Annals of Health Law

Volume 16
Issue 2 *Summer 2007*

Article 6

2007

Electronic Healthcare Data Collection and Pay-for-Performance: Translating Theory into Practice

Ramesh C. Sachdeva
Marquette University Law School

Electronic Healthcare Data Collection and Pay-For-Performance: Translating Theory into Practice

Ramesh C. Sachdeva, M.D., Ph.D., D.B.A., J.D., F.A.A.P., F.C.C.M.[*]

INTRODUCTION

This paper focuses on the practical implementation issues related to data collection and interpretation of comparative outcomes evaluation for purposes such as public reporting and pay-for-performance in health care.

This paper is divided into five sections. The first section provides an overview of pay-for-performance in health care.

The second section focuses on the practical issues and challenges related to data collection, such as obtaining the necessary relevant clinical outcomes data to meaningfully satisfy the pay-for-performance requirements. This section is divided into three parts. First, this article discusses the issue of electronic data quality control from the standpoint of reliability and validity. Second, challenges resulting from limitations on the availability of meaningful clinical outcomes data in electronic healthcare data systems are evaluated. Third, it introduces the concept of severity-risk adjustment as a mechanism for providing meaningful comparisons of hospital and physician performance.

The third section identifies the potential legal issues that may emerge as the currently available performance data is used. Legal issues may arise because of: 1) the limitations in data quality, and 2) the meaningfulness of results from a clinical outcomes standpoint due to limitations in the selection of outcomes indicators and the lack of severity-risk adjustment.

---

[*] Dr. Sachdeva currently serves as the Boden Research Fellow and Adjunct Associate Professor of Law at Marquette University Law School. He also holds the position of Executive Vice President and Chief Operating Officer for the National Outcomes Center of Children's Hospital and Health System and Vice President of Quality and Outcomes at the Children's Hospital of Wisconsin. Dr. Sachdeva also is an Associate Professor of Pediatrics at the Medical College of Wisconsin. He would like to thank Lisa Ciesielczyk and Caroline Hackstein for their assistance in preparation of this manuscript.

The fourth section identifies solutions that are being implemented in the healthcare sector. Here I discuss my general experience in the pediatric critical care arena from a national implementation perspective, as well as my specific experience within a pediatric children's hospital from a local healthcare delivery system perspective. This discussion highlights the potential translation for application to other areas within health care.

Finally, the fifth section concludes the discussion and identifies the key lessons to consider while implementing data collection processes for systems such as pay-for-performance. If these lessons are applied, they may serve as a catalyst for the entire healthcare system to achieve transparency, which will provide patients and society with greater choice and added value in health care.

I. PAY-FOR-PERFORMANCE IN HEALTH CARE

Much has been written about the theories related to pay-for-performance, the growing trend in public reporting of healthcare quality information, and the use of information technology to improve patient safety and other challenges facing the healthcare system in the U.S.[1] Such theories and theses are extremely valuable as they provide the foundation for the work that needs to be done in this area. However, the electronic healthcare data collection used to measure the clinical performance of hospitals and physicians has not been fully explored. There are many practical challenges when comparing performance and measuring quality and outcomes of health care, which directly impact the goals of pay-for-performance. This paper focuses on these practical issues and describes some potential solutions that have been successfully implemented in discrete areas of health care.

1. *See, e.g.*, Stacy L. Cook, *Will Pay for Performance Be Worth the Price to Medical Providers? A Look at Pay for Performance and Its Legal Implications for Providers*, 16 ANNALS HEALTH L. 163, 163-212 (2007) (discussing the emergence, structure, and operation of pay-for-performance programs and their legal effects); Ramesh C. Sachdeva, *The New Era of Quality Improvement*, The WIsper (American Academy of Pediatrics—Wisconsin Chapter), March 2006, at 1 (discussing the historical landmarks in the national endeavors to measure healthcare quality and performance, recent efforts that have been successfully implemented into practice, the quality framework proposed by the Institute of Medicine, and the professional responsibility of physicians to measure and improve quality of healthcare); Peter K. Lindenauer et al., *Public Reporting and Pay for Performance in Hospital Quality Improvement*, 356 NEW ENG. J. MED. 486, 492-95 (2007) (study concluding that hospitals already engaged in public reporting achieved modest increases in quality when combined with the financial incentives of pay-for-performance); Helene Nelson, *The Promise of "eHealth"*, 105 WIS. MED. J. 28, 28-29 (2006) (highlighting the U.S. President's August 2006 Executive Order to advance health care through the adoption of health information technology and Wisconsin's plan to do the same via a 5-year plan).

There is a growing body of research and literature in the area of pay-for-performance. Medicare's recent adoption of pay-for-performance lends this emerging research and literature even greater significance. A recent study identified several factors that may have triggered the growing movement toward pay-for-performance.[2] These factors include the increasing availability of quality and outcomes indicators and electronic data systems, the growing national focus on quality of health care and patient safety, and the experiences from other healthcare systems that have successfully implemented policies to link physician reimbursement to performance.[3]

It is important to recognize that pay-for-performance from a policy standpoint is not intended to merely reward or punish certain actions.[4] Rather, pay-for-performance tends to serve as a catalyst for implementing necessary changes and transformations within healthcare systems.[5]

Examples of comprehensive performance indicators for pay-for-performance implemented in the United States include clinical measures, patient ratings, and the adoption of information technology.[6] Consider the Integrated Healthcare Association, which is a collaboration of six health plans in California that serves eight million enrollees.[7] Under its performance indicators, clinical measures can account for as much as 50% of the total performance score.[8] Patient ratings make up 40% of the ultimate score leaving 10% to reflect the group's effective use of information technology.[9]

However, from the patients' and physicians' perspectives there is a need to select the right mix of criteria for quality and ask questions like, "What are the appropriate clinical indicators that should be measured to evaluate performance?"[10] These considerations constitute the focus of this article: the practical issue of selecting performance outcome measures that are meaningful from a clinical standpoint. This issue is integral to the long-term success of any pay-for-performance implementation policy.

In the Pacific Care Health System, which is a large California and Pacific Northwest based healthcare system, a comparison of the impact of pay-for-performance was performed with a control physician group in which pay-

2. *See* A.M. Epstein et al., *Paying Physicians for High-Quality Care*, 350 NEW ENGL. J. MED. 406, 406-10 (2004).

3. *Id.* at 406.

4. *Id.*

5. *Id.*

6. *Id.* at 407; *cf.* Nelson, *supra* note 1, at 28-29 (discussing Wisconsin's 5-year plan to improve delivery of safe, high-quality health care by focusing on better use of health information technology).

