systematic coding of the qualitative data. (e.g., Jasinski and Huff, 2002). Second, qualitative researchers also use the database software to analyze relations in the database in much the same way as quantitative researchers use the SPSS or SAS software packages (e.g., Malina and Selto, 2001). Using this type of software for data analysis helps bridge the gap between the qualitative and quantitative research perspectives. This use of qualitative database software is discussed in method 1.

[4] The term qualitative method is often used synonymously with field research and case research method. The proper distinction between qualitative method and other methods is by the nature of the data themselves. Qualitative method generally uses interview and observational data rather than quantities or quantified survey responses. See Ragin (1987) for a discussion of these alternative methods.

The key step in creating or analyzing the database is the development of a coding scheme. The coding scheme used in this study is consistent with the interview protocol and reflects the underlying theoretical priors that form the foundation for the study. The coding scheme marks occurrences of discrete KSF and performance driver themes in the data. However, the researchers also drew on their subsequent experience in the field to ensure that any variables unanticipated in the initial development of the interview protocol were not omitted in the coding process. The coding scheme is *not* designed to reflect causality or other associations between codes. This would have predetermined causal links; rather the coding scheme enables the researchers to identify key constructs that can form elements of a KM.

An advantage of using the computer software for coding is that it helps to ensure that all data are coded and thus reduces the potential for researcher bias when selecting data for analysis. Psychology studies, for example, find that people place more weight on confirming evidence than on disconfirming facts and tend to ignore or forget information that does not follow their line of reasoning (Nisbett and Ross, 1980). The use of the computer software for coding can reduce the occurrence and perception of data-selection and reporting bias.

A further advantage of the computer-assisted approach is the creation of an auditable and easily accessed qualitative database. Through the coding scheme, researchers express theoretical constructs and additional field knowledge they have gained. The data codes provide an index that enables researchers and others to retrieve all data relating to each code. For example, if one wanted all data relating to one of the dimensions of a construct of interest, say "empowerment," it is trivially easy to retrieve all data associated with the appropriate code. Without the software, researchers must expend significant effort to search for all data relating to this particular dimension and inevitably run the risk of omissions caused by fatigue and available time. Without such an approach, those not directly involved with the data have no feasible way to replicate the qualitative analysis, and must rely entirely on the reputation of the researcher(s) and the rhetoric of the report to assess its validity.

Because all three qualitative methods used in the study rely on the coded interview data, establishing coding reliability is critical to the validity of all the analyses and findings. This was done using the codes of two of the researchers. After coding the first interview, the researchers discussed the suitability of the initial coding scheme and refined the set of codes (Table 3).[5] The researchers then re-coded the first interview and coded the remaining interviews. Comparing the ratio of agreements and disagreements in coding all interviews by both researchers measures the degree of inter-rater reliability. An agreement occurs when both researchers use the same code for approximately the same section of text. A disagreement occurs when either the researchers did not code a section of text or they coded the section differently. Coding reliability averaged 83.2 percent, ranging from a low of 79 percent to a high of 88 percent. The average falls within the normally accepted range of at least 80 percent (Miles and Huberman, 1994). The two researchers discussed all disagreements then agreed upon one code for each. These consensus-coded interviews support the three data analysis methods.

Table 3

Method 1: Computerized discovery of causal links

As mentioned previously, an advanced feature of *Atlas.ti* and some other software packages is the ability to thoroughly analyze relations among the coded data. The researcher can use the software to create and test relational maps that aid in developing theoretical models and assessing the relative strength of the relations among the variables of the models. The use of the computer software to build a model of relations among qualitative data items reflects an attempt to achieve validity by establishing that the discovered relations are reliable and not purely subjective. Another researcher would be able to use the methods to develop the same model from the data. Interpretations of the model can vary, but that is also the case with quantitative research. Studies that utilize this model-building feature of qualitative data software include Malina and Selto (2001) and Friese (1999).

A disadvantage of computer-assisted data analysis is the potential to report causal links that only represent the query rules used for establishing the links (e.g., proximity of codes) rather than valid causal relations. Thus, additional, subjective evaluations of software-discovered links usually are necessary. This computer-assisted approach also can be incomplete if the coding scheme does not reflect all of the relevant underlying theory and field experience gained. As in

[5] As recommended by Miles and Huberman (1994), the researchers found that definitions become sharper when two or more researchers code the same data set and discuss their initial difficulties.

other research perspectives, the use of multiple methods can be a valuable addition to qualitative research.

The first KM derived in this study results from the relatively objective computerized analysis of the coded interview transcripts, following the method described by Malina and Selto (2001). The researchers used the database software to discover associations and possible causal relations among the coded sections of text from the interviews. The following criteria are the query rules used to identify associations between variables:

- Coded quotations of one type *enclose* coded quotations of another type
- Coded quotations of one type *are enclosed by* coded quotations of another type
- Coded quotations of one type *overlap* coded quotations of another type
- Coded quotations of one type *are overlapped by* coded quotations of another type
- Coded quotations of one type *precede* coded quotations of another type by no more than *one* line
- Coded quotations of one type *follow* coded quotations of another type by no more than *one* line

The software uses these rules to count the number of associations between pairs of codes. The researchers inferred causation between the elements of a model based on these frequencies and the subjective evaluation of the theoretical coherence of these links by the researchers. (Miles and Huberman, 1994). The researchers set a quantitative threshold to identify a likely casual link in the model at seven associations, the mean number of all such occurrences, excluding zeros. Within this reduced set, each observed linkage was evaluated by reading the relevant transcript sections to subjectively assess whether the observed linkage was coherent or spurious. Only those links with at least seven coherent links were retained in the model. This constraint is an acceptance probability compromise of 50 percent, between including all observed, coherent links and setting a standard confidence interval about the mean (e.g., 95 percent).

These linked codes form the basis of the first version of the department's KM, which is shown in figure 2. Figure 2 follows the basic framework of the *naive* model but adds KSF for each of the major components. This KM reflects frequently observed interactions among the constructs within the "people and technology" component. For example, frequent interactions exist between "employee training" "teamwork/networking", "employee retention", "employee empowerment" and "employee satisfaction". Note that only "employee training" and "teamwork/networking" have direct impacts on the constructs identified within the "process" component of the model, namely "patient care" and "patient flow". These two constructs, in turn, have downstream effects on "clinical outcomes" and "department-level financial outcomes." People and process effects on outcomes such as "patient satisfaction" are indirect. Importantly for subsequent additions and future use as a communication and control device, this KM tells a credible story about how performance in this medical program is causally related.

Figure 2 also includes external factors, which all respondents mentioned repeatedly. For example, the ability to recruit and train staff is directly dependent on availability of trained nursing and medical staff; economic and political factors influence the hospital's overall budget and hence influence ability to hire additional staff; bed shortages within the hospital and outside the hospital influence process changes (i.e. processes designed to improve patient flow so that patients can be discharged earlier are influenced by availability of nursing home beds). For economy of presentation, the links among external factors and KSFs are not elaborated in this or subsequent models.

Figure 2

Figure 2 reflects the most conservative KM of those derived from this study because of the conservative quantitative threshold used. While figure 2 is not as rich descriptively as those that follow, it does build on the *naive* model, identifying KSF in this particular hospital setting. It tells a concise, coherent story about building physician and nursing capital, improving clinical processes, affecting patient outcomes, and influencing department-level financial outcomes. Method 1 and figure 2 demonstrate that the qualitative data support at least this core model and that the richer models that follow are not purely subjective interpretations. Subsequent models build on the core KM in figure 2, although neither the researchers nor other participants had access to this model during the development of models using methods 2 and 3.

