



## **Building a Better Delivery System: A New Engineering/Health Care Partnership**

Proctor P. Reid, W. Dale Compton, Jerome H. Grossman, and Gary Fanjiang, Editors, Committee on Engineering and the Health Care System, National Academy of Engineering, Institute of Medicine

ISBN: 0-309-65406-8, 276 pages, 8 1/2 x 11, (2005)

**This PDF is available from the National Academies Press at:**  
<http://www.nap.edu/catalog/11378.html>

Visit the [National Academies Press](http://www.nap.edu) online, the authoritative source for all books from the [National Academy of Sciences](http://www.nap.edu), the [National Academy of Engineering](http://www.nap.edu), the [Institute of Medicine](http://www.nap.edu), and the [National Research Council](http://www.nap.edu):

- Download hundreds of free books in PDF
- Read thousands of books online for free
- Explore our innovative research tools – try the “[Research Dashboard](#)” now!
- [Sign up](#) to be notified when new books are published
- Purchase printed books and selected PDF files

**Thank you for downloading this PDF. If you have comments, questions or just want more information about the books published by the National Academies Press, you may contact our customer service department toll-free at 888-624-8373, [visit us online](#), or send an email to [feedback@nap.edu](mailto:feedback@nap.edu).**

**This book plus thousands more are available at <http://www.nap.edu>.**

Copyright © National Academy of Sciences. All rights reserved.

Unless otherwise indicated, all materials in this PDF File are copyrighted by the National Academy of Sciences. Distribution, posting, or copying is strictly prohibited without written permission of the National Academies Press. [Request reprint permission for this book.](#)

# **Engineering Tools and Procedures for Meeting the Challenges**



# Systems Engineering: Opportunities for Health Care

**Jennifer K. Ryan**  
*Purdue University*

Systems engineering involves the design, implementation, and control of interacting components or subsystems. A system consists of interacting, interrelated, or interdependent elements that form a complex whole, a set of interacting objects or people that behaves in ways individuals acting alone would not. The overall goal of systems engineering is to produce a system that meets the needs of all users or participants within the constraints that govern the system's operation. The objectives can generally be divided into two broad categories: service and cost. Service can be measured by a variety of criteria, such as availability, reliability, quality, and so on. Cost is usually measured by how much costs can be reduced or at least controlled.

A final objective of systems engineering is to gain a better understanding of system behavior and the problems associated with it. Models enable us to study the impact of alternative ways of running the system—alternative designs or controls and different configurations and management approaches. In short, systems engineering models enable us to experiment with systems in ways we cannot experiment with real systems.

Systems engineers generally prefer to work with analytical or mathematical models rather than with conceptual models because they are generally better defined, have more clearly defined assumptions, and are easier to communicate, manipulate, and analyze. We begin with a graphical representation of the system, which often includes a diagram showing the flow of information and resources. We then create a mathematical description that includes objectives, interrelationships, and constraints. The components of the mathematical model can be divided into four categories: (1) decision variables, which represent our options; (2) parameters or givens, which are the inputs to the decision-making process; (3) the objective function, which is the goal, the function to be optimized; and (4) the constraints, which are the rules that govern operation of the system.

When dealing with large complex systems, we often deconstruct it into smaller subsystems that interact with one another to create a whole. The decision-making structure provides natural breaks in the system. We model and analyze the subsystems and then connect them in a way that recaptures the most important interdependencies between them.

Systems engineering requires a variety of quantitative and qualitative tools for analyzing and interpreting system models. We use tools from psychology, computer science, operations research, management and economics, and mathematics. The quantitative tools include optimization methods, control theory, stochastic modeling and simulation, statistics, utility theory, decision analysis, and economics. Mathematical techniques have the capability of solving large-scale, complex problems optimally using computerized algorithms.

Mathematical models clarify the overall structure of a system and reveal important relationships. They enable us to analyze the system even when data are sparse. Models, combined with analyses, reveal the most critical parameters and enable us to analyze the system as a whole. Sensitivity analysis involves testing out trade-offs. Before we can convert a model solution to an implementable solution, we must test and validate the model to ensure that it actually predicts the behavior of the system.

A logistics system can be defined as a network of suppliers, manufacturing centers, warehouses, distribution centers, retail outlets, and end consumers. The system includes raw materials, work in process, inventory, finished products, all of the materials in the system, all of the information that flows within the system, and all of the resources in the system (e.g., people, equipment, etc.). Logistics-systems engineering can be defined as the planning, implementation, and control of the system to ensure the efficient, cost-effective flow and storage of all materials and information from point of origin to point of consumption for the purpose of meeting customer requirements. Our goal is to ensure that the right

amount of materials or resources is in the right place at the right time at minimum cost.

We deliberately leave the definition of service (i.e., meeting customer requirements) somewhat vague so we can define the needs and requirements of different customers in different ways. Logistics-systems engineering involves the difficult problem of simultaneously improving customer service and quality, improving timeliness, reducing operating expenses, and, if possible, minimizing capital investment. We are also interested in answering strategic questions, such as where we can expand capacity or what types of collaboration with customers or suppliers would be most beneficial.

Systems engineering problems have some common characteristics. They tend to be interdisciplinary, involving both technical and nontechnical fields. They require multiple, high-level, or strategic metrics or performance measures, often measurements of nonquantitative factors (e.g., customer satisfaction). They involve many participants with different value systems and many decision makers; therefore, we have to find optimal solutions that meet conflicting criteria. The systems and issues tend to be hierarchical and complex, but the systems also evolve and change over time; they generally involve significant uncertainties. Much of the current research in logistics is driven by the needs of public and private organizations, such as health care systems, that operate in environments characterized by intense competition, constant change, and a strong focus on customer needs.

Health care delivery systems, for example, consist of a variety of health care organizations, caregivers, and patients. State and federal governments are involved, as well as a variety of other organizations. These complex systems also involve a large number of interconnections between the components and the system—multihospital systems and provider networks with linkages between hospitals, physician groups, insurers, and others. There are also many decision makers who often have conflicting criteria, and there are complex

interactions between participants. The effective organization and management of a health care delivery system requires careful management of resources to ensure that the necessary staff and equipment are in the right place at the right time. The problem is complicated by uncertainties and system complexity.

Some aspects of the health care delivery system, such as government intervention, the level of uncertainty, and the nature of the demand, appear to be unique to health care. But similar problems can be found in other industries, such as the telecommunications and electricity industries, which also have to factor in government intervention. The nature of the uncertainties may be different, but they have similar effects on the system. Both the telecommunications and electricity industries have used logistics models to their advantage.

Systems engineering models can provide structured, quantitative methods of studying alternative control policies and system designs for almost any industry. The methods can be used to help coordinate information systems, operations, and capital investment; develop control policies; predict and evaluate outcomes; and evaluate the benefits and costs of a given program or system design.

The elements included in the model depend on the question or problem to be solved. For the output of the model to be useful, it must mimic the expected behavior of the real system. To control the behavior of one part of the system, the incentives driving that aspect of the system must be built into the model.

A good deal of literature is now available on research in this area. Operations research tools and systems engineering tools have been used to address a wide variety of problems, from the operation of a hospital to higher levels of complexity, such as incentives, efficiency, and payment schemes. Quantitative models can provide important input for making decisions that involve complex societal, ethical, and economic issues.

# Supply-Chain Management and Health Care Delivery: Pursuing a System-Level Understanding

Reha Uzsoy  
*Purdue University*

In recent years, effective supply-chain management has emerged as a significant competitive advantage for companies in very different industries (e.g., Chopra and Meindl, 2000). Several leading companies, such as WalMart and Dell Computer, are differentiated from their rivals more by the way they manage their supply chains than by the particular products or services they provide. A supply chain can be defined as the physical and informational resources required to deliver a good or service to the final consumer. In the broadest sense, a supply chain includes all activities related to manufacturing, the extraction of raw materials, processing, storing and warehousing, and transportation. Hence, for large multinational companies that manufacture complex products, such as automobiles, machines, or personal computers, supply chains are highly complex socioeconomic systems.

The ability of successful firms to make the effective management of supply chains a source of competitive advantage suggests that there may be useful knowledge that can provide a point of departure for the development of a similar level of understanding of certain aspects of health care delivery systems. Similar to the supply chains in manufacturing and other industries, the health care delivery system is so large and complex that it has become impossible for any individual, or even any single organization, to understand all of the details of its operations. Like industrial supply chains, the health care “supply chain” consists of multiple independent agents, such as insurance companies, hospitals, doctors, employers, and regulatory agencies, whose economic structures, and hence objectives, differ and in many cases conflict with each other. Both supply and demand for services are uncertain in different ways, making it very difficult to match supply to demand. This task is complicated because demand for services is determined by both available technology (i.e., available treatments) and financial considerations, such as whether or not certain treatments are covered by insurance. Decisions made by one party often affect

the options available to other parties, as well as the costs of these options, in ways that are not well understood. However, almost all of these complicating factors are also present, to one degree or another, in industrial supply chains; the progress made in understanding these systems in the last several decades is a cause for hope that some insights and modeling tools developed in the industrial domain can be applied to at least some aspects of health care delivery systems.

In general, a centralized approach to controlling the entire system is clearly out of the question, although centralized decision models may be useful for coordinating the operations of segments of the larger system controlled by a single decision-making body. Designing decentralized models of operation that render the operation of the overall system as effective as possible is the main challenge for both health care delivery and industrial supply chains.

In the following section, I shall briefly discuss how the study of industrial systems has evolved from individual unit processes to considerations of complex interactions among many different components of an industrial supply chain. I shall then describe some examples of modeling approaches that have been applied to supply chains and close with some comments on how these tools might be adapted for the health care delivery environment.

## FROM UNIT PROCESSES TO SUPPLY CHAINS

If we examine how industrial operations, particularly manufacturing operations, have evolved since the beginning of the nineteenth century, we can see that many efforts were motivated by a desire to understand and optimize individual unit processes (see, for example, Chandler, 1980). These efforts led to many innovations, among them the development of improved machine tools and fixtures, a significantly better understanding of the chemistry of processes (e.g., steel-making), and through the work of the early industrial

engineers, such as Frederick Taylor and Frank and Lillian Gilbreth, the optimization of interactions between workers and their environment.

As the understanding of unit processes developed, engineers began to consider larger and larger groupings of unit processes, trying to understand interactions between them and optimize the performance of entire systems, sometimes to the detriment of individual components. Hence, from considering individual unit processes, we progressed to considering departments of factories that perform similar operations, entire manufacturing processes from raw materials to finished products, and eventually, the operations of entire firms, as well as their suppliers and customers. It has often been observed that most significant new opportunities, both for cost reduction and the generation of new products and services, have been based on an understanding of interactions between different subsystems, or different agents, operating in the supply chain.

Among today's leading companies, examples abound. Many automotive companies, for instance, have developed joint ventures with transportation firms; the objective is to optimize the interface between the production and distribution functions and facilitate the just-in-time operation of automakers' final assembly plants. Software companies that provide supply-chain planning software for multilocation companies is another strong indicator of the advantages companies perceive will accrue to them by the effective management of the various elements of their supply chains. The strong trend in industry to outsource noncritical functions has increased the need for companies to effectively manage and clearly understand their relationships with other companies. As a final example, we can point to the collaborative forecasting, planning, and replenishment initiative in the retail sector; retailers work closely with major suppliers to develop demand forecasts for products through information-sharing and joint planning processes.

Clearly, the basic process of improving a system by a detailed understanding of the most fundamental unit processes, in other words the "atomic" elements of the system, and steadily extending that knowledge to interactions among larger and larger groupings of these elements is directly applicable to health care delivery systems. The individual unit processes in this case include the processing of a patient in an emergency room, the process by which a medical insurance claim is approved, and the scheduling of hospital operating rooms to optimize their performance. The need for a better understanding of how the operations of individual elements affect each other is apparent; these interactions can be quite complex because of long time lags between cause and effect. For example, the decision by a regulatory agency to disallow a certain kind of preventive procedure for infants may result in the emergence of an unexpectedly large number of children with special needs in the elementary school system several years later. The same kinds of problems are present to some degree in industrial supply chains, and a

significant body of knowledge has been developed over the years to address them.

Based on the history of industrial enterprises, we know that the development of today's enterprises required substantial organizational innovations, such as capital budgeting to allocate scarce capital between competing activities, cost accounting to develop an understanding of factors contributing to product costs, and the development of multidivisional corporations with complex structures of management incentives and coordination mechanisms. An important development in recent years has been the recognition of the need for a cross-functional view of supply-chain operations. All aspects of a firm's operation, from the design of a product to the specific timing of marketing promotions, have a direct effect on the operation of the supply chain. Therefore, different functional specialties must actively collaborate to develop solutions to optimize the performance of the overall system. Similarly, in health care delivery a number of different constituencies, such as doctors, government agencies, insurance providers, and patient groups, are all involved in the operation of the health care delivery supply chain.

## KNOWLEDGE OF SUPPLY-CHAIN MANAGEMENT

In the domain of industrial supply chains, it is probably safe to say that we have developed a fairly good understanding of the operation and economics of individual unit processes, including functions such as transportation, distribution, warehousing, and information processing. In particular, we have developed a substantial understanding of the often complex dynamics of capacity-constrained systems subject to variability in both demand and process (Hopp and Spearman, 2000). However, in general we are only beginning to learn how to integrate the solutions to these individual elements to reach a reasonable understanding of the operation of the overall supply chain.

Integrated planning models based on linear and integer programming have been applied to the segments of the supply chain controlled by a single company for at least four decades (e.g., Johnson and Montgomery, 1974). Although these models have been successful in many instances, they have not been effective in addressing the needs of a supply chain that involves many different companies with potentially conflicting objectives. In recent years, considerable efforts have been made to use some of the tools of economics, such as contracts, as a mechanism for coordinating the operation of complex supply chains (Tayur et al., 1998). However, these models are generally subject to long-run, steady-state assumptions that can be carefully evaluated relative to market conditions.

Conventional Monte Carlo simulation techniques (Law and Kelton, 1991) have proven extremely effective for systems in which the operational dynamics can be described at a high level of detail, such as segments of manufacturing processes or hospital operations. The difficulty with these

models is that for large-scale systems the level of detail required to unequivocally model system behavior accurately becomes prohibitive in terms of both data collection and computation time. Systems dynamics models used to model large systems work by establishing input-output relationships for their components and simulating their operation through time using techniques based on the techniques used for the numerical solution of differential equations (Sterman, 2000). Although these techniques are capable of modeling large, complex systems, they usually do so by specifying aggregate input-output relationships for large subsystems, which must be validated and whose parameters must be estimated carefully. Nevertheless, these models can capture many critical aspects of supply-chain behavior, such as the “bullwhip effect,” in which variability in orders is amplified as it passes down the supply chain from the consumer towards the producers of raw materials (Forrester, 1962).

## RESEARCH NEEDS AND FUTURE DIRECTIONS

At the risk of overgeneralizing, it appears that most of the tools required for analysis of the individual unit processes in health care delivery, such as efficiency of hospital facilities, have been developed in the engineering literature and have, in fact, been applied intermittently to a variety of systems over the last several decades (e.g., Pierskalla and Brailer, 1994). However, if our experience with industrial supply chains is any guide, only limited improvements in health care delivery can be obtained by these means. Repeated experience has shown that far greater improvements can be obtained by a thorough understanding of the interactions between different elements of the system and restructuring them in a way that leaves all parties better off. This brings the modeling issues squarely into the region where current supply-chain research is weakest (the effective coordination of socioeconomic systems consisting of multiple, independent agents); but this is also the area that is developing most rapidly. The development of novel models at the intersection of conventional engineering and economics promises to provide a wide range of challenging research problems for many years to come.

To support this agenda, the most pressing research need is for techniques that can be used to model systems at the

aggregate level, where one can accept some level of approximation to obtain computationally tractable models that achieve the correct qualitative behavior and provide useful insights into interactions between systems. This means that the aggregate models must capture the often nonlinear relationships between critical variables correctly, which has not always been the case in supply-chain modeling. The literature on systems dynamics may be a good starting point for this initiative, but it must be complemented by a variety of other techniques, such as economic models of competition and collaboration and agent-based techniques for modeling complex systems.

It is important to bear in mind that the purpose of these models is far more likely to be descriptive than prescriptive, that is, models are far more likely to be used, and arguably far more useful, to inform debate between the various parties involved in health care delivery than to deliver decisions to be executed. Hence, the development of large-scale computational simulations of different scenarios with different actors and interaction protocols between the actors appears to offer interesting research challenges. These tools would be extremely beneficial to decision makers in health care delivery.

## REFERENCES

- Chandler, A.D. 1980. *The Visible Hand: The Managerial Revolution in American Business*. Cambridge, Mass.: Belknap Press.
- Chopra, S., and P. Meindl. 2000. *Supply Chain Management: Strategy, Planning and Operations*. Englewood Cliffs, N.J.: Prentice-Hall.
- Forrester, J.W. 1962. *Industrial Dynamics*. Cambridge, Mass.: MIT Press.
- Hopp, W., and M.L. Spearman. 2000. *Factory Physics*. 2nd Edition. New York: McGraw-Hill/Irwin.
- Johnson, L.A., and D.C. Montgomery. 1974. *Operations Research in Production Planning, Scheduling and Inventory Control*. New York: John Wiley & Sons.
- Law, A., and W.D. Kelton. 1991. *Simulation Modeling and Analysis*, 2nd edition. New York: McGraw-Hill.
- Pierskalla, W.P., and D.J. Brailer. 1994. Applications of Operations Research in Health Care Delivery. Vol. 6, pp. 469–505 in *Handbooks in OR & MS*, S.M. Pollock, M.H. Rothkopf, and A. Barnett, eds. Amsterdam, The Netherlands: Elsevier Science.
- Sterman, J.D. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. New York: McGraw-Hill.
- Tayur, S., M. Magazine, and R. Ganesham, eds. 1998. *Quantitative Models for Supply Chain Management*. Amsterdam, The Netherlands: Kluwer Academic Publishers.





# The Human Factor in Health Care Systems Design

Kim J. Vicente  
*University of Toronto*

The simplest way to think about the discipline of engineering is that engineers design things that are useful to society and satisfy important needs based on what we know about the physical world. When a bridge fails, we do not usually blame the bridge. We look to its design, trying to find a mismatch between what we know about the physical world and the outcome.

