Peer Review Benchmarking

Pursuing Medical Staff Excellence

Robert J. Marder, MD

Peer Review Benchmarking: Pursuing Medical Staff Excellence

Robert Marder, MD

HCPro
a division of BLR

Peer Review Benchmarking: Pursuing Medical Staff Excellence is published by HCPro, a division of BLR.

ISBN: 978-1-68308-107-4

Robert Marder, MD, Author
Karen Kondilis, Managing Editor
Erin Callahan, Vice President, Product Development & Content Strategy
Elizabeth Petersen, Executive Vice President, Healthcare
Matt Sharpe, Production Supervisor
Vincent Skyers, Design Services Director
Vicki McMahan, Sr. Graphic Designer
Jake Kottke, Layout/Graphic Design
Reggie Cunningham, Cover Designer

Arrangements can be made for quantity discounts. For more information, contact:
HCPro
100 Winners Circle Suite 300
Brentwood, TN 37027
Telephone: 800-650-6787 or 781-639-1872
Fax: 800-785-9212
Email: *customerservice@hcpro.com*

Visit HCPro online at *www.hcpro.com* and *www.hcmarketplace.com*

Contents

About the Author

Robert Marder, MD

Robert Marder, MD, is president of Robert J. Marder Consulting, LLC (*www.robertjmarderconsulting.com*). He brings more than 30 years of healthcare leadership, management, and consulting experience to his work with physicians, hospitals, and healthcare organizations nationwide. A highly respected speaker, consultant, and author, he has helped hundreds of hospital medical staffs evaluate and improve their approach to peer review and physician performance measurement. He has authored numerous books on peer review for HCPro, including all three editions of *Effective Peer Review.* He is an editorial advisory board member for *Medical Staff Briefing.*

Acknowledgments & Dedication

I would like to thank all of the participants in the case studies for this book, whom I cannot name individually due to their organizations' requests for anonymity. I also wish to thank The Greeley Company for the opportunity it provided me to engage in this work for a substantial proportion of my consulting career.

I would like to dedicate this book to my loving wife of 39 years, Susanne, through whom God's grace and love is even more evident as we continue on our journey together facing her affliction with Alzheimer's disease.

Introduction

Apart from enjoying the benefits of Medicare and senior citizen discounts for movies and golf, getting old tends to provide time for reflection on the past. Personally, I began to contemplate whether my efforts over the past 15 years of consulting with medical staffs on how to improve peer review have led to real improvements in physician performance and patient care and whether these improvements were sustainable.

My concerns were heightened about a year ago when I ran into a well-known physician quality leader at a national meeting. As we discussed various aspects of physician performance evaluation, he made a somewhat broad statement that was reasonable but felt unsettling. His view was that, since healthcare quality is now about systems and is actually a team sport, individual physician performance evaluation (i.e., peer review) has become irrelevant. Having spent a substantial portion of my career in this field, his comment gave me pause and led me to reflect on whether perhaps I had wasted my time. And then I thought about baseball.

Baseball is clearly a team sport, but it relies heavily on the competence of individuals. Unlike football players, baseball players are positioned far apart from each other, so it is easy to evaluate individual responsibilities: who threw the pitch, who hit the ball, and who caught the ball. Unlike in soccer, each play has a distinct beginning and end, so it is easy to measure the individual's impact on that play. Although we do ultimately care about the team's won-loss record, we pay careful attention over the course of a season

to individual player performance through both specific numbers (e.g., home runs, runs batted in, errors, games saved) and rates (e.g., batting average, on base percentage, earned run average).

Although baseball certainly requires coordination among multiple players, such as when completing a double play from shortstop to second to first, the individual players' judgment and skill are mainly what result in team success and are the focus of extensive ongoing evaluation by coaches and teammates as well as by reporters and fans. When a player makes a mistake or when negative trends occur, these issue are usually addressed in the clubhouse, hopefully to improve the player's performance, as well as at the water cooler and in the media. Ultimately, even if the team has a winning season, it is an individual player's performance that determines whether he or she remains on the team.

Does this sound familiar? Contemporary peer review evaluates the performance of individual physicians in the context of being members of the healthcare team. It looks at individual events through case review and rule indicators, and at trends through OPPE data. Fortunately, individual errors are not replayed over and over again on a giant screen in center field or on a sports channel blooper highlights show. But the process does result in feedback to individual physicians on how he or she might improve and ultimately determines who stays on the team.

Let's return to the statement regarding the relevance of peer review today. Of course it is appropriate to question the status quo in a changing environment like healthcare. However, as long as we credential individual physicians for specific privileges that require their individual judgment, knowledge, and skill, peer review is both helpful and necessary. Credentialing is not performed simply to comply with regulations: It is practiced because individual competency still matters in patient care. Although patients want hospitals to have great patient care systems in place, they above all demand competent

individuals to care for them. Our legal system also has not yet abandoned the concept of individual physician accountability.