Method 2: Ethnographic analysis of interview data

The second method of qualitative analysis reflects a traditional, ethnographic interpretation of interviews and interview context. This is perhaps the most frequently used method for analyzing qualitative data. Ethnographic interpretation allows the researcher to drive the creation of a KM through his or her understanding of the context (from current and prior experience) and the use of the entire interview transcript to identify the causal relations. These relations might not be captured in software-discovered frequencies of association based on proximity of comments. This use of the entire transcript provides a means of interpreting the perceived importance of the causal links from interviewees' comments. Use of the transcript data in this way can increase the likelihood that the ultimate model reflects reality rather than associations based only on software-discovered proximity of themes. Disadvantages of ethnographic interpretation are the tendency to focus on confirming evidence and the possibility of an incomplete analysis because of the cognitive complexity of the task.

To build the model ethnographically, three researchers, who did not participate in the development of the computer-assisted model or see it beforehand, used the basic code-and-retrieve program of *Atlas.ti* to collect all text attached to each thematic code. Inference of causality associated with these themes required reference back to the transcripts to ensure that:

1. Extracted segments were not taken out of context, and
2. All relevant text segments faithfully reflected the causal connections described by interviewees.

The researchers drew ovals to represent the KSF and arrows to capture causal links among the constructs based on their interpretation of the coded transcript. KSF and causality linkages for all four interviews were highly convergent among the three researchers. Researchers resolved differences in terminology, linkages, or levels of aggregation by consensus. This process resulted in the development of consensus KMs for each of the four interviews. These four models were then combined into one overall KM as shown in figure 3. Method 2's model reflects the most inclusive collection of KSFs and relations revealed by analysis of each interview. For example, if one model included "training," "empowerment," and "communication", but others included only "training" and "empowerment," figure 3 includes all three factors. Note that in some cases the constructs are combined into an oval. This occurred when the researchers always grouped them in their intuitive models to describe the links among the constructs. For example, "employee empowerment", "satisfaction & morale" and "retention" had the same relation with the other constructs.

Figure 3

Figure 3 describes a more complex KM that reflects the researchers' interpretations of the importance of expressed causal linkages, beyond the more restrictive frequency of comment in method 1. Figure 3 includes a number of additional factors, such as performance reviews, research, and clinical trials, that at least one interviewee expressed as important but did not generate sufficient frequency of comment to be captured in method 1. Finally, observe that figure 3 also identifies numerous external factors.

Method 3: Interactive modeling by experts

The third method actively engaged the expert participants by asking them to map causal relations among KSF that researchers had identified via the coding scheme using data collected from the first round of interviews. This method allows the experts to arrange KSF cards according to their experiences of causality (Daniels et al., 1995, 2002; Sirsi et al., 1996). This method is particularly advantageous for creating KM in early stages of investigating complex processes, when more objective data are either unavailable or pose an undue risk of dominating model building (Homer-Dixon, 1996). It captures how informants themselves construct causal patterns when given the discrete themes or constructs extracted from their prior interviews. This method presumes that the interviewees have insight into the causal relations within their own processes, which might not be the case in all situations. Thus, this method requires the participation of experts (e.g., knowledgeable and experienced in the phenomena), as this study achieved.

This method can be described as a visibly aided, post-encoding sorting or construction task. The type of task used here can be significantly more successful in eliciting participants' complex, relational knowledge than unaided memory tasks (Wattenmaker, 1992). [6] Furthermore, to help subjects organize their tacit knowledge of performance by relevant features, the researchers framed the task as modeling causality among the revealed KSF (e.g., Attarwala and Basden, 1985; Luft and Shields, 1999; McEarlean et al., 1999). Thus, this method did not require subjects to recall all previously stored or encoded information, a cognitive task that can be impeded by memory storage and retrieval processes (Spector and Davidsen, 2002).

The constructs given to participants used the terminology of their individual interviews. Self-stick labels reflected each KSF, and blank self-stick arrows allowed for connections. The

[6] The researchers expected the participants to be intimately familiar with the performance factors and possible relations among them, which should enable successful recall of encoded information. However, the study's objective is not to test experts' abilities but to reliably elicit their knowledge. This objective and the consistent results from several decades of research on memory and recall indicate the benefit of providing as much assistance as possible without leading participants to predetermined results.

researchers met with the participants for a second time, and during this second meeting, the researchers:

a. Provided each participant with the set of self-stick labels developed from the transcript of his or her prior interview,
b. Explained that these represented KSF or activities based on their comments in the prior interview,
c. Carefully defined each factor by using the specific comments and examples from their prior interview,
d. Established mutual understanding that the specific comments reflected the theme on the self-stick label, and
e. Requested their input on how these factors were inter-related (e.g., "We would like you to look at these factors or activities, see how they fit together – whether there are relationships among them – and position them in time sequence – whether there are some things that you do at one point in time that influence other factors later.")

Each participant positioned the self-stick labels on an A3 piece of paper and placed or drew arrows between them as appropriate to reflect their causal knowledge. At the same time, the accompanying discussion was tape-recorded to capture the rationale behind the relations discussed. Participants were able to revise their positioning of the labels and causal arrows as often as they wished during the interviews. Once the participants were satisfied that the model reflected their perceptions of the elements of the model and the linkages among the elements, the researchers firmly affixed the labels with tape. Researchers made no attempt to finalize any model until the participant declared it finished and a good representation of the relationships among the KSFs.

The research team overlaid the experts' models into a composite KM shown in figure 4. As in figure 3, figure 4 represents the most general inclusion of KSFs and relations obtained from the participative method. When the few modeling conflicts occurred, the majority judgment ruled while retaining minority insights if at all possible. For example, three participants placed "recruitment" as a leading indicator of the major construct, "people," but one placed it within the "resources" construct. Figure 3 shows "recruitment" as a leading indicator within the "resources" construct. Figure 4 describes a KM that is more complex than figure 2, less complex than figure 3, but still tells a coherent story of department performance.

Figure 4

The level of complexity of figure 4 undoubtedly reflects the number of cues provided to the participants derived from the consensus scheme shown in table 3. This deliberate research design judgment makes the task descriptive but also keeps the cognitive complexity of the task within reasonable bounds, although the maximum feasible number of cues for the task was not apparent. Participants were free to create additional self-stick labels beyond those provided by the research team, but three participants each added only one. Participants appeared to be comfortable working with the cues provided; whether they could have worked effectively with more cues is unknown.

Figure 4 adds several important features to the previous models. First, clinical trials can be an important source of employee training and well being, and can lead to process improvements. Second, research contributes to improvements in both processes and department status.

Triangulation of qualitative methods to prepare the final KM

Each method of eliciting the department KM yields important information and slightly different results. The computer-aided approach catalogues all frequent associations between pairs of KSFs. This approach, however, cannot capture the perceived intensity or importance of relations among factors beyond frequency of code associations. The ethnographic and expert-participation methods, however, can reflect these important, although more subjective, assessments. At this early stage of model development, all information should be used. Therefore, the research team overlaid the models from each of the three methods to create a composite model from all the elicited KSFs and relations among them, using common constructs as "glue points" to orient the models (Clarke et al., 2001). Once again, the research team created the most general composite model, effectively layering figures 3 and 4 upon figure 2. Figure 5 displays this composite model, which reflects the full set of counted, inferred, and elicited KSF and relations among them. The models – computer, ethnographic, or expert – are complementary. While this is not the same as cross-validation from independent data, the triangulation method uses three independent assessments of data. This triangulation method effects a more inclusive extraction and articulation of the causal linkages implicit in the participants' initial causal models than any single approach. The method enhances the possible descriptive validity of the model in figure 5.