We should apply this same logic to people. But when a system is poorly designed, we often blame the person using it rather than the flaws in the system. For example, when we design a mechanical lathe, we must place the mechanical controls in a way that respects what we know about human bodies. But sometimes, if a lathe is poorly designed, we blame the user rather than the design.

Although we know a great deal about teamwork and about human behavior at the organizational and political levels, that knowledge is not always taken into account by designers of health care systems and devices. Clearly, improvements could be made, and not just in terms of safety. The lack of respect for human nature in the design of health care systems causes injuries and deaths, but it also costs money.

Contrast that to the field of aviation. Despite September 11, 2001 was not a bad year for aviation safety. The average number of deadly crashes for the previous decade was 48 per year. In 2001, however, there were only 34 deadly crashes—worldwide, not just in the United States. That's the lowest number since 1946 when there were far fewer flights.

One reason for the improvement is that aviation engineers pay attention to the human factor. A familiar example is the rather high rate of crashes in a certain type of aircraft that occurred because pilots tended to raise the landing gear as the plane was landing, causing the airplane to scrape along the runway. When Al Chapanis, an aviation engineer, studied the problem, he found that the controls for the landing gear and the wing flaps were right next to each other and that they looked and felt identical. He realized that pilots could easily grab the wrong control, but he also realized that he

could not redesign the whole cockpit. He came up with an idea, now called shape coding. He did not move the controls, but he altered the feel of the landing gear control. The controls are still right next to each other, but the change eliminated the errors. It was as simple as that.

Can we apply the same type of thinking to health care systems? Patient-controlled analgesic devices, which allow patients to self-administer analgesics (usually morphine), are a case in point. A number of parameters are programmed into these devices by the nurse, the most important being drug concentration. These devices rely strictly on the programming and cannot independently verify either the concentration or even the type of analgesic in the syringe. Therefore, errors in programming can mean underdoses or overdoses; and errors have enduring effects, that is, the problem lasts until the programming is corrected.

For the particular device that we studied, programming errors were associated with five to eight *reported* patient deaths. Adverse drug events and adverse events in general in medicine are severely underreported—roughly only 1.2 to 7.7 percent are reported (Vicente et al., 2003). In other words, adverse events may be 13 to 83 times higher than the reported rate. We calculated that programming errors had lethal results for this particular device at least 65 times, and perhaps as many as 667 times, over a 12-year period. To put these numbers in context, the manufacturer reports that the device was used safely over 22 million times.

We then examined the existing design using traditional human-factor principles to see if there was room for improvement. We also talked to nurses, the users of this device. One serious problem we found was that the layout of the buttons on the interface was confusing and counterintuitive. So we came up with a new design by resegmenting the buttons and changing some of the labels. The new design offered the same functionality but changed the mode of interaction between the programmer and the pump. The system now provided more feedback and gave the user an overview

of the programming sequence. The redesigned device told the programmer the drug concentration, what was coming up next, how to program the mode, and then showed the settings. In essence, the new programming sequence was much less convoluted.

We tested the redesigned interface in a laboratory setting with professional nurses who had more than five years of experience programming the commercial device. With the commercially available design, there were eight programming errors for drug concentration, three of which were undetected. With the new interface, there were no errors in drug concentration. They were eliminated.

Given the epidemiological data, the change was obviously important for safety reasons. But it was also important in terms of cultural attitudes. If the problem had originated with the person programming the device, then changing the interface should have made no difference in the error rate. In fact, changing the design did eliminate the errors. Therefore, we concluded that the problem was not with the people, or, at least, not only with people.

Surprisingly, we had a great deal of difficulty getting this research published. One journal refused it because the editor took for granted that what we had scientifically demonstrated was not true. We went through some pretty hard times, both in terms of getting the work published and dealing with the

response from the public. One reviewer even suggested that a lawyer look at the research because of potential legal action by the manufacturer. We had chosen the particular device because it was relatively new, but soon after our research was completed, the media began to report some deaths as a result of errors in programming the device.

This example shows three important points. First, we know how to design technology that works for people because we know a lot about people at many levels—physical, psychological, team, organizational, and political. We do not always make the most of this knowledge when we design health care devices, but lack of understanding is not the problem. Second, not making the most of that knowledge results in a tremendous loss to society. Tens of thousands, perhaps even hundreds of thousands, of people are injured or die every year unnecessarily. Finally—the most difficult lesson—change is important and necessary, but there is a great deal of resistance that must be overcome before we can make progress.

## REFERENCE

- Vicente, K.J., K. Kada-Bekhaled, G. Hillel, A. Cassano, and B.A. Orser. 2003. Programming errors contribute to death from patient-controlled analgesia: case report and estimate of probability. *Canadian Journal of Anesthesia* 50(4): 328–332.

# Changing Health Care Delivery Enterprises

Seth Bonder  
*The Bonder Group*

The health care delivery (HCD) system in the United States is in crisis. Access is limited, costs are high and increasing at an unacceptable rate, and concerns are growing about the quality of service. Many, including the Institute of Medicine, believe the system should be changed significantly in two ways: (1) HCD enterprises should be reengineered to make them more productive, efficient, and effective; and (2) substantially more effort should be devoted to a strategy of prevention and management of chronic diseases instead of the current heavy reliance on the treatment of diseases. Although operations research can make substantial contributions to both areas, the focus of this paper is on: (1) reengineering HCD enterprises, particularly areas in which operations research can provide valuable support to senior health care managers; and (2) enterprise-level HCD simulation models to determine the reengineering initiatives with the biggest payoffs *before implementation*.

HCD enterprises are very large, complex operational systems comprised of large numbers of people and machine elements. Tens of thousands of people are involved as providers, patients, support staff, and managers organized into specialties, departments, laboratories, and other organizations that are considered independent service units (“stovepipes”). Machines include durable medical equipment, information technologies, communications equipment, expendable supplies, rehabilitation equipment, and so on. These elements are affected by many clinical and administrative processes (e.g., arrivals, testing, diagnosis, treatment, scheduling, purchasing, billing, recruiting, etc.), most of which are probabilistic (i.e., uncertain) and change significantly over time.

Perhaps most important, these processes involve large numbers of *interactions* within units, among units, and across processes. Decisions by enterprise managers regarding one unit may have second, third, and fourth order effects, which may be more significant than the first order effect. HCD enterprises are driven by endogenous and exogenous human decisions made by providers, patients, insurers,

administrators, politicians, government employees, and others. Demand and supply issues have complex feedback effects. A great many resources are required for the development and operation of an HCD enterprise. For example, the University of Michigan’s budget for its HCD enterprise is more than \$1 billion; the Henry Ford Health System’s budget is \$2.5 billion, and these are relatively small HCD enterprises. Billions of dollars have been spent on cost containment initiatives over the past 15 years by the Agency for Healthcare Research and Quality (formerly the Agency for Health Care Policy and Research), the U.S. Department of Defense, the Veterans Administration, National Institutes of Health, foundations, universities, and others to reengineer the HCD system. Nevertheless, costs continue to rise at double-digit rates.

We need better ways of analyzing systems of this magnitude. The operations research community has been involved with HCD enterprises for more than 40 years working on a wide range of problems, such as inventory for perishables; management of intensive care units; laboratory and radiology scheduling; relieving congestion in outpatient clinics; nurse staffing, scheduling, and assignments; and layouts for operating and emergency rooms. These efforts have focused on the small, stovepipe units, referred to by Don Berwick as clinical and support “microsystems,” and have produced some useful information for unit managers but have not addressed enterprise-level reengineering and planning issues (the so-called “macrosystem”). Macrosystem issues have interactive effects across the enterprise and have large cost, access, and effectiveness impacts. Some of these interrelated issues are listed below:

- the mix of health *services* necessary to support a given population
- the *staff* required (e.g., specialties, numbers, locations) to provide necessary services
- the impacts of changing demands (e.g., aging populations, effects of preventive measures)

- the impacts of new HCD models (e.g., home health care, task performance substitution)
- the effects of centralized radiology services
- the impacts of primary care outreach
- facility capacity for the next 20 years and the best way to provide it
- operational changes to adapt to regulatory changes (e.g., Medicare)

These and other macrosystem issues can be addressed quantitatively using enterprise-level simulation models that represent all of the elements, units, and processes in the enterprise *as well as the interactions among them*. Because analyses of these issues are necessarily prospective, the models must be *structural* rather than *statistical*. Statistical models, which are usually used in economics and the social sciences, use existing system data to develop aggregated statistical relationships between system inputs and outputs (i.e., the model). Statistical models are used primarily retrospectively, that is, for making *inferences* and *evaluations*. In contrast, structural models are usually developed in the engineering and physical sciences by modeling the detailed physics of each process and activity. Structural models are used prospectively, that is, for *predictions* and *planning*. Statistical models are less appropriate to prospective analyses of future systems because the data used to develop statistical models are intrinsically tied to the existing system.

Figure 1 provides an overview of a particular enterprise-level HCD simulation model. The figure shows the elements in the Healthcare Complex Model (HCM), which was developed seven years ago and has been continually updated in a prototyping process by Vector Research Incorporated (now the Altarum Institute). HCM simulates individual patient episodes in a network of facilities for a population of patients. The network of facilities, with its entities and processes, is referred to as a “complex” (synonymous with an

enterprise). Complexes usually have one or two major medical centers (where much of the tertiary care is provided), five to ten hospitals, and many clinics. The model can be adapted to represent specific features of any HCD enterprise.

Inputs to the model include demographics of the population that receives care. A model preprocessor converts the demographics into a stream of patients entering the complex; each patient’s condition is described by an International Classification of Diseases, ninth edition (ICD-9) code. Patients can enter the enterprise at a clinic, a hospital, or a medical center. They can be referred physically or via telemedicine consults from clinics to hospitals or to a medical center. Providers of various types are located at each facility in the complex. The *care protocols* represent practice guidelines and patient pathways, define what service patients receive next, where patients receive the service, and the type of personnel who will provide it. The model keeps track of the resources used and estimates costs using related cost models. Each protocol is a tree with many probabilistic branches to simulate that different providers may provide patients having the same condition with different medical services. The care protocols may be tailored for simulations of specific enterprises and facilities. The model represents various ancillary personnel (e.g., nurses, nurse assistants, medical technicians, etc.) and various ancillary resources (e.g., laboratories, pharmacies, beds, CAT scans, MRIs, and durable medical equipment). Finally, the model represents various clinical (e.g., computerized patient record system) and administrative (e.g., billing, scheduling) information technologies and communications systems.

Because the HCM explicitly simulates all of the entities, processes, and activities in the system, any one or combination of them can be changed, and the impact on various output costs and access metrics can be observed. For example, HCM can determine how a change affects the cost of running the enterprise, a hospital, or a particular unit in a

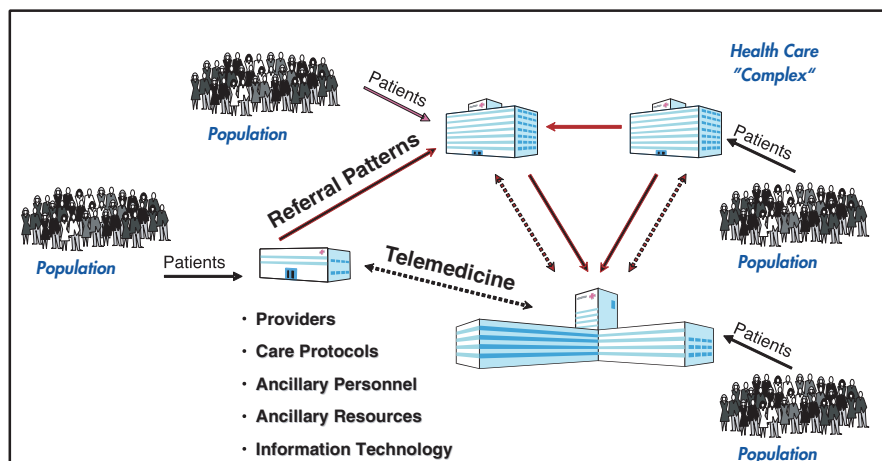


FIGURE 1 Overview of the Healthcare Complex Model.

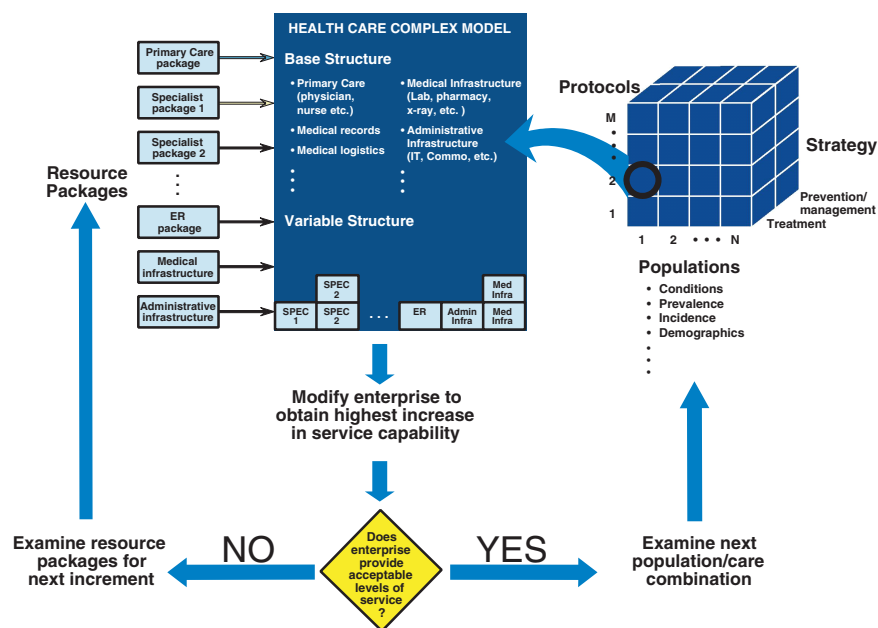


FIGURE 2 Zero-based HCD enterprise design.

hospital. It can calculate the impact on access metrics for the enterprise, a hospital, or a unit in a hospital. Because the model is being enhanced continually via a prototyping process, consideration has been given to simulating false positive and false negative statistical errors and their effects. Although these are not outcomes, they would provide useful quality information about the simulated HCD enterprise.

HCM has reasonable fidelity at this stage in its development. It contains more than 1,200 ICD-9 code conditions (e.g., acute appendicitis, asthma, cellulite, open chest wound, viral hepatitis, low back pain, etc.) and more than 1,500 clinical tasks/procedures (e.g., preoperative anesthesia, computer tomography for staging/radiation, EEG, interpretation of angiogram, administration of antibiotics, etc.). The model simulates 60 different kinds of health care providers, 17 types of ancillary resources (e.g., x-ray, ultrasound, pathology, dialysis unit, etc.), 6 different inpatient beds, and 23 combinations of telemedicine equipment. And its fidelity improves with every study.

The model was tested on one of the smaller regional HCD enterprises in the military health system (MHS). The enterprise has one major medical center, two hospitals, two clinics, and a managed care support contractor that provides additional capacity for the region. Together they handled about 1.6 million outpatient visits in fiscal year 1999. The model was adapted to represent the facilities, workforce, ancillary resources, information technologies, and clinical protocols used by the regional complex. Using population demographics provided by the government, regional operations for the year 1999 were simulated a number of times (because of the probabilistic nature of the protocols) to develop stable average outputs. These were compared to the historical values

from the enterprise's 1999 operations with encouraging results. Total outpatient visits differed by 0.11 percent, same-day surgeries by 1.02 percent, inpatient admissions by 2.99 percent, emergency room visits by 6.04 percent, and average length of stay by 0.94 percent. More detailed comparisons of outpatient visits by individual facility and individual specialty all differed by less than 4 percent. Although this was not a true validation study (which would require implementing model-suggested changes and comparing predicted impacts with actual results after the changes), it did show that simulation models can represent the complex dynamics of health care enterprise operations and can generate useful information and insights for enterprise managers.

HCM has been used in a number of other studies including the geographic distribution of primary care providers for a large, dispersed enterprise; telemedicine needs for a MHS regional complex; centralization of radiologists to service a 20-facility enterprise; and determining return-on-investment for information technologies. HCM is currently being used to determine capacity requirements for an enterprise that would experience increased demand following a bioterrorist attack.

Enterprise-level simulation models like HCM can be used to address a broad range of issues facing enterprise executives. Here is one challenge that could be posed: Given a population of patients, how can operations research determine an efficient set of resources to provide an acceptable level of services to that population. Assuming the HCD enterprise is a shell with no existing medical services, models like HCM can be used to address difficult issues, such as designing a system from scratch to serve a given population (sometimes referred to as "zero-based" design). A schematic drawing of the analysis process is shown in Figure 2. For purposes of this

discussion, we assume that an acceptable level of service can be defined in terms of some access/quality metrics, cost of enterprise operations, and cost of the resources.

The resources required to service the specified population depend not only on characteristics of the population (e.g., conditions, prevalence, incidence, etc.), but also on the protocols, as well as the degree to which the enterprise strategy for servicing the population focuses on treatment or prevention/management of medical conditions. The three-dimensional structure shown on the right side of Figure 2 allows the analysis team to select a population, a protocol set, and a mixed treatment/prevention strategy as input to the analysis process. (The protocols are obviously related to the strategy and designed to reflect the strategy.) Figure 2 shows that input (1, 2, T), representing population 1, protocol set 2, and a treatment-focused strategy is used to begin the analysis.

Regardless of the input set, the enterprise will need a “base structure” consisting of a primary care package, medical records, medical logistics, a medical infrastructure package, and other base resources, as shown in the figure. Enterprise operations with the base-level resources, protocol set 2, and strategy T can then be simulated for a period of time to see if it provides an acceptable level of service to the selected population (#1). If the answer is no (as shown by the decision diamond), the analysis can then try adding individual resource packages to see which provides the most improvement in service capability to the population. Resource packages are designed by the user team (e.g., pediatrician/internist/obstetrician/ENT package, which can be substituted for a primary care package; a gastroenterologist/orthopedist package; an oncologist/urologist package; a cardiologist/thoracic surgeon package; an emergency room package; and other resource groupings). Enterprise operations are simulated for each package to determine the improvement in service capability above the base level. The resource package with the most improvement on the margin is added to the enterprise (as shown under the variable structure).

