MEASURING HEALTH CARE QUALITY:
A STATISTICAL QUALITY MANAGEMENT VIEWPOINT

James C. Benneyan
114 Marston Hall
Industrial Engineering / Operations Research
University of Massachusetts
Amherst, MA 01003 U.S.A.
(617) 522-0616
benneyan@ecs.umass.edu

ABSTRACT

This paper discusses and illustrates the evaluation of health care systems from the perspective of statistical quality management, in general, and the philosophy of the late quality pioneer, Dr. W. Edwards Deming, in particular. Although many health care organizations have adopted some form of continuous quality improvement (CQI), the use of statistical quality management (SQM) and basic statistical process control (SPC) has yet to become widespread. Particular attention therefore is given to the Deming approach to managing for quality, the role of SPC within this philosophy, and the motivation for applying SPC to health care quality indicators, clinical pathway variances, so-called "report cards", and various other metrics required by regulatory bodies and measurement systems.

INTRODUCTION

Statistical quality management and SPC can compliment traditional approaches - both philosophically and statistically - in the study of almost all medical systems, including administrative processes, clinical processes, risk management indicators, infection control, and laboratory processes, where failing not only to use correct quality methodologies, but to use them correctly, can result in considerable cost, liability, and human consequences. For example, medical facilities should use statistical control charts to study mortality rates, Cesarean-section births, lengths of stay, infection rates, needle sticks, mammography imaging, laboratory accuracy, prescription errors, patient falls, hospital admit rates, HCFA case indices, and other metrics required by regulatory bodies and "report cards".

The remainder of this paper summarizes the application of statistical quality control to health care, with particular emphasis on the role of process-oriented measurement systems, as advocated by Dr. Deming, in contrast to current trends to evaluate and rank individual health care providers via comparative outcome metrics. Many aspects of quality management and SPC, of course, are beyond the scope of an introductory paper. Unfamiliar readers therefore are urged to explore additional materials on the principles and philosophies of quality management, Dr. W. E. Deming, and statistical process control. Understanding these topics is absolutely essential.

THE DEMING QUALITY PHILOSOPHY IN HEALTHCARE

Health care organizations aspiring to adopt quality management must recognize that the Deming approach is not simply an additional program which can be implemented in a short span of time, instead requiring continual and ceaseless work, study, and improvement throughout the entire future existence of the organization. The Deming philosophy, in fact, requires a fundamentally different ap-
approach to viewing business and clinical processes, as well as a dramatic shift in the primary roles and responsibilities of all layers of management, from the board of directors to line supervisors. While full discussion of its application to health care is not possible in this paper, additional information on Deming’s approach to managing for quality can be found in several of the listed references [12,3,1].

Although numerous health care organizations claim to operate according to quality management methodologies, few actually have demonstrated a thorough comprehension and total commitment to this approach. Deming’s philosophy, in fact, challenges traditional approaches to managing any organization, health care or otherwise, and impacts every aspect of the organization, including business objectives, the use of data, the role and authority of employees, job design, compensation and reward structures, competition, cooperation, and, perhaps most importantly, management responsibility.

For example, within a Deming-oriented approach, one primary responsibility of health care managers is to continually ensure process improvement and remove barriers which prevent employees from doing their jobs better. These management responsibilities and barriers to health care quality are summarized in Dr. Deming’s well-known “14 Points” and “7 Deadly Management Diseases” [3], which are equally applicable to health care, service, manufacturing, or any other type of process. In the United States, for example, the Joint Commission on Accreditation of Health Care Organizations (JCAHO), the National Committee for Quality Assurance (NCQA), HCFA, and other regulatory bodies require that hospitals, HMOs, and other providers be engaged in CQI activities, including the application of statistical methods such as SPC to critical clinical and non-clinical processes.