7. Epstein, *supra* note 2, at 407.

8. *Id.*

9. *Id.* at 408.

10. *Id.*

for-performance was not implemented.[11] The clinical measures used to estimate quality included cervical cancer screening, mammography, and HbA_{1C} (a laboratory measure related to diabetes control).[12] All of these measures are process based measures reflecting the preventive care aspects of health care. Overall, there was not a significant or consistent level of improvement due to the program.[13] For example, during 2003-2004 year, the health plan allocated almost $12.9 million for pay-for-performance, yet they distributed only $3.4 million.[14] The authors of the study suggested that the incentive design was flawed in this case because the incentives stimulated the underperforming group, yet failed to achieve a similar result elsewhere.[15]

From a practical standpoint, although preventive care is an important aspect of healthcare delivery and can be easily measured, it does not capture the full spectrum of healthcare delivery provided by physicians, particularly in the inpatient hospital setting. A significant amount of health care delivered in the United States is hospital-based, so this type of care requires its own set of performance and outcome measures.[16] Because measuring the performance and outcomes of acute care delivery using electronic data systems is challenging, much of the pay-for-performance clinical indicators are primarily focused on the preventive aspects of health care. However, with the growth of pay-for-performance in the future, it will become imperative that clinical indicators be rapidly expanded from the existing preventive measures to more of the classical clinical outcome measures.[17] These clinical outcome measures should be related to the care provided by physicians in the acute care hospital setting, as well as the ambulatory setting, for measuring long-term patient outcomes.

In an attempt to expand process-based clinical indicators to outcome-based clinical indicators, a significant limitation has been the availability of

---

11. *See* M. B. Rosenthal et al., *Early Experience with Pay-for-Performance: From Concept to Practice*, 294 J. AM. MED. ASS'N 1788, 1788 (2005).

12. *Id.*

13. *Id.*

14. *Id.* at 1792 ("First, groups with baseline performance already above the targeted threshold appeared to understand that they needed only to maintain the status quo to receive the bonus payments. More surprising, perhaps, is that low-performing groups improved as much as they did, given that their short-run chances of receiving a bonus were likely to be low.").

15. *Id.*

16. Steven R. Machlin & Kelly Carper, Agency for Healthcare Research and Quality, *National Healthcare Expenses in the U.S. Civilian Noninstitutionalized Population, 2004*, 1 (Nov. 2006), http://www.meps.ahrq.gov/mepsweb/data_files/publications/st149/stat149.pdf (attributing 31.4 percent of total healthcare spending to hospital inpatient expenses); *cf.* Claudia A. Kozinetz et al., *Health Status of Children with Special Health Care Needs: Measurement Issues and Instruments*, 38 CLINICAL PEDIATRICS 525, 525-33 (1999) (indicating that the outcome measures for children in the ambulatory setting are unique).

17. *See* Rosenthal, *supra* note 11, at 1789.

the clinical outcomes data.[18] With the rapid growth of healthcare information technology, along with the movement toward achieving transparency in federally-administered healthcare programs, outcomes data will increasingly become more available for use.[19] However, as electronic data systems facilitate the rapid evaluation of large amounts of healthcare data, new practical challenges are likely to emerge. Three key practical challenges are prominent: first, issues related to data collection itself from the standpoint of data quality control; second, selection of the appropriate clinical outcomes measures; and finally, the need for severity-risk adjustment to allow for meaningful comparison of the performance of hospitals and physicians. This paper focuses on these practical issues involved in translating theory into practice from the perspective of physicians and hospitals.

II. TRANSLATING THEORY INTO PRACTICE

As discussed in Section I of the paper, three key practical issues must be addressed when translating the theories of information technology-generated electronic data into the practice of health care. These include issues related to data quality, identification of outcomes measures (including the need for patient-generated outcomes measurements), and risk adjustment of clinical outcomes data.

A. Data Quality Control

Data quality control has two facets—validity and reliability.[20] Data validity relates to the accuracy of the measurement relative to what is being measured (i.e. how well does the measurement capture and address the question for the purpose of the measurement?).[21] Data reliability relates to the reproducibility of the same measurement over time (*i.e.* how reproducible are the results?).[22] Both of these aspects, validity and reliability, form the basis of data quality that is integral to the measurement

18. W.S. Weintraub et al., *Can Cardiovascular Clinical Characteristics be Identified and Outcome Models be Developed from an In-patient Claims Database?*, 84 AM. J. CARDIOLOGY 166, 166 (1999) (demonstrating the limitations of administrative databases to measure clinical outcomes).

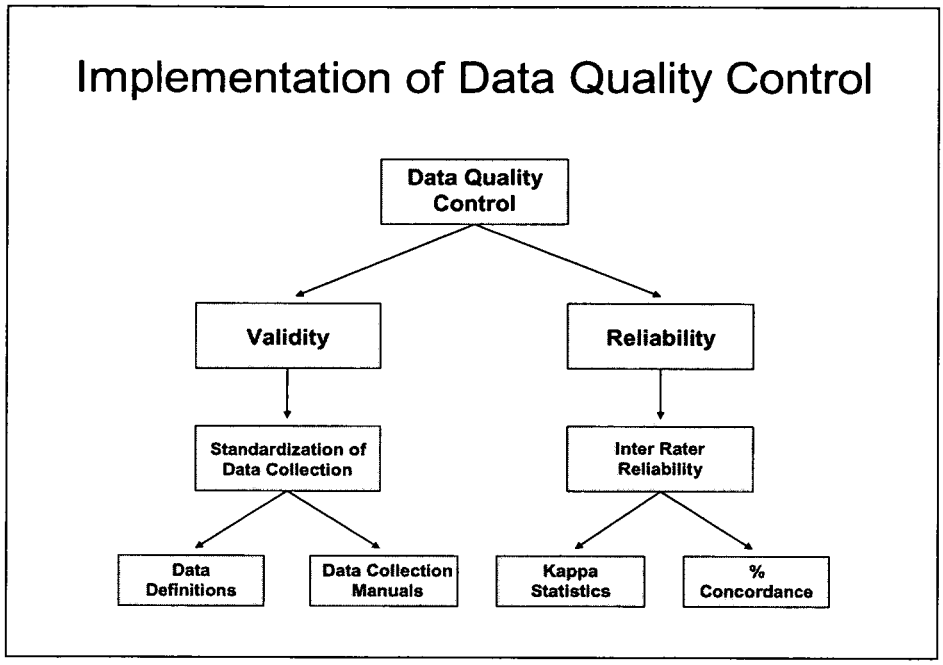
19. *See, e.g.*, Exec. Order No. 13,410, 71 Fed. Reg. 51,089 (Aug. 28, 2006) (stating that "[a]s each agency implements, acquires, or upgrades health information technology systems used for the direct exchange of health information between agencies and with non-Federal entities, it shall utilize, where available, health information technology systems and products that meet recognized interoperability standards.").