This process is repeated, and resource packages with the most marginal improvement to the enterprise are added until an acceptable level of service is reached. (Mathematical programming techniques would likely make this iterative search process more efficient.) When this process is complete, the sum of the base and variable resources constitute an *efficient*

*set of resources* that provide an *acceptable level of service* (measured by access/quality and cost metrics) to the *designated population* using the *specified protocols*. The effect of different protocols on the resource requirements, as well as resource requirements for other populations, can be determined in a similar way. This process could be used to design a “versatile” set of resources that would provide a capability to serve multiple populations using different protocols.

Operations research could address some of the important enterprise-level issues but would require cultural changes on the part of enterprise management, as well as the operations research community. Enterprise management would have to encourage centralized planning for enterprise design and resource allocation issues, simultaneously maintaining decentralized operations. Higher order (and usually large) effects of interactions across stovepipes can only be identified at this level. Enterprise management would have to encourage a culture of prospective analyses to identify necessary changes that would be useful and would provide a high return on investment. (Retrospective analysis is an expensive trial-and-error process to learn what doesn't work). Enterprise management would have to establish a “requirements-pull” process for equipment and IT decisions, rather than the existing “technology-push” process, which is based on what is available from industry rather than what is needed. Management would also have to require that processes be reengineered when implementing new technologies (technology changes overlaid on existing processes produce zero value).

The health operations research community would also have to make important cultural changes. It needs to begin addressing enterprise-level issues, which should not remain in the purview of health econometricians who have failed to solve the cost, access, and quality problems that have beleaguered health enterprises and the nation. The operations research community would have to start working with enterprise-level structural models and begin using them for prospective analyses. Health operations research practitioners must become integral partners with senior enterprise managers in their business planning. They should use their 40 years of tactical-level support as an *entr e* and then demonstrate (and market!) the value of enterprise-level analyses to enterprise managers.

# Transforming Current Hospital Design: Engineering Concepts Applied to the Patient Care Team and Hospital Design

Ann Hendrich

*Ascension Health National Clinical Excellence Operations Office*

Health insurance premiums and the cost of hospital services and care have risen significantly over the past few years. Public and private data recently analyzed by PricewaterhouseCoopers (2003) for the American Hospital Association and the Federation of American Hospitals confirmed that, from 1997 to 2001 spending on hospital care increased by \$83.6 billion. Increased volume, the most important reason for this increase, accounted for 55.4 percent; 33.4 percent was attributed to increased use and 21.0 percent to population growth. Since 1996, adjusted admissions increased at least 3 percent every year except 1998, when the increase was 2 percent. Other factors included an aging population; lack of effective care management and patient education; less restrictive benefit plans; and new, more expensive technologies.

Spending on hospital services increased 61 percent over the last 10 years and is still the largest component of rising national expenditures on health care (31.7 percent in 2001). Increased compensation is the most significant driver of the rising cost of goods and services purchased by hospitals. Nearly three-fifths of hospital expense goes to the wages and benefits of caregivers and others. Furthermore, labor costs accounted for 38.8 percent of the increase in spending on hospital care between 1997 and 2001. The study also determined that improved hospital efficiency accounted for \$15 billion in savings between 1997 and 2001. These initiatives resulted in shorter hospital stays, less inpatient capacity, higher productivity, and consolidations.

Labor costs (related to the nursing shortage) are anticipated to account for the largest share of the current increase in spending on hospital services. Between 1995 and 2000, hospital wages exceeded increases paid in private industry, and, as a result, financial margins eroded. In addition to wages, hospitals have absorbed other expenses to retain or recruit nurses, such as tuition reimbursement, sign-on bonuses or referrals, loan repayments, and financing of child care centers. This has put great financial pressure on

hospitals to be more efficient, which in turn has put significant stress on the workforce. The lack of significant, sustained efforts at improvement, coupled with efforts to reduce labor costs, have led to caregivers spending less time with patients and lower job satisfaction. These statistics suggest that we have an enormous opportunity to improve efficiency, safety, and environmental designs to counteract increases in labor costs and inflation.

My presentation is divided into three sections: (1) a study of how health care workers spend their time; (2) a study of current and future hospital designs, with a focus on the patient room (about 400 new hospitals are currently being built from the ground up, many of them designed the same way they have been designed for 100 years or more raising concerns about their sustainability); and (3) the results of changes in design.

## BACKGROUND

In Methodist Hospital, a large time-and-motion video study of patient care processes and the patient care team, with Ann Hendrich as the principal investigator, was done to determine how improvements could be made (Hendrich and Lee, 2003a). Four video cameras were installed in hospital patient units: one in the nursing unit hallway, one on each side of the nursing station, and one in each patient room. (This was an informed Institution Review Board consent study.) The four cameras fed video data into a quad screen for data review and analysis. About 1,000 hours of continuous work were studied in a hospital nursing unit very similar to units in most hospitals in this country. Almost 4,000 events in the patient room and thousands more in the nursing station and the nursing hallway were tracked and “trended” to measure how health care workers spend their time.

We found that in this typical unit a nurse executive budgets for about five-and-one-half to six hours of direct nursing hours per patient day. But patients received less than



10 percent (about 20 to 40 minutes) of direct care in their rooms. Nursing-acuity systems cannot account for the waste and inefficiency we were able to measure in design, distance, transfers, and differences among units. We concluded that the built environment (new or transformed) enabled by technology is a nearly untapped opportunity for improving the cost, quality, and access to hospital care. A main reason nurses are unhappy in their professional roles is that most of their time is spent doing things other than professional nursing. For the most part, their time is not spent with patients on healing, intervention, care, or teaching. It is spent instead on what I call “hunting and gathering”—hunting and gathering paper, supplies, medical records, equipment, trays, carts, linen, and so on. Thus hour by hour, much more time is spent in the nursing unit hallway and the nursing station than in the patients’ rooms.

In addition, many patients are moved two to five times during short hospital stays, which adds to waste, inefficiency, and the workload index. Patients are moved from unit to unit for two reasons: (1) the head wall and technology; and (2) nursing skills. Admittedly, these are very important reasons, but if hospitals address these issues, a whole new level of care and efficiency could be provided.

In a separate patient-transport study, patient-placement data (the chance of transfers, waits, and delays) were entered into a simulation model to show actual patient flow (Hendrich and Lee, 2003b). This study affirmed the need for changes in the current hospital design to reduce waste and inefficiency, improve safety, increase meaningful work for caregivers, and align facilities with future needs. The need for flexible, acuity-adaptable rooms for current and future hospital designs is imperative. The need for comprehensive care and progressive-level care will continue to increase with anticipated changes in demographics and technology. The model clearly demonstrated the high cost and inefficiency of running hospitals the way they are run now and the potential improvements of doing things differently. The model suggested that we have a multimillion dollar opportunity to reduce waste for both patients and caregivers.

## A NEW DESIGN

Based on the internal and external trends revealed in these studies, a demonstration unit was established at Methodist Hospital, shortly after it was consolidated with University Hospital and Riley Hospital for Children. Additional bed space was needed for the cardiovascular consolidation, but we chose not to replicate the familiar nursing unit design. A coronary critical-care unit was combined with a coronary medical unit into a future-state patient room. The head wall was acuity adaptable, and patients were admitted and discharged from the same room. The unit was called the comprehensive coronary critical care unit (Hendrich et al., 2003).

The simple change in the head wall required minimal investment (approximately \$100 dollars per room) to provide

the pounds per square inch necessary to handle multiple gases (oxygen and suction) up through a multilevel tower. Other monitoring technologies would cost more and could be added when needed. Private rooms with acuity-adaptable head walls, adequate space for family, and lighting and temperature controlled by the patient could help reduce infection rates and bed placement times. This design offers maximum flexibility for hospitals of the future.

Hospital patient flow also requires a major transformation. The demonstration unit showed the value of not moving patients from unit to unit. When patients are moved, not only do we lose their dentures, but we also make serious clinical errors because of communication gaps. Every time a patient is moved from one nursing unit to another, the patient comes into contact with another 25 or so caregivers.

The new room design balanced privacy with high observation and created a healing environment for the patient. The windows facing the interior hallway were electronically charged. With the flip of a switch located on the wall, the window in front of the decentralized nursing station could become clear or opaque. (The same effect could be provided with an inexpensive blind.) The nurses used an infrared tracking system to reduce hunting and gathering time to find each other on the unit. The phone was modem capable for family or patient use, and blood analysis modules were in each patient room, so routine blood tests could be done quickly, at the point of care, to reduce lead time for physicians and caregivers.

As electronic medical records become more prevalent, hospitals should think about changing how they use the space of a centralized nursing station. This centralized space could become a business/care center for interdisciplinary practice (nurses and physicians), which would in turn make physician office and department practices more efficient. The nursing stations could be decentralized to reduce travel time and workload index and increase direct-care time. Problems relating to cultural change and human factors (nurses are most familiar with centralized stations) can be resolved with concerted effort. The data are clear—decentralized stations reduce the waste and inefficiency of typical work patterns of hospital nurses (see Figures 1 and 2).

When we consolidated the two units (coronary critical-care and the coronary medical unit), we had a definite moment in time for comparison because patients from both units were moved to the new unit on the same day. We were able to compile true pre-baseline data, and, with this case-control comparison, we were able to measure the impact of change on a variety of levels (clinical, cost, satisfaction). The case-mix was unchanged in the new unit. We measured sentinel events, length of stay, cost of care, medication errors, nursing turnover, and patient falls. The decrease in errors and adverse events was a direct result of the changes in design and care model. Patient dissatisfaction decreased greatly and more rapidly in this unit than in any other unit in the hospital. Nursing hours returned to 1997 levels—patient-care time



FIGURE 1 Typical critical-care patient room.



FIGURE 2 Comprehensive cardiac critical care: acuity-adaptable room for single stay.

was increased, not decreased. Direct-care contact was increased, and hunting and gathering time was decreased.

Previously these two units had transported 200 patients a month back and forth between them; the number dropped to fewer than 20. Remember that the average time for a transport is 25 minutes to 48 hours in most acute-care facilities. Theoretically, we had predicted that acuity-adaptable rooms would be more efficient and that there would be less need to move patients; this was demonstrated in the outcome data. Although the total number of beds was reduced by seven, there were dozens more patient days handled on fewer beds. When the data were entered into the simulation model, the results showed millions of dollars in efficiency improvement. This suggests that smaller, more efficient facilities would bring some relief from workforce shortages and growing demand in the future.

At the heart of the hospital capacity and flow problem (or the cause and effect) is the tension between medical and surgical care specialties and critical care. Many patients don't require critical care, but because progressive beds are usually full, they are often assigned to a critical care bed. Emergency departments and operating room recovery areas are often backlogged with patients waiting for the "right" bed. Thus, patients who are between the critical care and medical-surgical care levels ("tweeners") create a bottleneck in hospital flow. Physicians and nurses tend to err on the side of safety and "hold them" until critical care beds become available. This bottleneck phenomenon tells us something

about future demands for care and the necessity of migrating the middle section of care to the "next generation" of care delivery (Hendrich and Lee, 2003c).

The built environment, enabled by technology, provides an enormous untapped opportunity for reducing waste and improving care when non-value-added analysis is used to improve caregiver work spaces. The development of new care-delivery models to match new hospital environments will be an imperative for the future. This demonstration unit, which provided a healing, patient-centered design to support the patient and caregivers, improved both clinical and fiscal outcomes.

## REFERENCES

- Hendrich, A.L., J. Fay, and A.K. Sorrells. 2003. Cardiac comprehensive critical care: the impact of acuity adaptable patient rooms on current patient flow bottlenecks and future care delivery. *American Journal of Critical Care*. Accepted for publication.
- Hendrich, A., and N. Lee. 2003a. A Time and Motion Study of Hospital Health Care Workers: Tribes of Hunters and Gatherers. Manuscript in progress.
- Hendrich, A., and N. Lee. 2003b. The Cost of Inter-Unit Hospital Patient Transfers. Manuscript in progress.
- Hendrich, A., and N. Lee. 2003c. The Cost of Current Hospital Patient Flow: A Simulation Model. Manuscript in progress.
- PricewaterhouseCoopers. 2003. *Cost of Caring: Key Drivers of Growth in Spending on Hospital Care*. Washington, D.C.: American Hospital Association and Federation of American Hospitals.

# Discrete-Event Simulation Modeling of the Content, Processes, and Structures of Health Care

Robert S. Dittus, M.D., M.P.H.  
*Vanderbilt University and  
Veterans Administration Tennessee Valley Healthcare System*

The Institute of Medicine (IOM) report, *Crossing the Quality Chasm*, challenged health care providers to deliver care that is safe, timely, effective, equitable, patient-centered, and efficient (IOM, 2001). To meet these challenges, health care providers must redesign, implement, and continually improve current health care systems, including: (1) the content of care (what is being delivered); (2) the processes of care (how care is delivered—the microsystems of care); and (3) the structures of care (how delivery systems are organized and financed—the macrosystems of care). Although biomedical and clinical researchers will continue to identify potentially modifiable risk factors for disease and improve methods for diagnosis and treatment through observational and experimental studies, such advances alone cannot address the IOM challenges.

## CONTENT OF CARE

The content of care will be shaped largely by advances in biomedical and clinical research. In colorectal cancer, for example, new chemotherapeutic agents have recently been developed that can prolong life for patients with advanced colorectal cancer (Rothenberg, 2004). In addition, new diagnostic modalities have been developed, such as radiographic “virtual colonoscopy” and a fecal DNA test, to detect early colorectal cancer (Winawer et al., 2003).

Traditional clinical research designs can address the efficacy and effectiveness of treatment and the sensitivity and specificity of diagnostic tests, but cannot easily address many important clinical management questions. Clinical research cannot readily examine the cost-effectiveness of screening colonoscopy at different ages, the most cost-effective time for surveillance colonoscopy among patients who have had a polypectomy, or the combination of age and morbidity at which colorectal cancer screening should be stopped. The myriad of possible solutions to these questions precludes comparing alternatives using traditional research designs and

the size of a clinical study for adequate power would be prohibitive. In addition, the time required to gather study results would be measured in decades because of the slow growth of adenomatous polyps, the precursor of colorectal cancer. However, simulation modeling is a study design that could effectively address these questions (Banks et al., 2004; Law and Kelton, 2000).

## PROCESSES OF CARE

Biomedical research can contribute little to improvements in the processes of care. Clinical observational and experimental studies on the processes of care could be helpful, but little work has been done to date in this area. In the past decade, management science methods have been introduced into clinical medicine more formally and extensively than in the past. A set of such methods, often referred to as continuous quality improvement, have been used worldwide to reduce variations in care delivery. Because health care is generally operating far from the efficiency frontier, these reductions in variation are often accompanied by improvements in quality and reductions in cost. However, the “plan-do-study-act” incremental approach to improvement is not always applicable because external forces, such as governmental or professional regulations, may require significant sudden change. Simulation modeling can be used to explore the implications and consequences of alternative processes of care. Simulation modeling can also generate new insights into underlying systems of care and identify new approaches that might not otherwise be apparent.

## STRUCTURES OF CARE

The structures of care will also require substantial modifications. For example, financing systems are not designed to align incentives to improve the quality and efficiency of care delivery. Even though care delivery systems have

changed over the past decades, they are still based on the same general structures as they were a century ago. For example, the relationships and tasks among health care workers have changed very little. In the past two decades, questions have been raised about the effects of long hours (usually more than 80 hours per week) put in by residents on the quality and safety of care. In response to these concerns, the Accreditation Council for Graduate Medical Education recently established work-hour restrictions for residents. However, it is difficult for residency programs and hospitals to make small, incremental changes to their residency programs. Changes are generally made once a year, and implementing a poor system can affect a program's reputation and subsequent resident recruitment. In this situation, simulation modeling can again be an effective way of examining the potential impact of alternative systems of resident scheduling on both residents and the quality of care.

## TWO SIMULATION MODELING PROJECTS

In this paper, I will describe two simulation modeling projects that highlight the benefits of this systems approach to improving health care. Both projects have been previously published. The first project is a disease-based simulation model that examines the content of care for colon cancer; the project also demonstrates how the model can affect the structure of care. The second project is a hospital-based scheduling simulation that examines the structure of care; the results of this simulation led to improvements in both the structure and processes of care.

### Disease-Based Simulation Model

Colorectal cancer is currently the second leading cause of death from cancer in the United States (Jemal et al., 2003). There are more than a million deaths per year from colorectal cancer, predominantly among the elderly; mortality rates rise logarithmically with age. There is no cure for unresectable disease, although when discovered at an early stage the disease is curable through resection. Several different screening tests are available for early detection, and studies have shown that screening decreases mortality by 15 to 30 percent and that the removal of adenomatous colorectal polyps (e.g., during colonoscopy) decreases the incidence of cancer by 70 to 90 percent (Winawer et al., 2003). Based on these data, a single screening colonoscopy at an appropriate age might be an appropriate diagnostic and therapeutic strategy. Our objective was to develop a decision model and examine the cost-effectiveness of one-time colonoscopic screening for elderly patients (Ness et al., 2000).

A discrete-event network simulation model was used as the platform. The model included the biology of the disease, risk factors for incidence and prognosis, and the health care system that screens for and treats the disease. Input parameters for the model were described as distributions with

characteristics, including distribution shapes, and fit to the data. To measure the cost-effectiveness of alternative screening strategies for colorectal cancer, the outcomes of colorectal cancer had to be described; to measure quality-adjusted life years (QALYs), a standard metric for cost-effectiveness analyses; utilities (as morbidity weights) needed to be measured for each outcome (Gold et al., 1996). Two clinical studies were conducted to create these outcomes, develop a utility instrument, and measure the utilities associated with the outcome states (Ness et al., 1998, 1999). Next, a comprehensive review of the literature (more than 2,500 citations) was conducted. Cost information for diagnosis and treatment were derived from a variety of sources. Once the process was conceptualized and the model formally constructed, verification and validation tests were conducted (Ness et al., 2000).