Figure 5

The KM in figure 5 reflects the complicated nature of the clinical program studied here, which (like most business units) is a complex entity. Furthermore and at this stage of the research, including all relevant complexity is preferable to premature pruning of the model. However, complexity does not equate to construct validity. A judgment of whether this KM captures critical KSF of clinical success requires additional validation, which is the object of ongoing research. Furthermore, whether a validated KM can form the basis of a *cost-effective* performance management model, awaits evidence from planned field experiments and implementations.

Conclusions and Future Research

This study presents a general qualitative approach to identifying a plausible and coherent KM, which can be the first step in establishing a causal PMM in knowledge-intensive organizations. This study is set in a large clinical program of a major hospital, recognizing the importance of improving healthcare performance to the public, governments, and medical personnel. This study's primary contribution is the triangulation of multiple qualitative methods to study causal performance modeling. The study's secondary contribution is to document a KM in a hospital setting where knowledge workers perform complex processes, the outcomes of which are difficult to quantify. This approach demonstrates (1) using mental data to portray the knowledge of experts in the field and (2) using qualitative approaches to convert field experts' accumulated tacit knowledge to a KM of explicit KSF and their interrelations. This approach for developing a KM is adaptable to other knowledge-intensive organizations.

The research effort is motivated by the twin realities in knowledge-based organizations of (1) opportunities to improve performance management via a causal model and (2) insufficient archival data on which to base such a model. The derived KM is unconstrained by currently available data or measures; instead, it is a conceptual model of performance and a foundation for performance measurement. Validating the measurement model is a later stage, because currently available archival measures might not be valid representations of model factors.

The study uses interviews as a common source of mental data but triangulates three independent approaches to the analysis of the data to enhance the validity of the use of qualitative methods and the resulting KM. Most qualitative studies rely on one of several available methods, but a single qualitative method might not identify all of the organization's important performance factors and causal relations. This study triangulates the results of (1) computer-assisted modeling, (2) ethnographic modeling, and (3) interactive system modeling by experts. The integration of the three qualitative approaches leads to a model of system performance. To our knowledge, no previous study has triangulated methods as a means of validating the causal modeling of qualitative data.

None of the three methods used in this study, by itself, revealed the complexities of activities and their relationships that are reflected in the composite KM in figure 5. Each method added information and cross-validation to the modeling effort. The computer-modeling approach might offer comfort to more quantitatively oriented researchers that the ultimate model has a relatively objective core. A limitation of the computer analysis is that it uses an unweighted numerical threshold to identify the "core model" of possible causal links and does not consider expressed or perceived importance of causal relations beyond simple counting. Both the ethnographic modeling (using researchers' perceptions of causality) and the expert modeling (eliciting experts' causal knowledge add this crucial qualitative (although more subjective) input.

The balanced scorecard (BSC) also motivates this study. However, the results of this study do not depend on acceptance of the balanced scorecard *per se*. Indeed, the resulting KM is an open model containing factors not usually included in reported BSC models. Of particular importance are the complex interactions among KSFs within the "resources" component of the performance model. Understanding these interactions is important in the final design and implementation of the PMM. The KM presented here also recognizes the importance of external factors. While these largely uncontrollable factors might not form part of a PMM, as such, they can important in forecasting and in explaining variances that occur between expected and actual performance.

It might not be feasible or desirable to implement a PMM that includes proxies for all of the KSFs identified in Figure 5. Sufficient data usually are not readily available, and one should resist the temptation to use whatever data are at hand. Data-driven development of a PMM can ignore performance drivers, such as staff empowerment and satisfaction, that are critical to success but are currently immeasurable. However, given an understanding of the relation of factors within each major component, it might be possible to select a reliable proxy for the leading KSF related to that construct. For example, sick leave statistics (which are relatively easily obtained from an organization's HR system) might be used to reflect the KSF, staff satisfaction. Despite the availability of suitable proxies, one also must guard against using excessive metrics for performance measurement, which can lead to information overload or selective focus on the most easily achieved objectives (e.g., Feltham and Xie, 1994).

Future research

Figure 5 shows a complex model of the drivers of clinical program performance, but it is not necessarily either a valid or cost-effective PMM. The research program also will expose the causal logic of the KM to validity tests by experts in the field and statistical validity tests using available data. Subsequent deficiencies in the PMM might be traced to new measures, implementation, or use of the PMM, rather than the KM foundation. This two-stage validation is not possible without the conceptual definition of a KM. The identification of KSF separate from their measurement also allows for a less data-driven model. Although critical people-focused measurements, for example, might be desirable because of their salience or attention-directing qualities, performance drivers such as "empowerment" might be best controlled by mechanisms other than measurement (e.g., Ouchi, 1977; Simons, 2000).

The KM in figure 5 forms the foundation for a PMM but the development of a feasible PMM requires identification and measurement of the most critical drivers in the KM. Ideally, the evolution of figure 5 into a successful PMM can proceed as follows:

1. Identification of suitable measures or proxies for each KSF in the model.
2. The assembly of time-series data representing each of the KSFs in the model and covering a sufficiently long period of time to capture the temporal effects between the identified drivers and performance outcomes.
3. Identification of the time lags inherent in the causal relations depicted in the model by documenting the time-lagged correlations among pairs of causally-linked variables.
4. Statistical identification of the most critical performance drivers among those identified in the model.
5. Consideration of the costs and potential accuracy of on-going measurement of these KSF.
6. Selection of the measures to be included in the PMM based on their established validity as KSF, and the economics of routine performance measurement.

The data currently available are insufficient to fit the full KM or properly test more than a few of its individual linkages (see the appendix to this paper). The limited feasible tests do identify several key measures that are likely to be key components of an eventual PMM. Further measurement development and testing will form the basis for transforming the conceptual KM into a functioning performance management model. This further development effort conceivably will require aggregation and pruning of some aspects of the conceptual KM but that is left to ongoing research.

The project also will test the effectiveness of the proposed PMM through field experiments similar to Lipe and Salterio (2000) and Luft and Shields (1999). As the existing data are not sufficient to fit the proposed PMM, the project also commenced the design and collection of new data. The development of a PMM based on the analysis of field experiments and these data represents ongoing longitudinal research. The project anticipates abundant data availability within several years.

Appendix: Preliminary Statistical Analyses of Clinical Knowledge Map

The researchers first sought the assistance of interviewees to identify KM proxies. For example, interviews revealed that hospital personnel regard sick leave and non-paid leave as indications of nursing job satisfaction. The second step was to work with hospital personnel to identify operating and cost information from the hospital's databases that matched the proposed proxies. Table 1A presents the constructs in the model, proxies derived from interview data, and the corresponding hospital data currently available for testing portions of the model. Available data include monthly operating statistics, staffing (EFT and dollar amounts), and budget figures for 34 consecutive months. EFT refers to effective full-time employees. LOS refers to length of stay. AWOP refers to absent without pay (non-paid leave). WEIS is caseload weighted by intensity of care. All staffing variables reported as EFT & dollars. These data were combined to make a more parsimonious dataset; for example, data for all classes of permanent nurses were combined. It was apparent at this stage that the data were insufficient for validation of the model, identification of time lags or reliably identifying key causal factors. As is evident in Table 1A, little of the currently available data matches the requirements of Figure 5. Most of the routinely collected hospital data satisfies required regulatory reporting and does not fit the model.

Table 1A

At this stage, the best use of the limited data available is the illustrative testing of minimal elements of the KM. The reduced model which is testable at this stage is in Figure 2A and includes several people, process and performance constructs, proxies, and measures. One exogenous variable, department budget, is also included. To avoid confirming obviously mechanical or size-driven relations, such as the internal budget's driving training (EFT or dollar) just as a function of size, the EFT and dollar measures are deflated by department totals to obtain rates of training, sick leave, unpaid leave, turnover, and temporary staffing.