20. Claudia A. Kozinetz et al., *Health Status of Children with Special Health Care Needs: Measurement Issues and Instruments*, 38 CLINICAL PEDIATRICS 525, 531 (1999).

21. *Id.* at 526, 531.

22. *Id.*

process.[23] Data quality control can be established using both paper and electronic healthcare data. The implementation of the data quality control process into practice is illustrated in the figure below:



This implementation is further discussed in Section IV of this paper. If adequate data quality control cannot be established, it significantly impacts the results and willingness of key stakeholders in relying on the conclusions.[24]

B. Limitations of clinically relevant outcomes measures to evaluate performance

Another issue that arises from the use of electronic data in health care is the limitations on clinically relevant outcomes measures used to evaluate performance. First, most of the currently available electronic healthcare data is obtained from administrative databases. Administrative databases,

23. Ramesh C. Sachdeva, *Measuring Quality of Legal Services—Implementing Outcomes Research in Law*, 11 J. MED. & L. 1, 20 (2007).

24. *See* Ramesh C. Sachdeva, Mixing Operational Research Methodologies to Achieve Organizational Change—A Study of the Pediatric Intensive Care Unit 100-01 (2005) (unpublished D.B.A. thesis, University of Strathclyde, Glasgow, U.K.) (on file with the University of Strathclyde Library).

although easily available in most hospitals and healthcare settings, are primarily intended for purposes of billing and reimbursement.[25] For this reason, this data lacks the clinical rigor necessary to measure the true underlying quality of care being delivered.[26] In contrast, clinical databases provide the necessary richness to capture the quality of care being delivered.[27] These clinical details are increasingly available in large healthcare databases, which make comprehensive process-based measures of quality possible for even large patient populations.[28] While some recent attempts to evaluate administrative databases for purposes of measuring the outcomes of healthcare delivery have been successful, there remains a general skepticism, particularly among physicians, for the use of this data for purposes of quality and outcomes measurement.[29] At least one study supports this concern that existing administrative electronic data sources are incomplete, estimating that up to 20 percent of the essential data elements may be missing.[30] Although many comprehensive clinical databases are being rapidly developed, and such data is likely to be more readily available in the future, the current lack of databases does represent a significant limitation.[31]

Second, currently available electronic data is limited with respect to measuring long-term patient outcomes that are clinically relevant. When measuring the quality of medical care, the performance measures include both process and outcome measures.[32] Process measures include preventive

25. *Cf.,* Christine A. Beck et al., *Administrative Data Feedback for Effective Cardiac Treatment: AFFECT, a Cluster Randomized Trial*, 294 J. AM. MED. ASS'N 309, 314-15 (2005) (positing that the administrative data was not effective in the study at improving care because it was perceived as invalid or irrelevant to practice).

26. *E.g., id.* at 314 (2005) (finding that administrative data is not effective in improving the quality of acute myocardial infarction care).

27. *See* Arthur J. Hartz & Evelyn M. Kuhn, *Comparing Hospitals That Perform Coronary Artery Bypass Surgery: The Effect of Outcome Measures and Data Sources*, 84 AM. J. PUB. HEALTH 1609, 1609 (1994).

28. R.H. Palmer, *Process-Based Measures of Quality: The Need for Detailed Clinical Data in Large Health Care Databases*, 127 ANNALS INTERNAL MED. 733, 736-37 (1997).

29. *E.g.,* Ramesh C. Sachdeva & Cinthia S. Christensen, Presentation at the Annual Meeting for the National Initiative for Children's Healthcare Quality (NICHQ): Implementation of an Innovative Organizational Policy for Clinical Program Evaluation and Strategic Planning—A Programmatic Approach for Enhancing Quality of Care in Children (Mar. 2005); *see* Beck, *supra* note 25, at 314; *see* Sara Bornstein Voit et al., *Electronic Surveillance System for Monitoring Surgical Antimicrobial Prophylaxis,* 116 PEDIATRICS 1317, 1318-21 (2005).

30. *See, e.g.*, Leon G. Fine et al., *How to Evaluate and Improve the Quality and Credibility of an Outcomes Database: Validation and Feedback Study on the UK Cardiac Surgery Experience*, 326 BRIT. MED. J. 25, 26 (2003).

31. Palmer, *supra* note 28, at 737 (arguing that quality of care determination requires measuring outcomes of care or processes that have shown to lead to good health outcomes, and that there is currently a limitation to even obtain the necessary process based quality indicators from databases).

32. Avedis Donabedian, *Evaluating the Quality of Medical Care*, 44 MILBANK

healthcare measures such as immunization rates.[33] Although important, process measures do not provide an assessment of the real clinical quality of health care. Rather, clinical quality of health care needs to be evaluated using clinical outcomes measures, which can include both objective and subjective measures.[34] Objective measures include mortality or survival, and length of hospital stay, while subjective measures include patient functional outcomes and quality of life measurements.[35] More recent research indicates that such subjective outcome measures can be relatively objectively measured, with a fair degree of accuracy, for application into clinical practice.[36] Currently, most electronic healthcare data sources capture the short-term objective measures such as mortality or survival and length of hospital stay.[37] However, such electronic data sources are limited in that they do not capture long-term measures such as functional status and quality of life over time.[38]

Another limitation of the available outcome measures is that they are typically developed by healthcare providers.[39] Although these providers attempt to capture the true patient outcomes, these measures are limited because they do not identify what is important from patients' perspectives.[40] Some developing efforts attempt to expand the comprehensiveness of

---

MEMORIAL FUND Q. 166, 166-69 (1966) (discussing an spproach using the structure-process-outcome model for the evaluation of medical care).

33. *Id.* at 169.

34. *See, e.g.*, Ramesh C. Sachdeva, *Statistical Basis and Clinical Applications of Severity of Illness Scoring Systems in the Intensive Care Unit,* 5 CURRENT OPINION CRITICAL CARE 180, 180 (1999) ("The development of severity-of-illness systems in ICU's has historically included a subjective and objective approach."

35. *See id.* ("The subjective approach uses varying degrees of expert group consensus in contrast to the objective approach, which uses formal statistical techniques.").