In constructing the model, an attempt was made to match polyp prevalence data measured through autopsy series and cancer incidence data measured through cancer registries, under the assumption that all adenomas progress to cancer. However, matching the adenoma prevalence rate and the cancer incidence rate required using a dichotomous population of "slow-growing" and "fast-growing" polyps, with mean transition times from adenoma to carcinoma of 52 years and 26 years, respectively. As a result, it was revealed that adenomas progress to cancer at substantially different rates and that some, perhaps many, adenomas regress without treatment. Subsequent data have also suggested that adenomas may regress. As this experience shows, modeling can not only lead to insights into the effectiveness and efficiency of alternative strategies of care, but can also inform the basic biomedical sciences and generate hypotheses regarding the pathophysiology of disease.

The main study results revealed that, among men who had not previously been screened for colorectal cancer (unfortunately, a significant percentage of the population), one-time screening colonoscopy between the ages of 55 and 59 not only reduces the incidence of colorectal cancer, but is also less costly overall than no screening (Ness et al., 2000). In a hypothetical cohort of 100,000 40-year-old men, a screening colonoscopy between the ages of 55 and 59 reduced the overall incidence of colorectal cancer from 5,672 to 2,060 and reduced deaths from colorectal cancer from 2,177 to 654. One-time screening colonoscopy thus was demonstrated in this model to reduce the incidence and mortality of colorectal cancer by approximately 65 to 70 percent. At the same time, the cost of care (colorectal cancer screening, follow-up, and treatment) for these 100,000 men was reduced by 15 percent, from \$75 to \$63 million. If the screening was done five years earlier, between the ages of 50 and 54, the incidence and mortality were reduced even more, but at a slightly higher cost. The marginal cost per QALY was less than \$4,000, which is generally considered a very favorable cost-per-quality ratio. Similar findings were demonstrated for women. The results of this study thus informed

changes in the “content” of health care, that is, the specific, recommended care.

A clinical trial to compare the costs and effectiveness of screening different age groups would be prohibitively expensive and take a very long time. A simulation model is feasible and, in addition, can also examine other features of these strategies of care, such as differential risk patterns among subgroups for the formation of adenomas or the speed of transformation from adenoma to cancer. The impact of differential sensitivities and specificities of diagnostic tests and new diagnostic modalities can be examined quickly. The model can also be used to examine the timing of a repeat “surveillance” colonoscopy after a polyp has been identified and removed. The frequency of surveillance colonoscopies can have a significant impact on the effectiveness and costs of a screening strategy. Given the current lack of capacity in this country to meet the need for colonoscopy under current recommendations, any strategy that reduces demand (such as lengthening the interval for surveillance colonoscopy) can be important. The simulation model can also examine the importance of compliance with certain elements of the strategies on the overall effectiveness and cost-effectiveness of care.

Clinical trials, observational studies, and decision analyses, such as the one described above, have since been used to inform Medicare payment policy. Prior to 2001, Medicare did not reimburse for screening colonoscopies. When cost-effectiveness models demonstrated the overall impact and potential cost savings of screening compared to not screening, this policy was changed. With potential reductions of 70 percent in deaths from colorectal cancer and simultaneous reductions in costs, the “structure” of health care was improved significantly, in this case by a financing change.

### Workforce-Scheduling Simulation Model

Outside of healthcare, simulation modeling has been most commonly used to address facility design, inventory management, scheduling, and workforce deployment. Simulation modeling has also been used in a variety of settings to examine and design new structures and processes of health care (Klein et al., 1993). The second project described in this paper addressed issues related to workforce scheduling.

As a result of a variety of pressures to improve patient safety and reduce resident fatigue, many residency programs began in the 1980s to review and implement changes in house staff work schedules. The initial focus was on the frequency of in-hospital call and the amount of resident sleep time. In the 1970s, first-year residents in internal medicine in some programs were on call either 5 nights out of 7 or every other night, with the norm being every third night, and the work hours regularly exceeded 100 per week. Over time, the frequency of call has been reduced, and recently, the work week has been limited to 80 hours by professional training regulations. In addition, the number of continuous work

hours and the quantity of work, such as patient volume, have also been regulated. As a result, residency programs have been forced to redesign their resident work hours and, at the same time, hospitals have had to redesign their workforces to make up for the reduction in resident work. Resident work scheduling remains an ongoing problem for academic health centers.

In 1989, simulation modeling was used to examine resident scheduling in a county hospital affiliated with an academic medical center (Dittus et al., 1996). A goal of the project was to show whether a discrete-event simulation model of an internal medicine service constructed from easily obtainable information could make valid predictions of residents’ experiences; the major focus was on the amount of sleep residents experienced while on call. A two-stage study was conducted. First, a network simulation model of the internal medicine service of the teaching county hospital was constructed, parameterized, verified, and validated using readily available hospital data and physician surveys. Second, the model was used prospectively to predict the effects of changes in the resident work schedule; the changes were made the year after the model was built.

The setting for the study was a 450-bed municipal teaching hospital with an average daily census of 90 patients on the internal medicine service (78 ward patients and 12 intensive care unit [ICU] patients). Each week, approximately 91 new patients were admitted. The service averaged eight admissions per night, one-third of which went to the ICU. To care for these patients, the medicine service had six teams; each team included a faculty member, a second or third year resident (resident), two first-year residents (interns), a senior student, and several junior medical students. In the baseline call schedule, two of the six teams were on call each night—one ward resident and his or her two interns and senior student, as well as a consulting resident and two interns from another team. Interns were on call every third night and residents every sixth night. Interns averaged 97 hours per week in the hospital.

To model the service, a discrete-event network simulation model was constructed using the INSIGHT simulation language (Roberts, 1983). The model characterized hospital schedules, such as the on-call schedule, the nighttime cross-coverage plan, clinic and conference schedules, and weekday versus weekend work schedules. The model described patient arrivals based on both scheduled and emergency admissions either to the ward or the ICU and characterized 38 house staff activities (residents and interns), including routine patient care, patient-initiated requests for care, and other activities. A decision-priority list established the order in which tasks would be addressed by the house staff following completion of any task. Twenty preemption levels described the prioritization of new tasks added to the work list, which described the interruption of a task prior to completion when a more urgent request was received. Because tasks were time sensitive, their preemption levels could change

over time. The baseline model was constructed and validated against observational data not used in the parameterization of the model.

In contrast to other types of decision analyses, a discrete-event network simulation model is flexible enough to accommodate such representations. The model also allowed for complete flexibility in the description of the input parameters. A flexible distribution system was used to characterize and parameterize input data elements by mean, variance, skewness, and kurtosis.

The model was used to inform a change in the call system from four interns on call every third night to three interns on call every fourth night. To test the predictive validity of the model following this change, a second phase of the project, a prospective work-measurement study, was conducted. Senior medical students were assigned to track the house staff and record the time for the beginning and end of each task. In a pilot study, we measured interobserver variability among the medical students, and, after making clarifications, more than 96 percent agreement was established. The predictive validation study was conducted on 18 house staff days and 6 house staff nights during which house staff were followed and their tasks recorded. We then programmed the simulation model to reflect the change in call schedules and replicated the timing and number of admissions to the hospital to reflect the actual workload managed during the observed time periods.

The simulation model was able to make accurate predictions of the observed house staff work and very close predictions of house staff sleep time, the principal objective of the study (Dittus et al., 1996). For example, in the work-measurement study observations, interns spent 32 percent of their time during the day providing ward and ICU care; the model had predicted 31.5 percent. Residents spent 22.6 percent of their time on ward and ICU time; the model had predicted 23.5 percent. The observed measurements were then compared to the model prediction for total house staff sleep time when on call. The measurement study observed that each member of the house staff spent 3 hours and 30 minutes sleeping; the model had predicted 3 hours and 27 minutes. Thus, the model appeared to be a valid representation of the actual work. Once validity was established, the model was used to improve work and care delivery.

One advantage of the model is that it can examine a number of parameters and monitor outcomes. For example, a quality-of-care metric might be based on the percentage of care provided by “tired” house staff members, the percentage of emergency or urgent care delivered by “tired” house staff members, or the average time taken to complete a care request. The model allows for a very flexible definition of “tired” (e.g. the total number of minutes of sleep over a past period of time and/or the total number of minutes of uninterrupted sleep, etc.). In addition, the model could track the percentage of time that an emergency or urgent care request was managed by a member of the patient’s true team, and

not a covering member of another team, who wouldn’t know the patient as well. The new and old call schedules were compared against these quality metrics using varying definitions of “tired.” The results showed that the new call schedule, although designed to reduce house staff fatigue, resulted in significantly less sleep on call because the house staff teams were busier during their nights on call. As a result, the quality metrics deteriorated.

The model also allowed for the examination of potential improvements in the “processes” of care. An examination of the causes of interruptions of sleep time revealed a common demand for starting intravenous lines and drawing blood at various times during the night. The model illustrated that relieving the house staff of these jobs would result in a substantial increase in uninterrupted sleep time. As a consequence, a phlebotomy and intravenous placement team was hired by the hospital, which had an important impact on the quantity and quality of house staff sleep time.

## CONCLUSION

As the colorectal cancer and house staff scheduling models demonstrate, discrete-event network simulation modeling can be used to analyze and improve the content, processes, and structures of health care. Continued advances in computational speed and modeling software should make this technology increasingly accessible to health care leaders and managers. The incorporation of such models into the routine planning, examination, and improvement of health care systems holds promise for helping health care become increasingly safe, timely, effective, equitable, efficient and patient-centered.

## REFERENCES

- Banks, J., J. Carson, B.L. Nelson, and D. Nicol. 2004. *Discrete-Event System Simulation*. Upper Saddle River, N.J.: Pearson Prentice Hall.
- Dittus, R.S., R.L. Klein, D.J. DeBrotta, M. Dame, and J.F. Fitzgerald. 1996. Medical resident work schedules: design and evaluation by simulation modeling. *Management Science* 42(6): 891–906.
- Gold, M.R., J.E. Siegel, L.B. Russell, and M.C. Weinstein, eds. 1996. *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press.
- IOM (Institute of Medicine). 2001. *Crossing the Quality Chasm: A New Health System for the 21<sup>st</sup> Century*. Washington, D.C.: National Academy Press.
- Jemal, A., T. Murray, A. Samuels, A. Ghafoor, E. Ward, and M.J. Thun. 2003. Cancer statistics, 2003. *CA: A Cancer Journal for Clinicians* 53(1): 5–26.
- Klein, R.L., R.S. Dittus, S.D. Roberts, and J.R. Wilson. 1993. Simulation modeling and health care decision making. *Medical Decision Making* 13(4): 347–354.
- Law, A.M., and W.D. Kelton. 2000. *Simulation Modeling and Analysis*. Boston, Mass.: McGraw-Hill.
- Ness, R.M., A. Holmes, R. Klein, J. Green, and R.S. Dittus. 1998. Outcome states of colorectal cancer: identification and description using patient focus groups. *American Journal of Gastroenterology* 93(9): 1491–1497.

- Ness, R.M., A. Holmes, R. Klein, and R.S. Dittus. 1999. An assessment of patient utilities for outcome states of colorectal cancer. *American Journal of Gastroenterology* 94(6): 1650–1657.
- Ness, R.M., A. Holmes, R. Klein, and R.S. Dittus. 2000. Cost utility of one-time colonoscopic screening for colorectal cancer at various ages. *American Journal of Gastroenterology* 95(7): 1800–1811.
- Roberts, S.D. 1983. *Simulation Modeling and Analysis with INSIGHT*. Indianapolis, Ind.: Regenstrief Institute.
- Rothenberg, M.L. 2004. Current status of second-line therapy for metastatic colorectal cancer. *Clinical Colorectal Cancer* 4(Supp.1): S16–S21.
- Winawer, S., R.H. Fletcher, D. Rex, J. Bond, R. Burt, J.T. Ferrucci, T.G. Ganiats, T. Levin, S.H. Woolf, D. Johnson, L. Kirk, S. Litin, and C. Simmang. 2003. Colorectal cancer screening and surveillance: clinical guidelines and rationale—update based on new evidence. *Gastroenterology* 124(2): 544–560.





# Measuring and Reporting on Health Care Quality

**Dana Gelb Safran**  
*New England Medical Center*

I will address a crucial question in this talk—what brought us to the point that we mistrust or question our doctors and our insurance companies? The answer is complex. Research has shown that medical practice varies greatly across the country, raising the question of how much of medical practice is really science. Research has also been done on health care spending and cost inflation—but efforts to contain spending have raised concerns about compromising quality. Gradually, the idea of accountability through measurement and reporting is gaining support.

Several definitions of “health” and “health care quality” have been proposed over the years. Back in 1952, Lembeke proposed this definition:

The best measure of quality is not how well or how frequently a medical care service is given but how closely the result approaches the fundamental objectives of prolonging life, relieving distress, restoring function, and preventing disability.

In 1948, the World Health Organization defined health in the Declaration of Human Rights: “Health is a state of complete physical, social, and mental well-being, not merely the absence of disease and infirmity.” We are just beginning to measure health in these terms and to study the impact of medical care on functional health status and well-being.

One of the many difficulties in measuring health care quality is determining how overall health relates to health care spending. Managed care raises the question of where we are on the hypothetical curve that economists propose reflects the relationship of health care to health (Figure 1). Will spending more or providing more care lead to better health? If we are on the ascending part of the curve, then more care or more spending will lead to better health. But if we are on the flat part of the curve, and that is the theory of many managed care organizations, then we can afford to cut back on care without doing harm. This is a fundamental

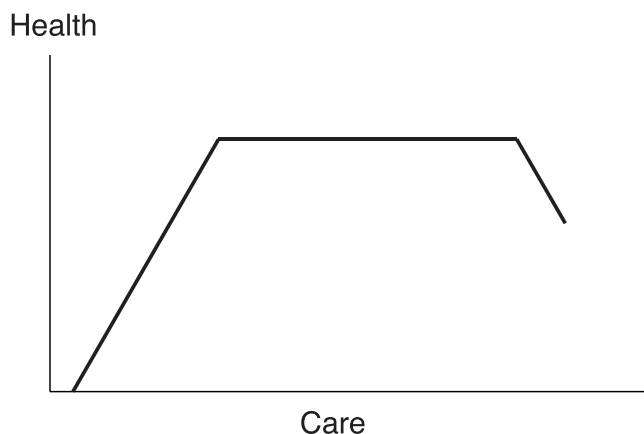
question in health services research, and the answer depends on who you ask.

The most widely used instrument for measuring health in the multidimensional terms outlined by the World Health Organization is the Short Form 36-Item Health Survey (SF-36), which measures eight dimensions of health—physical function, the physical component of role function, bodily pain, general health perceptions, vitality or energy, social function, the emotional component of role function, and mental health. These eight dimensions combined yield two global assessments—one of physical health and one of mental health.

The Health Care Financing Administration and the National Committee on Quality Assurance are using the SF-36 to study the health outcomes of Medicare beneficiaries in Medicare HMOs. This Medicare Health Outcomes Survey is the first to follow data on patients’ health longitudinally. The goal is to be able to hold systems accountable for patients’ health as defined by the World Health Organization. There has been tremendous resistance to this approach. Many have questioned how doctors can be expected to affect multiple aspects of patient functioning. My answer is that, until they try, they probably can’t.

From 1986 to 1992, I had the privilege of participating in the Medical Outcome Study, a large-scale, longitudinal study by leading scientists at New England Medical Center; the study was performed in conjunction with RAND. Two goals of the study were: (1) to determine where we are on the health care curve; and (2) to assess how differences in health care delivery and specialty care are reflected in health outcomes (Tarlov et al., 1989). This study could start a new dialogue about health care and the way patients think about their health.

In the early 1990s, we began measuring and reporting on the performance, or quality, of health care plans. The impetus for the study was a demand for data by large employers



**FIGURE 1** The health care curve.

in the United States who needed information about health plans for themselves and their employees. The website for the National Committee for Quality Assurance ([www.ncqa.org](http://www.ncqa.org)) now provides report cards (by zip code) rating health maintenance organizations (HMOs) in the following categories: Access, Service, Qualified Providers, Staying Healthy, Getting Better, Living with Illness. Each HMO is given an overall rating of Commendable or Excellent.

Unfortunately, this two-dimensional system of information is not very helpful to patients. First, an increasing share of Americans don't have a choice of plans. Second, even those who do have a choice find it difficult to process this much information. Third, the information is not presented in a user-friendly way; it does not allow the user to prioritize the aspects of care or to add dimensions. One of the things we learned from quality measurement on the health plan level is that plans don't vary much in a given market, especially in the provision of care.

However, it is important to note that our analyses and others have demonstrated considerable variability across markets—suggesting that where you live has an important bearing on the quality of care you can expect to receive. But assuming that the public will not use health care quality data to make “relocation” decisions—but rather to make health care decisions, the question remains as to what level of data are most appropriate and most relevant.

The next level of quality measurement that has been attempted, after health plans, is performance at the medical group level. And, indeed, within markets there appears to be considerably more variation among medical groups than among health plans. However, our own analyses reveal that the variability *within* groups is considerably greater than the variability between them. In other words, a medical group's overall performance does not provide an accurate representation of the performance of individual physicians within that group. So knowing how a group performs, on average,

doesn't tell me very much about the quality of care I will receive from an individual physician in that group whom I might select. And for patients, choosing a doctor—not a group—is almost always the relevant choice. There are exceptions in a few select U.S. health care markets, but for the most part, patients choose a specific physician to take care of them, not a group. Indeed, recent studies confirm that the kind of health care quality data that is a priority for U.S. adults is information that will help them choose a doctor (Tumlinson et al., 1997).

So, what do we know about measuring performance at the individual physician level? First, although there is some momentum for measuring physician performance, there is also tremendous resistance. Our research group developed a tool to measure each attribute of primary care based on the definitions of the Institute of Medicine Committee on the Future of Primary Care (IOM, 1996):

Primary care is the provision of integrated, accessible health care services by clinicians who are accountable for addressing a large majority of personal health care needs, developing a sustained partnership with patients, and practicing in the context of family and community.

We measured access (a defining characteristic of primary care), continuity, and comprehensiveness of care, including knowledge of the patient. We measured two aspects of clinical interaction, the quality of communications and the thoroughness of physical exams. We measured the quality of interpersonal relations and trust, both of which relate to sustained partnerships. We then used these measurements to get an understanding of the organizational and individual characteristics that predict performance in these areas and to determine whether performance is a meaningful predictor of outcomes.