This new management approach is based on, among other things, knowledge of statistics, philosophy, psychology, and human motivation and requires that the day-to-day operating philosophy throughout an organization radically change toward focusing on continuous improvement in quality and service. For health care providers, this means, among other things, transitioning from traditional quality assurance orientations focused on inspection, reporting, and regulatory adherence, to quality improvement orientations focused on process study, continual improvement, designing and re-designing better systems, and reduced bureaucracy. An essential management role within CQI, therefore, is to significantly refocus health care personnel, capital, and other resources in order to develop an understanding - physically, statistically, and otherwise - of the performance of critical processes.

ROLE OF PROCESS-FOCUSED DATA

Within this quality management and quality improvement philosophy, all health care systems should be recognized as processes which exist and inherently vary across time. When viewed longitudinally in this fashion, these processes will exhibit various amounts of temporal natural (i.e., consistent) and unnatural (i.e., inconsistent) variability. One ultimate objective of quality management, in fact, is to study, control, and reduce this variation - and to otherwise improve process performance - via the various methodologies within the fields of quality engineering and statistical process control. For example, in his many speaking engagements Dr. Deming repeatedly stated

“If I had to reduce my message to management to just a few words, I’d say it all had to do with reducing variation”.

Deming also emphasized the importance of developing “an appreciation of the system, and the optimization thereof”. The role of measurement and analysis, therefore, must change significantly, with data now being gathered much more frequently in order to develop fact-based information about process performance, over time, almost as if monitoring its pulse, rather than basing decisions on intuition or anecdote. These data then are evaluated via SPC, in the manner discussed below, in order to understand and improve process performance. This collection and use of data is a fundamental change away from traditional purposes of reports, documentation, and evaluation of individuals.
In studying Deming's approach to managing for quality, a common theme emerges that it is management's job to study process performance and to change situations which inhibit process improvement, rather than to try, naively, to evaluate individuals and reward or penalize them for circumstances often beyond their control. Dr. Deming, in fact, was quite clear that a fundamental purpose of collecting and examining data is to understand and improve a process, and that these data should not be used for scrutinizing the performance and value of individuals, an historic management behavior which deters optimal process performance. This general quality management viewpoint has significant implications for the collection and analysis of health care data.

HEALTH CARE OUTCOME METRICS VERSUS PROCESS QUALITY MANAGEMENT

Currently, a typical manner of reporting health care data might be to summarize several aggregate figures, such as shown in Table 1. Similar tables historically have been the general manner of reporting risk management, infection control, and other required data to various regulatory and accreditation organizations. These figures then might be used in an attempt to compare hospitals, departments, or specific physicians, perhaps using traditional static hypothesis tests. Sometimes a certain amount of blame may be cast, either directly or indirectly, on those with lower figures than others. In fact, an increasing number of standardized health care metrics, popularly referred to as "report cards", are being developed for these purposes.

These include NCQA's Health Plan Employer Data and Information Set (HEDIS), JCAHO's similar hospital-focused metric set, Henry Ford Health System's Consortium Research in Indicators of System Performance (CRISP), as well as measures being developed by the Jackson Hole Group, the Foundation for Accountability (FAacct), HCFA, the Maryland Project, and numerous others. The evolution and current state of many of these measurement systems are discussed by Reiley [17] and Carey [11]. Some of these indicators focus, or attempt to focus, specifically on individual physician performance. For example, US Healthcare issues performance reports on each physician's appointment access, office waits, bed-side manner, staff courtesy, and overall medical care (!). Additionally, the basic idea of the much-debated Coronary Artery Bypass Graft (CABG) mortality metrics, developed in Pennsylvania and New York state, is that by publishing survival rates of individual heart surgeons, the public can somehow identify the "good" doctors versus the "bad".

