



Published in final edited form as:

*Med Care*. 2007 January ; 45(1): 73–79. doi:10.1097/01.mlr.0000244510.09001.e5.

## Measuring Performance Directly Using the Veterans Health Administration Electronic Medical Record:

### A Comparison With External Peer Review

Joseph L. Goulet, PhD, MS<sup>\*,†</sup>, Joseph Erdos, MD<sup>‡</sup>, Sue Kancir, RN<sup>†</sup>, Forrest L. Levin, MS<sup>†</sup>, Steven M. Wright, PhD<sup>‡</sup>, Stanlie M. Daniels, RN<sup>‡</sup>, Lynnette Nilan, RN, EdD<sup>‡</sup>, and Amy C. Justice, MD, PhD<sup>\*,†</sup>

<sup>\*</sup>Yale University School of Medicine, New Haven, Connecticut

<sup>†</sup>VA Connecticut Healthcare System, West Haven, Connecticut

<sup>‡</sup>Veterans Health Affairs Central Office (VACO), Washington, DC

### Abstract

**Background**—Electronic medical records systems (EMR) contain many directly analyzable data fields that may reduce the need for extensive chart review, thus allowing for performance measures to be assessed on a larger proportion of patients in care.

**Objective**—This study sought to determine the extent to which selected chart review-based clinical performance measures could be accurately replicated using readily available and directly analyzable EMR data.

**Methods**—A cross-sectional study using full chart review results from the Veterans Health Administration's External Peer Review Program (EPRP) was merged to EMR data.

**Results**—Over 80% of the data on these selected measures found in chart review was available in a directly analyzable form in the EMR. The extent of missing EMR data varied by site of care ( $P < 0.01$ ). Among patients on whom both sources of data were available, we found a high degree of correlation between the 2 sources in the measures assessed (correlations of 0.89–0.98) and in the concordance between the measures using performance cut points (kappa: 0.86–0.99). Furthermore, there was little evidence of bias; the differences in values were not clinically meaningful (difference of 0.9 mg/dL for low-density lipoprotein cholesterol, 1.2 mm Hg for systolic blood pressure, 0.3 mm Hg for diastolic, and no difference for HgbA1c).

**Conclusions**—Directly analyzable data fields in the EMR can accurately reproduce selected EPRP measures on most patients. We found no evidence of systematic differences in performance values among these with and without directly analyzable data in the EMR.

### Keywords

veterans; quality of care; medical records systems; quality measurement

---

The Veterans Health Administration (VA) is the largest integrated healthcare delivery system in the United States and is a leader in healthcare quality and efficiency.<sup>1,2</sup> Several strategies have enabled this achievement: emphasizing outpatient and primary care,

instituting clinical performance measurements, and development of an integrated national electronic information system (EMR).<sup>3,4</sup> An additional component was the development and measurement of clinical practice guidelines as a means of improving patient care outcomes.<sup>5-7</sup>

To assess guideline adherence, the VA Office of Quality and Performance (OQP) created the External Peer Review Program (EPRP), administered by an independent contractor at a substantial cost per year, to track disease-specific and general prevention measures.<sup>8,9</sup> In VA fiscal year 2005, EPRP assessed 55 quality measures in 9 areas: within each, several inpatient-based and outpatient-based targets were evaluated. Although abstractions began as reviews of paper and electronic data, they have largely become reviews of the “paperless” EMR.

Although portions of the VA medical record have been electronic since 1985, the Veterans Health Information Systems and Technology Architecture (VistA) now provides a single graphic user interface for providers to review and update patients' medical records and to place orders, including pharmacy, procedures, and laboratory tests. This improved data entry efficiency and helps clinicians comply with legislative mandates and clinical practice guidelines. VistA now comprises over 100 applications that clinicians and researchers can access.<sup>10,11</sup>

Given the transition to a “paperless” EMR, the only substantive difference between the EPRP and the EMR is information abstracted from text fields (eg, progress notes). In contrast, laboratory results and diagnostic codes data can be directly accessed and examined using directly analyzable fields. Although EPRP full medical record reviews provide detailed and comprehensive information, they are time-consuming and expensive. Therefore, they are conducted on a sample of patients sufficient only to estimate performance for each Veterans Integrated Service Network (VISN; eg, VISN1 comprises all New England states). As a result, the EPRP sample is insufficient to determine performance for a facility, clinic, or provider.<sup>12</sup> Finally, reviewers do not differentiate among data that vary in accuracy: patient self-reported cholesterol is counted equally toward the achievement of the cholesterol control performance measure as a test result recorded directly by the laboratory.