To continue the baseball analogy, there are more aspects of coordination in healthcare today than there were in the past, but physicians are still positioned far enough apart in the field that their individual judgment, knowledge, and skill can be observed, evaluated, and, if needed, improved. Consider hospitalists, for example. Although a great deal of the patient care that hospitalists provide involves coordination between caregivers, and although hospitals often contract with hospitalists as a group, no organization grants privileges to the hospitalist group as a whole. The whole team's performance may be evaluated by contractual goals, so if the whole team does not perform well, the organization can essentially fire the manager (i.e., the owners of the hospitalist contract) and bring in a new management team and players. However, privileges are still granted to and removed from individual hospitalists based on individual competency and performance.

The real question regarding the relevance of peer review for the 21st century is not whether we should continue to evaluate individual physician performance but how we do so in a manner that works with the changing healthcare environment. The book and subsequent movie *Moneyball* describe how baseball metrics advanced to explain previously unappreciated aspects of player performance. Similarly, peer review has evolved in the past 15 years to address at least three fundamental concepts that more fully define physician competency and how to evaluate it:

1. How to define competency
2. How to reduce bias
3. How to measure performance

Work on each of these concepts has led both to important advancements and continued challenges.

Returning to that unsettling question from the beginning of this introduction, my personal conclusion is that I don't believe that I have wasted my time working to make peer review relevant for today's healthcare needs. If anything, there is still work left to be done.

So why did I write the book? To move forward, it is often important to assess the past. As such, I decided to embark on a six-month research journey, conducting interviews at organizations with which I worked to make significant changes to their peer review programs. My goal was to find out what has worked and what has not and to share this with others through this book.

I will note, unapologetically, that this book has biases. First, this book is not intended to represent a scientific assessment of the field and all models of peer review—rather, it uses the case study approach, which is described in Chapter 4. Chapter 1 will provide an overview of various peer review models, but the case studies focus on organizations that have taken a multi-specialty approach to peer review. However, these organizations did more than just change to a multi-specialty peer review model; they engaged in a total makeover of their peer review processes for case review and OPPE as well. So even if you do not use a multi-specialty peer review structure, the lessons learned should translate to other structures.

There is also clearly a bias regarding organizations selected, as all of them had consultant assistance in the design and, to varying degrees, the implementation process. That said, once the consultant leaves, the questions regarding value and sustainability of that initial effort still remain. There is an advantage to this bias in that these organizations represent a cohort that have used similar principles and procedures in attempting to make peer review more relevant for their medical staffs, which should make them easier

to compare. However, you will find variations between the organizations, either initially or over time and for good or for ill, which I hope will be of value in redesigning or fine-tuning your peer review program should you be using or trying similar approaches.

Through this book, my hope is that the insights and challenges that these organizations have been willing to share will provide you with new ideas for improving your own peer review program, affirming your path, or letting you know that you are not alone in your struggles.

1

Peer Review Redesign in the 20th and 21st Centuries: A Personal Historical Perspective

Decade and century demarcations are usually artificial in their exact delineation of when a change process began or ended. For example, even though it dominated the 1950s, rock and roll did not exactly begin at the start of that decade. Still, the beginning of a new century provides a more dramatic way to focus our attention on a particular type of change. So while the germination phase of many current peer review ideas and concepts began in the late 20th century (i.e., the 1980s and 90s), the 21st century has seen these ideas bloom.

On a personal note, I have had a ringside seat for many of these changes. First, as the project director of The Joint Commission (TJC) Agenda for Change hospital indicator project from 1988 to 1991, I was responsible for directing six specialty task forces commissioned to develop national standardized hospital performance indicators. (Note: It has gone through several name changes over time, but for the remainder of this chapter, I will use the current name—The Joint Commission (TJC)—regardless of its official name during any given time frame.) I then became medical director for quality and assistant vice president for quality at Rush Presbyterian St. Luke's Medical Center in Chicago from 1991 to 1998, which was an early adopter of Total Quality Management (TQM). The focus of this book is the case studies about

medical staffs that redesigned their peer review programs between 2001 and 2014. At the beginning of that period, I began my work as a consultant with The Greeley Company, for whom I worked full-time until 2012 and then independently until 2015 and had the opportunity to facilitate many of these changes, whether personally through education programs, books, or supervision of other consultants.

I share this in the hope that it will add some credibility to the observations that follow. This chapter will attempt to describe some of the milestones that affect the evolution of hospital-based peer review in the 20th century and that set the stage for 21st century peer review. Its main goal is to provide some broader historical context for the case studies that you will read in Chapter 5. This book is not intended to analyze the effectiveness of all peer review models; rather, it is designed to look at one model and its variations to see whether, given the extent of the efforts surrounding it, those efforts were worthwhile.

In undertaking this review, I need to state that it is not designed to be a scholarly text. It is intended to be reasonably accurate and fair, but it will contain personal observations and conclusions based on my full-time involvement in the field of hospital quality and medical staff functions beginning in the late 1980s through today. I realize that others may disagree with my views, and I hope that any individual or group with a different point of view will feel comfortable contacting me to discuss the issue. Although I can't modify this book once it is published, I write a quarterly column for *Medical Staff Briefing* and would be happy to devote a column to any revisionist history issues necessary to set the record straight.