36. *See* Reinould J.B.J. Gemke et al., *Long Term Survival and State of Health After Pediatric Intensive Care*, 73 ARCHIVES DISEASE CHILDHOOD 196, 199 (1995).

37. *See, e.g., id.* at 196 ("Success of intensive care is usually presented as mortality rate . . . disregarding long term survival and functional outcome.").

38. *See, e.g., id.* (stating there is a growing concern for long term prospects of pediatric ICU patients, and accordingly "longitudinal assessment of morbidity change and health related quality of life have become important supplementary outcome measures").

39. This is based upon my experience at Children's Hospital of Wisconsin and also my experience at the Texas Children's Hospital in Houston, Texas (1993-99). Further, I was the recipient of the research grant from the NIH (NIH-CAP Award) in 1996 at the Baylor College of Medicine in Houston, and this grant was aimed at studying the long term outcomes of children during and after discharge from the hospital. An additional aim of this grant was to evaluate the feasibility of developing patient generated outcomes. This research effort has evolved into practice as part of my quality improvement leadership role and activities at the Children's Hospital of Wisconsin.

40. *See* Danny A. Ruta et al., *A New Approach to the Measurement of Quality of Life: The Patient-Generated Index,* 32 MED. CARE 1109, 1110 (1994) (stating clinicians' perception of disease are reflected by clinical measures of health status that include limited aspects of a patient's life, which concentrate on physiological and physical measurements, and do not correlate well with the patient's perceptions of health status).

available outcomes measures by directly involving patients or their families in the outcomes measurement development process.[41]

C. Severity-risk adjustment of data

Another practical and significant limitation of using the currently available electronic healthcare data to measure performance and outcomes of care is that such data are typically not severity-risk adjusted. Severity-risk adjustment is the concept of statistically adjusting data to compensate for differences in patients' clinical severity.[42] This allows for more accurate and meaningful comparison of the differences within results or outcomes for patients based upon the interventions performed by healthcare providers.[43] Severity risk adjustment can be accomplished by using statistical models in which the severity of illness score is used along with other variables to provide a predicted probability of an outcome. The observed to predicted outcome probabilities are then statistically compared to measure performance.[44]

The process of performing severity risk-adjustment is scientifically valid and well established.[45] It involves the development of clinically relevant severity scoring systems that may include clinical factors such as multiple physiologic and laboratory parameters, as well as other underlying non-clinical factors such as age, gender, and cultural, ethnic, and socioeconomic factors.[46] It is a crucial step in analyzing outcomes analysis across healthcare facilities, because without such an adjustment the comparisons of outcomes can become meaningless.[47] To further illustrate this point, consider a hypothetical hospital or physician practice group providing care for highly acute patients with significant underlying clinical risk. These patients are likely to have worse outcomes as compared to another hospital

41. *See* Council for Quality Clinical Program Performance Reports (2006) (on file with the Children's Hospital of Wisconsin, Milwaukee, Wis.).

42. *See, e.g.*, RISK ADJUSTMENT FOR MEASURING HEALTHCARE OUTCOMES 3 (Lisa I. Iezzoni ed., Health Admin. Press 3d ed. 2003) ("Risk adjustment aims to account for differences in intrinsic health risks that patients bring to their health care emergencies.").

43. Sachdeva et al., *Mixing Methodologies to Enhance the Implementation of Healthcare Operational Research,* 58 J. OPERATIONAL RES. SOC'Y 159, 161 (2007) (supporting that available healthcare data is typically not severity risk-adjusted and it is important to make results meaningful to clinicians).

44. Ramesh C. Sachdeva, Risk Adjustment of Clinical Performance Quality Data to Facilitate Outcomes Comparisons of Physicians and Hospitals, Conference on Empirical Legal Studies (Oct. 2006).

45. *See* Iezzoni, *supra* note 42, at 4 (noting the more than two decades of intensive research that have produced "credible risk-adjustment methods for certain outcomes in widely divergent contexts").

46. *Id.* at 4-5 (diagramming patient factors that combine with treatment effectiveness and random events to produce a range of possible outcomes).

47. *Id.* at 3.

or provider group that provides care for less acute patients. By performing severity risk-adjustment, the emerging results can be better standardized to account for differences in levels of patient acuity and how they translate to differences in patient clinical risk and outcomes.

While a large body of health services research dealing with severity-risk adjustment exists, this research is not typically part of standard electronic databases. Although the DRG System (a risk stratification scheme) may be available in electronic databases, the use of this process is limited as the allowance for risk stratification still does not include formal severity-risk adjustment.[48] Although the DRG system is not a physiological scoring system when compared to severity scoring systems such as the Pediatric Risk of Morality (PRISM), the DRG classification scheme can be successfully used to risk stratify patients and take a step towards full severity risk-adjustment.[49]

A landmark comparison of the outcomes of referral tertiary care hospitals with non-tertiary care hospitals in Oregon highlighted the important role of severity-risk adjustment.[50] An initial evaluation of the data suggested that the outcomes in the tertiary care intensive care units, as measured by mortality (and survival), was significantly higher than that seen in the non-tertiary care setting.[51] However, after performing the appropriate severity-risk adjustment, this difference was eliminated.[52] Further, the actual risk adjusted survival chances of a patient in a tertiary care facility was found to be significantly higher than that of a non-tertiary care facility, contrary to the initial results prior to the severity-risk adjustment.[53]

In order to perform such severity-risk adjustment comparisons, formally validated severity-risk scoring systems are needed.[54] This is beyond what is currently available in electronic data sources through the DRG System. Although providing a mechanism to stratify and group patients into different risk categories, the DRG System does not allow for a formal statistical severity-risk adjustment of the nature described above. The use of severity-risk adjustment is crucial in order to perform quality

48. Ramesh Sachdeva & Julie Pedretti, Using Clinical Program Quality Improvement to Enhance Your National Brand, NACHRI National Meeting (Oct. 11, 2005).

49. *See* Council for Quality Clinical Program Performance Reports, *supra* note 41.

50. *See* Murray M. Pollack et al., *Improved Outcomes from Tertiary Care Center Pediatric Intensive Care: A Statewide Comparison of Tertiary and Non-tertiary Facilities*, 19 CRIT. CARE MED. 150, 150 (1991).

51. *Id.*

52. *Id.* at 151.

53. *Id.* at 157.

