Clin Orthop Relat Res (2015) 473:2455–2459 / DOI 10.1007/s11999-015-4383-x

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

Published online: 5 June 2015

Not the Last Word

Not the Last Word: Codman Was Right—Spread The Word

Joseph Bernstein MD

Second only to the discovery of insulin by Sir Frederick G. Banting, the development of an "End Result System" by Ernest A. Codman MD was the greatest contribution to general medicine by an orthopaedic surgeon. You may know Codman more for his eponymous Codman's Triangle (made by the elevated periosteum adjacent to a long bone tumor) or for his classification of proximal humeral fractures (which is the basis of the current practice of defining such fractures in terms of "parts"), but the End Result System outweighs them both, by far.

The End Result System is based on a simple idea, radical in its day—that we can determine how well a treatment works by assessing how well a treatment works. (The pre-Codman approach was to consider how close a treatment aligned with the reigning theories of disease. For example, if depression is caused by excess black bile (melancholia, literally), a good treatment would be one that drains it, independent of the actual clinical costs and benefits). Codman stressed recording patient treatment outcomes—"end results"—as a means of discerning clinical effectiveness from clinical futility.

For taking his theories to their logical conclusion, that surgeons should be evaluated by their outcomes, Codman lost his staff position at the Massachusetts General Hospital. Nonetheless, Codman's approach has taken root. We may disagree regarding the nuances of evidence-based medicine, but few of us would ever employ a treatment without any empirical evidence in hand. That is Codman's legacy, and countless patients have benefited accordingly.

It may come as a surprise, then, that Dr. Donald Berwick recently suggested in the *Journal of the American Medical Association* (*JAMA*) [2] that maybe Codman was wrong. Dr. Berwick was dismayed by two studies [5, 11] that reported on complication and inpatient mortality rates by comparing hospitals that participated in the American College of Surgeons National Surgical Quality Improvement Program (NSQIP) to those that did not. To Dr. Berwick, the observation that the improvements in complication and inpatient mortality rates at NSQIP hospitals were no better than that seen at nonparticipants somehow undermines the Codman approach. I disagree. A non-NSQIP hospital may show improvement without an explicit monitoring program of its own because good practices can be contagious.

Take note: The *JAMA* studies that vex Dr. Berwick did not demonstrate a failure of improvement at those hospitals that collect data; rather, what was shown that the NSQIP hospitals' improvement was manifest at nonparticipants as well. The simple explanation for this is that the good ideas, generated at NSQIP hospitals, travel to other institutions. After all, Landon and colleagues [10] have

The author certifies that he, or any member of his immediate family, has no funding or commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.
All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.
The opinions expressed are those of the writers, and do not reflect the opinion or policy of *Clinical Orthopaedics and Related Research*® or the Association of Bone and Joint Surgeons®.

J. Bernstein MD (✉)
Department of Orthopaedic Surgery, University of Pennsylvania, 424 Stemmler Hall, Philadelphia, PA 19104, USA
e-mail: orthodoc@uphs.upenn.edu

Not the Last Word

demonstrated the existence of "information-sharing networks of physicians" that cross the lines of formal organizations. Using these social connections, doctors at a NSQIP hospital who discover that a certain practice is helpful can inform colleagues from non-NSQIP hospitals in short order.

The contagion of ideas within orthopaedics was demonstrated in a study presented at a recent meeting of the American Academy of Orthopaedic Surgeons. This research examined the rate of clavicle fracture surgery in Ontario, Canada. The researchers found from 2002 through 2006, about 6% of all clavicle fractures were treated surgically; from 2006 to 2007, there was an increase to a rate of about 10%; thereafter, the rate rose to more than 20% by the end of 2011 (and is rising still). These data are germane to the issue of idea contagion because the major trial demonstrating the benefits of clavicle fracture surgery was published in 2007. That is to say, the "response" to this trial's findings actually anteceded the study's publication by at least a year, and probably more. It must be the case, therefore, that surgeons receive information from other sources besides the scientific literature. (This observation will come as no surprise to device manufacturers and pharmaceutical firms who spend large sums on sales representatives). Ideas cannot be contained.