20th Century Peer Review

For those of my generation, describing something as "20th Century" implied that it was modern. Therefore, describing peer review from that era as old fashioned caused me to pause briefly to accept the inevitable passage of time and move on.

In the first half of that century, hospital peer review was not clearly organized and was primarily driven by professionalism. In the latter half of the 1900s, hospital peer review began to emerge as an organized activity. This change was driven mainly by the ascent of voluntary hospital accreditation standards in the 1950s through TJC. The original peer review standard required hospitals to review mortalities, delays in diagnoses, and discrepancies between pathology and clinical findings.

The need for peer review was reinforced in the 1960s as a regulatory requirement, either directly through the Centers for Medicare & Medicaid Services (CMS) *Conditions of Participation* (*CoP*) or via deemed status for TJC and, subsequently, a few other accreditation bodies. I would like to trace the remainder of the story of 20th century peer review as two parallel movements that, like two strands of rope, were ultimately woven together in the 21st century to drive the change to multi-specialty peer review and physician performance improvement. These two strands are the medical staff structure and functions movement and the quality measurement and process standardization movement.

The medical staff structure and functions movement

In the 1960s, the *CoP*s required hospitals to perform peer review to accomplish the credentialing of their medical staff, but how they were to do so was not prescribed. Peer review could be performed solely by an individual, such as a department chair, or by a group, such as a committee. The one element that appeared to be consistent across the field was the definition of a peer:

Medical staffs generally adhered to the definition that a peer was someone from the exact same specialty. Thus, peer review was conducted in the sacrosanct confines of one's peers, meaning the medical staff departments.

For that reason, it was natural that physicians organized peer review around specialties. In academic hospitals with powerful department chairs, the chair was often entirely responsible for peer review. In community hospitals, department committees did it. As the field of credentialing and privileging began to mature in the 1970s, with more systemic approaches to credentialing and privilege delineation, peer review became even more closely linked to departments and specialties.

During this time period, TJC also increased the number of medical staff functions to be conducted by committees, presumably to ensure that the medical staff would engage in its self-governance. Medical staffs, which had little work to do in the past, now became concerned about the time commitment involved to serve on the increasing number of committees needed to meet the regulations.

Because many hospitals had educational morbidity and mortality (M&M) conferences to discuss interesting cases, TJC decided in the late 1970s that these meetings could also qualify as the peer review committee if the medical staff would record the findings and use them in the credentialing process. Unfortunately, a negative consequence of this desire for efficiency was that it allowed peer review to occur without clear conflict of interest standards. This is because, in order to achieve its primary educational goal, the M&M conference model does not restrict the participation of those under review and actually encourages it. The other negative consequence was that the M&M was no longer strictly educational, since the findings were used in credentialing. This tension often led to a constraint on physician candor during the M&M discussions. As a result, ironically, the M&M results were less likely to be useful for credentialing. As will be noted in the case studies in Chapter 5, the effect of this decision lingers even today.

In the 1990s, physicians experienced greater economic pressures and no longer had an abundance of time to lend to the hospital medical staff activities. Many medical staffs sought ways to reduce the complexity of the medical staff structure, especially in regard to the burden of meeting accreditation standards. For example, one of the thought leaders at that time, Hugh Greeley, proposed eliminating the medical staff department structure entirely. The logic was that since many of TJC's requirements were departmental, TJC required functions like credentialing and peer review to be replicated by each department. By moving to a single medical staff department structure with clinical sections, medical staffs could reduce the number of committees. With his typical showmanship flare at education programs, Greeley would stack up cardboard bricks representing medical staff departments and committees and then knock them over to show the impact of this new model.

Although a number of medicals staffs converted to this model, it raised some uncertainty as to how the clinical sections were to operate in terms of leadership and authority. On the positive side, now that they were freed of department peer review requirements, medical staffs began to consider using other models to perform peer review, including a single multi-specialty peer review committee to handle all medical staff peer review. In addition to Hugh Greeley and his subsequent consulting group, The Greeley Company, the law firm Horty-Springer, which in the 1990s offered joint presentations with Hugh Greeley, also began to promote this peer review approach.

In a different approach designed to address the medical staff time commitment concerns in community hospitals, a consultant named Richard Thompson developed a model for non-academic hospitals where the department chair reviewed and rated all peer review cases. This model could best be characterized by the quote from Plato's Republic that "the best form of government is the benevolent dictator." Unfortunately, while efficient, this model led to peer review programs not functioning well. In this model, community

hospitals' department chairs often were reluctant to perform their duties, due to either the volume of work or the pressure of the perception of personal bias. A number of requests that I received to help medical staffs redesign their peer review program in the following decade came from hospitals that had tried unsuccessfully to implement this model.

In the 1970s, external peer review began to arise as a factor in hospital-based care, with the implementation by Medicare of Professional Standards Review Organizations (PSRO) and Peer Review Organizations (PRO) occurring subsequently in the 1980s. These organizations were designed to address both the quality and cost of care for Medicare patients through retrospective audits. On the commercial payer side, pre-certification programs with specific criteria were implemented to address medical necessity. Finally, in the 1990s, some of the PROs and independent companies began to offer external peer review services to medical staffs either to resolve internal differences of opinion, fill in gaps where there was a lack of expertise, or simply reduce the burden of case review on the medical staff.