We have studied three outcomes so far: (1) a patient's adherence to a doctor's advice; (2) a patient's voluntary disenrollment from a doctor's practice; and (3) a patient's functional health outcome. We found that two attributes of primary care predict a patient's adherence to the doctor's advice: the patient's trust in the doctor and the patient's feeling that the doctor has “whole-person knowledge” about him or her. Lack of trust and lack of comprehensive knowledge of the patient are strongly correlated with patients voluntarily leaving a doctor's practice. Over a three-year period (from 1996 to 1999), 11 percent of patients who had the most trust in their doctors voluntarily left the practice; 37 percent of patients with the least trust in their doctors voluntarily left. For an individual doctor, that translates to a loss of 400 patients over a three-year period, a lot of patients to replace. We need to find ways to improve the interpersonal dimensions of health care and thus close the gaps in performance in these attributes.

So far, we have not improved on them. Our three-year follow-up study in Massachusetts showed that patients who had stayed with their primary care doctors had noticed an

erosion in their relationships with their doctors. Physician satisfaction and physician morale had also declined, especially in terms of professional autonomy, time spent with patients, and time for family and personal life (Murphy et al., 2001; Safran et al., 2001).

We have also attempted to assess the current medical environment, specifically the experiences of clinicians and nonclinicians who work together. We developed an instrument to gauge medical care culture by job classification. First, the data suggest that physicians are dissatisfied, and, if physicians are dissatisfied, the feeling probably cascades to everybody else in the health care setting. When we measured several aspects of the quality of the medical workplace, including job demand, job control, leadership quality, supervisor support, interactions with physicians, interactions with patients, team culture, overall mood, and job benefits, the results were alarming. A survey of residents, for example, showed that they were satisfied with their interactions with other residents but very dissatisfied with their interactions with nurses. In addition, because of the complexity of health care and time constraints on clinicians, patients must rely on teams for their care. Our study showed that both patients and

team members were dissatisfied with team care. To engender a true team culture, we will have to change the way physicians and other caregivers are educated.

## REFERENCES

- IOM (Institute of Medicine). 1996. *Primary Care: American's Health in a New Era*. Washington, D.C.: National Academy Press.
- Lembeke, P.A. 1952. Measuring the quality of medical care through vital statistics based on hospital service areas. Part 1. Comparative study of appendectomy rates. *American Journal of Public Health* 42: 276–286.
- Murphy, J., H. Chang, J.E. Montgomery, W.H. Rogers, and D.G. Safran. 2001. The quality of physician-patient relationships: patients' experiences 1996–1999. *Journal of Family Practice* 50(2): 123–129.
- Safran, D.G., J.E. Montgomery, H. Chang, J. Murphy, and W.H. Rogers. 2001. Switching doctors: predictors of voluntary disenrollment from a primary physician's practice. *Journal of Family Practice* 50(2): 130–136.
- Tarlov, A.R., J.E. Ware, Jr., S. Greenfield, E.C. Nelson, E. Perrin, and M. Zubkoff. 1989. The Medical Outcomes Study: an application of methods for monitoring the results of medical care. *Journal of the American Medical Association* 262(7): 925–930.
- Tumlinson, A., H. Bottigheimer, P. Mahoney, E.M. Stone, and A. Hendricks. 1997. Choosing a health plan: what information will consumers use? *Health Affairs* 16(3): 229–238.



# Archimedes: An Analytical Tool for Improving the Quality and Efficiency of Health Care

David M. Eddy and Leonard Schlessinger  
*Care Management Institute, Kaiser Permanente  
and Kaiser Permanente Southern California*

The practice of medicine has become extraordinarily complex, and it promises to become even more complex as the pace of innovation accelerates. Managing that complexity requires good information about the effects of different courses of action on health, logistic, and economic outcomes. The preferred method of obtaining that information is through empirical clinical research. Unfortunately, in medicine the ability to conduct clinical research is severely limited by the high cost of enrolling and following patients, the long follow-up times, the large number of options to be compared, the large number of patients, unwillingness of people to participate (e.g., to be randomized or to follow a specified protocol), and unwillingness of the world to stand still until the research is done. A typical clinical trial comparing just two options requires thousands of patients, costs tens or hundreds of millions of dollars, takes 3 to 15 years, and is likely to be outdated before it is completed.

In other fields, mathematical models have been used to help make decisions and design systems. However, the variability of human biology and behavior, the size and complexity of health care systems, and the wide variety of important questions to be addressed all place special demands on health care models. We have designed a new type of model, called Archimedes, to try to address these special demands. This paper describes the basic structure and scope of the model, the modelling methods, how we can validate the model, and its potential uses.

## STRUCTURE AND SCOPE

Archimedes has three main parts. At the core is a model of human physiology that describes the pertinent aspects of anatomy, physiology, pathophysiology, occurrence of signs and symptoms, effects of tests and treatments, and occurrence of health outcomes. The second part consists of care process models; these describe what providers do when a person seeks care or what providers can do to prevent a

person from needing care. The third part, system resources, includes such things as personnel, facilities, equipment, and costs. The full Archimedes model is applied in a specific health care setting defined by specific care processes and specific system resources.

A complete description of all the objects and their attributes, functions and interactions is not possible here. But to give you a sense of the model's scope, I will describe some of the main classes of objects and give examples of their attributes and functions.

*Patients.* We use the term "patient" to mean anyone who might receive health care from the system, including people when they are well. The attributes of patients can be as detailed as required; they can include age, sex, risk factors, behaviors, education level, type of employment, and insurance coverage. All patients have physiologies, which include all pertinent organs and biological variables. As governed by the equations, patients can get diseases, which can modify the functions of their organs and can cause signs, symptoms, and health outcomes. Patients have perceptions, memories, and behaviors that determine how they respond to signs and symptoms and how they adhere to interventions. Their risk factors, physiologies, and behaviors can respond to interventions, which in turn can affect the occurrence and progression of their diseases. As in reality, each patient is different, and the spectra of physiologies, behaviors, and other characteristics correspond to the spectra seen in reality.

*Health Care Providers.* All pertinent types of personnel involved directly or indirectly in providing health care are included. Examples are nurses, pharmacists, physicians, telephone operators, and case managers. Within each of these types are the appropriate subtypes to model a particular problem (e.g., physicians → surgeons → cardiac surgeons → pediatric cardiac surgeons). Health care providers have attributes (e.g., ages, skill levels, behaviors), as well as functions (e.g., cardiac surgeons can perform bypasses, but telephone operators can not).

*Interventions.* Archimedes includes two main types of interventions. “Tests and treatments” encompass *what* care is delivered. This type includes: changes in risk factors and preventive treatments; tests that provide information about the existence, severity, or prognosis of a disease; “curative treatments” that directly affect the progression and outcomes of a disease; and “symptomatic treatments” that affect the symptoms of a disease, without affecting its progression. The other type of intervention, “care processes,” determines *how* tests and treatments are delivered. Examples are: use of case managers, creation of a registry to increase compliance with a performance measure, and development of criteria for referrals to specialists. For either type of intervention it is possible to specify the types of providers who can deliver it, the types of facilities or locations where it can be provided, and the types of equipment and supplies it requires. In the model, such things as the use, effectiveness, and cost of an intervention can vary depending on many factors, such as patient characteristics, type of provider, skill of provider, time of day, delivery site, and random factors.

*Policies, Protocols, and Regulations.* The use and effectiveness of any intervention can be determined by a set of policies and protocols that describe such things as: who delivers it, where it is delivered, the criteria for determining which patients should get it, the sequence of events for implementing it, and the decision rules applied at different steps. Clinical practice guidelines, performance measures, strategic goals, and the “what-to-do” parts of disease management programs are examples of policies that affect tests and treatments. Continuous quality improvement projects, nursing protocols, instructions to telephone operators, and the “how-to-do-it” parts of disease management programs are examples of policies that affect care processes. The accuracy with which any of these is applied can allow for variations and random factors that mimic the variations and randomness of real practice. For example, adherence to a particular guideline can be different for a primary care physician than for a specialist, for a physician who has attended a continuing medical education class within the last 12 months, or for a physician who sees more than 50 patients a year who are candidates for the guideline.

*Facilities, Equipments, and Supplies.* Archimedes can include all types of facilities, equipment, and supplies that are involved in the management of a disease. Any type of any of these classes can be expanded to any level of detail (e.g., bed → monitored bed → monitored bed in the emergency department).

*Logistics and Finances.* Archimedes can record the cost, location, time, and any other important circumstance of every event. Thus virtually any type of budget, table of accounts, utilization report, or forecasting report can be calculated.

## METHODS

The mathematical foundations of the Archimedes model are described elsewhere (Schlessinger and Eddy, 2002).

Briefly, it is written in differential equations and programmed Smalltalk, an object-oriented language. The most difficult part of the model is the representation of physiology. We conceptualize the physiology of a person as a collection of continuously interacting objects that we call “features.” The concept of a feature is very general, but features correspond roughly to anatomic and biological variables. Examples in the current Archimedes model are systolic and diastolic blood pressures, patency of a coronary artery, cardiac output, visual acuity, and amount of protein in the urine. Features can represent real physical phenomena (e.g., the number of milligrams of glucose in a deciliter of plasma), behavioral phenomena (e.g., ability to read an eye chart), or conceptual phenomena (e.g., the “resistance” of liver cells to the effects of insulin).

The model is largely driven by the trajectories of features—their values as continuous functions of time. They register the effects of patient characteristics, interact continuously with each other, determine the occurrence and progression of diseases, trigger the onset and determine the severity of signs and symptoms, are measured by tests, respond to treatments, and cause health outcomes. Specifically, differential equations are used to define the progression of each feature as a function of patient attributes as well as other features. At any given time, the values of features can be measured by tests, subject to both random and systematic errors. Equations define clinical events, such as signs, symptoms, and health outcomes, as functions of the magnitudes and trajectories (e.g., rate of change) of various combinations of features. Diseases, which in reality are human-made labels for constellations of biological variables, are defined in the model in the same way. For example, in the model as in reality, a person is said to have “diabetes” if the fasting plasma glucose exceeds 125 mg/dl or the oral glucose tolerance test exceeds 199. Treatments are included as parameters in the equations for features, being able to change their values, rates of progression, or both. In the model, treatments do this at the level at which their actual mechanisms of action are understood to occur. For example, in the model the drug Metformin affects the equation that determines the amount of glucose produced by the simulated liver cells. Finally, the signs, symptoms, and behaviors caused by changes in features set in motion all the logistic events and use of resources that occur in a health care system.

In general, several dozen features and 10 to 30 equations are necessary to calculate the occurrence of any particular outcome (e.g., the rate of heart attacks in a specified population). The model currently includes the features pertinent to coronary artery disease, congestive heart failure, diabetes, and asthma. Features relating to other diseases are being added continually. Other formulas describe the clinical, logistic, and economic events. These formulas are typical of decision trees, flow charts, and accounting models. All of the formulas can include person-to-person differences, random variations, and uncertainty.

The level of detail of the model is determined by the intended users. We build the physiology part of the model to the level of detail clinicians tell us they consider necessary for their decisions. As a result, the physiology model corresponds roughly to the level of biological detail found in patient charts, general medical textbooks, and the designs of clinical trials. Care processes, logistics, resources, and costs are modelled at an equally high level of detail, as determined by administrators. For example, there are 37 different types of outpatient primary care visits.

## BUILDING THE MODEL

Archimedes is built from existing basic research, epidemiological studies, and clinical trials of treatments (Schlessinger and Eddy, 2002). When person-specific data are available, they can be used to derive equations for features as functions of other features. When person-specific data are not available, aggregated data, such as those routinely published for registries, population-based studies, and clinical trials, can be used. In general, the results of any well designed study can be used to build the part of an Archimedes model that addresses biological phenomena, outcomes, and interventions that were investigated in the study.

The data to describe care processes are not routinely collected or published. In practice, we develop our models of care processes through examination of administrative data, existing protocols, interviews, and on-site observations, checked against any available data. Pilot studies can be conducted as needed for processes that are determined through sensitivity analysis to be critical.

## VALIDATION

*Methods.* Ultimately, the value of a model depends on how accurately it can represent reality. The deep level of physiological detail coupled with the care processes in the Archimedes model provide a rigorous way to test this. The validation strategy is to identify an epidemiological study or clinical trial, conduct a “virtual study” or “virtual trial” in the virtual world of the model, and then compare the results. The basic steps are: (1) Have the model “give birth” to a large population of simulated people. Imagine a large city of simulated people with a representative spectrum of characteristics (e.g., age, sex, race/ethnicity, and genetic background) and medical histories. They are all unique, and most will never get the disease to be studied in the trial. (2) Run the model to let them age naturally until they reach the age range of the people who were candidates for the real trial. (3) Identify those who would meet the inclusion criteria for the trial, and select from them a sample that corresponds to the sample size of the real trial. (4) Randomize the simulated participants into groups, as was done in the real trial. (5) Have simulated providers give the patients the treatments

according to the protocols described for the real trial. (6) Run the model for the simulated duration of the trial, with the simulated providers applying whatever follow-up and testing protocols were used in the real trial. (7) Count the outcomes of interest that occurred to the participants in the simulated trial. (8) Compare them to the results observed in the real trial. We use Kaplan Meier curves to make the comparisons because they contain the most information about the outcomes in all of the arms of a trial at all time periods.

All of this is done at whatever level of detail is necessary to simulate what was done in the real trial, using whatever descriptions are available from publications. For example, if “hypertension” is defined as “a finding on at least two of three consecutive measurements obtained one week apart . . . of a mean systolic blood pressure of more than 135 mm Hg or mean diastolic blood pressure of more than 85 mm Hg, or both,” that is what we have the simulated physicians do.

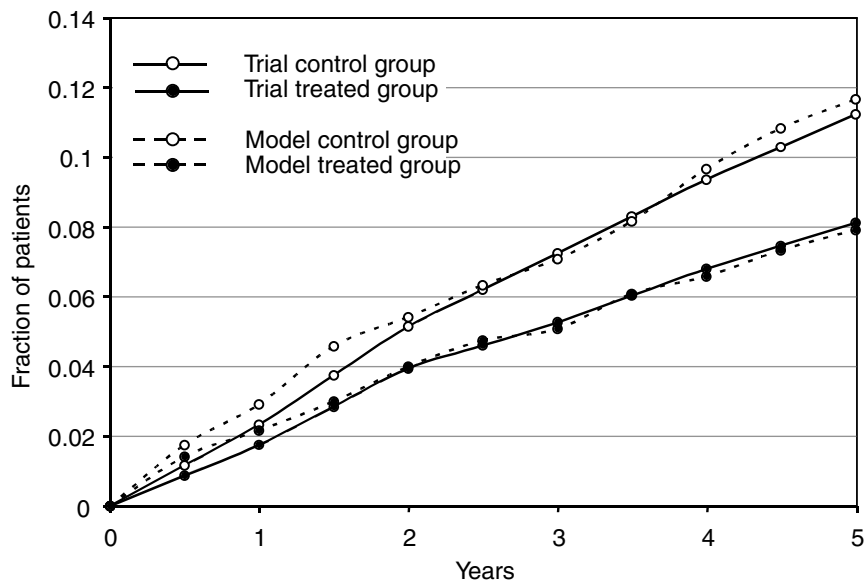
Each trial that is simulated in this way provides a sensitive test of the model. For each, the simulated results come from thousands of simulated individuals, each of whom has a simulated liver, heart, pancreas, and other organs. Each liver produces glucose, each coronary artery can develop plaque or thrombus at any point, and each kidney clears urine. The progression of the pathological process is different in every person, just as in reality. The simulations also include simulated physicians following simulated practice patterns or guidelines, with different degrees of compliance . . . on through to the performance of tests, reporting of results, making of errors, giving of treatments, use of facilities and equipment, and generation of costs. All told, each simulation tests scores of equations in every patient and hundreds of other equations that all have to work correctly in concert.

At the end of a simulation, the results of the virtual study should closely match the results of the real study, within the bounds of random variation related to sample size. We say there is a “statistical matching” of results if there is no statistically significant difference between the model’s results and the real results.

To help probe different parts of the model and to check its validity for different populations, organ systems, treatments and outcomes, we test the model in this way against a variety of different trials. Each validation exercise uses the same model with the same parameter values; parameters are not set to “fit” one trial and then reset to fit another trial. The trials are chosen by an independent advisory committee, which also reviews the results.

In some cases, some information from a trial is needed to help build some part of the model. When this occurs, the information from the trial is used to help derive only one equation out of the 10 to 30 used to calculate the outcome of interest in the population of interest. Thus a validation exercise involving such a trial not only confirms the equation it helped build, but also provides an independent validation of the other equations. Furthermore, the equation built with help of any particular trial is independently tested by all of the





**FIGURE 1** Comparison of model and trial of fraction of patients having major coronary events in the Heart Protection Study (2002).

validation exercises involving other trials. Out of the 18 trials used to validate the model thus far, 8 were used to help build the model, 10 were not.

*Validation Results.* Using these methods, the Archimedes diabetes model has been validated against 17 epidemiological studies and 18 clinical trials thus far. The example shown in Figure 1 compares the model results with the trial results for the Heart Protection Study (2002). This trial randomized about 25,000 high-risk people to receive either a placebo or a cholesterol-lowering drug, Simvastatin. People were defined as being at high risk if they had coronary artery disease, occlusive arterial disease, or diabetes. The primary outcome was the fraction of people who developed heart attacks. No information from this trial was used to help build the model.

Counting the different arms and outcomes of the 18 trials, a total of 74 validation exercises have been conducted to date. (Figure 1 illustrates 2 of the 74.) In 71 of the exercises, the model's results statistically matched the real results. For the three exercises that were not a statistical match, in one case the difference in results was just barely statistically significant ( $p = 0.04$ ), which is to be expected in 74 exercises. In the other two, the difference was due to the model underestimating the underlying rate of the outcome in the trial population by about 35 percent. (The model estimated the effect of the treatment accurately.) The advisory committee concluded that this discrepancy was most likely due to a risk factor in the trial population that was not described in the publication and therefore could not be included in the model. Considering all 74 exercises, the correlation between the

model's results and the real results is  $r > 0.99$ . Considering only the 10 trials that were not used to help build the model, the correlation was still  $r > 0.99$ .