To be sure, the contagion of ideas is not always beneficial. Just as good practices may spread by word of mouth, so too may less desirable ones. For example, the injection of platelet-rich plasma to treat tendinopathy seem to have ensconced themselves in mainstream practice long before the appearance of the evidentiary support Codman would have demanded.

For better or for worse, memes and themes can percolate through our social networks, often without our explicit awareness. Sometimes, it takes a long while for these ideas to fully spread. Codman, for instance, did not live to see how popular his approach has become. At other times, comments from the podium at medical meetings (such as the clavicle study mentioned above), articles in the lay press, or even casual conversations in the operating room corridors may send ideas into broad circulation with great speed.

Because ideas both good and bad can spread so readily, we are ever more required to apply the methods of Codman, and rigorously assess our performance. Only through diligent and meticulous collection of data, and the sharing of validated results, can we disinfect ourselves of those ideas that may harm patients.

David C. Ring MD, PhD

Department of Orthopaedic Surgery

Massachusetts General Hospital

Codman was correct that evidence rather than eminence, should determine what doctors and patients do, and that transparency about our shortcomings and adverse events improves care. However, Codman died alone and poor, was buried in an unmarked grave, and is most appreciated posthumously because he did not cultivate effective communication strategies for his progressive ideas. Part of the lesson of Codman, and other pioneers such as Semmelweiss, is that it is not enough to have a good idea or even solid evidence. Culture change is a better achieved by working with, rather than against, the system. Captivating stories and compelling data will have a greater impact than heavy-handed critique.

I have a slightly different take on the NSQIP experience than Dr. Bernstein. NSQIP exists because the idea that measuring outcomes and using them to benchmark and provide feedback improves performance became compelling. NSQIP was part of a culture change that was already occurring in medicine. Participation in NSQIPs may be more a measure of being part of the network of the originators and having the resources to pay people to collect and report specific data—and not so much a measure of how much one believes that measurement and analysis as well as checklists and planned, evidence-based care—help us perform better.

Not the Last Word

We must look beyond NSQIP, the Surgical Care Improvement Program (SCIP) [12], and similar programs and consider just how difficult it is to come up with meaningful things to measure. The measurement of, and public reporting of, the use of beta blockers after myocardial infarction was necessary to get doctors to adopt useful treatments where the evidence ran counter to gut feelings. "Operate through your initials" and "sign your site" still have not been adopted as widely as these great ideas merit. In contrast, the data behind programs like SCIP that suggest stopping antibiotics and removing a urinary catheter after 24 hours are not quite as compelling.

In science, it is not only important that we ask the right questions, but we must also ask them in the right way. Testing the hypothesis that NSQIP participation improves care risks finding out that it does not. Since we know that people might misinterpret this finding as indicating that these efforts are not worthwhile—an interpretation that fits many surgeons' biases and agendas—scientists should consider how the findings of a given experiment will be interpreted before they begin. If the hypothesis is whether or not surgeons are now more aware of quality improvement projects, public reporting of data, and even tying financial incentives to performance because of initiatives such as NSQIP, SCIP, and the Joint Commission Universal Protocol [13]; whether the culture in the hospital and operating room is forever transformed; and whether having these discussions feels more natural than the "illusion of perfection" that many of us were raised in, I think the experiment would have a positive result and a useful interpretation.

Nicholas H. Osborne MD, MS

Center for Healthcare Outcomes and Policy

University of Michigan

The NSQIP has provided hospitals and surgeons with quality reports on a semi-annual basis since 2005. These reports were intended to provide hospitals with the quality feedback necessary to improve outcomes longitudinally. As referenced by Dr. Bernstein, Codman pioneered this principle of measuring patient outcomes to improve treatments. Although Codman emphasized the measurement and sharing of data regarding the treatment and outcomes, it is implicit that the data be used to improve care.