In summary, by end of the 20th century, hospital medical staff peer review was primarily performed by department-based committees, but some new models were emerging to attempt to make it more efficient and useful.

The quality measurement and process standardization movements

Measuring and improving quality underwent dramatic changes in the second half of the 20th century. Both the concepts and methods for quality measurement and the tools to use that information to improve care evolved. Outlined below are some key aspects of how these changes affected peer review.

Quality measurement and peer review: From clinical audits to performance improvement

Much of peer review in the 1960s and 1970s was based on the clinical audit approach. This model typically took a clinical outcome, such as mortality or readmissions, and audited the medical records to determine whether medical standards of care were met. It was considered to be a minimalist approach and was a remnant of the 1950s peer review method from TJC's standards when accreditation was voluntary. It coincided with the quality culture of that era, in which hospitals did not see the need for extensive quality evaluation systems because they "knew" that they delivered good quality.

In 1979, TJC promulgated new standards that called for a systematic hospital quality assurance (QA) program, including the approach to peer review. In terms of the performance improvement field, while this was an advance over the audit system and probably a necessary next step, QA represented an inspection model that focused peer review on meeting a minimal standard for competency. A component of the QA model was the requirement that medical staffs use indicators or criteria, such as mortality or readmissions, to more consistently identify cases for peer review.

Recognizing that much peer review was conducted by departments, the standards required that each department determine its peer review indicators—but TJC did not tell them which indicators to choose or how many. This approach represented an example of an ongoing philosophic conflict in the field regarding how prescriptive TJC standards should be: Many hospitals wanted the standards to be broad to give them the opportunity to best meet them within the context of their individual organizations. Other hospitals saw broad standards as time consuming, as they required the hospital to "reinvent the wheel" by defining criteria that others had already figured out and allowed variability of interpretation by surveyors. Some of these hospitals saw accreditation as merely a requirement to pass with the minimum

amount of effort, like high school gym class, and not highly relevant to daily patient care. They would often beg TJC to at least be more specific regarding the number of indicators that were required per department. In response, TJC addressed this concern by creating the minimum requirement of two indicators per department, which has persisted in the minds of peer review support staff long after it ceased to be a requirement.

During the 1980s, legal concerns began to develop regarding the protections for medical staffs conducting peer review. In part as a response to the Patrick case in Astoria, Ore. (Patrick v. Burget, et al.), the Healthcare Quality Improvement Act (HCQIA) was passed in 1986. This act was intended to provide clear national legal standards for immunity for physicians performing peer review. It also established the National Practitioner Data Bank, which required reporting of corrective actions. In addition, many states adopted laws protecting peer review from discovery in medical legal proceedings. These actions helped to somewhat alleviate medical staff concerns regarding whether participating in peer review committees or corrective actions would result in personal harm to the reviewer or the reviewee.

As the world became more enamored with TQM (the concepts of process standardization promoted in manufacturing industries by W. Edwards Deming, Joseph Juran, and Walter Shewhart in the 1980s), the interest in how to define healthcare quality measures increased. The need for better healthcare performance measurement tools also became a significant part of TJC's project called the Agenda for Change. This ambitious project was an attempt by TJC to reinvent itself to be more relevant to the healthcare field and to address congressional concerns regarding retention of deemed status, by moving from quality assurance to quality management. It involved a significant overhaul of the standards and the survey process as well as the creation of a national healthcare performance measurement system. The latter component was expected to define clinical performance measures, called clinical indicators.

The behind-the-scenes story of the Agenda for Change clinical indicator project was that, in 1987, the initial physician project director recommended that the indicators be based on statistical analysis outcome measures derived from large-scale claims databases, such as MedPar files. This recommendation met significant political opposition from TJC's board and from the field based on concerns of data inaccuracy and clinical relevance of the information in claims data.

Then, in 1988, TJC declared that the indicators would be mainly process measures developed by expert specialty task forces, which would define more clinically relevant data elements requiring varying degrees of specific hospital abstraction. These indicators would then be field tested in hospitals for collectability and validity, including the development of software to collect the data. The stated goal was that, following field testing, the useful indicators would be mandated for the hospitals to collect by 1990 and eventually used in the accreditation process but would not be directly available to the public.

With this change in direction from claims data analysis, the initial physician project director left, and I was recruited to work with the research and field-testing staff to help develop materials to define potential indicators and facilitate the initial task forces. By 1989, the project was on track and had produced sets of 20 indicators for each specialty, and other task forces were initiated on non-specialty functions such as medication use. Although the indicators were designed to focus on hospital performance, many of them also had implications for physician performance, just as core measures do today (we will return to those later in the book).

Approximately 100 hospitals were recruited to test the indicators and the software. Here was where reality set in: The hospitals pushed back on the number of indicators for each area and the amount of time for data abstraction. The data collection software also had problems. As 1990 came and went, TJC first delayed the mandate to collect data and reduced the number of indicators

required to five per specialty set. These changes did not sufficiently appease many hospitals because, although they argued that their concern was validity of the measures, the real issue was the need to add a full-time equivalent (FTE) to collect data. TJC also was unprepared to be in the software database business.