Figure 2A

Table 2A reports the results from testing the restricted model. The analysis in table 2A finds several statistically significant results consistent with the KM. After eliminating collinear variables, there is a positive and significant relation between two proxies that capture the people variable, namely rates of non-paid leave (AWOP; a= 0.055) and temporary staffing (a = 0.036) and the proxy for the process variable, namely, average length of stay (LOS). Increases in either of the explanatory variables would reduce the overall capabilities of the nursing staff and thus adversely influence patient flow. Secondly, department level performance, proxied by total number of cases in process, is negatively affected by the average length of stay (a < 0.000). In other words a decrease in process flow has an adverse impact on unit performance. A lower rate of processing would increase the number of cases in process, and vice-versa. These measures (average length of stay, rates of non-paid and temporary staffing, and total number of cases in process) are candidates for inclusion in the department PMM, whereas other measures are less likely to be useful. As indicated in table 2A, regression tests using lagged independent variables (up to six months) reinforce the importance of these measures, but not others in table 4 (a = 0.05). First-difference (changes) models identify turnover rate, total LOS, and WEIS as other possibly important PMM elements (a = 0.05).

Table 2A

Several factors might explain why other predicted relations are not supported by the data. The KM might be misspecified, although this seems unlikely given the extensive efforts to extract and use experts' knowledge. It is more likely that the available measures are not valid proxies of the KM constructs. Omitted variables, noisy monthly data, and limited time series data also can explain these findings. Importantly, the results presented here provide valuable input into future statistical research at this and similar sites; namely, most existing data that are created for external reporting purposes are not likely to be adequate for building, testing, and using a PMM. Significant new measurement efforts are required. This finding reinforces the importance of building KMs on experts' knowledge rather than on the expediency of using currently available data.

References

Abernethy, M. and A. Lillis. (2001). Interdependencies in organization design: A test in hospitals. *Journal of Management Accounting Research* (13): 107-129.

Ambrosini, V. and C. Bowman. (2002). Mapping successful organizational routines. in Huff, A. and M. Jenkins (eds.). (2002). *Mapping Strategic Knowledge*. Thousand Oaks, CA: Sage Publications: 19-45.

Amir, E and B Lev. (1996). Value-relevance of non-financial information: The wireless communications industry. *Journal of Accounting and Economics* 22(1-3): 3-30.

Attarwala, F. and A. Basden. (1985). A methodology for constructing expert systems. *R and D Management* 15(2): 141-49.

Banker, R., G. Potter and R. Schroeder. (1995). An empirical analysis of manufacturing overhead cost drivers. *Journal of Accounting and Economics* 19(1): 115-37.

Banker, R., H. Chang and S. Majumdar. (1993). Analyzing the underlying dimensions of firm profitability. *Managerial and Decision Economics* 14(1): 25-36.

Banker, R., S. Lee and G. Potter. (1996). A field study of the impact of a performance-based incentive plan *Journal of Accounting and Economics* 21(2): 195-226.

Banker, R.., G. Potter, D. Srinivasan. (2000). An empirical investigation of an incentive plan that includes non-financial performance measures. *The Accounting Review* 75: 65-92.

Behn, B. and R. Riley. (1999). Using non-financial information to predict financial performance: The case of the US airline industry. *Journal of Accounting, Auditing and Finance* 14(1): 29-56.

Boje, D. (1991). The storytelling organization: A study of story performance in an office-supply firm. *Administrative Science Quarterly*, 36(1): 106-26.

Brown, J. and P. Duguid. 2000. Balancing act: How to capture knowledge without killing it. *Harvard Business Review*, May-June: 73-80.

Capettini, R., C. W. Chow and A. H. McNamee. (1998). On the need and opportunities for improving costing and cost management in health-care organizations. *Managerial Finance*, v24(8): 46-59.

Clarke, I. and W. Mackaness. (2001). Management 'intuition': An interpretive account of structure and content of decision schemas using cognitive maps. *Journal of Management Studies*, 38(2): 147-72.

Clarke, I., M. Horita, W. Mackaness. (2000). The spatial knowledge ofretail decision makers: Capturing and interpreting group insight using a composite cognitive map. *International Review of Retail, Distribution and Consumer Research*, 10(3): 265-85.

Daniels, K. G. Johnson, L. de Chernatony. (2002). Task and institutional influences on managers' mental models of competition. Organization Studies, 23(1): 31-62.

Daniels, K. L. de Chernatony, G. Johnson. (1995). Validating a method for mapping managers' mental models of competition. *Human Relations*, 48: 975-91.

Davenport, T., S. Jarvenpaa, M. Beers. 1996. Improving knowledge work processes. *Sloan Management Review*, Summer: 53-65.

de Geus, A. 1994. Modeling to predict or to learn? In Morecroft and Sterman (eds.) *Modeling for Learning Organizations*. Portland: Productivity Press.

Eden, C. (1992). On the nature of cognitive maps. *Journal of Management Studies*, 29(3): 261-5.

Epstein, M., P. Kumar and R. Westbrook. (2000). The Drivers Of Customer And Corporate Profitability: Modeling, Measuring, And Managing The Causal Relationships, *Advances in Management Accounting*, v9 (1), 43-72.

Evans, J.H. 1998. Cost management and management control in health care organizations: Research Opportunities. *Behavioral Research in Accounting,* 10: 78-93.

Feltham, Gerald A. and Jim Xie. 1994. Performance measure congruity and diversity in multi-task principal-agent relations, *The Accounting Review*, 69(3): 429-453.

Forgione, D. A. 1999. Healthcare accounting research: A review of the professional literature, models, data, and research opportunities. *Research in Government and Nonprofit Accounting*, v10(1), 163-200.

Forrester, J. 1994. Policies, decisions, and information sources for modeling. In Morecroft and Sterman (eds.) *Modeling for Learning Organizations*. Portland: Productivity Press.

Foster, G. and M. Gupta. (1990). Manufacturing overhead cost driver analysis. *Journal of Accounting and Economics* 12(1-3): 309-37.

Foster, G. and M. Gupta. (1999). The customer profitability implication of customer satisfaction. Stanford University and Washington University working paper.

Friese, S. (1999). *Self-concept and Identity in a consumer society: Aspects of symbolic product meaning*. Marburg, Germany: Tectum.

Handler, A., M. Issel and B. Turnock. (2001). A conceptual framework to measure performance of the public health sector. *American Journal of Public Health* 91(8): 1235-1239.

Hellstrom, T, U. Malmquist, J. Mikaelsson. 2001. Decentralizing knowledge: Managing knowledge work in a software engineering firm. *The Journal of High Technology Management Research*, 12: 25-38.

Homer-Dixon, T. (1996). Strategies for studying causation in complex ecological- political systems. *Journal of Environment and Development* 5: 132-48.

Huff, A. and M. Jenkins (eds.). (2002). *Mapping Strategic Knowledge*. London: Sage Publications.

Ittner, C. and D. Larcker. (1997). Quality strategy, strategic control systems, and organizational performance. *Accounting Organizations and Society* 22(3/4): 293-314.

Ittner, C. and D. Larcker, (1998). Are non-financial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research* 26(supplement): 1-34.

Ittner, C. and D. Larcker. (2001). Assessing empirical research in management accounting: A value-based management approach. *Journal of Accounting and Economics* 32: 349-410.

Jasinski, D. and A. Huff. (2002). Using a knowledge-based system to study strategic options and researchers' mental models. In Huff, A. and M. Jenkins (eds.). *Mapping Strategic Knowledge*. London: Sage Publications.