<table>
<thead>
<tr>
<th>Hypothetic Health Care Quality Indicator Metric</th>
<th>1995 Year-to-Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
</tr>
<tr>
<td>Cesarean-Section Births</td>
<td>281</td>
</tr>
<tr>
<td>Maternity Length-of-Stay</td>
<td>2077 total days</td>
</tr>
<tr>
<td>Needle Sticks</td>
<td>34</td>
</tr>
<tr>
<td>Surgical Nosocomial Infections</td>
<td>28</td>
</tr>
<tr>
<td>Clinical Pathway Variances</td>
<td>204</td>
</tr>
<tr>
<td>Prescription Errors</td>
<td>54</td>
</tr>
<tr>
<td>Patient Falls</td>
<td>173</td>
</tr>
<tr>
<td>Patient Satisfaction</td>
<td>4.7</td>
</tr>
</tbody>
</table>

Typical Traditional Semi-Annual Reporting of Health Care Data, Primarily for Reporting Purposes

Table 1
Although the increased use of outcome-based measurements has many advocates, several problems from a Deming perspective exist with their present use, which in the long run may hinder efforts to improve health care systems via CQI. Moreover, many agree that none of these systems are likely to measure quality very accurately, including Dr. Wennberg, Director of the Center for Valuative Clinical Sciences at Dartmouth, who recently stated [18] "I think some of the proposals are quite naive." Several survey results also have questioned whether purchasers actually believe these efforts are useful, with FAcect even acknowledging that "most of the measures in use today are not relevant or meaningful to purchasers and consumers" [11] and with a recent article in Modern Healthcare [13] stating

"The development of data that will enable employers and other healthcare purchasers to measure and compare treatment outcomes is akin to the search for healthcare's Holy Grail."

In the case of the CABG mortality measures, for example, raw unadjusted data can make the "best" heart surgeons at the best hospitals appear the "worst", as they might operate primarily on the toughest surgical cases. Several other authors [6,17,11] have discussed related problems with the current trend towards reliance on standards and simple outcome metrics. For example, epidemiologist David Birnbaum recently stated [9] "The proliferation of administratively simple but misleading report card metrics is a prime example of the worst in today's practices". Perhaps most importantly, all of these various objective metrics and "report cards" further the dangerous trend that health care quality assurance may become primarily outcome-focused, de-emphasizing the process focus and detailed knowledge necessary for continual improvement. Similarly, performance to external or internal standards, sometimes arbitrary - such as emergency room waits, falls, prescription errors, and access - also provide very little process information by which to improve. Even if these types of information might be useful for consumers, employers groups, and other purchasers, they are highly inadequate for providers wishing to improve and remain competitive.

If the current trend continues, therefore, considerable danger exists that these indicator systems will serve merely as means of accreditation and marketing, rather than to drive meaningful quality improvement [17]. There is a clear distinction between measuring for judgment and measuring for improvement, and many have suggested that the current state of report cards may be inadequate for helping health care providers measure any extent to which they may be improving or deteriorating. For example, Dr. Carey [11] recently stated that "The missing link in most [healthcare] CQI programs is measurement for improvement, as distinguished from measurement for judgment."

In other industries, this style of management historically has been referred to as defect detection (rather than defect prevention), management by blame, or the carrot and stick mentality, and Deming's argument appears equally applicable to health care. Unless the situation improves, this style of management will persist in health care, with a growing danger of simply using more process data in the old mindset, thinly disguised as a "QM program". Presumably, requiring performance report cards, as a way of somehow comparing one health care facility or provider to another, will drive competition and therefore, it would seem, improvement. However, a considerable body of research questions whether these efforts might in fact produce the opposite result. For example, Alfie Kohn, a leading researcher on the effects of competition, has stated that [14]

"The growing clamor to make hospitals... more 'competitive' raises once again the question of whether this goal has anything to do with reaching excellence, or whether we have simply blurred the two ideas... Whether the first leads to the second is a legitimate question, but the answer, contrary to conventional wisdom, is that it almost always does not. The trouble with our schools, for example, is that they are much too competitive, which helps explain why so little learning is taking place."