By 1991, TJC essentially abandoned the indicators by making them voluntary for hospital internal use but not to submit to TJC. Interestingly, by 1993, TJC pivoted 180 degrees, back to using claims data in its ORYX project. However, the field pushed back on its accuracy and validity because the measures were not risk adjusted and not very clinically oriented.

Finally, in the late 1990s, CMS implemented core measures, which required adding additional FTEs for data collection. It also defined outcome measures using claims data that pushed hospitals to invest in outcomes analysis software. Again, while CMS analyzed the measures at the hospital level, many of these measures were based on physician actions (although initially the data collection software produced by most vendors did not have the ability to collect data on physician attribution). Unlike TJC's approach, all of the data would be publicly available.

Thus, a decade after TJC tried voluntary field engagement to measure hospital and physician performance to improve care with a standardized data set, CMS made it happen via a regulatory requirement. Ironically, many of the core measures selected by CMS were similar to those that had been proposed by TJC's specialty task forces, which the field had opposed. Also, instead of TJC's approach that restricted public access to data, CMS made both the process and outcomes data publicly available. This data transparency had immediately improved compliance with the core measures, far beyond that seen with voluntary compliance with medical specialty society recommended practice guidelines. As a result, it solidified the future use of publicly available data for all types of healthcare performance measures.

The sad lesson of this story is that by resisting voluntary investment in standardized healthcare performance measurement in 1990, the hospital field lost a decade of potential improvements in patient care and paved the way for mandatory government-defined data. The good news is at least the CMS mandate paved the way for 21st century hospital peer review to go beyond case review as the sole tool to evaluate physician performance by accessing the physician relevant data for rule and rate indicators.

Another influential factor that further accelerated interest in hospital quality and peer review data at the end of the 1990s was the Institute of Medicine report "To Err Is Human," published in 1999. This report described the large number of potentially preventable deaths, between 45,000 and 90,000, occurring in hospitals. Again, rather than acknowledging the opportunity for improvement, much of the hospital field and organized medicine spent several years debating the methodology regarding the exact number of deaths. Eventually, however, this report provided the foundation for increased voluntary collaborative improvement efforts through the Institute of Healthcare Improvement, mortality data transparency through CMS and commercial hospital rating information, and focus in peer review on preventable outcomes.

Process standardization and peer review: Clinical pathways and evidence-based medicine (EBM)

An important influence of TQM in the early 1990s that was not regulatory driven was the increased focus on clinical process standardization. This influence was manifested in the proliferation of clinical pathways and practice guidelines designed to be based on EBM.

Pathways and related clinical order sets were a multi-disciplinary collaborative tool involving all caregivers for a given disease or procedure. They were particularly popular with nursing, which sought to engage physicians to make care processes more consistent. Part of my role at Rush was to facilitate the development of clinical pathways as part of the organization's TQM

initiative. Rush developed and implemented more than 30 clinical pathways over six years, but the effort was not without physician resistance.

During that time frame, I observed three main sources of such resistance. The first was the concern that standardizing care that would lead to "cookbook medicine," a concern that was most strongly expressed by medical departments. The second was the question of how noncompliance with pathways and order sets would apply to peer review and ultimately to credentialing. The third was the willingness of physicians performing the same procedure to create a common pathway for nursing vs. individual physician or physician group pathways.

Ultimately, pathways tended to work best for procedure-related care, like total hip replacement, and were reasonably well accepted by surgeons because expected recovery time frames were definable. Pathways for medical conditions, such as pneumonia, were less accepted because of the large variation in patient comorbidities and response to therapy and often were composed of admission and discharge order sets. At Rush and at most other hospitals, pathways were considered optional, and compliance was a voluntary effort by medical staff departments who were conceptually in agreement with process standardization. Whether the standardized or an individualized pathway was used for a procedure depended on the organizational and medical staff's political will to push standardization.

While often used in clinical pathways, EBM was a broader movement. Although it was philosophically supported by organized medicine, it was quite controversial in how it was to be applied. Like pathways, physicians and their specialty organizations often used phrases such as "cookbook medicine" to raise their concern about standardizing medical judgment. The researchers, the AMA, and specialty societies made significant efforts to clarify standards for the types of evidence that might be available. Detailed descriptions of these efforts are beyond the scope of this chapter, but its current use

in peer review will be addressed later. Suffice it to say, while there is much greater acceptance of the use of EBM today (I haven't heard the words "cook-book medicine" in quite a while), the battle over what evidence is valid still results in some heated discussions.

21st Century Peer Review

As mentioned earlier, the changes in medical staff structures and functions and the desire for better measurement began to intertwine in the first decade of the 21st century. This resulted in increases in four areas related to medical staff peer review:

1. Multi-specialty peer review
2. Standardization of case review
3. Aggregate physician performance measures
4. Movement toward a positive peer review culture

In contrast to the above discussion of 20th century peer review, this section will be organized more as a summary of the impact of these four changes than as an historical progression. Although the case studies are based on organizations that implemented multi-specialty peer review along with the latter three concepts, it should be noted that many medical staffs pursued those other areas without using a multi-specialty model.