## USES OF AN ARCHIMEDES MODEL

Archimedes is meant to create a virtual world at the level of detail at which real clinical and administrative decisions are made. Once created and validated, the virtual world can be used to explore a wide variety of scenarios and questions, much as a flight simulator can be used to simulate different types of flying conditions and emergencies. Applications include: (1) designing and testing clinical management tools, such as guidelines, performance measures, strategic goals, disease management programs, priorities and continuous quality improvement programs; (2) evaluating and performing cost-effectiveness analyses of clinical and administrative programs; (3) designing and interpreting clinical research, including setting priorities for new trials, planning trials (e.g., sample size, duration, clinical costs), projecting long-term Phase 3 results from short-term Phase 2 results, estimating outcomes in subpopulations, and extending the results of a trial (e.g., predicting 15-year outcomes from 3-year outcomes, predicting outcomes that were not initially measured); (4) estimating outcomes for specific patients who are contemplating different treatment options; and (5) creating a "living library"—a place where the current body of knowledge about a disease is not only organized and stored, but is also integrated in a quantitative way that can be used for the other types of applications just described.

## DISCUSSION

Archimedes is distinguished from other models by several features. It is a person-by-person, object-by-object simulation. It covers a broad spectrum, spanning features from biological details to the care processes, logistics, resources, and costs of health care systems. It is written at a deep level of biological, clinical, and administrative detail. It is continuous in time; there are no discrete time steps, and any event can occur at any time. Biological variables that are continuous in reality are represented continuously in the model; there are no clinical “states” or “strata.” It includes many diseases simultaneously and interactively in a single integrated physiology, enabling it to address comorbidities, syndromes, and treatments with multiple effects. Finally, it has been validated by simulations of a wide range of clinical trials.

Archimedes is not intended to replace reality. If a question can be answered with a well designed empirical study, that approach is always preferable. Our goal is to provide a trial-validated method that can be used to address problems that can not be feasibly addressed through empirical studies, because of high cost, long follow-up times, large sample size, unwillingness of providers or patients to participate, large

number of options, or the rapid pace of technological change. In the way that a flight simulator provides valuable experience, shortens the time needed in real planes, and simulates experiences that are too dangerous or rare to attempt for real (like severe wind shear), the Archimedes diabetes model should be a useful tool for sharpening our understanding of diseases and their management.

The model, which was developed and is owned by Kaiser Permanente, is currently being prepared to be made accessible to individuals and organizations, over the Web, through a friendly interface on a nonprofit basis. The website is expected to be completed by the end of 2005. In the meantime, the authors can be contacted by e-mail about access (eddyaspen@yahoo.com).

## REFERENCES

- Heart Protection Study Collaborative Group. 2002. MRC/BHF Heart Protection Study of antioxidant vitamin supplementation in 20,536 high-risk individuals: a randomized placebo-controlled trial. *Lancet* 360(9326): 23–33.
- Schlessinger, L., and D.M. Eddy. 2002. Archimedes: a new model for simulating health care systems—the mathematical formulation. *Journal of Biomedical Informatics* 35(1): 37–50.



# Applying Financial Engineering to the Health Services Industry

**John M. Mulvey**  
*Princeton University*

The primary goal of operations research (OR) is to improve the efficiency of public and private organizations. There have been many significant success stories since the field began during the Second World War. For example, military war planners employ OR methods to assist in the logistics of moving people and equipment to designated locations and time points; the last two wars in Iraq show the critical benefits of efficient logistical planning. In another application, both the airline and telecommunication industries rely on optimization for scheduling and planning purposes. OR methods improve the efficiency of these complex logistical decisions.

The goal of financial engineering is to analyze, manage, and transfer risks efficiently within and across organizations. To achieve this goal, we focus on modeling uncertain elements as stochastic systems of equations. Financial engineering has been used to price options, design structured securities, employ dynamic portfolio theory for investors, and manage asset-liability for institutional and individual investors. In contrast to traditional OR, financial engineers must model risks and create instruments for transferring risks. Financial engineering addresses both tactical and strategic decisions. At the strategic level, we optimize complex organizations (enterprises) in the face of uncertainties.

There are differences, of course, between traditional engineering and service-sector engineering (such as financial engineering). Traditional engineers typically design physical objects, machines, and networks. Financial engineers design financial products and services. Traditional engineers build upon physical reality, whereas financial engineers attempt to solve problems using advanced mathematics build objects that are not physical in nature, such as novel securities. Traditional engineers typically take on professional responsibility; they are personally liable if harm results from a failure. Financial engineers have little personal liability at this time. And finally, perhaps because of personal liability and related issues, traditional engineers

concentrate on design failures (e.g., why a bridge collapsed). For financial engineers, the deep study of failures is just beginning.

Regulations have an enormous impact on both domains. With the 1936 Flood Control Act, the government created regulations that required government projects to meet minimum economic standards. Those regulations led to methods enabling the government, and companies, to compute cost-benefit analyses of proposed projects and to engage in projects only when the overall benefits outweighed the overall costs. Similarly, the 1974 ERISA Act helped U.S. pension plans analyze their assets and liabilities and compute annual pension plan surpluses. When there are deficits, contribution rules are based on these calculations. In addition, the “prudent-man rule” (required by ERISA) has had a substantial impact on how decisions are made.

Despite these significant regulations, severe difficulties can arise. In the past three years, many large U.S. companies have seen ample surpluses turn to large losses as the equity markets have plunged and interest rates have declined. The 1974 regulations should be revised to prevent this type of difficulty from recurring. Financial engineering can play an important role in developing more efficient regulations for the pension industry.

## **IMPROVING EFFICIENCY IN THE INSURANCE INDUSTRY**

The U.S. insurance industry, another highly regulated domain, is regulated mostly through the 50 state insurance commissions. In 1998, the National Association of Insurance Commissioners completed its revisions of statutory accounting standards, the code of standards that requires insurance companies operating in the United States to evaluate their assets and liabilities according to regulatory standards. Annual assessments are made so that a company’s assets and liabilities can be applied to surplus calculations. There are

several methods for determining an insurance company's surplus, including GAAP, statutory surplus, and economic (market value) surplus. Each of these values helps determine the health of an insurance company in terms of its ability to pay future liabilities and make a profit along the way.

An "optimal" insurance company would not only be safe in terms of protecting itself against adverse circumstances, but would also be reasonably profitable so that shareholders benefit and the cost of capital is relatively low. An optimal insurance company would satisfy all of its policy holders, provide relatively inexpensive products and services, pay shareholders a profit, and have a low chance of bankruptcy. Company employees and customers would both be pleased with this optimal environment.

Are there insurance companies that satisfy all these criteria? Most existing insurance companies, including health care insurers, fall short on several counts. Many primary insurers have low profitability, and customers may be unhappy with existing rates. Financial engineering can play an important role in improving the efficiency of insurance companies.

An example of an efficient company is the Renaissance Reinsurance Company of Bermuda, which operates primarily in the area of catastrophic risk. This reinsurance company takes in money by selling reinsurance to insurance companies that sell catastrophic insurance—mainly for earthquakes and hurricanes. Major decisions for a reinsurance (or insurance) company are: (1) how to invest assets (called asset allocation); and (2) which businesses to insure. Other decisions include who the policy holders are, how much is charged, and how diversification is done. Once assets and liabilities are decided, the question becomes how much insurance the company buys for itself—thus insuring its own risk. This leads to what is called retrocessional insurance.

For large insurance companies, the capital structure is an important factor in these decisions. Capital acts as a buffer that protects the company against loss. How large the capital should be depends on how much risk the organization is willing to accept. The amount of risk capital generally depends on the size of losses at the left tail of the profit/loss probability distribution. As we will see, diversification reduces the tail losses thereby lowering the capital charge.

## GLOBAL INSURANCE COMPANIES

In many cases, multidivisional insurance companies can operate more efficiently than single product companies. For example, AXA, the global insurance company based in Paris, France, allocates capital according to a system that projects scenarios into the future and estimates profit under each scenario for each division. The company then allocates risk capital through its headquarters (see Figure 1). This approach saves total enterprise capital because benefits are diversified and profits are gained by lower capital requirements.

Would this work for an insurance company in the health-care industry? A company with a life-insurance division and

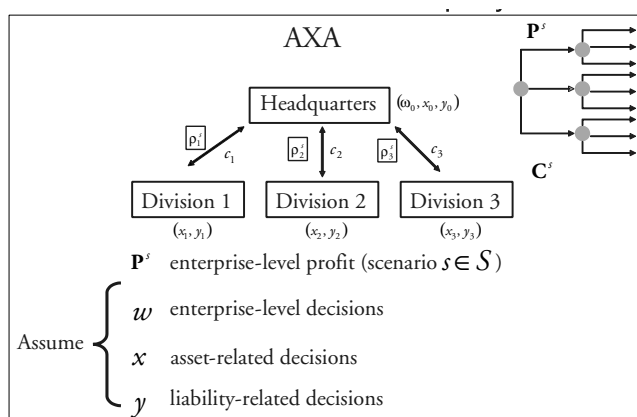


FIGURE 1 Allocation of risk for a global financial company.

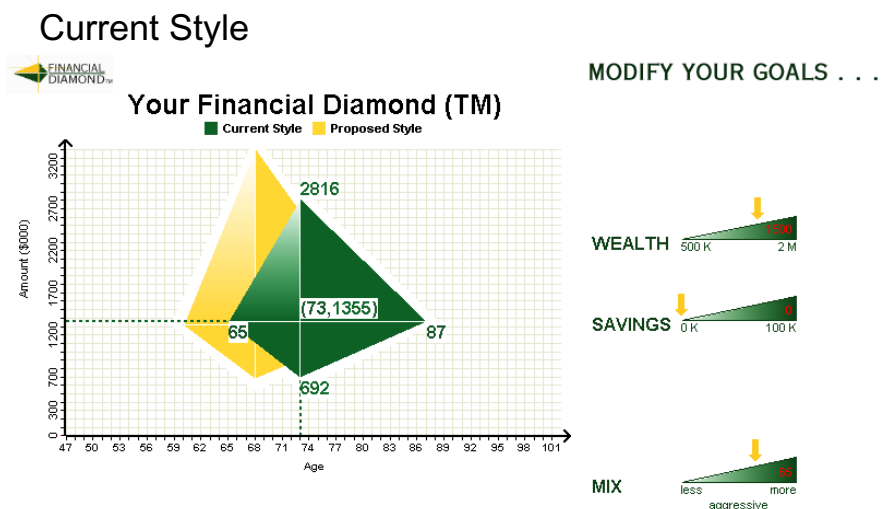
a health-insurance division would have to determine how capital should be spread between the two divisions. Would it be more or less efficient to separate risks? The factors that determine risk are related to the work of both the life insurer and the health insurer. For instance, a health insurer makes less profit on elderly patients, but a life insurer makes more profit if the clients live longer. So in some sense, for an efficient operation, the enterprise risks would be lower for a merged health/life insurance company than for two separate companies. However, current regulations discourage the single organization structure and the sharing of risks between life insurance and health insurance.

## PLANNING TO ACHIEVE FUTURE GOALS AND OBLIGATIONS

Financial engineering methods can be used by individuals planning their personal investment, consumption, and savings decisions. The first step is determining an individual's financial goals, for example, establishing an account for retirement purposes or saving for the purchase of a house or setting up a health care expense account.

Someone aiming for a long-term target, for example, a million dollars for retirement, could use various formulas to calculate how much to set aside each year to reach that goal. There are several ways to simplify this process, such as assigning a market value to cash flows and discounting them back to the present with a risk-adjusted rate to determine the individual's surplus value (similar to a pension plan). However, because cash flows and discount rates are generally uncertain, this approach is not usually used for future savings. Also, the decision involves long periods of time and information regarding the chances of meeting the goal at the designated time is not available.

A more comprehensive approach to help individuals make investment decisions for their future retirement would be to



**FIGURE 2** A graphical representation to help individuals make financial decisions.

project the investor’s wealth for a set of plausible scenarios via a stochastic Monte Carlo simulation. In particular, we can evaluate factors, such as interest rates or inflation, and then simulate each of these factors, along with the respective performance of the assets over a planning horizon. This approach is the basis for a multiperiod asset and liability management (ALM) model for an individual. The planning model helps estimate how much capital is needed to protect against any particular set of circumstances, such as, if the equity market returns are lower than their historical 10.5 percent average annual values. As we did for the pension plan, we evaluate the assets and liabilities (and now goals) under a common framework. This analysis is similar to the analysis insurance companies use to make risk-based capital decisions—how much money the company needs to protect against adverse losses when future targets are uncertain.

### LIMITATIONS OF INDIVIDUALS MAKING DECISIONS UNDER UNCERTAINTY

Financial engineering promises to help individuals make investment decisions under uncertainty. Unfortunately, individuals may not always make the most economically efficient decisions. For example, Princeton University allows employees to set aside up to \$5,000 a year before taxes in a health care expense account; unspent monies are lost at the end of each year. We have observed that many Princeton faculty and staff, despite their greater than average intelligence, rarely make the best decision relative to the costs of overage and underage. In fact, most individuals do not evaluate the opportunity cost of saving from the expense account as rendering an error and, therefore, do not set aside adequate funds in the health-expense account. Even in simple cases, individuals do not address risk consistently. Nevertheless,

the U.S. government is encouraging individual responsibility for life-choice decisions.

Information to assist individuals with investment decisions can be provided based on the ALM model. One graphical approach is called the Financial Diamond™ (see Figure 2). This illustration provides an intuitive way of thinking about the risks of achieving future goals. It allows an investor planning for retirement to set a target goal and then simulate scenarios to arrive at a range of time periods to obtain the goal. A portfolio of assets, such as stocks and bonds, is simulated in conjunction with savings strategies to determine the chances of meeting the goal. The shape of the Financial Diamond™ determines the range of likely outcomes in the future, given the proposed investment and saving strategies. Thus, individuals can evaluate alternative strategies and *see the results*.

Individuals require sound, intuitive methods to understand stochastic outcomes from investment/consumption/savings strategies. Training is also important for selecting the best strategy for an individual. A barrier to improving the health-care system is getting individuals to think about health-related decisions consistently and cost efficiently. Like pension planning decisions, health-care decisions often involve long time periods and substantial uncertainties.

### IMPLICATIONS FOR THE HEALTH CARE INDUSTRY

Two primary challenges relate to financial engineering and the U.S. health care system. The first involves the design of an efficient insurance industry. Deciding how risks across organizations should be diversified involves: (1) structural decisions for the enterprise; and (2) the creation of effective risk-transfer mechanisms. Because capital allocation helps determine an insurance company’s future profitability, a well

diversified company will not only be more profitable but will also be safer than a single-line company. Of course, the combined company must manage its risks and price its products via profitable risk-adjusted values for the enterprise. Thus, the overall structure of an insurance company should be optimized as a single enterprise.

Improving the environment for health care delivery will require determining a market mechanism, choosing a way to transfer risk across different kinds of activities, seeking stability across time, and planning under uncertainty. The insurance industry should be restructured to make greater diversification of risks possible, thus improving the profitability of companies and reducing their overall risks. Given the large losses in the past few years, insurance companies in the health care field are ripe for restructuring. Of course, a primary issue involves the rapid increase in costs for U.S. health care. Even a highly efficient insurance company cannot overcome the barriers created by rapidly rising costs. Nevertheless, well managed insurance companies can improve the environment of the health care industry.

The trend today is to give individuals greater responsibility for managing their affairs in general. The emergence of defined-contribution pension plans (Keoghs, IRAs, etc.) over defined-benefit pension plans is a significant example. A similar pattern may be emerging in health care, with

proposals for increasing rollover health care expense accounts and related arrangements. Individuals will have to make significant investment decisions that may affect their future health. Unfortunately, individuals are not always equipped to make wise decisions when faced with financial choices involving uncertainty.

Financial engineers can assist by creating understandable decision-support systems. Education will also be important; for example, courses on decision making under uncertainty in health care could show how to find the best compromise between costs, efficiency, and possible states of future health. Financial engineers face similar issues on a regular basis. Many insights and methods from financial engineering can be directly applied to the health care industry.

## RELATED READING

- Doherty, N. 2000. *Integrated Risk Management: Techniques and Strategies for Managing Corporate Risks*. New York: McGraw-Hill.
- Laster, D., and E. Thorlacius. 2000. Asset-liability management for insurers. *Swiss Re, sigma* 6: 7–11.
- Mulvey, J.M., and W.T. Ziemba. 1999. *Asset and Liability Management Systems for Long-term Investors: Discussion of the Issues*. Pp. 3–38 in *Worldwide Asset and Liability Modeling*, W.T. Ziemba and J.M. Mulvey, eds. New York: Cambridge University Press.
- Zaik, E., J. Walter, and G. Kelling. 1996. RAROC at Bank of America: from theory to practice. *Journal of Applied Corporate Finance* 9(2): 83–92.

# Engineering Tools and Methods in the Delivery of Cancer Care Services

**Molla S. Donaldson**  
*National Cancer Institute*

For several reasons, cancer care is an especially interesting and challenging field. First, cancer is a major cause of mortality. Second, we have a large, rapidly increasing evidence base of what works, promoted in part by strong patient advocacy groups. Third, as more patients survive for longer periods of time, cancer is changing from an acute condition to a chronic condition. Fourth, despite the existence of comprehensive cancer centers, we need new models of care delivery based on the consistent use of evidence about ways to deliver care that meet the needs and expectations of patients and their families.

This year, 1.3 million new cases of cancer will be diagnosed in the United States. Cancer is the second leading cause of death in the United States, accounting for slightly more than 23 percent of all deaths; large disparities in incidence and mortality rates have been found for different racial and ethnic groups, despite the strong evidence base that has been developed for cancer screening, diagnosis, and treatment (DHHS, 2001). Randomized controlled trials (RCTs)—cancer’s working models of care—are the gold standard in cancer care. RCTs compare, for example, the best known treatments with new approaches. Based on a few simplified assumptions and a very restricted set of variables, RCTs test the efficacy of new agents or combinations of agents. Based on the results, they put forward hypotheses about how well a model will work and its effectiveness in real-world practice. Only 2.5 percent of adults with cancer are ever involved in clinical trials, and participation in trials varies by age (Sataren et al., 2002). One estimate is that more than half of children younger than 15 are in clinical trials and that findings are quickly translated into pediatric oncology practice (Bleyer et al., 1997).