A quality management view of a process recognizes that the majority of the inherent randomness of the overall system is a direct result of the complexity of the process itself, and not the result of poor performance by any particular employee. All individuals working within the organization, including
physicians, nurses, managers, supervisors, laboratory technicians, and the like, therefore must be free to go about their jobs in an environment that is free of fear, evaluation, rewards, and punishment, whether arbitrary or not. Similarly, the imposition of higher quality standards without any change in the process is contrary to managing for quality and contributes to an environment which is, as Dr. Deming declared [12], "... unfair [and] destructive. We cannot afford this nonsense any longer."

MEASURING HEALTH CARE PROCESS QUALITY: THE ROLE OF SPC

As an alternative, all health care systems should be evaluated, as dynamic processes which vary over time, via well-designed quality control charts. This more engaged, informative, and pro-active use of process data is at the very foundation of the philosophy of statistical quality management. Quality control charts are chronological displays of process data used to help understand, control, and improve a system, here health care processes. The three general objectives and uses of control charts, described in greater detail in textbooks by Montgomery [16], Banks [2], and Benneyan [3], focus on:

1. **Establishing** consistent and predictable health care processes by identifying and removing causes of *unnatural* process variation and bringing them into states of statistical control;

2. **Maintaining** standardized, consistent, and reliable health care processes by identifying and preventing future causes of *unnatural* process variation; and

3. **Continually improving** processes by studying their physical and statistical behavior, identifying and removing causes of unacceptable performance, and reducing *natural* variation.

Without getting into statistical specifics, several common types of control charts exist, each being appropriate for different types of processes and each constructed using different formulas. Process data now are collected in the format shown in Figure 1, much more frequently and closer to the continuous manner in which they are produced, rather than infrequently as previously shown in Table 1. These observed process data, such as the number or rate of Cesarean-section births per week, then are plotted soon after they become available, on an appropriate control chart, in the manner shown in Figure 2. Three horizontal lines also are plotted, called the center line (CL), the upper control limit (UCL), and the lower control limit (LCL), which define the central tendency and range of natural variation, and are used to detect if the underlying process performance has changed.

![Longitudinal Data Collection Format for Statistical Process Control](image)

*Figure 1*
Values outside the control limits, process shifts, and other non-random events indicate that non-systemic causes exist which should be investigated, identified, and removed in order to achieve a single consistent process. Furthermore, if no such special cases exist, then no occurrences should be considered as positive or negative exceptions, regardless of any standards superimposed by management or an outside regulatory body. For example, in the above maternity process, only weeks 10 and 32 should be considered as possibly exceptional from the norm, whereas all other births appear to be part of the same delivery process. Maternity length of stay might be examined in a similar fashion.

Note also that a minimum of at least 25 to 35 subgroups of data, collected as shown in Figure 1, are necessary in order to conclude that a process is in statistical control, which has radical implications on how data should be gathered and analyzed by quality-oriented health care organizations. Even using quarterly data, a provider may have to wait 6 to 9 years to determine if a critical process is stable or contains abnormal events. The typical use of data, by many health care organizations today, therefore is almost worthless for the purpose of statistical quality control. Annual, semi-annual, and quarterly reporting of large amounts of aggregated outcome-oriented data must be replaced with very frequent and much smaller subgroups of process-oriented data on statistical control charts.

SOME OTHER EXAMPLES

Many other key health care issues are prime candidates for analysis via SPC. As one example, hospital accrediting and governmental bodies are urging that quality improvement methodology be applied to surgical wound infections, pneumonia, catheter infections, and other types of nosocomial infections acquired during hospitalization, which together result in approximately 8.7 million additional U.S. hospital days and 20 thousand deaths nationwide per year. In fact, a paper by several authors from the U.S. Center for Disease Control [15] recently stated that

"Many of the leading approaches to directing quality improvement in hospitals are based on the principles of W. E. Deming. These principles include the use of statistical measures designed to determine whether improvement in quality has been achieved. These measures should include nosocomial infection rates."