Increase in multi-specialty peer review

Over the past 15 years, an increasing number of hospitals have adopted some form of multi-specialty peer review. Before discussing this shift, there is a need to distinguish between two terms that are sometimes confused: multi-specialty and multi-disciplinary. Multi-specialty means that voting

committee members are physicians from various specialties. Multi-disciplinary means that voting members are from different disciplines (e.g., physician, nurse, physical therapist). Although multi-specialty peer review committees can greatly benefit from having a few nonvoting representatives from other disciplines to reduce professional bias, the model being discussed is multi-specialty peer review.

How prevalent in hospitals is multi-specialty peer review? Anecdotally, when I first started leading seminars on peer review for The Greeley Company in 2001, only about 10% of the attendees were using this model. By 2015, it was more than 70%. Although this is not meant to be a statistically accurate poll, it does reflect the significant shift toward this model, which I would estimate may represent in 40%–50% of hospitals today. In 2009, a survey of more than 1,800 physician hospital leaders by Edwards et al. reported that a multi-specialty process is the norm for peer review, and 27% note that it is used at least occasionally.

What has driven this change? First, as mentioned earlier, medical staffs were looking for ways to improve the efficiency of peer review by reducing the number of committees. Second, medical staff leaders were often frustrated by the variation in effectiveness of the department peer review committees. Third, as hospitalized patients became sicker with more complex needs, many medical staffs recognized that patient care requires greater collaboration among specialties, and department-based committees created a lack of resolution and delays in sending cases back and forth.

Who has driven this change? A number of consulting groups and individual consultants have worked with hospitals directly. In addition, many hospitals have implemented, or attempted to implement, some of this model through educational seminars and publications using variations on this model or by obtaining information from another hospital that has made this transition. The variability in design and implementation of the model tends to be greater

when done through the internally directed approach rather than a consultant-led approach.

Multi-specialty peer review has a number of advantages that address the common concerns of traditional specialty-based models, such as the following:

- **Minimizes bias**: Because this model is multi-specialty, there is less individual and specialty bias. It brings multiple perspectives to the table during the evaluation process instead of adding them after the fact via a second department review or appeal to an oversight committee.
- **Improves efficiency**: It reduces the number of physician participants required for peer review in an age with declining physician involvement in medical staff functions. It is also more efficient for support staff because they have fewer committees to support.
- **Increases reliability**: With fewer committees, there are fewer reviewers to train, and as they work together over time, they normalize the evaluation process and increase reliability. It also removes the issue of variability because there is only one committee.

There are two main ways multi-specialty-based peer review can be structured:

1. Single, central multi-specialty committee
2. Multiple multi-specialty committees

Although the single committee model is most common, I have seen recently an increase in the use of the multiple committee model. Outlined below is a brief description of each. The case studies in Chapter 5 will provide examples of both.

Single multi-specialty peer review committee

The single, multi-specialty peer review committee performs all case reviews and may also perform the oversight functions related to other measures of physician performance. Typically, the committee is composed of seven to nine physicians from different specialties or departments. The number of members can be greater, but the goal is not to have representation from every specialty. Rather it is to create a cadre of dedicated, clinically credible, and respected peer reviewers who are well trained in the peer review process. Thus, the terms of appointment tend to be longer (e.g., in the three-year range), with a staggering of the terms to provide more continuity of the process.

Peer review committee members act as the initial physician reviewers and typically are assigned cases on a rotating basis regardless of specialty, but if additional clinical expertise is needed, the reviewer or committee can request assistance from any specialist on the medical staff who the committee thinks would be most appropriate. The committee, however, makes all final determinations of appropriateness of care. As a single, centralized committee, it can look at all physicians involved in the care and their interactions in a single review. Additionally, many physicians enjoy the cross-disciplinary dialogue and learning, which increases participation.

In this model, the peer review committee typically makes recommendations regarding improvement strategies to the department chair, who is responsible for working with the physician under review on the actual improvement approach. This separates the evaluation portion of the peer review process from the action or implementation portion. This distinction is important because it preserves the responsibility to take action for elected medical staff leaders, including the department chairs and the officers of the medical executive committee (MEC), that is usually required by medical staff bylaws.

Although the single-committee model was originally thought to be best-suited for smaller hospitals, many medium-size hospitals (300–400 beds) and some

large tertiary-care academic medical centers have also found it effective, as will be described in some of the case studies. There are pros and cons, as in any model, and a more detailed discussion about them can be found in *Effective Peer Review* (HCPro, 2013) and will be also addressed in the case studies in Chapter 5 of this book.