54. *See, e.g.*, Ramesh C. Sachdeva, *Functional Outcomes in Pediatric Models*, 3 CURRENT OPINION CRITICAL CARE 179, 179-80 (1997) (stating that when predicting short term mortality risks in the ICU setting, it is important to account for those significant variables in each patient which predict outcomes).

comparisons of clinical programs and physicians, and it is vital for the necessary buy-in of the results by physicians. This is further discussed in Section IV of this paper.

III. POTENTIAL LEGAL IMPLICATIONS

Two broad sets of legal issues potentially emerge as data quality, identification of outcomes measures, and severity-risk adjustment are considered. First, the Federal Anti-Kickback Statute,[55] Federal Anti-Referral (or "Stark") Law,[56] and HIPAA[57] all present traditional legal concerns in the context of implementing a pay-for-performance program. Additionally, the lack of data quality control or inadequate severity-risk adjustment further magnifies these pre-existing legal issues. Also, physicians remain concerned about traditional standard of care violations that may result in civil medical malpractice litigation. The risk arises because data quality might be used to establish the standard of care in a given case, but a lack of severity-risk adjustments could change the comparisons being used to quantitatively establish the standard of care.[58] While a relatively novel concept in medical malpractice litigation, the issues surrounding violations of the law and medical malpractice were recently addressed at Loyola University Chicago School of Law's Sixth Annual Health Law and Policy Colloquium, *Diagnosing the Data*.[59]

Second, although not readily obvious, data quality control and severity-risk adjustment can result in potential violations of informed consent. An important decision by the Wisconsin Supreme Court in *Johnson v. Kokemoor* demonstrates this point.[60] In *Johnson*, the plaintiff sued her physician for not disclosing the associated risks of surgery to treat her brain aneurism.[61] The patient alleged that the physician failed to obtain her informed consent as the law requires.[62] Ms. Johnson argued that her physician had a duty to provide the quantitative outcomes data related to risks and benefits of proceeding with the surgery, and specifically the mortality rates, because it may have impacted her decision whether to

55. 42 U.S.C. §1320a-7b(b) (2006) (criminalizing the knowing or willful solicitation or receipt of remuneration for referring persons for care covered by a federal healthcare program).

56. 42 U.S.C. §1395nn (2006) (prohibiting certain referrals to entities with which physicians have a "financial relationship").

57. 42 U.S.C. §201 *et seq.* (2005).

58. *See* Michelle M. Mello, *Using Statistical Evidence to Prove the Malpractice Standard of Care: Bridging Legal, Clinical, and Statistical Thinking*, 37 WAKE FOREST L. REV. 821, 852 (2002).

59. 16 ANNALS HEALTH L. 2 (forthcoming June 2007).

60. Johnson v. Kokemoor, 199 Wis. 2d 615 (1996) (discussing the application of sharing outcomes data as a requirement for obtaining informed consent in health care).

61. *Id.*

62. *Id.*

obtain surgery from Dr. Kokemoor or a different physician.[63] Whether Dr. Kokemoor met his duty depended on the availability of comparative data related to outcomes of this surgical procedure.[64] Ultimately, the jury found the defendant liable, concluding that "a reasonable person in the plaintiff's position would have refused to consent to surgery by the defendant if she had been fully informed of its attendant risks and advantages."[65]

The process of data collection, analysis, and interpretation can lead to unintended consequences and place physicians in a paradoxical situation. Inadequately performed data quality control or poorly performed analysis, including inadequately performed severity-risk adjustment, can result in imprecise results and conclusions related to hospital and provider performance. Nevertheless, as electronic health systems become more widely available, they will likely be groomed to produce performance data related to hospital and physician outcomes. Should these results be made public and induce reliance by patients, physicians will have a duty to disclose such data as part of informed consent.[66] The court in *Johnson* defined such a disclosure as one that "requires assessment of the gravity of the patient's condition and the probabilities of success."[67] Strictly imposing this duty is problematic if doctors were required to disclose inherentlyflawed data, and patients would likely rely on this misinformation. Fortunately, the court did not outline a blanket duty requiring the disclosure of comparative risk evidence to obtain informed consent, but rather cautioned that its decision "will not always require" this duty.[68]

IV. SOLUTIONS TO ADDRESS THE LIMITATIONS OF ELECTRONIC HEALTHCARE DATA—IMPLICATIONS FOR LAW

This section includes three parts with a focus on implications for law. First, the Virtual Pediatric Intensive Care System (VPS) is described as an example of how data quality control and severity-risk adjustment have been translated from theory into practice, while continuously adjusting to an ever-increasing network of acute care hospitals in the United States. Second, to highlight how data quality control can be incorporated into the pediatric setting using electronic health systems, I describe my experience helping to create a unique department, Clinical Data Management, at the Children's Hospital of Wisconsin. Third, the article molds these

63. *See id.* at 623-27.

64. *Id.* at 623.

65. *Id.* at 620-21.

66. *See Johnson*, 199 Wis. at 646.

67. *Id.* at 632 (quoting Scaria v. St. Paul Fire & Marine Ins. Co., 68 Wis. 2d 1, 11 (1975)).

68. *Id.* at 646.

experiences into a template for application in other hospitals in the United States, to support the goal of successfully implementing programs such as pay-for-performance and managing the public reporting of performance data.

A. The National Experience

The Virtual Pediatric Intensive Care Unit System (VPS) is a national consortium of pediatric institutions created to improve the care of critically ill children by fostering research and quality improvement strategies.[69] The VPS includes three entities that provide the overall leadership for this national initiative: the National Outcomes Center from the Children's Hospital and Health System in Milwaukee, Wisconsin; the National Association of Children's Hospitals and Related Institutions in Washington, D.C.; and the Children's Hospital of Los Angeles in California.[70] These three entities developed the VPS national initiative to standardize data sharing and benchmarking among PICUs.[71] All participants in the VPS system collect information on patient and hospital measures, diagnoses, interventions, discharge, organ donation, pediatric severity of mortality scores, and discharge data.[72] Currently, there are nearly 100,000 patients in this electronic dataset, which was developed by and for pediatric intensivists and includes both medical and surgical diagnoses.[73] In our experience, the VPS system includes three unique features—a strong emphasis on data quality control in the data collection process, selection of indicators by clinicians so that these indicators are clinically relevant, and severity-risk adjustment of the data so that the results are meaningful.

Data quality control is established through a formal process that addresses data validity and reliability.[74] Data validity is established by having standardized data definitions for all the fields within the VPS electronic data system.[75] This minimizes subjective interpretations of various data elements, which could otherwise result in significant

69. VPS, LLC, *Critical Care (PICU) Research and Collaboration*, https://myvps.org (last visited Apr. 21, 2007).

70. Lesley Sedehi & Kristine Shulz, *Changing Children's Health Care Through Collaboration*, CHILDREN'S HOSPS. TODAY 36, 28 (Fall 2006), *available at* http://www.childrenshospitals.net/AM/Template.cfm?Section=Search&template=/CM/HTMLDisplay.cfm&ContentID=23007.