Kaplan, R. and D. Norton. (1996). *The balanced scorecard*. Boston, MA: Harvard Business School Press.

Kaplan, R. and D. Norton. (2001). The *strategy-focused organization*. Boston, MA: Harvard Business School Press.

Lipe, M. and S. Salterio. 2000. The balanced scorecard: Judgment effects of common and unique performance measures. *The Accounting Review*, 75(3): 283-298.

Lorino, P. and J-C. Tarondeau. (2002). From resources to processes in competence-based strategic management. in Morecroft, J., R. Sanchez, A. Heene (eds.). (2002). *Systems Perspectives on Resources, Capabilities, and Management Processes*. Advanced Series in Management. Amsterdam: Pergamon: 127-52.

Luft, J. and M. Shields. (1999). Accounting classification of expenditures on intangibles: Cognitive causes of managerial myopia. Working paper, Michigan State University.

Malina, M. and F. Selto. (2001). Communicating and controlling strategy: An empirical study of the effectiveness of the balanced scorecard. *Journal of Management Accounting Research*: 47-90.

Malina, M. and F. Selto. (2003). Management control and performance measurement: An empirical test of causal relations. Naval Postgraduate School working paper.

McErlean, F., D. Bell and J. Guan. (1999). Modification of belief in evidential causal networks. *Information and Software Technology* (41): 597-603.

Miles, M., and A. Huberman. (1994). *Qualitative Data Analysis*. Thousand Oaks, CA: Sage Publications.

Morecroft, J., R. Sanchez, A. Heene (eds.). (2002). *Systems Perspectives on Resources, Capabilities, and Management Processes*. Advanced Series in Management. Amsterdam: Pergamon.

Morecroft, J.D.W. and J. Sterman (eds.). (1994). *Modeling for Learning Organizations.* Portland: Productivity Press.

Nisbett, R. and L. Ross. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1): 14-38.

Nonaka, I. and H. Takeuchi. (1995). *The Knowledge-Creating Company*. New York: Oxford University Press.

Otley, D. (1999). Performance management: A framework for management controls systems research. *Management Accounting Research* 10 (4): 363-382.

Ouchi, W. (1977). The relationship between organizational structure and organizational control. *Administrative Science Quarterly*, 22: 95-113.

Perera, S., G. Harrison and M. Poole. (1997). Customer-focused manufacturing strategy and the use of operations-based non-financial performance measures: A research note, *Accounting Organizations and Society* 22(6): 557-572.

Porac, J., Y. Mishina, and T. Pollock. (2002) Entrepreneurial narratives and the dominant logics of high-growth firms. in Huff, A. and M. Jenkins (eds.). (2002). *Mapping Strategic Knowledge*. Thousand Oaks, CA: Sage Publications: 112-136.

Ragin, C. (1987). *The comparative method: Moving beyond quantitative and qualitative strategies*. Berkeley, CA: University of California Press.

Rucci, A., S. Kirn and R. Quinn. (1998). The employee-customer-profit chain at Sears," *Harvard Business Review* 76(Jan-Feb): 82-97.

Sanchez, R. (ed.) (2001). Knowledge *Management and Organizational Competence*. Oxford: Oxford University Press.

Sedatole, K. 2003. The effect of measurement alternatives on a nonfinancial quality measure's forward-looking properties. Forthcoming *The Accounting Review*.

Simons, R. (2000). *Performance measurement and control systems for implementing strategy*. Upper Saddle River: Prentice-Hall.

Sirsi, A., J. Ward and P. Reingen. (1996). Microcultural analysis of variation in sharing of causal reasoning about behavior. *Journal of Consumer Research* (22): 345-72.

Spector, J. and P. Davidsen. (2002). Cognitive complexity in decision making and policy formulation: A systems dynamic perspective. in Morecroft, J., R. Sanchez, A. Heene (eds.). (2002). *Systems Perspectives on Resources, Capabilities, and Management Processes*. Advanced Series in Management. Amsterdam: Pergamon: 155-71.

Stewart, L. and A. Lockamy III. (2001). Improving competitiveness through performance measurement systems. *Healthcare Financial Management*, Dec: 46-50.

Vennix, J. and J. Gubbels. 1994. Knowledge elicitation in conceptual model building: A case study in modeling a regional Dutch health care system. In Morecroft and Sterman (eds.) *Modeling for Learning Organizations*. Portland: Productivity Press.

Wattenmaker, W. (1992). Relational properties and memory-based category construction. *Journal of Experimental Psychology* 18(5): 1125-38.

Widener, S. (2003). An empirical investigation of the relation between the use of strategic human capital and the design of the management control system. Rice University working paper.

Yin, R. (1994). *Case Study Research: Design and Methods*, *2nd ed*. Applied Social Research Methods Series, Volume 5. Thousand Oaks, CA: Sage Publications.

TABLE 1

Expert Qualifications

Participant	Qualifications (Condensed personal statements. The identities of the departments have been disguised to protect respondents' anonymity.)
1. Nurse manager 1:	I'm the Nurse Unit Manager of (X department). It's a 26-bed ward, which (has clinical and surgical components). I have a staffing of 32 EFT (*effective full time staff*). My role is to coordinate admissions and discharges and ensure that a safe level of care is delivered to all of those patients. I have been in charge of this ward 12 months. Before that I did surgery (in this clinical specialization) for 25 years, and was in charge of departments (within this specialization) for the last 13 years. Before my current assignment, I was at the University of Y lecturing at the School of Post Graduate Nursing.
2. Nurse manager 2:	I am the Nursing Manager of the (Y specialist sub-unit) and a 16-bed acute care unit, with a day procedure centre with a facility for six patients on a Monday to Friday basis. My role encompasses caring obviously for my staff, recruiting, retaining staff, control over a budget and a cost centre. Also overseeing things like education, certainly directives that come down from my Director of Nursing which have come from the Executive before that.
3. Physician specialist, sub-unit manager:	I'm Director of (Y specialist sub-unit) at the X medical center and a clinician with an interest in (specialization Y). All of my in-patient work is in the (specialization) service. I have administrative responsibilities for that service. I also have an outside practice. I have research interests in health services at the delivery of care and also in decision support and its relation to improving functional outcomes.
4. Director of surgery within department:	The work involved as Director is small; my main job is being employed as a surgeon. As director, I'm a liaison with the rest of the hospital and management, but management is a very small component of my role. Most of my time all year is spent doing surgery. I have been associated with this hospital in this role about six years.

TABLE 2

Interview Protocol

A. Record time, place, number and identity of interviewee (separate from taped interview)

B. Obtain signed consent form

C. Speak into recorder: This is the first interview number XXX

D. Ask the following questions in order:

1. Could you describe your current position and work at the hospital?
2. Does your unit have overall goals? Follow-up:
    a. Such as being the best or most efficient unit of its type?
    b. Such as fitting well into the overall network?
3. Can you describe what would be good performance outcomes for your unit?
4. If you could tell a story about how your unit can succeed, what would that story be like? Can you tell that story to us?
5. What factors are most important in determining whether your unit meets its goals, achieves good performance, or meets performance targets? Follow up, ask for examples:
    a. Investments in people? Improved personnel capabilities, competencies?
    b. Improvements in clinical technology or processes?
    c. Improvements in patient management and relations?
    d. Improvements in
        i. Clinical outcomes,
        ii. Efficiency
        iii. Control of costs
6. Are these factors linked in any way? Follow up:
    a. Does an improvement in one area lead to improvements in other areas?
    b. If so, what sort of time lags do you expect?
7. Do you think these factors can be measured?
8. If that were possible, would a combination of these measures help you manage your unit to reach its goals?

