Special Article – Plastic Surgery

An Assessment of Artificial Intelligence Software Measurement of Physician Performance: Not Quite Ready for Prime Time

Subhas Gupta*, MD, CM, PhD, FRCSC, FACS, Rachel Ford, MD and Saira Gupta

Department of Plastic Surgery, Loma Linda University, USA

*Corresponding author: Gupta SC, Loma Linda University Department of Plastic Surgery, Loma Linda, California, USA

Abstract

With an increase in hospital-based employment of plastic surgery graduates, understanding how hospitals evaluate physician performance is essential, as data could have important career implications. At our hospital (Loma Linda University Medical Center), the software program The CRIMSON Initiative is used to provide data related to physician performance. Upon review of this data, our Plastic Surgery department had a 30 days readmission rate that was higher than the hospital average. We were interested to see how accurate this number actually was, and so we closely examined the 30 days readmission rates for our department for an entire year. Using the CRIMSON Initiative software, two separate major searches were used (by "Attending Physician" and by "Performing Physician"). Searching either way revealed a 30 days readmission rate that was higher for our department than the hospital average. However, after manually sorting through each patient chart, there were cases that were not readmissions due to the patient's plastic surgery. Once these cases were excluded from the original calculations, it brought our department's readmission rate down to slightly lower than the hospital average for each search. This discrepancy highlights the margin of error of such automated physician performance programs, and brings to light pitfalls that physicians should be aware of concerning similar programs at their own institutions. As plastic surgery residency graduates are increasingly employed by hospitals, it would behoove them to be cognizant of this issue, and to be empowered to question the data being used to assess their performance.

Introduction

Data compiled from graduates of plastic surgery residencies shows an interesting trend, that an increasing number of plastic surgery graduates will be employed by a hospital following graduation. For example, in 2018, 77% of plastic surgery graduates will be employed by a hospital. Presumably, this number will increase in years to come due to factors such as the desire for a more academically oriented practice, the current economic crisis, the built-in referral base, and an environment, which has the benefit of mentorship from more senior physicians. With an increase in hospital-based employment of plastic surgery graduates, it is prudent to understand how hospitals evaluate physicians, as ongoing physician evaluation is a requisite for hospital accreditation by The Joint Commission, and the results of such an inquiry can have consequences for one's career. Additionally, physician performance data has been increasingly available on the internet for public review.

Within the past several years, there has been a push, called the Health Care Transparency Movement, to make physician performance transparent to consumers. It has been a significant part of the reform of our country's healthcare system. Goals of this movement are to provide consumers with quality data to assist and to empower them when choosing healthcare providers [1,2], and to encourage healthcare providers to continually enhance the quality of their patient care [1,3].

Aside from the data made public for the Transparency Movement, hospitals use computer software to evaluate physician performance within their institution. However, the way this information is collected, reported, analyzed, and the extent to which that information is utilized to reward or to penalize hospital physicians, varies with each establishment. Since this type of information is likely to play a more significant role over time due to the increase in physician employment in hospitals, and also since more residency graduates (at least in our specialty of plastic surgery) will likely be evaluated by these types of systems, the validity of the data being reported must be questioned. One would hope that data with such important ramifications would be accurate, valid, reliable, and cost-effective.

In a paper entitled "The crucible of physician performance reports," Dr. Hawkin Woo, an Internal Medicine physician at UCLA, illustrates his personal experience with inaccurate performance reports caused by faulty data collection methods [4]. Consequently, he had to spend countless hours delving through numerous patient records to document dates for preventative medicine procedures (that were supposedly missed according to his performance report, but were, in fact, completed).

In this paper, we closely examined the 30 days readmission rates reported for our Plastic Surgery department by our hospital's physician performance software, The CRIMSON Initiative. The 30 days readmission rate was an area within our department that was

shown to have a higher percentage, when compared to the hospital as a whole, and we were interested to see how accurate this number actually was, once we examined each individual patient medical record that was flagged for a 30 days readmission.

Methods

The CRIMSON Initiative is the Advisory Board Company [5] acquired a physician performance software program that was founded in 2003 by CRIMSON Software, Inc., and in 2008. Since its inception, the software has compiled data using information from more than 200 hospitals, accounting for over 25,000 physicians [5]. While examining the data reported by this software at our hospital (Loma Linda University Medical Center, Loma Linda, California), we were surprised to see that our Plastic Surgery department had a 30 day readmission rate that was higher than the hospital average. We had a concern that, perhaps, some readmissions for staged procedures were being included in this calculation. While they are still technically readmissions, they are obviously planned, and not due to any unforeseen complications.