The evidence base on effective cancer treatment and management has been used as the basis of guidelines that include descriptions of the strength of the evidence for treatment and supportive care for most tumor sites by stage. The guidelines developed by the National Comprehensive Cancer Network,

for example, are reviewed annually by standing panels, for a large set of tumor types and are readily available to oncologists (NCCN, 2001) and patients ([www.nccn.org](http://www.nccn.org)). Yet, when researchers studied oncologists’ compliance with these guidelines, they found a lot of room for improvement. For example, the appropriate use of guidelines depends on accurate staging, yet many patients are not accurately staged, not staged at all, or staging information is not available to treating clinicians.

The evidence base is also growing because of major advances in basic biology. The implication of the genome project is that oncologists will no longer classify cancers by tumor site (e.g., lung, prostate, pancreas, etc.) but by genetic transcription errors in the germ line (i.e., in the genetic makeup) or in somatic cells. Previously unexplainable differences in patient responses to therapy for tumors that look alike to pathologists are beginning to be understood in terms of the chemical pathways that produce various proteins. Recent advances have raised hopes that molecular profiles and individual phenotypes can be matched to the most effective therapy, something like matching antibiotics to specific bacteria, but at the molecular level.

With earlier diagnosis and more effective treatment, survival times have increased, sometimes making cancer care more like treating a chronic condition than an acute condition; thus, coordinated follow-up care and the late effects of treatment are becoming a central interest. New therapies may also require sustained treatment. Molecular therapies may mean less toxic and more targeted interventions, but they may also mean that patients will have to take pills for a very long time, perhaps even for a lifetime. Successful treatment will also mean that survivors will live much longer, which will shift the emphasis to follow-up care. Like care for other chronic diseases, long-term follow-up care is complex and requires multidisciplinary, multisetting, coordinated services. In addition, early detection may require long-term chemoprevention. Long periods of time may pass during



which cells change before genetic defects become evident as tumors, and the distinction between prevention and therapy may disappear as detectable genetic errors are treated long before they are expressed as lesions.

The achievements and promise of genomics, proteomics, and molecular discoveries, however, have not been matched by advances in the organization and delivery of services. When patients are diagnosed with cancer, they often find navigating the medical care system a nightmare. A colleague I had not seen for a while said to me, “When I was diagnosed with Stage 3 melanoma, I thought everyone in the health system would swing into action and take care of me. I didn’t realize until much later that no one could or would. It was up to me to make sure things happened and that my doctors knew about it.” She is a patient in a world-class medical center in the Baltimore-Washington area. Despite her education, her considerable resources, her excellent insurance, and her husband who took full-time leave to help her, she was not able to make the system work.

The processes by which a patient accesses care (because of a symptom or for screening), receives a diagnosis, makes decisions, and plans for care in a hospital or outpatient facility or arranges for services from community service and support groups or home care may include initial treatment (such as surgery), follow-up treatment (such as adjuvant chemotherapy or radiation therapy), palliative care, education and information about community services, monitoring as a survivor, and treatment for recurrent disease, continuing primary care, and if needed, timely and appropriate end-of-life care in a hospital, hospice, or home. It may also involve genetic screening, rehabilitation, and support for family and others during and after serious illness. It is easy to understand why when Lee Atwater, campaign manager for Ronald Reagan, was diagnosed with a brain tumor and began treatment, he is reported to have exclaimed, “I need a campaign manager.”

One hears the same complaints from the medical side of health care. *Ensuring Quality Cancer Care*, a report by the Institute of Medicine National Cancer Policy Board, states emphatically, “There is no national cancer program, care program or system of care in the United States” (IOM, 1999). A pediatric oncologist commented, “In the standard model of delivery of care to pediatric cancer patients, the onus of negotiating all aspects of treatment falls on the patient and his or her family” (Wolfe, 1993).

Figure 1 shows a very common model of health care for cancer. In this distributed model, with oncologists practicing in the community, the patient goes from one doctor and laboratory to another trying to integrate sometimes conflicting information. In addition, oncologists have difficulty obtaining information, which results in waste, duplication of effort, and delays; and the primary care physician often has little information about the patient’s treatments. Care is provided in multiple settings, not only at the time of diagnosis and primary treatment, but also over time through later

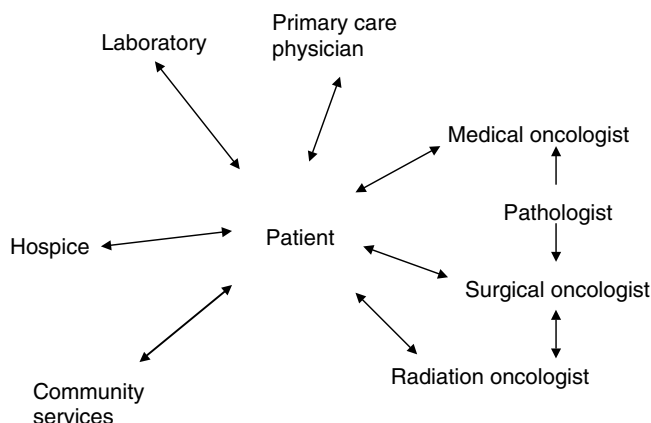


FIGURE 1 Distributed model of health care for cancer.

treatments and follow-up, as needed. Recently, interest has grown in the use of “patient navigator” programs to help patients schedule appointments and keep up with their treatment and progress, but I am not aware that such programs have been evaluated for effectiveness (American Cancer Society, 2002; Christensen and Akcasu, 1999).

Figure 2 shows a different model based on care in a comprehensive cancer center, such as M.D. Anderson, Memorial Sloan Kettering, or Dana Farber, where oncologists and other caregivers are grouped together in one facility. Even in these settings, patients may still go from one caregiver to another, and their records may be quite separate. A care coordinator, such as a nurse oncologist, might help the patient coordinate his or her care, and patients in these centers are more likely to enter clinical trials with stringent protocols and follow-up. In this model, tumor boards or multidisciplinary conferences among oncologists and pathologists develop a plan for patient care. Such conferences, which may be held

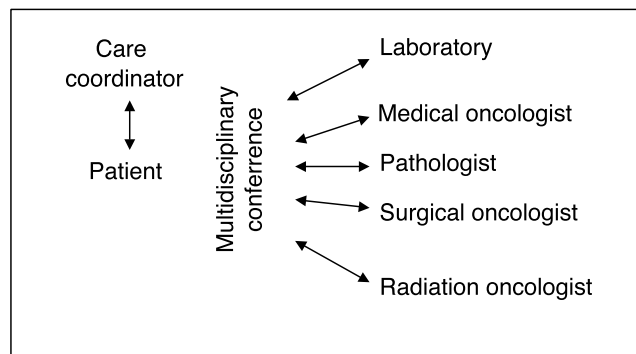


FIGURE 2 Comprehensive cancer center model.

periodically after primary treatment, may, but usually don't, include the patient and his or her family (Joishy, 2001).

Figure 3, a pediatric multisite model, was developed by Dr. L.C. Wolfe and his colleagues when he was at the New England Medical Center (Wolfe, 1993). The model attempts to remedy the boundary problems at the transitions between settings, particularly between the hospital and home, home care and some outpatient care, and outpatient care and inpatient care. When something goes wrong, people do not always know what to do or who to contact.

The model addresses these problems by having the oncologist and the nurse spend time in the hospital together with the patient and then in the outpatient setting and then, as a team, continuing to care for patients who had been in the hospital. To ease the boundary problems between hospital and home care, Wolfe devised an electronic system that enables families to transmit problems and questions to their doctors. O'Connell and colleagues (2000) have critiqued other models of care that try to integrate the hospital-community interface.

Only a few efforts to design better health care delivery systems have been reported. Last week, I attended the annual meeting of the American Society of Clinical Oncology, which drew 25,000 participants from all over the world. Of the more than 3,000 abstracts published, only two reported on programs for improving care. One was a report from France on the number of cancer patients who had attended a nutritional workshop; the other was on the costs and satisfaction of palliative care service in a hospital.

This points up a stark contrast. The knowledge base for the science of cancer care has undergone a radical transformation, but little attention has been paid to ensuring the consistent translation of this knowledge to the health care setting—not just for patients in cancer centers on protocols, but for all cancer patients all the time. Indeed, the assumption seems to be that the results of clinical trials will be translated into practice without error and without specifying how services should be organized and delivered.

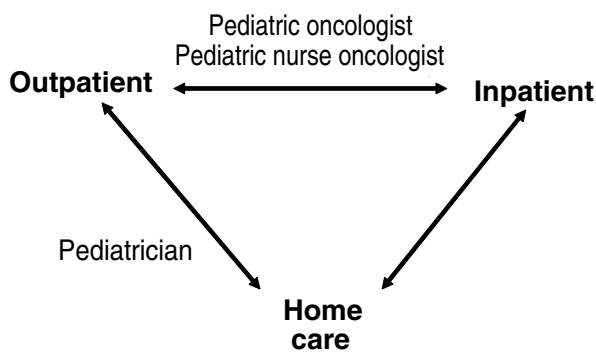


FIGURE 3 Pediatric multisite model. Source: Wolfe, 1993.

The lack of well designed systems can result in the loss of benefits to patients. In many systems, failures can and do occur that could have been addressed by operational engineering. One of the most common consequences is the failure to screen patients. A research project involving health maintenance organizations found that only 50 to 83 percent of women who were expected to have mammographies in a particular year actually had them (Taplin et al., 2002). In Colorado, a risk-management study of lawsuits for failure to diagnose breast cancer found that the average length of delay from symptom to detection or detection to diagnosis was 13.4 months (Marjie G. Harbrecht, M.D., personal communication, April 2001). There were many reasons for the delay, but most of them were system problems. In some cases, the primary care clinics did not have systems for tracking or follow-up. In many cases, individuals thought someone else was following up with the patient. Sometimes a lump found in an exam was not visible on a mammogram, and there was simply no follow-up. Failure to diagnose was also found in the United Kingdom, where there was on average a seven-month delay between detection and definitive diagnosis.

A study in New York hospitals on women who clearly should have had adjuvant breast therapy after treatment for early-stage breast cancer found that in hospitals that were part of the Mount Sinai system, only 18 to 33 percent of these patients, depending on the hospital, received their indicated adjuvant therapy for early-stage breast cancer (Bickell and Young, 2001; Bickell et al., 2000). This was not because of a lack of knowledge. After going through the medical records of these patients and talking to the surgeons, the study found that the surgeons simply did not know what had happened to these patients, they had simply "fallen through the cracks."

Another serious problem is failure to use the evidence base. Dr. Ezekiel Emanuel (2001) at NIH recently reported on an excessive use of chemotherapy for patients in the last months of life. He found that in the last six months, three months, and one month of life, as much chemotherapy was given for tumors that are known to be unresponsive to chemotherapy as for tumors that are responsive to chemotherapy.

Other losses of benefits include: failure to ensure that the necessary information is available at the time of decision making and at the point of care; failure to help with transitions following active treatment; failure to monitor and manage symptoms, including pain; and failure to support dying patients and their families.

A few health systems have reported their attempts to develop an integrated model of care—financially, organizationally, and in data management (Clive, 1997; Demers et al., 1998; Glass, 1998). Other reports include the development of disease-management models of inpatient and outpatient oncology care (Hennings et al., 1998; Piro and Doctor, 1998; Sagebiel, 1996; Uhlenhake, 1995), breast cancer centers (Frost et al., 1999; Kalton et al., 1997), psychosocial support services (McQuellon et al., 1996), support for

long-term cancer survivors (Hollen and Hobbie, 1995), and quality improvement teams (Frank and Cramer, 1998).

A remarkable example of what can be accomplished is the use of logistical engineering in the United Kingdom for cancer services (Kerr et al., 2002; NHS Modernisation Agency, 2001; H. Bevan, personal communication, May 2001). The story began with a major comparative study that showed that survival rates in the United Kingdom were low compared to rates in the rest of Europe and the United States. The study also found that therapy was initiated at a much more advanced stage of disease than expected, which resulted in low five-year survival rates. One reason was the seven- to eight-month delay between (1) detection and (2) diagnosis and staging. Patients were also not able to get the radiation therapy they needed, even though 20 to 50 percent of the appointment slots were not used. By the time patients were seen, the plan of care was often outdated or no longer appropriate. Although the patients' needs were predictable, they did not know what to do once they left the hospital. Further, the percentage of patients referred for abnormal exams or test results who will, in fact, have cancer can be predicted. Hence, services could be designed according to a known demand function.

Using such information, the National Health Service (NHS) made improvements in cancer care services a priority. The program began with 50 teams from nine cancer networks; the program has now been expanded to all 34 networks. The project teams tested more than 4,400 changes in the first 12 months and implemented nearly 550 of them. They instituted multidisciplinary teams that meet regularly to manage the experiences of families and caregivers. They revamped services to meet patient and family needs. For example, tests that used to require three separate hospital visits are now done in one visit.

As a result of this initiative, there was, on average, a 50-percent reduction in time to first appointment and a 60-percent reduction in radiology waiting times. The NHS believes the five-year cancer survival rate can be improved by 10 percent and is reengineering systems accordingly.

Engineering can play a major role in accelerating improvements in the quality and efficiency of cancer care. The unique skills of practicing engineers should be applied in six major arenas of cancer care:

1. Redesign care processes using engineering tools, such as the 80/20 rule, continuous flow, mass customization, production planning, and supply-chain manufacturing.
2. Use information technology to make medical information and patient-specific information available when needed. The goal is to ensure that timely, accurate information is available to clinicians and patients when they need to make decisions.
3. Redesign care to include the patient and family in decision making.

4. Encourage the continuous acquisition of knowledge and skills by all health care workers to support multidisciplinary work. The health care workforce must have the expertise to manage complex tasks, which may require changes in training, education, and protocols and rules about which tasks are permitted. Human factors analysis, which has been used in other industries for crew resource management, shift management, ensuring patient and worker safety, and ensuring high-level, reliable performance in dynamic, high-risk settings, should be applied to the health care setting.
5. Care should be coordinated across settings and over time using any engineering tools available.
6. Measurement of performance and outcomes should be used to improve care. This entails measuring the results of practice and removing the distinctions between research and clinical practice environments so that all patients and patient care can increase our knowledge.

## REFERENCES

- American Cancer Society. 2002. Harlem cancer screening clinic embraces the medically underserved. ACS News Today. Available online at: [http://www.cancer.org/docroot/nws/content/nws\\_1\\_1x\\_harlem\\_cancer\\_screening\\_clinic\\_embraces\\_the\\_medically\\_underserved.asp](http://www.cancer.org/docroot/nws/content/nws_1_1x_harlem_cancer_screening_clinic_embraces_the_medically_underserved.asp). Accessed 9/04/03.
- Bickell, N.A., and G.J. Young. 2001. Coordination of care for early-stage breast cancer patients. *Journal of General Internal Medicine* 16(11): 737–742.
- Bickell, N.A., A.H. Aufses, Jr., and M.R. Chassin. 2000. The quality of early-stage breast cancer care. *Annals of Surgery* 232(2): 220–224.
- Bleyer, W.A., H. Tejada, S.B. Murphy, L.L. Robison, J.A. Ross, B.H. Pollock, R.K. Severson, O.W. Brawley, M.A. Smith, and R.S. Ungerleider. 1997. National cancer clinical trials: children have equal access; adolescents do not. *Journal of Adolescent Health* 21(6): 366–373.
- Christensen, J., and N. Akcasu. 1999. The role of the pediatric nurse practitioner in the comprehensive management of pediatric oncology patients in the inpatient setting. *Journal of Pediatric Oncology Nursing* 16(2): 58–67.
- Clive, R.E. 1997. Update from the Commission on Cancer. *Topics in Health Information Management* 17(3): 10–14.
- Demers, R.Y., R.A. Chapman, M.H. Flasch, C. Martin, B.D. McCarthy, and S. Nelson. 1998. The Henry Ford Health System. *Cancer* 82(10 Suppl): 2043–2046.
- DHHS (U.S. Department of Health and Human Services). 2001. 2001 Cancer Progress Report. NIH Publication No. 02-5045. Bethesda, Md.: National Institutes of Health. Also available online at: <http://progressreport.cancer.gov>.
- Emanuel, E. 2001. Use of Chemotherapy for Advanced Disease. Presentation at the Annual Meeting of the American Society for Clinical Oncology, San Francisco, California, May 2001.
- Frank, J., and M.O. Cramer. 1998. The development of an oncology services performance improvement team. *Journal for Healthcare Quality* 20(6): 26–32.
- Frost, M.H., R.D. Arvizu, S. Jayakumar, A. Schoonover, P. Novotny, and K. Zahasky. 1999. A multidisciplinary healthcare delivery model for women with breast cancer: patient satisfaction and physical and psychosocial adjustment. *Oncology Nursing Forum* 26(10): 1673–1680.
- Glass, A. 1998. Delivery of comprehensive cancer care at Kaiser Permanente. *Cancer* 82(10 Suppl): 2076–2080.