TABLE 3

Qualitative Interview Codes (*descriptions*) and Frequencies

	Interviews				
CODES	1	2	3	4	Total
1 Interviewee					
1.1 Physician	0	0	1	1	2
1.2 Nurse Head of Department	1	1	0	0	2
2 Learning and Growth					
2.1 Training (*formal training programs, incorporation of best practices*)	12	9	11	14	46
2.2 Empowerment (*influence, control, meeting personal vision*)	5	3	4	2	14
2.3 Recruitment (*efforts and success of hiring qualified personnel*)	7	1	6	3	17
2.4 Teamwork/Networking (*internal teamwork, interactions with peers outside the unit*)	25	18	18	6	67
2.5 Communication (*quality, frequency, effectiveness of communication within and outside the unit*)	7	7	2	0	16
2.6 Employee satisfaction (*job, professional satisfaction*)	13	3	4	2	22
2.7 Employee retention (*success in retaining qualified personnel*)	12	1	4	3	20
2.8 Performance reviews (*use of evaluations to motivate individuals, improve unit*)	3	0	3	2	8
2.9 Learning & Growth(L&G) constraints (*uncontrollable factors affecting L&G; e.g., nurse shortages*)	9	3	2	6	20
3 Processes					
3.1 Patient flow (*rate and smoothness of patient flow through unit*)	16	2	8	13	39
3.2 Patient care processes (*clinical processes affecting patient care*)	12	14	16	9	51
3.3 Process constraints (*uncontrollable factors affecting clinical processes; e.g., nursing home bed shortages affecting discharge rate*)	3	5	1	14	23
4 Technology (*information, medical technology used or desired*)	0	1	4	1	6
5 Clinical Trials (*participation in new drug or new process testing sponsored by external parties*)	0	0	6	0	6
6 Financial outcomes					
6.1 Internal funding of L&G, processes, technology (*funding from hospital*)	6	1	4	8	19
6.2 External funding of L&G, processes, technology (*funding from other providers*)	3	0	5	2	10
7 Process outcomes					
7.1 Clinical outcomes (*quality of medical outcomes*)	5	5	5	12	27
7.2 Patient satisfaction (*quality of care and outcomes from patients' view*)	8	3	4	2	17
7.4 Hospital-level financial results (*financial outcomes for the entire hospital*)	1	0	1	7	9
7.5 Department-level financial results (*financial outcomes for the unit*)	7	4	5	6	22
7.6 Department status (*level of prestige associated with the unit*)	0	0	4	4	8
Total code frequencies	155	81	118	117	471