One of the benefits of this particular software program is that it allows one to locate each specific medical record number that was flagged and included in the data analysis. That made it possible for us to examine each medical record that was counted as a 30 days readmission rate. We were interested to see if the data reported by CRIMSON was correct, and if not, where the discrepancies were, in an effort to possibly make this program more accurate in the future. We examined each medical record that was reported by the software as having a 30 days readmission rate for the entire year of 2010, for the Plastic Surgery department to see if, indeed, the patient was readmitted within 30 days of discharge, or if not, what triggered the flag for readmission.

Using The CRIMSON Initiative software, two separate major searches for the entire year of 2010 were used, and each individual patient chart was reviewed, to determine the reason they were readmitted. The two different ways to search readmissions using The CRIMSON Initiative are by "Attending Physician" and by "Performing Physician." Between the two, there are subtle differences. The Attending Physician is defined as the one physician who provided the most care to the patient during their stay (typically the Attending of the primary team). It is a billing-specific definition. Searching through this manner, therefore, misses readmits that were operated on by our faculty, but where the Plastic Surgery team was not the primary team taking care of the patient. The Performing Physician is defined as the physician who actually operated on the patient. This type of search still misses the few cases where patients had two subsequent readmissions, if no surgery was done during the first readmission. In this scenario, the second readmission would not be flagged as another readmission under this search. In addition, it should be noted that since these are *re*admissions that are being captured by the system, the software would not flag patients that get outpatient surgery and then are subsequently admitted for complications. This is because these patients were never admitted to begin with.

Results

For the year 2010, searching by Attending Physician, the software reported that our Department had a 9.30% 30 day readmission

rate, which is high compared to the hospital average of 7.28%. This number corresponds to 49 out of 527 total cases. However, after manually sorting through each patient chart, there seems to be 12 cases that were not readmissions due to the patient's plastic surgery. Six cases were staged procedures and therefore planned readmissions (which Crimson does try to filter out), two patients were admitted to a rehabilitation facility, one was admitted as a complication from another medical team, two were admitted for purely medical reasons unrelated to their surgery and one was not readmitted at all, they just had a same-day surgery within 30 days. Once these cases are accounted for and excluded from the original calculation, it brings our department's readmission rate down to 7.02%, which is actually slightly lower than the hospital average.

Searching by Performing Physician, for the year 2010, the software reported that our department had a 9.84% 30 days readmission rate, which was again high, compared to the 7.37% average hospital readmission under this search. This number corresponds to 54 out of 549 total cases. After sorting through each patient chart, there seems to be 14 cases that were readmissions due to the patient's plastic surgery. Five cases were staged procedures, two patients were admitted to a rehabilitation facility, two were admitted as a complication from another medical team, four were admitted for purely medical reasons unrelated to their surgery and one was not readmitted at all, they again just had a same-day surgery within 30 days. Once these cases are accounted for and excluded from the calculation, it brings our readmission rate down to 7.29%, which is again actually slightly lower than the hospital average.

Overall, the two searches yielded about 80% of the same patients, but each search has its limitations and loopholes where it misses a few cases, and there are patients that should not have been included in the reported readmission rate for our department. Granted, there were also actual readmissions for unrelated reasons that still classify as readmissions (and would be hard to exclude by an automated system). In an ideal world, they would not be included in a number that theoretically represents readmissions because of care by the Plastic Surgery department.

Discussion

The Joint Commission is an independent, not-for-profit accreditation body that certifies upwards of 19,000 healthcare programs and organizations across the country [6]. This association has many requirements that healthcare institutions must meet in order to be accredited. One of these requirements is the standard entitled Ongoing Professional Practice Evaluation (OPPE). The intent of this standard "is that organizations are looking at data on performance for all practitioners with privileges on an ongoing basis rather than at the two year reappointment process, to allow them to take steps to improve performance on a more timely basis" [6]. At Loma Linda University Medical Center, the software program The CRIMSON Initiative has recently been implemented to facilitate evaluation of physician performance, and data obtained from this program will likely be used to satisfy a portion of the OPPE requirement in the future.