71. *Id.*

72. VPS, LLC, *List of Data Elements* 1, https://myvps.org/Documents/VPS_List_of_Data_Elements.pdf (last visited Apr. 21, 2007).

73. VPS, LLC, *Bridging the PICU Continuum of Research, Quality Improvement, and Management* 1, https://myvps.org/Documents/VPS_Brochure.pdf (last visited Apr. 21, 2007).

74. VPS, LLC, *About the VPS Software*, https://myvps.org/software.aspx (last visited Apr. 21, 2007).

75. *Id.*; VPS, LLC, *List of Data Elements*, *supra* note 72.

inaccuracies in the electronic data capture. Reliability is established by an initial inter-rater reliability check and subsequent ongoing inter-rater reliability statistical quality control assessments.[76] Unless a hospital and its data collectors successfully demonstrate the minimum threshold for the inter-rater reliability (greater than 90% concordance), they are not allowed to enter data into the VPS System.[77] Accordingly, these standards have resulted in a rich electronic repository of clinical and administrative data, with a high degree of reliability and validity. This reliability, in turn, has enhanced the buy-in and willingness to use the results emerging from this electronic data source by hospital administrators, physicians and nurses, and communities.[78]

The indicators used to measure performance are clinically driven—the electronic data is clinically relevant and captures the outcomes of health care provided by physicians and hospitals.[79] This has required active collaboration in the development and ongoing maintenance of the VPS. New clinical measures continue to evolve and the healthcare electronic data must be updated in a dynamic manner to reflect the changing clinical environment.

As discussed in Section II above, severity-risk adjustment is a key component of forming outcomes analysis for comparing results across hospital and physician groups. Accordingly, two severity-risk systems are built into the VPS electronic data system. These include the PRISM Severity Scoring System and the Pediatric Index of Mortality (PIM Scoring System).[80]

The PRISM Scoring System is a clinical scoring system comprised of 14 elements derived from individual patients' physiologic and laboratory information.[81] It has been successfully tested in numerous studies in the U.S. and Europe.[82] The system seeks to measure physiologic instability, a predictor of mortality, but it also adjusts for other factors that may impact

76. VPS, LLC, *About the VPS Software*, *supra* note 74; Sachdeva, *supra* note 23, at 20-21.

77. Sachdeva, *supra* note 23, at 21.

78. Sachdeva, *supra* note 24, at 100-01 (describing controls on data quality as essential to encouraging physician participation).

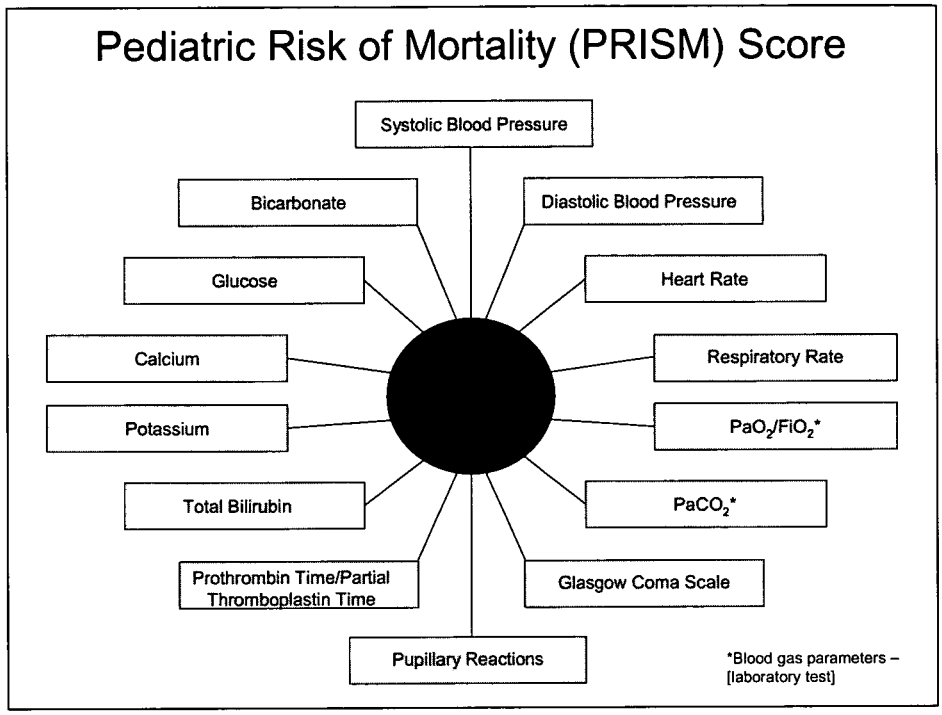
79. *See* VPS, LLC, *List of Data Elements*, *supra* note 72.

80. *Id.*

81. *See* Murray Pollack et al., *The Pediatric Risk of Mortality (PRISM) score*, 16 CRITICAL CARE MED. 1110, 1110 (1988).

82. *See, e.g.*, A.L. Davis et al., *Comparisons of French and USA Pediatric Intensive Care Units*, 17 RESUSCITATION 143, 143 (1989).

this value.[83] The 14 elements of the PRISM scoring system are illustrated in the following figure: [84]



The PRISM score provides weighted numbers that relate to the patient level system dysfunction.[85] The cumulative raw PRISM score is a measure of the level of patient severity.[86] When used along with age and operative status of the patient in a logistic regression model, the PRISM Scoring System provides the predicted risk of mortality for that patient during that admission.[87] By measuring the observed survival and mortality, and comparing it to the predicted risk of survival and mortality over a large number of patients, an estimate of the performance of hospitals and physician groups can be statistically computed.[88]

83. *See* Pollack, *supra* note 81, at 1110-11.

84. *See id.* at 1110.