FIGURE 1

Preliminary Conceptual Model

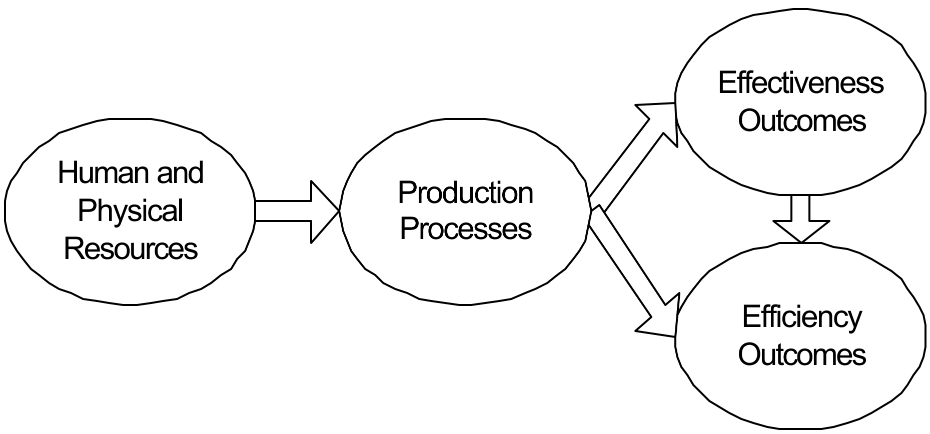


FIGURE 2

Software-Generated Knowledge Map

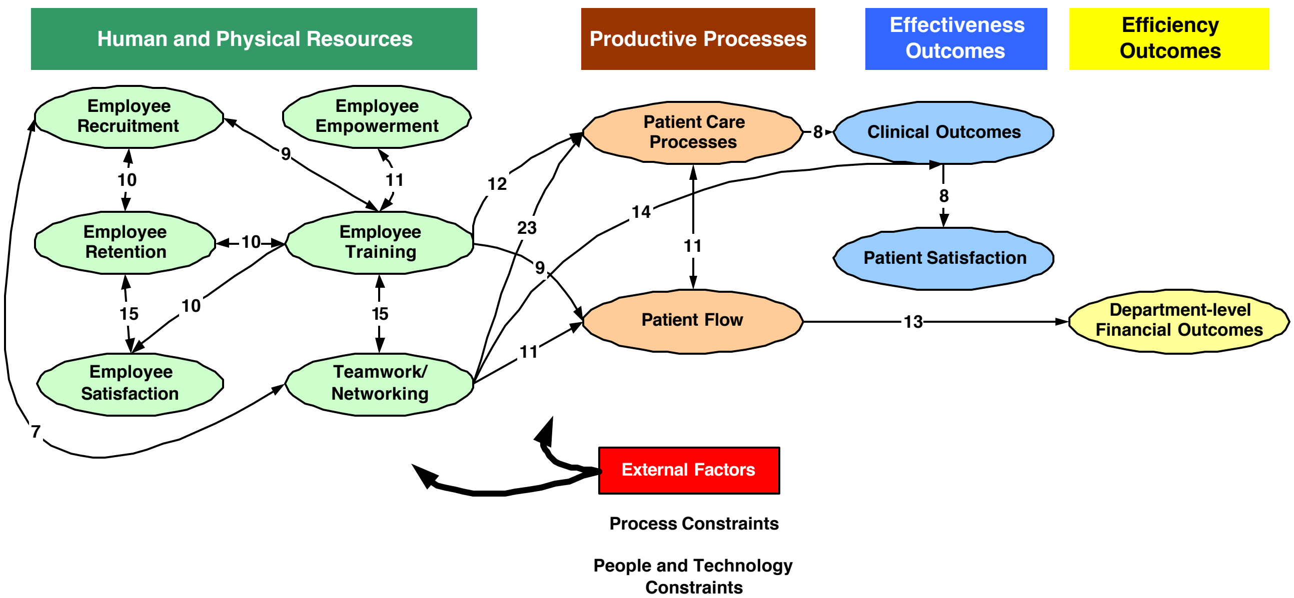


FIGURE 3

Researchers' Consensus Knowledge Map

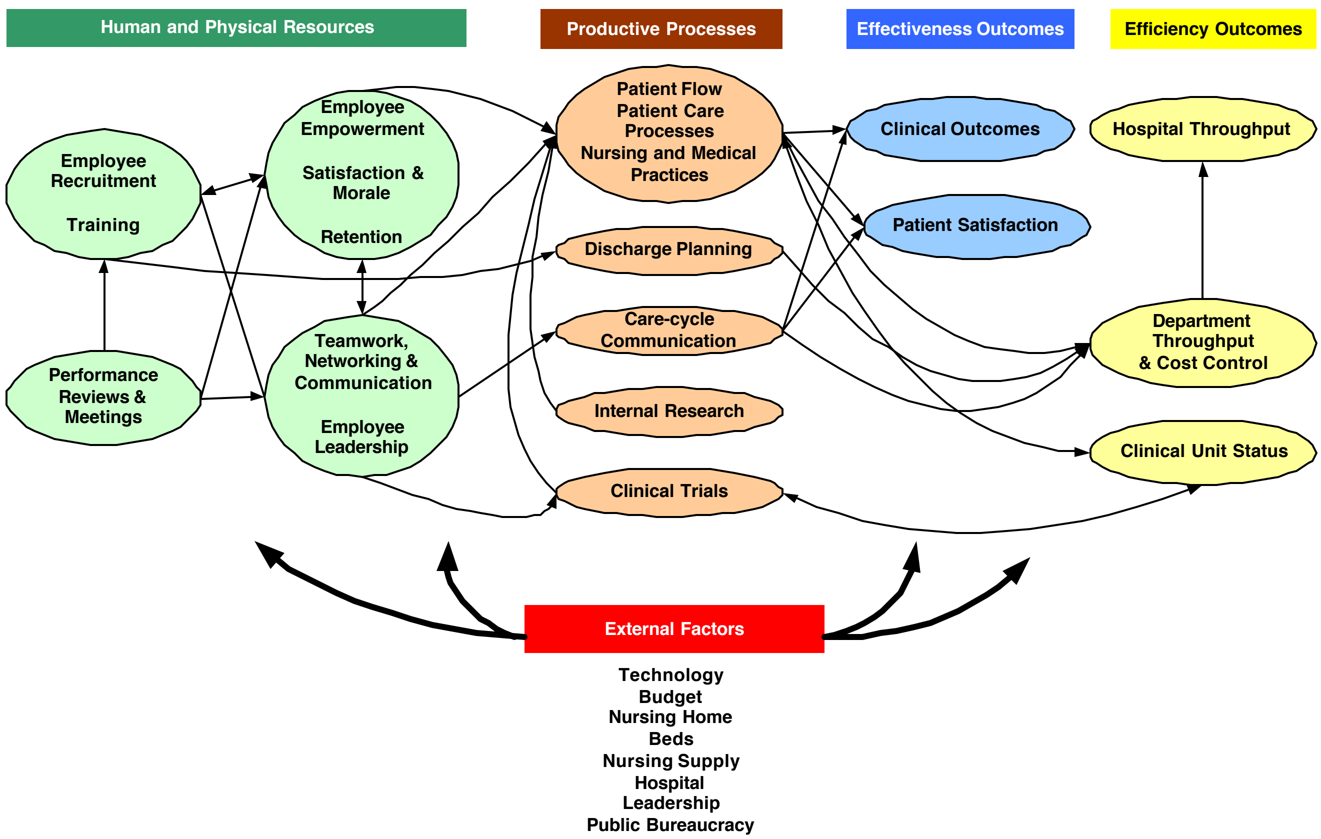


FIGURE 4

Experts' Composite Elicited Knowledge Map

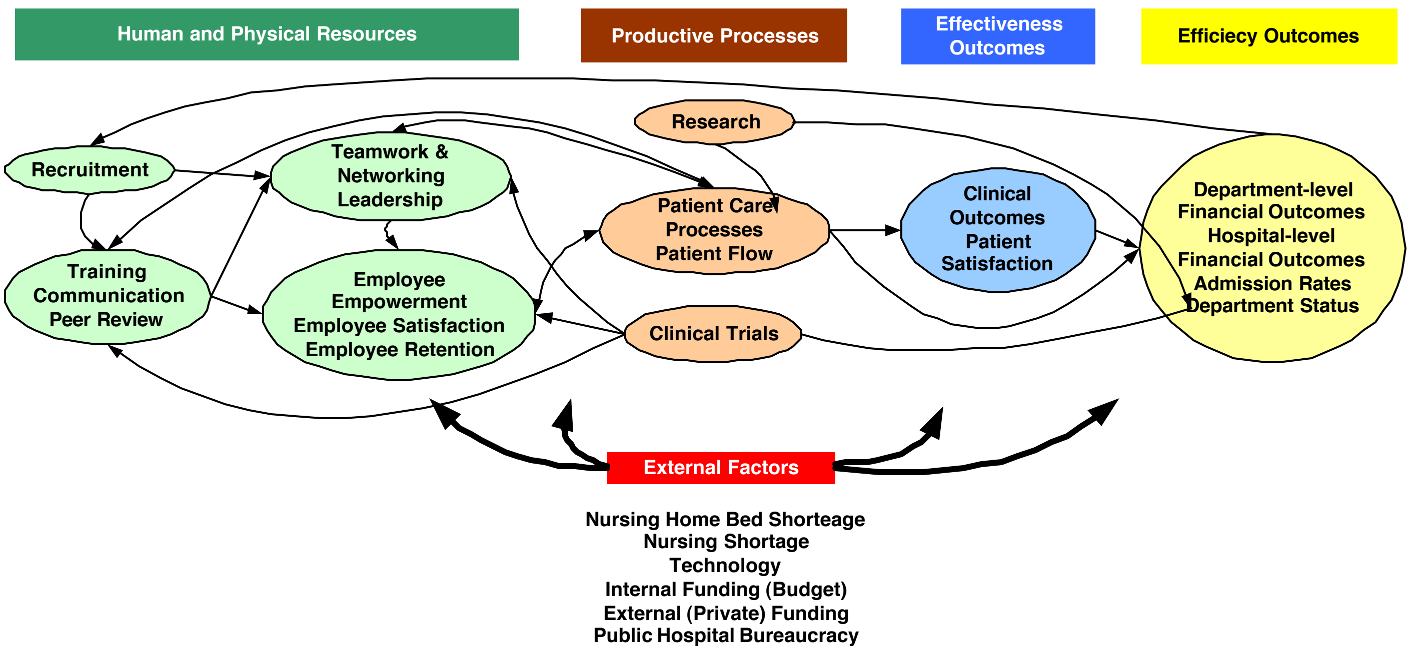


FIGURE 5

Triangulated, Final Conceptual Knowledge Map

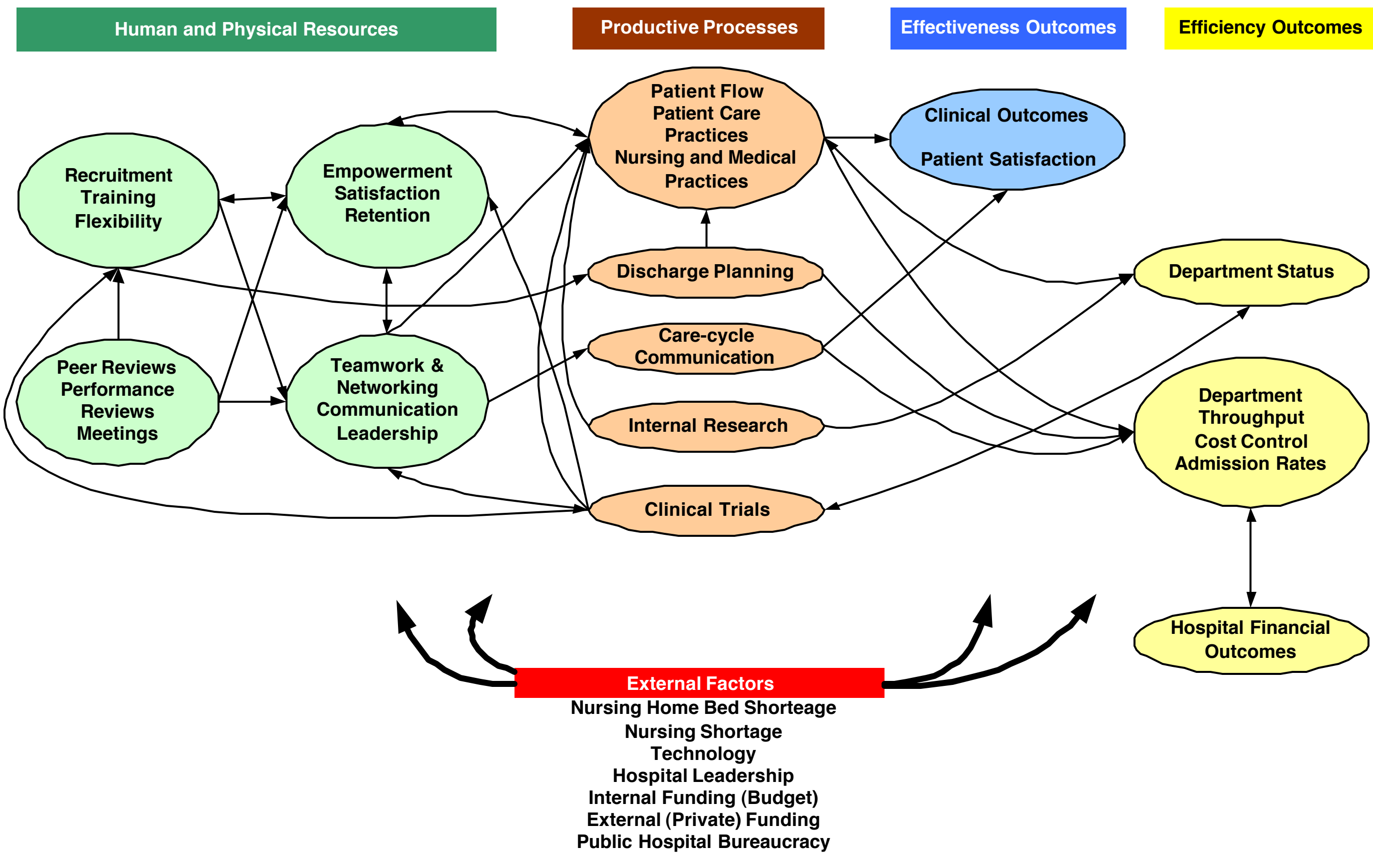


TABLE 1A

Currently Available Proxies for KM Constructs

KM Variable*	Archival Proxies	Explanation
Recruitment	None available	
Training	Training (EFT** and $)	Monthly allocations of nursing time and salaries to training activities
Flexibility	None available	
Peer reviews	None available	
Performance reviews	None available	
Meetings	None available	
Empowerment	None available	
Satisfaction	Sick leave (EFT and $), AWOP (EFT and $)	Monthly allocations of nursing time and salaries to sick leave and non-paid leave. Less satisfied nurses will tend to take more sick and unpaid leave.
Retention	Turnover (EFT) Temporary staff (EFT and $)	Monthly permanent turnover in nursing personnel (inverse of retention) Monthly temporary hires to fill nursing vacancies
Teamwork and networking	None available	
Communication	None available	
Leadership	None available	
Patient flow	Average length of stay	Average monthly cycle time per case (LOS)
Patient care practices	None available	
Nursing and medical practices	None available	
Discharge planning	None available	
Management talent	None available	
Care-cycle communication	None available	
Internal research	None available	
Clinical trials	None available	
Clinical outcomes	None available	
Patient satisfaction	None available	
Department status	None available	
Department throughput	N of cases, Total length of stay, WEIS+	Monthly number of cases in process, Total patient-days, and intensity of treatment given