- Hennings, M.N., D.S. Rosenthal, D.A. Connolly, M.N. Pollach, and M.M. Lynch. 1998. Collaborative approaches to purchasing and managing oncology services for a prepaid population. *Cancer* 82(10 Suppl): 2026–2034.
- Hollen, P.J., and W.L. Hobbie. 1995. Establishing comprehensive specialty follow-up clinics for long-term survivors of cancer: providing systematic physiological and psychosocial support. *Supportive Care in Cancer* 3(1): 40–44.
- IOM (Institute of Medicine). 1999. *Ensuring the Quality of Cancer Care*, M. Hewitt and J.V. Simone, eds. Washington, D.C.: National Academy Press.
- Joishy, S.K. 2001. The relationship between surgery and medicine in palliative care. *Surgical Oncology Clinics of North America* 20(1): 57–70.
- Kalton, A.G., M.R. Singh, D.A. August, C.M. Parin, and E.J. Othman. 1997. Using simulation to improve the operational efficiency of a multidisciplinary clinic. *Journal of the Society for Health Systems* 5(3): 43–62.
- Kerr, D., H. Bevan, B. Gowland, J. Penny, and D. Berwick. 2002. Redesigning cancer care. *British Medical Journal* 324(7330): 164–166.
- McQuellon, R.P., G.J. Hurt, and P. DeChatelet. 1996. Psychosocial care of the patient with cancer: a model for organizing services. *Cancer Practice* 4(6): 304–311.
- NCCN (National Comprehensive Cancer Network). 2001. *Practice Guidelines in Oncology*. Available online at: [http://www.nccn.org/physician\\_guids/index.html](http://www.nccn.org/physician_guids/index.html).
- NHS Modernisation Agency. 2001. *The Cancer Services Collaborative: Twelve Months On*. Presented to the National Patients' Access Team, Leicester, England, July 18, 2001.
- O'Connell, B., L. Kristjanson, and A. Orb. 2000. Models of integrated cancer care: a critique of the literature. *Australian Health Review* 23(1): 163–178.
- Piro, L., and J. Doctor. 1998. Managed oncology care: the disease management model. *Cancer* 82(10 Suppl): 2068–2075.
- Sagebiel, R.W. 1996. The multidisciplinary melanoma center. *Surgical Clinics of North America* 76(6): 1433–1439.
- Sateren, W.B., E.L. Trimble, J. Abrams, O. Brawley, N. Breen, L. Ford, M. McCabe, R. Kaplan, M. Smith, R. Ungerleider, and M.C. Christian. 2002. How sociodemographics, presence of oncology specialists, and hospital cancer programs affect accrual to cancer treatment trials. *Journal of Clinical Oncology* 20(8): 2109–2117.
- Taplin, S., M. Manos, J. Zapka, A. Geiger, M. Ulciskas-Yood, S. Weinman, S. Gilbert, and J. Mouchawar. 2000. Frequency of Potential Breakdowns in Screening Implementation among Women with Invasive Cervical and Late Stage Breast Cancer. Presented at the 8<sup>th</sup> Annual HMO Research Network Conference, Long Beach, California, April 9–10, 2002.
- Uhlenhake, R. 1995. Is patient-focused outpatient cancer care on target? *Journal of Ambulatory Care Management* 18(4): 32–42.
- Wolfe, L.C. 1993. A model system: integration of services for cancer treatment. *Cancer* 72(11 Suppl): 3525–3530.



# Patient Trajectory Risk Management

Charles Denham  
*HCC Corporation and  
Texas Medical Institute of Technology*

This paper addresses the notion of risk trajectory of individual patients and the resultant aggregate risk trajectory of the healthcare enterprise caring for populations of patients. It also describes the use of various engineering concepts applied to medicine.

In the late 90's, working with a team from the Institute for Healthcare Improvement (IHI) and Premier Inc. a group purchasing organization of 1,800 hospitals we focused our attention on medication management. The project involved collaborators from the Cleveland Clinic, Partners System, Harvard Medical School, Mayo Health System, a number of frontline hospitals and leading experts. Our goal was to identify the idealized design for medication management to reduce adverse drug events, a major cause of preventable death and disability in U.S. hospitals. To do that, we first had to identify achievable world-class performance, then the "is state" of frontline hospital performance, and finally processes and technologies that would enable us to close the gap between the two. We were surprised by our findings and gratified by the opportunities they revealed.

Engineers are used to using process impact evaluations, risk analyses, and pattern recognition methods, however these are new to the practice of medicine at frontline institutions. Clearly, medicine has much to gain from engineering, and many benefits have yet to be realized.

The Institute of Medicine report, *Crossing the Quality Chasm* (IOM, 2001), proposes that we must redesign healthcare so that it is patient centered, evidence based, and systems focused. As such we must have a much better understanding of "integrated performance"—i.e. operational, clinical, and financial processes and outcomes—of an individual patient's care delivery through a healthcare episode. We must look at the performance/risk trajectories of common patient treatment process paths and examine the contributive impact to enterprise wide performance. Hospital administrators must step back from their traditional vertical business unit view and take into account their patient

populations as they move through those vertical units so that they can recognize operational innovations that can eliminate process segment failures.

The game of golf provides a powerful metaphor. The desired outcome is to deliver the ball to the hole. For a given link one golfer may take eight strokes and another might take three. Both reach the goal if the outcome measure was just "ball in hole," however one expended more energy and time than the other. The golfer taking eight strokes has increased the risk of having mishap along the way. In a similar way, if a patient requires two or three extra days of care, the risk of having an adverse event is greatly increased due to greater exposure to the inherently dangerous hospital environment.

To come up with an ideal design for medication management, we first mapped the clinical and operational processes involved in medication use. Next, we considered the products, services, and technologies involved that enable best or better practice (technologies might include process reengineering tools, for example). Then we identified their impact on the risk of adverse events and whether they closed the gap between typical performance and best achievable performance.

Traditionally administrators and clinicians have been trained to define a medication error by violation of one or more of the "five rights"—the right patient, the right drug, the right time, the right dose, the right route. Such errors occur with virtually every patient admitted to hospital. Dr. David Classen a noted patient safety expert on our team demonstrated that the overlap between error and harm minimal using this definition of error—only a small fraction of harm is caused by error as defined by the "five rights." A great number of errors do not cause harm, and more importantly a number of adverse drug events that cause death, disability, or require treatment would not normally be counted using the classical "5 rights" framework.

During the idealized design process, we worked with a

number innovative healthcare technology suppliers; 70 to 80 percent of them were attuned to error. Few focused on harm. The deeper we explored adverse drug events it became more and more apparent that distinguishing between error and harm was critical. We focused on the most common causes of adverse drug events including transition zones between care teams and high impact intravenous infusion events. We did not ignore errors without harm, but we did not focus on them. After completing about 80 percent of a thorough, evidence-based review of integrated care and operational processes, with the guidance of a number of experts, the opportunities for mitigation started to become clear.

Subsequently IHI led a number of very successful hospital collaborative initiatives using a “trigger tool” medical record review framework that helped identify adverse drug event (ADE) risk and performance gaps.

We studied smart the Alaris smart infusion pumps that have now have the ability of capturing and even preventing the most serious IV adverse events, clearly a technology advance that will deliver dramatic speed to impact in reduction of ADEs.

To illustrate the error-harm gap and the notions of patient trajectory and hospital risk trajectory we used the example case of anticoagulation management with our teams and collaborative groups. Anticoagulant drugs are often very poorly managed by clinicians and patients resulting in severe adverse drug events. In fact this is the area of the most common drug related malpractice claims and awards.

Certain engineering concepts have great application to medicine. When engineers evaluate airplanes, they examine and discuss its performance envelope. We applied this concept to the management of anticoagulation. Warfarin is an anticoagulant drug used to manage patients. Its danger lies in the fact that the therapeutic envelope of safety relating dose to effectiveness and complications may change or shift. The patient’s diet (i.e., wine or vitamin K consumption), or liver function can shift the therapeutic window. The therapeutic envelope is always changing, posing huge risk to patients for overdose or under dose leading to clotting or bleeding disorders. Currently physicians try to manage patients undergoing anticoagulation by trying to interpolate and extrapolate the relative patterns of multiple lab values and historical factors. Application of the performance envelope delivers terrific pattern recognition opportunities.

We also demonstrated the use of other aviation tools to communicate performance. For instance we created a mock up “digital dashboard,” illustrating how clinicians could recognize patterns, access relevant protocols, and in the case anticoagulation decide how to manage the patient.

In collaboration with one of the nations leading anti-coagulation experts we presented an example case study of a young adult admitted for treatment of a defective heart valve who experienced 11 typical and different adverse drug events, none of which was caused by a medical error (using the 5 right classification) and none of which would have

been picked up by the typical methods we use to catch medical errors. Dose adjustments unique to the patient’s condition and omissions due to missed laboratory values would not typically be classified as a medication error. The patient eventually has a stroke. In this case, the potential for recognition of the risk for adverse events would have been picked up by a computerized physician order entry (CPOE), which integrates order entries with laboratory and historical information. We know from other studies that CPOE can reduce adverse events dramatically.

In the future, we will have a decision-support systems that enable clinicians who are not specialists in anticoagulation to put that part of the treatment in the hands of a pharmacy team while being able to monitor potential adverse events. That is precisely what an information integrating device that pilots use called a flight director does. Flight information is provided as an input, the crew makes sure all the instrumentation is synchronized and the director follows the plan. If the workload becomes too heavy, the autopilot can be turned on.

Today, 16 different types of specialists prescribe anti-coagulants; none are specialists in anticoagulation. Orthopedists, internists, and cardiologists are all administering the drug and are responsible. The risk trajectories such patients are not being managed well and adverse events such as preventable strokes and bleeding related complications are occurring in epidemic proportions.

We used a mockup of the digital dashboard to study the young adult described earlier. His medical history and his recent history revealed a number of health problems that predisposed him to a bleeding and clotting disorder that made anticoagulation drugs extremely dangerous for him. When we asked what might have been done differently, we found that when the care data is reconstituted in a graphic it would allow us to recognize a pattern. Had the data presentation been like that presented in aircraft instrumentation we would have seen the window of safety narrowing and prevented catastrophe. Instead, we are caught by surprise driving from a view through the rear view mirror.

Clinicians could be assisted by innovations that make patterns simpler to recognize. The average doctor in an intensive care unit can interpolate three or four trends. A patient on a respirator who is very ill might have could have 60 pertinent trends. Our slowest cognitive capability is in processing data, which is exactly what computers do well.

Before retiring to focus full time on emerging technologies, I was a radiation oncologist with a very large practice, and I managed all of my patients all the way through therapy. I had a high volume of patients with common diseases, including colon, breast, lung, and prostate cancer. I had to navigate between the response of the tumor to radiation therapy and the response of normal tissue. I had to manage that patient through a safety window that would become narrower and narrower as we proceeded through care. As the dose was increased, the risk for a host of complications would increase

and continue intensify through out treatment. We knew that every treatment decision had a risk-benefit balance to it. Every patient had a unique trajectory based on historical data and how certain factors had impact as therapy progressed. These patients were managed based on tacit knowledge—we could tell when a patient was headed for trouble, we could link this to certain parameters.

In working with healthcare technology suppliers, we have found that an evidence-based, patient centered, and systems performance targeted approach to “enabling” best or better practice allows innovations to be developed that improve clinical performance and reduce risk. In addition, they often deliver improved enterprise wide performance as a by-product of improved patient specific performance.

If we had continuity of information with pattern recognition support we could examine the risk trajectory of patients with very complex disorders and create scenarios and real time forecasts, as we do in aviation. In the future, we might ask a medical student to use a computer model to run

scenarios for a specific patient. We could graphically portray patterns and risk trajectories to assist in decision making before patients get into trouble. Is the patient’s cardiac function adequate? Will his kidneys clear everything? What-if scenarios can be run before events cascade.

Engineers already provide wonderful computational support and pattern recognition solutions for many industries. These technologies will offer physicians a terrific opportunity to “think through” treatment scenarios. With an appropriate decision-support system, we could apply the lessons learned in other industries, such as aviation and aerospace, to complex medical problems. The principles of data analysis from engineering could be tremendously beneficial for health care.

## REFERENCE

IOM (Institute of Medicine). 2001. *Crossing the Quality Chasm: A New Health System for the 21<sup>st</sup> Century*. Washington, D.C.: National Academy Press.





# Deploying Resources for an Idealized Office Practice: Access, Interactions, Reliability, and Vitality

Thomas W. Nolan  
*Associates in Process Improvement and  
Institute for Healthcare Improvement*

The goal of our initiative is to create an idealized design of clinical office practices (IDCOP) that offers the best possible solutions to the health care practice needs of our customers. When implemented, these solutions should lead a visiting patient to say, “They give me exactly the help I want (and need) exactly when I want (and need) it.” To accomplish this goal, we have to improve measures associated with: clinical outcomes; patient satisfaction; finance; and staff satisfaction. To simplify and further systematize the systems that emerge from IDCOP, we have developed a framework of four “themes” to guide the redesign processes as a whole: access, interactions, reliability, and vitality.

**Access.** Timing is an essential component of health care. *When* things happen is almost as important as *what* happens. Of all forms of timing, patients almost certainly value most the timing of entry into the system—getting *to* care when the care is needed. *Care* in this context does not mean only encounters or visits. It means all appropriate forms of interaction, including access to information, support, dialogue, reassurance, treatment, and supplies, as well as all possible routes of delivery—not just face-to-face meetings, but also electronic, print, and other media of exchange.

**Interactions.** Health care is fundamentally interaction. Interaction is not the price of or vehicle for care; it *is* the care. Those who regard health care as a list of resources—people, medications, machines, technologies, and so forth—are merely listing the “inert” ingredients that become care only when they are combined in interactions between patients and the system. The quality of care is the quality of interaction among resources, not the quality of the resources per se.

**Reliability.** Reliability involves ensuring an exact match between knowledge and activity in the IDCOP practice. Ideally, “all and only” effective and helpful care is given. The IDCOP practice, therefore, aims *always* to give care that can

help a patient and *never* to give care that harms or cannot help a patient. Reliability is the conscious attempt to avoid the defects in health care that the Institute of Medicine Roundtable on Quality summarizes as “overuse, underuse, and misuse” of care. (The Roundtable defines misuse as errors in care and threats to patient safety.)

**Vitality.** IDCOP aims for a sustainable design. The new system would be financially viable and would provide a great workplace. In other words, the demanding performance standard is not realized at the expense of those who work in the practice and depend upon it for their livelihood. Vitality also implies renewal—continual innovation and improvement. The IDCOP practice is not a fixed, solved system; it is a learning organization with the capability, agility, resilience, and will to change over time as desires, environments, and knowledge change.

Each of these themes or aspects of IDCOP requires certain activities, some familiar and some new. One of the initial steps to redesigning the system as a whole is the systematic examination of the current premises and beliefs concerning the activities performed and the people who perform them. Meeting each of the goals requires some resource deployment and scheduling. To achieve excellent *access*, the demand for visits and other interactions must be estimated beforehand, and capacity, for example for appointments, must be available to meet the demand. Conceiving of care as *interactions* between the patient and the system via multiple media means that resources must be deployed to enable these interactions. *Reliability* requires an exact match between knowledge and activity in the practice, knowing the activities that will meet the needs of patients and ensuring that these activities are performed in an orderly manner and at the proper time. The activities that contribute to the *vitality* of a practice, such as training and process redesign, might easily be put off in the face of pressing daily demands, but

these activities are essential. Hence, time must be scheduled for them.

Besides helping with the daily deployment of resources, the development of a master schedule for the practice will facilitate the fundamental rethinking of the design of the practice. The following three tasks serve as a guide to the deployment of resources consistent with the IDCOP themes:

1. Understand and define the *work* involved in caring for persons who depend on the practice.
2. Assemble a *team* of people and resources to match the work.
3. Develop a *repetitive master schedule* to optimize the use of resources relative to the needs of the population.

Defining the work involves describing activities in the practice and then assessing them in terms of the four themes. The activities can then be adjusted to ensure that the practice has all four characteristics and the appropriate clinician matched with the work. Once the work and appropriate team have been identified, the practice can match the work to the members of the team on specific days of the week using a repetitive master schedule.

## REPETITIVE MASTER SCHEDULE

The work of a clinical practice is varied and complex—no two patients are alike, insurance companies have different requirements, and the external environment is changing rapidly. Designing an IDCOP practice is impossible unless some sense of order is established in the midst of increasing demands and varying conditions. Developing and using a repetitive master schedule is one method of establishing order.

Although the work varies, every practice has a natural rhythm—the length of time after which the work begins to repeat. Staff in a primary care practice often cite one week as the repetitive period. Up to a point, the work done in one week is similar to the work done the next week. Of course, the rhythm in a practice is also influenced by shorter periods, such as days, and longer seasonal periods that must also be taken into account.

The practice must first establish the period for which a master schedule will be designed. For purposes of discussion, let's assume the period is one week. That means that a master schedule for a "typical" week can be used with minor adjustments for any week. The definition of the repetitive period simplifies the task of deploying the resources of the practice because the schedule is built only for a short period of time.

Once the period has been chosen, a master schedule can answer the questions of what work will be done, who will do it, when they will do it, and where they will do it. An IDCOP practice calls for forms of interaction in addition to one-on-one visits with the doctor. Who will be using e-mail? Who will provide chronic disease management and review registries? When will training and staff development take place? The master schedule should provide answers to these questions.

The slogan for a master schedule with a period of one week is "do today's work today." Although there is some overlap in each day's work, Tuesday's work will not be exactly the same as Thursday's. The practice may hold a group visit on Tuesday, for example, and review the chronic disease registries on Thursday. Daily work should be completed on the day it is scheduled.

"Open access" requires that patients be scheduled within the master schedule cycle. Hence, practice-patient interactions are a very large component of the master schedule. Backlogs are defined as work that is not scheduled or completed within the master scheduling period. Consider a patient's initial appointment in a behavioral health practice. Because the initial appointment requires that multiple providers see the patient during the visit, a practice may designate one morning a week for initial appointments. The "open access" philosophy requires that new patients be seen within a week. Backlogs of two or more weeks for new patients are inconsistent with the repetitive master scheduling approach.

Open access and repetitive master scheduling are based on the general concept of "continuous flow," which requires that the amount of work be predicted and resources deployed to complete the work in a specified period of time without backlogs. Continuous flow principles apply to weekly scheduling and even daily scheduling. The physician who sees a patient and completes the chart before moving on to the next patient within the specified activity cycle time is using continuous flow.

Many practices already use some aspects of master scheduling. Practices with open access to visits and phone calls are well along in the development of a repetitive master schedule. For practices that wish to develop a master schedule the following steps should be considered:

1. Implement an open access system for visiting patients.
2. Define the care process for each of the top diagnoses to use as input to the master schedule. Include in the definition the desired time between when a patient first presents with the problem and when an effective plan of treatment is begun.
3. List the services required to accomplish the themes and the internal processes required to support these services.
4. Devise a master schedule of one to two weeks that addresses who, what, where, and when for the services and processes enumerated above.
5. Use the following metrics to assess success in executing the master schedule:
  - a. the degree of completion of the schedule and the reasons for not achieving it
  - b. the percentage of time physicians are doing work that only they can do or that only they are legally allowed to do
  - c. the time from patient presentation to treatment for the top 10 diagnoses