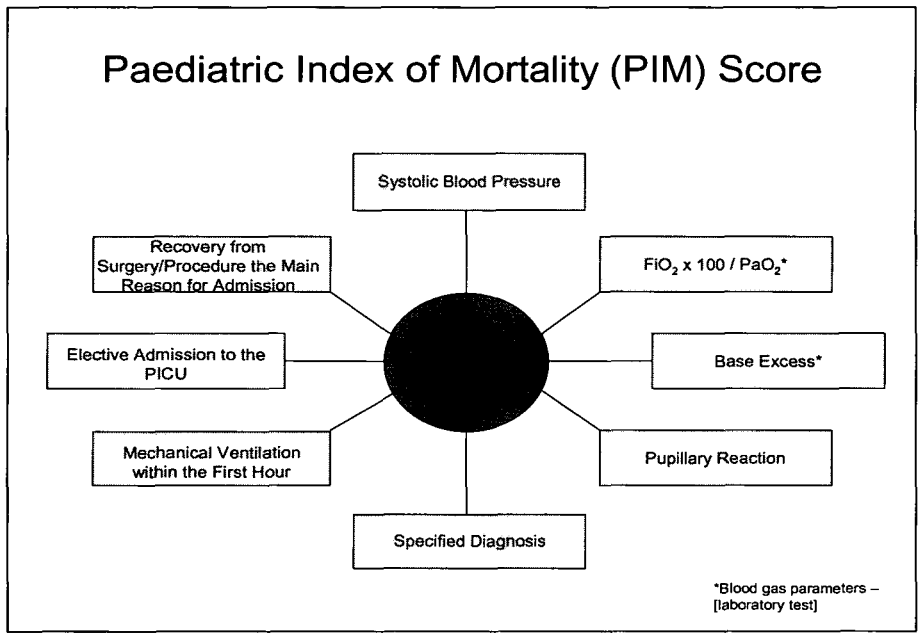
85. *Id.*

86. *See id.* at 1010-12.

87. *See id.*

88. *See* Council for Quality Clinical Program Performance Reports, *supra* note 41 (measuring and comparing the national clinical performance using electronic healthcare data and conducting severity-risk adjustment for pediatric intensive care units in sixty-seven hospitals in the U.S.).

The PIM scoring system, developed in Australia, is similar to the PRISM scoring system.[89] The PIM system has been used routinely in intensive care units in Australia and New Zealand to predict the mortality rates of groups of patients since 1998.[90] Since then, research based on a more recent data set (incorporating data from fourteen intensive care units, eight in Australia, four in the United Kingdom, and two in New Zealand) has led to an expanded new model with wider application: the PIM2.[91] An initial validation of the PIM Scoring System in the U.S. suggests that this is a reliable severity-risk adjustor;[92] however, the formal validation of this severity scoring system is currently in progress. The elements of the PIM scoring system are illustrated in the figure below:



The importance of data quality control in the data collection phase and severity-risk adjustment in the analysis and reporting phase cannot be over-emphasized. These attributes have resulted in the successful buy-in of the emerging results from the VPS data system by hospital administrators and

89. Anthony Slater et al., *PIM2: a Revised Version of the Paediatric Index of Mortality,* 29 INTENSIVE CARE MED. 278, 279 (2003).

90. *Id.* at 279.

91. *Id.* at 279, 282 ("PIM2 is derived from a larger, more recent and more diverse data set than the one used for the first version of PIM.").

92. *See* Council for Quality Clinical Program Performance Reports, *supra* note 41.

healthcare providers.[93] To illustrate this point, the National Outcomes Center has been working with the VPS network to develop outcomes reports for hospitals.[94] As of the end of 2006, the National Outcomes Center has developed 33 such reports.[95] These reports have been met with a high level of acceptance from various hospitals and physician groups, despite the fact that in many instances such hospitals and groups are performing below the national average.[96] In the absence of having adequate data quality control and severity-risk adjustment, a frequent concern—particularly in instances where the observed performance is unfavorable—is the unwillingness of hospitals and physician groups to accept the data behind these results.[97] Many hospitals, such as the Children's Hospital of Wisconsin, have gone a step further to make such data fully transparent, which means sharing the results that emerge from electronic data systems such as the VPS with other healthcare providers and with patient families. This change is a tremendous step towards achieving full transparency of healthcare quality using electronic data systems, as encouraged by the President's Executive Order in August, 2006.[98] However, supporting openness and transparency is not possible without a high level of acceptance of the available electronic data being used to perform the analyses. By having a significant focus towards data quality control and severity-risk adjustment the VPS system helps accomplish acceptance by hospitals and physician groups.

The impact of the VPS system is felt beyond the hospital walls and has implications for the law. The VPS experience shows that electronic healthcare data developed with an emphasis on data quality control, selection of clinically relevant indicators, and performance data that is severity-risk adjusted can successfully result in a high level of physician and hospital buy-in, even in situations when the results are unfavorable. Further, this allows confidence among physicians to share the outcomes data openly and with transparency, despite the legal risk of loss of quality improvement protection should discovery and litigation ensue. Data quality control and severity-risk adjustment will likely minimize legal challenges related to data accuracy and applicability of performance measures for a specific clinical practice or hospital when implementing a pay-for-performance policy.

---

93. *See* Sachdeva, *supra* note 24, at 100-01.

94. VPS, LLC, *Bridging the PICU Continuum of Research, Quality Improvement, and Management*, *supra* note 73.

95. *See Clinical Program Performance Reports*, *supra* note 41.

96. *See id.* Additionally, Dr. Sachdeva serves as the Vice President of Quality and Outcomes at the Children's Hospital of Wisconsin, and facilitates its Council for Quality.

97. Sachdeva, *supra* note 24, at 100-01.

98. Exec. Order No. 13,410, *supra* note 19.

B. The local experience

The example illustrated above highlights a national application in a critical care healthcare setting across many hospitals in the U.S. An example relating to the experience of a single pediatric institution representing a healthcare delivery system is discussed below. In my role of overseeing the quality and outcomes function at the Children's Hospital of Wisconsin, I facilitated the development of a new unique department to address many of the challenges related to healthcare data. The Department of Clinical Data Management was created in 2005 at the Children's Hospital of Wisconsin. This department is an integral component of the Division of Quality at this institution. As part of the creation of this department, a full departmental staff infrastructure was developed and integrated into the organizational operational strategy.

This group allows a team to explicitly focus on data quality control issues for purposes of internal and external reporting. This work includes public reporting of clinical performance and healthcare quality, and also preparation for the pay-for-performance movement that is rapidly growing within the country. Recognizing that no electronic data system is perfect, the Department of Clinical Data Management focuses on identifying the limitations of existing electronic healthcare data sources and making such limitations explicit. By formally identifying the limitations of the electronic data system, the emerging analyses and results can be more meaningfully interpreted and acted upon. This department also works closely with other hospital departments such as marketing and public relations to ensure that the external reporting of such data is provided in an accurate and reliable manner.