As discussed by Dr. Berwick in his editorial in *JAMA* [2], the idea of simply measuring outcomes may not lead to positive change, or more simply, "weighing a pig does not make it fatter" [2]. The two papers referenced by Dr. Berwick both lend credibility to the concept that we must take steps beyond the simple measurement of outcomes to improve quality [5, 11]. Interestingly, both studies demonstrated similar findings with different data and approaches. Although NSQIP hospitals improved during the study periods, no participating hospitals were improving at a similar rate over time. These findings emphasize that measurement of outcomes alone may not be enough to lead to improved quality.

In contrast, Bernstein has interpreted this data to suggest that perhaps the NSQIP has led to quality improvement efforts that have "spilled over" to other nonparticipating hospitals. It is certainly plausible that hospitals might share quality improvement efforts, but this requires both a collaborative infrastructure and a willingness to share outcomes and best practices. What is troubling is that hospitals participating in the NSQIP have not improved above and beyond other hospitals, and participating hospitals have not shown demonstrable cost-savings. Unfortunately, although the NSQIP has taken the first step toward quality improvement by providing hospitals with high quality data to enable hospitals to develop quality improvement projects, the American College of Surgeons (ACS) has not provided an infrastructure for quality improvement. Historically, individual hospitals embarked upon quality improvement efforts in isolation of each other, an often expensive and difficult task. However,

Not the Last Word

some regional collaboratives, such as the Michigan Surgical Quality Collaborative, have already created this infrastructure and meet on a quarterly basis to discuss outcomes, and share best practices [3, 4]. More recently, the ACS and other societies have emphasized this collaboration in addition to participation in the NSQIP alone. These societies now encourage the formation of regional and disease-specific collaboratives to share outcomes and approaches. Meeting together as surgeons, sharing our good and bad outcomes, and discussing our practices to improve care represent the true End-Result System described by Codman.

Randale C. Sechrest MD

Medical Director

Montana Spine and Pain Center

Plato is usually considered the originator of the idea of "knowledge" as justified true belief. "True" and "belief" are relatively noncontroversial concepts; "justified" has always been the fly in the ointment meant to salve the afflictions of dogma, ignorance, and self-delusion [8].

The modern scientific method reached a point of maturity that fundamentally changed the accumulation of knowledge during the Middle Ages. The production of "justification" accelerated as scores of the most notable names in history vigorously applied the scientific method to discover new truths, as well as new questions to ask. The scientific method, by all accounts, has enjoyed a pretty good run.

In his book, *The Island of Knowledge: The Limits of Science and the Search for Meaning* [7] Marcelo Gleiser directly ties the expansion of knowledge to our ability to artificially extend our human senses that give us empirical input of objective reality. What science elucidates—what we "know"—changes with the increasing capability of the tools we use to examine the external world. The telescope and the microscope, for example, each allowed us to see more clearly; each tool enhancing and extending the capabilities of our human senses to perceive the world around us quite literally through a different lens. But, our tools also include the cognitive extenders that we rely on to formulate and communicate our ideas and conceptualize what we understand as objective reality: Language, math, and (now) computers [7].

Ultimately, science is a human enterprise ensconced in human culture. Thomas Kuhn, in *The Structure of Scientific Revolutions* [9] clearly articulated the challenge we still face in distinguishing "knowledge" from "belief." Kuhn introduced the concept of the "paradigm shift" into our lexicon, arguing that even in science we are hampered in our ability to truly think outside the box. The reigning paradigm constrains even the questions we are likely to ask until we are presented with so much cognitive dissonance that we simply have no choice but to move towards a different understanding of the empirical evidence.

John Ioannidis MD at Stanford University may be our contemporary equivalent to Thomas Kuhn in medicine. Ioannidis has shown how all-too-human frailties can lead to the corruption of the scientific method, polluting the river of justification that is necessary to feed an ever-expanding ocean of credible knowledge [6]. While the credibility and the effectiveness of the scientific method may be valid beyond question, it must be applied with rigor, integrity, and objectivity in order to achieve its promise of delivering the goods [6].