Multiple multi-specialty peer review committees

Organizations that are attracted to the concept of multi-specialty review but concerned that the complexity of their services may make it too burdensome for a single committee, or that wish to enhance care collaboration in a more focused manner, have adopted the multiple multi-specialty committees model. Such peer review committees have multi-specialty representation but may be organized by service lines (e.g., cardiovascular, maternal/child, trauma), diseases (e.g., respiratory), units (e.g., intensive care, emergency services), or a combination of those categories. It also may result from combining many small departments into a larger committee (e.g., all medical specialties and subspecialties meet as a single committee). This model does result in some hospital-based departments (e.g., emergency medicine and anesthesiology) participating in several committees, such as a cardiovascular, maternal/child, or respiratory diseases peer review committee.

One variation of this model is to consolidate multiple committees into the same general specialty (e.g., a single surgical peer review committee for all surgery departments and subspecialties). Although this may appear to be multi-specialty on the surface, it is actually just a more efficient version of the specialty-based peer review committee model, since the "specialties" involved come from the same root specialty. It can become more truly multi-specialty if physicians from other root specialties are added to each committee (e.g., an emergency medicine physician is appointed to the surgery committee).

Combined department committees and multi-specialty committee model

Some medical staffs have been reluctant to abandon department committees but want to have a better forum for addressing cases involving multiple departments and ensuring consistency among the committees. To that end, they have established a multi-specialty oversight committee that can review cases involving more than one physician. Of course, this approach is not highly efficient, but it can be an initial step in moving toward a multi-specialty review committee in the face of strong department-based culture.

Increased standardization of case review

As noted earlier, a driving factor for change in peer review was the variability seen in the case review ratings. As a result, a key change described in the case studies was the development of a more standardized approach to case identification, case rating systems, and structured case review formats. Let's look briefly at each of these.

Case identification is typically based on either referrals or screening worklists to find the cases that are potentially appropriate for physician review followed by a screening process led by peer review coordinators. As the case studies will discuss, hospitals varied in how they accomplished this goal prior to redesign. The common elements resulting from the redesign project were clear written review criteria, with more explicit inclusion and exclusion criteria, greater engagement of referral sources, and more judicious use of worklists.

Prior to redesign efforts, case rating systems were typically two-level systems: standard of care met or not met. As medical staffs began to seek more refinements in case ratings, many, including the ones in the case studies, moved away from "standard of care" terms to reduce the potential legal impression of peer review. They also moved to a three-level system to provide committees with a "gray zone" (e.g., appropriate, questionable, or inappropriate) that

recognized that improvement opportunities existed despite the care not falling under the strict medical legal definition that the standard of care was not met. More recently, the terms have been modified to focus more on improvement (e.g., care appropriate/acceptable, minor improvement opportunity, and major/significant improvement opportunity). This shift has been well received by their medical staffs.

Some organizations used systems with 4 to 7 levels, which typically incorporated system issues as part of the rating of physician care. Interestingly, in the past few years, one consulting organization began advocating for a "no level" system of "care appropriate" or "referral to department chair for improvement opportunity," but in reality, this approach was a two-level system that just avoided the "standard of care" term.

An important factor in case review standardization is reviewer interrater reliability. The adoption of more explicit case review forms provided a means for keeping reviewers focused on the key issues of the case. As the case studies will show, hardwiring key elements into the case review form had a great impact on reviewer constancy. Some of these elements were as follows:

- Focused questions to the physician reviewer
- Justification of rating regardless of level
- Documentation rating separate from care rating
- Exemplary practice identification
- System issue identification
- Referral of topics to the CME committee for future programs

In addition to these changes, a number of medical staffs have incorporated into their peer review forms methods to evaluate causes of error, potential harm, and individual culpability. These approaches will be discussed in Chapter 2.

Increased aggregate physician performance measures

A third change was expanding the use of aggregate measures of physician performance instead of solely relying on case reviews. The goal was to reduce the burden of case review on the multi-specialty committees and to increase fairness for the type of issue under evaluation. There are two types of indicators that are useful for this approach: rule indicators and rate indicators. Below is a brief description of these two types.

Rule indicators

Rule indicators are based on the rules, standards, or generally recognized professional guidelines for the practice of medicine, and if such rules are not followed, the resulting noncompliance is unlikely to cause a patient harm. Rule indicators generally measure a process rather than an outcome because rules typically describe a policy or procedure to be followed. Non-physicians can obtain the information relatively objectively, so it does not require physician chart review. Physicians must be informed of every instance of noncompliance in a timely manner so that they have an opportunity to self-correct before a pattern develops. The medical staff establishes a target for the number of violations of a particular rule that would be considered excessive for the OPPE report and communicates that number, in advance, to all medical staff members.

The most common application of rule indicators has been to improve compliance with EBM-defined practices, either nationally via core measures or through medical staff adoption of standardized order sets or protocols. As core measure compliance nationally began to approach 95%, giving individual physicians their rate of compliance was not a useful way to increase compliance because they saw rates of 95% as implying that they were doing well when in reality a small number of cases of noncompliance could greatly affect the organization's ranking compared to national norms. The rule indicator approach was able to address those remaining misses on a more timely basis because

each instance of noncompliance was communicated directly to the physician and targets were set at a low number of instances rather than a rate.