Cost control	None available	
Admission rates	None available	
Hospital fin. outcomes	None available	
External funding	None available	
Technology	None available	
Internal funding	Budget $	Monthly budgets for all major department functions
Nursing supply	None available	
Nursing home beds	None available	
Hospital leadership	None available	
Public hospital bureaucracy	None available	

* Beginning at the top and left of figure 5.

** EFT = Equivalent full-time employees

+ WEIS is a measure of case load weighted by intensity of treatment

TABLE 2A

Currently Feasible KM Tests *

Dependent Variables	Independent Variables	Fit of Regressions, Variable Coefficients, and (Significance of Coefficients)
Training rate	Internal budget/100,000	$R^2 = 0.006$ - 0.000425 (0.662)
Sick leave rate	Training rate	$R^2 = 0.015$ 0.270 (0.493)
AWOP rate	Training rate	$R^2 = 0.001$ - 0.0369 (0.851)
Turnover rate	Training rate	$R^2 = 0.076$ 0.483 (0.126)
Temporary staffing rate	Training rate	$R^2 = 0.007$ -2.516 (0.633)
Average LOS	Sick leave rate AWOP rate+ Turnover rate** **Temporary staffing rate+**	Adj. $R^2 = 0.169$ - 3.677 (0.882) 19.621 (0.539) 73.739 (0.139) **3.224 (0.071)**
Average LOS (post hoc)	**AWOP rate+** **Temporary staffing rate+**	**Adj. $R^2 = 0.215$** **84.354 (0.055)** **3.508 (0.036)**
N of cases	**Average LOS+**	**$R^2 = 0.500$** **- 26.927 (0.000)**
Total LOS	Average LOS**	$R^2 = 0.002$ - 7.377 (0.817)
WEIS	Average LOS**	$R^2 = 0.004$ - 2.512 (0.725)

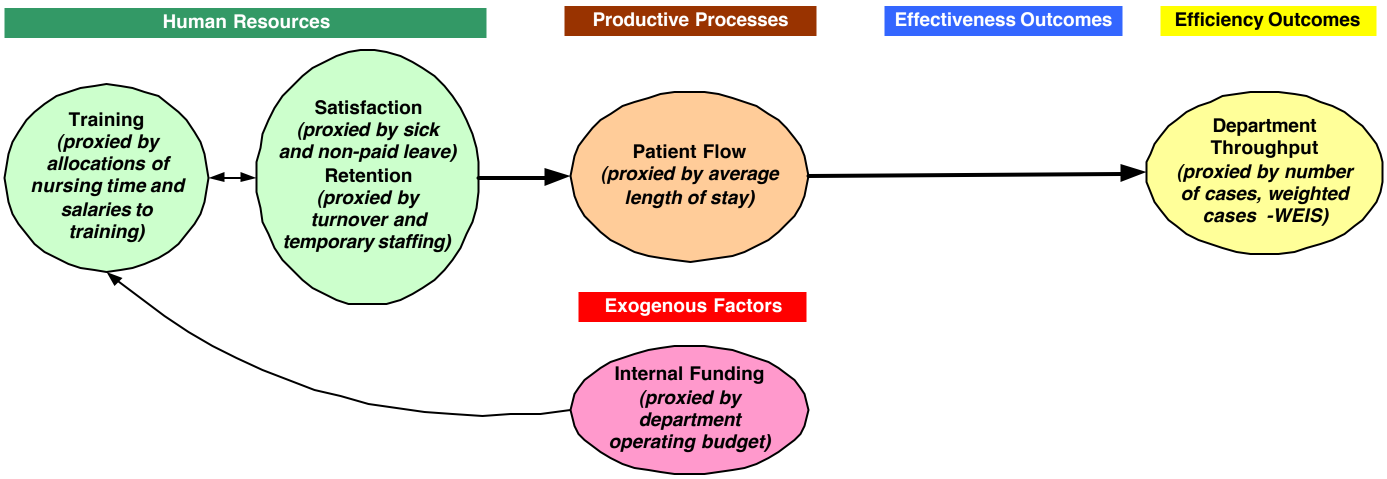
* Results shown are for EFT-based rate levels, which are indistinguishable from results using dollar-based rate levels.

+ Up to six-month lagged effects of independent variables also significant (a = 0.05) in predicted directions.

** First-difference (changes) models significant (a = 0.05) in predicted directions.

FIGURE 6

Currently Testable Model *



* Bold arrows signify statistically significant links.