To understand the legal implications of ensuring data quality control measures, consider the process of a national survey conducted by the *Child* magazine in United States.[99] The *Child* magazine survey is a data-driven, publicly available system that provides a rating to identify the top children's hospitals in the United States.[100] Because this information is so unique, it is very likely that parents and consumers would rely on the results when selecting hospitals for their children. Given the concerns of data completeness and accuracy in existing hospital electronic data sources, the Department of Clinical Data Management works closely with the public relations department at the Children's Hospital of Wisconsin to carefully interpret and answer the questions within the survey of the *Child* magazine.

99. Karen Cicero, *The 10 Best Children's Hospitals*, CHILD 81, 82 (Feb. 2007). The article described this process: "Our comprehensive 247-question survey . . . examines vital medical information including survival rates, the number of complex procedures and intricate surgeries conducted, volume of research studies, efforts to reduce medical errors, and the quality and training of the doctors and nurses—as well as child-friendliness, support for families, and community involvement." *Id.*

100. *Id.* at 81-82.

Consequently, the results reported in such surveys are generally accurate. This provides the right information to healthcare consumers for making informed choices and mitigates concerns of future legal challenges related to these data and results.

The Department of Clinical Data Management spans its activities across both the acute care setting within pediatrics (such as neonatal intensive care units and emergency medicine) and ambulatory outpatient settings (such as gastroenterology and neurosciences). Consequently, the Department of Clinical Data Management works with a large variety of electronic healthcare databases and data systems including the Pediatric Health Information System (PHIS), the Vermont Oxford Network, and The National Pediatric Trauma Registry, among others. A common feature that has allowed a high level of institutional acceptance of results emerging from these endeavors is these systems' focus on determination of data quality control.

C. Expanding the pediatric experience to the adult setting (including applications in recent national endeavors for physician board recertification)

The focus of the VPS and the Department of Clinical Data Management has been primarily on data quality control and severity-risk adjustment for purposes of overall quality enhancement and internal and external public reporting in pediatrics. However, the implications of these experiences in other emerging policies are important; sustaining the success of a policy, such as pay-for-performance, depends on its ability to withstand future legal challenges.

In the absence of data quality control checks and severity-risk adjustments, pay-for-performance can potentially open the floodgates for litigation, as discussed in Section III. The American Board of Pediatrics has begun to address this issue by adopting the VPS system to measure performance and quality of physicians seeking recertification.

The American Board of Pediatrics, a member of the American Board of Medical Specialties (which includes all adult specialty boards, such as surgery, internal medicine, and family practice), provides certification of all pediatricians in the United States and some physicians in other countries.[101] Along with many of the adult specialty boards within the American Board of Medical Specialties, the American Board of Pediatrics has recently taken the position to require the measurement of performance and quality of physicians seeking recertification. Currently, recertification is a requirement for all pediatricians and pediatric subspecialists on a seven year

101. *See* The Am. Bd. of Pediatrics, https://www.abp.org/ABPWebSite; then follow "About the ABP" hyperlink (last visited Apr. 21, 2007).

basis. However, in order to measure the quality and performance of physicians, there is a need for specific data.

Although there are many electronic healthcare data sources available that may provide insights for determining physician performance, tremendous skepticism and concern surround the use of electronic healthcare data sources both from the perspective of the American Board of Pediatrics and also physicians seeking recertification. This skepticism is due to concerns over data quality control and lack of severity-risk adjustment in the electronic healthcare data sources. In the event that such data were used for determining physician performance and competencies, there is a real likelihood of legal challenges from physicians who fail to obtain recertification. Accordingly, in fall of 2006, the American Board of Pediatrics implemented a pilot project in the U.S. using the VPS data system as a template.[102] A key attribute for using the VPS data system was the focus on data quality control and severity-risk adjustment as discussed in Section IV(A).

It is very likely that similar concerns from physicians and hospitals will emerge as the pay-for-performance movement continues to grow. Although current electronic healthcare sources may appear to provide the necessary data for successfully implementing policies such as pay-for-performance, the lack of formal quality control and severity-risk adjustment in most of the data systems puts these efforts at significant risk of low acceptance of results and increased legal challenges.

Experiences from the VPS in the acute care intensive care setting and the Department of Clinical Data Management at the Children's Hospital of Wisconsin in the pediatric setting can be expanded to other acute care hospitals and adult settings. Such application is valid because the underlying issues and their solutions related to electronic healthcare data collection and analysis in the pediatric setting are not different from the issues faced in the adult healthcare setting. Although the actual severity-risk adjusters may be different in children or adult patients, the underlying scientific paradigm for implementation remains robust. The adult acute care setting has severity-risk adjusters available that are similar to the PRISM and PIM severity scoring systems described above.

Recent collaborative efforts between the VPS and the adult Society for Thoracic Surgeons (STS) demonstrate this similarity of principles across adult and pediatric settings and further support this recommendation. The VPS and STS have been in close discussions over the past year and have agreed to actively partner in expanding the experience from the two groups—the VPS representing acute care (critically ill pediatric patients), and the STS representing pediatric and adult cardiology patients.[103] This

102. Sedehi & Shulz, *supra* note 70, at 26.

103. Collaborative meetings held August 25-26, 2005, in Chicago, Illinois, and January 9-10, 2006, in Dallas, Texas, solidified this partnership in principle. However, no formal

will result in the availability of electronic health data systems that span the entire age spectrum of patients with heart disease needing surgery. Furthermore, such data will represent clinically relevant outcomes measures with a high level of data quality control and appropriate severity-risk adjustment. These efforts will allow the objective evaluation of performance of hospitals and physician groups. Such evaluation may then be utilized to effectively implement public reporting of healthcare quality and pay-for-performance.

V. CONCLUSIONS

The experiences from the VPS at a national level and the Department of Clinical Data Management at a local level support that the three practical issues related to data collection and interpretation—quality control, identification of clinically relevant outcomes measures, and severity-risk adjustment—can be successfully addressed and implemented into practice by adopting scientifically sound approaches. Such programs have significant implications for law by minimizing the risk of potential legal challenges, which, in turn, results in a high level of buy-in from physicians and hospital administrators. Further, they foster the willingness of physicians and hospital administrators to make such data more transparent and readily available to patients and the communities they serve. Willingness to embrace this transparency in healthcare quality by physicians and hospitals is a catalyst for the successful implementation of policies related to pay-for-performance, and it also enhances the ability to develop meaningful and actionable information from electronic healthcare data.

partnership agreement currently exists. The VPS leadership and I made numerous presentations at both of these meetings.