Ironically, the initial software from vendors for collecting core measure data did not include a data field for physician attribution, despite the fact that many of the core measures were addressing physician ordering practices. This shortcoming resulted from the fact that the software was designed for reporting hospital performance to CMS, and physician-specific data was not to be reported. In the mid-2000s, hospitals began to recognize that they were limited in their ability to improve physician performance without physician-specific data and began to push vendors to add physician attribution fields where appropriate to the software.

Rate indicators

Rate indicators measure the number of events that have occurred compared to the number of opportunities for that event to have occurred. Thus, a rate indicator has a numerator and a denominator and can be expressed as a percentage, frequency, average, rank, or ratio. Rate indicators measure both processes and outcomes. Rate indicators ask whether the frequency of an adverse outcome or a failed process is different from established expectations and can reduce the bias of case review for events that have reasonable likelihood of occurrence for any physician. Thus, rate indicators level the playing field by adjusting for physician volume.

Rate indicators are more efficient and objective than case review because they do not require physician reviewers to comb through individual charts to obtain the data. Rather, hospital staff or hospital information systems collect the data for the appropriate physician leader or committee to evaluate. Finally, rate indicators lend themselves easily to statistical analysis and to risk adjustment for adverse outcomes.

Rule and rate indicators and OPPE

The use of aggregate data by medical staffs was being adopted as a best practice long before it became a TJC regulatory requirement in 2007. In fact, hospitals that had undergone peer review redesign prior to 2007 were often well positioned to incorporate OPPE into their peer review program. That said, OPPE clearly moved the use of aggregate physician performance data from voluntary best practice to a required system. As will be noted in the case studies, the limitations then and even today are related to data availability.

The other positive aspect of OPPE was TJC's adoption of the Accreditation Council for Graduate Medical Education (ACGME) framework for the six core competencies. This provided a common framework for evaluating physician competency. Prior to that time, most medical staffs had no framework for helping define relevant performance measures, although a number had adopted variations of the physician expectations framework introduced by the American College of Physician Executives.

With TJC's introduction of OPPE, the conflict within the hospital field regarding broad vs. proscriptive standards resurfaced in two areas. First, the OPPE standards required each hospital to figure out what indicators to put on its OPPE reports. This requirement was necessary because, despite CMS standardization of data collection at the hospital level for indicators such as core measures, the data vendors that hospitals were required to use to collect and submit data to CMS had markedly different systems to collect and analyze data at the individual physician level. Second, although it adopted the six core competencies from the ACGME, TJC had not tested what types of data applied to each competency for hospital attending physicians. As a result of the vagueness of the OPPE standards, TJC surveyors did not intensely evaluate the implementation of OPPE, and many hospitals have been able to get by without an effective application of this tool.

Since before the 21st century, peer review was almost exclusively based on case review, one question that had to be addressed in peer review redesign efforts was what group would oversee the use of aggregate data. In the early 2000s, when the use of aggregate data was gaining traction, the scope of the multi-disciplinary committee typically included oversight of rule and rate measures as well as conducting case review. However, at that time, there was a relatively small amount of rule and rate data. With the initiation of OPPE and the increase of performance data, this became more burdensome for these committees, and a number of them shifted the oversight of aggregate data for OPPE to the credentials committee or to a separate physician data council.

In the past 10 years, the proliferation of clinically relevant data systems through mining data either from claims data or the electronic health record clearly is the major factor driving the ability to produce the OPPE reports that TJC standards hoped for. Unfortunately, the promise of a data-driven environment for peer review is still unfulfilled for many hospitals.

Movement Toward a Positive Peer Review Culture

Over the past 15 years, physician leaders have often expressed the desire to make peer review less punitive. Key practices that allowed for this shift were the increased standardization of case review and use of aggregate data that increased the inherent fairness of peer review, and the hospital performance improvement and patient safety movements that provided new ideas for helping physicians improve. Chapter 2 describes in detail the influence of culture on peer review, and the case studies in Chapter 5 describe the impact of these cultural changes.

Despite Shakespeare's line from Romeo and Juliet "What's in a name? That which we call a rose by any other name would smell as sweet," words do

have an impact on culture. An interesting byproduct of medical staffs' desire for physicians to perceive peer review as less punitive was to adopt new names for the redesigned peer review committee. Examples of this approach included Physician (or Practitioner) Excellence Committee, Physician (or Practitioner) Evaluation Committee, and Professional Practice Committee.

Hopefully, the preceding overview will provide a common perspective on the remaining chapters of this book.

Peer Review Benchmarking

Pursuing Medical Staff Excellence

Looking for metrics and real-world processes to assess your peer review program? ***Peer Review Benchmarking: Pursuing Medical Staff Excellence*** offers the tools you need to analyze and redesign your peer review processes. Learn how to quantify your peer review data to measure the success of peer review at your organization and how to best share this information with various departments and committees. Once you use the data to determine what elements of your peer review process you need to redesign, read the case studies in this book from organizations who embarked on peer review redesign and learn from their triumphs and mistakes.

HCPro
a division of BLR

100 Winners Circle, Suite 300
Brentwood, TN 37027
www.hcmarketplace.com

EPR

ISBN-13: 978-1-68308-107-4
90000
9 781683 081074