In addition to satisfying accreditation requirements, physician evaluation hope fully serves as motivation for self-improvement, identifies inadequately performing physicians, and facilitates

validation for healthcare stakeholders [7]. However, there exists a formidable challenge to presenting comparable and comprehensive data to assess the care that physicians provide [8].

When automated quality-assurance tools are being used, it is wise to carefully examine the data being extracted, as it might differ from reality. The data presented by CRIMSON, at least the limited data examined for this article, seems to be accurate concerning capturing data per guidelines set forth for readmission rates. However, when one looks at the data for readmission rates specifically related to patient's plastic surgery care (and not readmissions for other medical issues, etc.), a different picture is painted by the data. Instead of our Plastic Surgery department having a higher readmission rate compared to the overall hospital average, our rate becomes lower than the hospital average once corrected. This specific discrepancy may not have huge implications, but it highlights the margin of error of such automated programs. By having accurate data collection methods, it ensures any subsequent consequences or decisions are based on actual differences in physician performance, rather than on any shortcomings of the data collection methodology [9].

So what can be done to help the reported data become more accurate? Physicians themselves actually play a big part. It is imperative that admission history and physicals be very detailed and accurate. One needs to be particularly vigilant to document any condition present upon admission (such as pressure ulcers, urinary tract infections, etc.), so these items are not later classified as a complication of their stay. In addition, it is also very important to document all co-morbidities and other illnesses, so any complications that do occur can be put into an appropriate frame of reference that illustrates how sick a patient really is. If hospital physicians are not careful about documenting past medical and surgical histories, it may appear to people examining the data that many complications happen to healthy people at a certain hospital. In actuality, the hospital may likely treat much sicker patients than what is portrayed by the data.

Conclusion

Although the physician evaluation software utilized by our hospital may be one of the best programs on the market; it still is not perfect, and has pitfalls that physicians should be aware of concerning similar programs at their own institutions. One should not just accept data provided by the hospital without careful critique, as the data could have important ramifications. Since it seems plastic surgery residency graduates are being increasingly employed by hospitals, it would behoove them to be cognizant of this issue. As mentioned in an article by Scott et al in 2011, physicians should be actively involved in choosing evaluation methods and specifying benchmarks, trained in the use of the evaluation methods, and fully aware of their drawbacks and limitations [7]. Even though automated systems are not perfect, Jeff Hanson, Vice President of Thomson Healthcare, gives good advice: "we should continue plodding along the bumpy road toward the best solution, doing the hard work and making the compromises that would result in a fair, constructive and useful measurement process" [10]. The hope is that this article serves to empower physicians to carefully dissect the tools that are being used to evaluate them, and to become proactive in making them better in the future.

References

1. Christianson JB, Volmar KM, Alexander J, Scanlon DP. A report card on provider report cards: status of the health care transparency movement. J Gen Intern Med. 2010; 25: 1235-1241.

2. Rothberg MB, Morsi E, Benjamin EM, Pekow PS, Lindenauer PK. Choosing the best hospital: The limitations of public quality reporting. Health Aff (Millwood). 2008; 27: 1680-1687.

3. Bernstein AB, Gauthier AK. Choices in health care: What are they and what are they worth? Med Care Res Rev. 1999; 56: 5-23.

4. Woo HE. The crucible of physician performance reports. J Gen Intern Med. 2011; 26: 226-227.

5. Hardy K. CRIMSON Initiative aims at performance transparency and improved care quality. Available at: http://www.healthcareitnews.com/news/crimson-initiative-aims-performance-transparency-and-improved-care-quality. Accessed March 19, 2012.

6. The Joint Commission. www.jointcommission.org. Accessed March 17, 2012.

7. Scott IA, Phelps G, Brand C. Assessing individual clinical performance: A primer for physicians. Intern Med J. 2011; 41: 144-155.

8. Higgins A, Zeddies T, Pearson SD. Measuring the performance of individual physicians by collecting data from multiple health plans: The results of a two-state test. Health Aff (Millwood). 2011; 30: 673-681.

9. Scholle SH, Roski J, Adams JL, Dunn DL, Kerr EA, Dugan DP, et al. Benchmarking physician performance: Reliability of individual and composite measures. Am J Manag Care. 2008; 14: 833-838.

10. Hanson J. Measuring physician performance. Building an effective physician performance system starts with transparency. Health Manag Technol. 2008; 29: 10-13.