Today, the next evolution in empirical science promising to accelerate the accumulation of "justification" involves computer networks, statistical algorithms, and "Big Data." This combination promises to uncover the patterns that rule our lives, including our health, allowing us to do what humans have always aspired to do—predict the future with greater certainty. International Business Machines' Watson (Armonk, NY, USA) has managed to overcome a chess grandmaster, best the greatest human trivia masters on "Jeopardy", and now is coming after the jobs of our oncologist colleagues [1].

Not the Last Word

The tools that we will use to extend both our senses and cognitive functioning are becoming more abstract. We rely on sensors that mediate and measure signals from the external world that our human senses have no way of confirming. These signals are in turn fed into algorithms that we can barely comprehend to give us answers—and we are asked to trust the math. Yet the predictions seem to be increasingly accurate, leading to an ever-greater comfort level with these black boxes that we have created and rely on to guide us in our decisions.

While the tools we use might be changing, becoming more complex and difficult to comprehend; and while there will always be cultural constraints in how we apply, imperfectly, the knowledge we gain from those tools, it would be hard to believe that we would question the value of empirical science to the practice of medicine.

References

1. Begley S. IBM's Watson to guide cancer therapies at 14 centers. Available at: http://www.reuters.com/article/2015/05/06/us-ibm-cancer-id USKBN0NR2K520150506. Accessed May 26, 2015.
2. Berwick DM. Measuring surgical outcomes for improvement: Was Codman wrong? *JAMA*. 2015;313: 469–470.
3. Campbell DA Jr, Kubus JJ, Henke PK, Hutton M, Englesbe MJ. The Michigan Surgical Quality Collaborative: A legacy of Shukri Khuri. *Am J Surg*. 2009;198(5 Suppl):S49–55.
4. Englesbe MJ, Dimick JB, Sonnenday CJ, Share DA, Campbell DA, Jr. The Michigan surgical quality collaborative: Will a statewide quality improvement initiative pay for itself? *Ann Surg*. 2007;246:1100–1103.
5. Etzioni DA, Wasif N, Dueck AC, Cima RR, Hohmann SF, Naessens JM, Mathur AK, Habermann EB. Association of hospital participation in a surgical outcomes monitoring program with inpatient complications and mortality. *JAMA*. 2015;313: 505–511.
6. Freedman DH. Lies, damned lies, and medical science. Available at: http://www.theatlantic.com/magazine/archive/2010/11/lies-damned-lies-and-medical-science/308269/. Accessed May 26, 2015.
7. Gleiser M. *The Island of Knowledge: The Limits of Science and the Search for Meaning*. New York, NY: Basic Books; 2014.
8. Ichikawa JJ, Steup, M. The analysis of knowledge. Available at: http://plato.stanford.edu/archives/spr2014/entries/knowledge-analysis/. Accessed May 26, 2015.
9. Kuhn T. *The Structure of Scientific Revolutions*. Chicago, IL: The University of Chicago Press; 1962.
10. Landon BE, Keating NL, Barnett ML, Onnela JP, Paul S, O'Malley AJ, Keegan T, Christakis NA. Variation in patient-sharing networks of physicians across the United States. *JAMA*. 2012;308:265–273.
11. Osborne NH, Nicholas LH, Ryan AM, Thumma JR, Dimick JB. Association of hospital participation in a quality reporting program with surgical outcomes and expenditures for Medicare beneficiaries. *JAMA*. 2015;313:496–504.
12. The Joint Commission. Surgical care improvement project. Available at: http://www.jointcommission.org/surgical_care_improvement_project/default.aspx. Accessed May 19, 2015.
13. The Joint Commission. Universal protocol. Available at: http://www.jointcommission.org/standards_information/up.aspx. Accessed May 19, 2015.