Benneyan [5] discussed several approaches to applying SPC to infection control and needle stick data, as well as some design issues, extensions, and alternative methods. For example, Figure 3 illustrates a
control chart of the weekly number of needle exposures, indicating that something atypical occurred in week 4, after which a run beneath the average is evident. Identification and removal of these special causes is the first step in improving the process of handling needles and reducing exposures.

Blumenthal, Laffel, and others [10,1] have suggested using control charts to examine mortality rates. Assuming a consistent level of care and presenting conditions, then a mortality \( p \) control chart of the rate of mortality for a particular hospital would appear somewhat as shown in Figure 4. Aside from the possible exception of subgroup 25, the mortality rate for this facility otherwise is fairly constant and in a state of statistical control. This lone special case should be reviewed for reasons for deviating from typical expectations. Experience also has shown, however, that the degree of stability can differ significantly between hospitals, having disturbing implications on the quality and constancy of care.

For example, Figure 5 indicates that variances from a particular standard procedure of care, or "clinical pathway", is unstable and unpredictable on a week-to-week basis. The rate of deviations from this protocol was beneath the lower control limit in week 6 and then above the upper limit the following
Clinical Pathway Variances
Figure 5

three weeks, followed by a pronounced increasing trend from weeks 11 through 23. This lack of standardized treatment again has implications on the quality and consistency of care. Under the philosophy of statistical process control, a first step in improving adherence to this clinical pathway is to bring the process into a state of statistical control so that it is operating with only natural variability.

Companion papers [4-7] provide further specifics on SPC and illustrate its application to several other health care and clinical laboratory processes, including prescription errors, patient falls, disenrollment, billing errors, admit rates, length of stay, chest x-rays, radiology, Pap smears, mammography, and HIV-1/2 and hepatitis testing. Special purpose control charts also have been developed for examining case mixes and laboratory diagnostic accuracy.

SOME CAUTIONS

Statistical quality management is an effective approach which should be used, in conjunction with traditional epidemiologic methods, to help study and improve health care process quality and, most importantly, the well-being of patients. Equally important, these tools should be used correctly, often beyond requirements of regulatory bodies. For example, improper use of quality methods has been a contributing factor in not detecting statistical changes in nosocomial infection rates, in determining liability in clinical laboratories, and in the manufacture of hospital equipment and biomedical devices.

In fact, several misunderstandings and over-simplifications of quality measurement and SPC seem endemic in health care, with literature ranging in technical detail and accuracy. An impression of particular concern is that concepts of natural versus unnatural variability are not applicable to health care. Related misperceptions include "management by average", reacting to natural variability, confusing SPC with software, and confusing quality with satisfaction. Common statistical errors include using incorrect control charts or insufficient data, misusing so-called "individuals" charts, and using "short-cut" control limit formulas. For further discussion, see [4,6].

Finally, in many situations care should be taken to properly adjust data for case mix, age, gender, severity, patient census, number at risk, etc. For example, significant differences exist between service-specific infection rates, such as for adult and pediatric intensive care units, surgical patients, and high-risk nursery patients, and separate control charts therefore might be applied to each of these catt
egories [5,15]. The earlier issue of evaluating individual surgeons also illustrates the importance of severity adjustments. Additionally, when combining data from non-homogenous processes, standard control charts may be incorrect for the composite metric and an alternative method should be used.

REFERENCES AND SUGGESTED BIBLIOGRAPHY

18. Interview with JE Wennberg in Hospital and Health Networks, August 5, 1993, p. 81.

This paper is based on excerpts from of an introductory book on successfully applying SPC in health care, Using Basic Statistical Quality Management and SPC to Improve Health Care [3].

Please address requests for reprints and additional information either to the address above or to Benneyan and Associates, 96 Dunster Road, Jamaica Plain MA 02130, U.S.A., (617) 522-0616, benneyan@ecs.umass.edu.