An EMR offers significant advantages over chart review.<sup>13</sup> All patient data can be analyzed because data from directly analyzable fields (eg, laboratory, and pharmacy) is standardized and more efficient. The main limitation of EMR-based measures is the difficulty in analyzing data contained in text fields, which are not directly analyzable using electronic means but are accessible to medical record reviewers.

We compared EMR-based performance measures (using directly analyzable data fields) with EPRP performance measures (based on all fields including text fields) using EPRP measures as the gold standard. We chose VA performance measures amenable to an EMR approach because they are based on distinct data elements such as vital signs (blood pressure) or a particular laboratory value (eg, cholesterol). Our focus on outcome measures reflects the fact that these established measures are widely used and amenable to this approach. Issues regarding the most appropriate measures of performance, although important, are beyond the scope of this current report. We use this set of measures to first determine the degree to which “EMR” data are missing when “EPRP” data are present (overall and by site of care) and, among those with both EMR and EPRP data, the degree to which these sources agree.

## METHODS

EPRP selects a random sample of patient records from VA facilities to monitor the quality and appropriateness of medical care. The EPRP sample includes patients who used VA health care at least once in the 2 years before the assessment period and who had at least 1 primary care or specialty medical visit in the month being sampled for the year being evaluated. Among eligible patients, a random sample of patients is drawn along with additional random samples of selected high-volume medical conditions based on International Classification of Diseases, 9th Revision codes (eg, diabetes). If a veteran is not randomly selected for review in a given month, that veteran is eligible for sampling in the next monthly sampling.<sup>14</sup>

Subjects in this study are those assessed by EPRP chart review during VA fiscal years 2001–2004 from VISN1 comprised of 8 medical centers and 37 community-based outpatient clinics (CBOCs) throughout the 6 New England states. In this analysis, patients in CBOCs were subsumed into the parent facility because it is the VA facility administratively responsible for a given CBOC. Only data on those patients eligible for assessment on performance guidelines for hypertension, ischemic heart disease, or diabetes mellitus were included. The measures included the most recent low-density lipoprotein (LDL) collected within that past 5 years, most recent diastolic and systolic blood pressures in the past year, and among diabetic patients, the most recent HgbA1c recorded in the last year.

After receiving a waiver of consent from the local VA Institutional Review Board and signing a data use agreement with the OQP, EPRP data were linked to EMR data elements for each EPRP subject from VistA laboratory (LDL cholesterol and glycohemoglobin) and vital signs (blood pressure). Then, patient identifying information was removed and a random study number was assigned for the analysis.

EMR- based performance measures were then constructed to mimic the EPRP algorithms (eg, blood pressure control <140/90 mm Hg, except for diabetic patients, in which case control is <130/80 mm Hg). Additionally, high blood pressure was defined as a systolic pressure  $\geq 160$  or diastolic pressure  $\geq 100$  mm Hg (stage 2 hypertension), because such patients should be appropriately evaluated by a healthcare provider, typically within 1 month or sooner if the clinical situation warrants.<sup>15</sup> Only readily analyzable EMR data elements were used: nonstandardized text or memo fields, the unstructured, narrative portion of a medical note in the EMR, were not used. For the blood pressure value, the EPRP chart reviewer uses the results from the most recent blood pressure, but may also abstract this data from a text note. For the EMR, the blood pressure measure was available as a single data element from the vital sign package and the last blood pressure measure from the EMR before the EPRP review date, but within the EPRP observation period, was used. In accordance with the EPRP protocol, if there were multiple blood pressure readings on that date, the lowest was used. If no readings were in the EMR, the EMR data were considered missing.

For the diabetes measures, we focused on: blood pressure, glycohemoglobin (HgbA1c), and LDL cholesterol. In accord with EPRP practice, the most recent blood pressure, HgbA1c, and LDL measures, relative to the EPRP review date, were used.

We performed only routine laboratory cleaning algorithms on the data we pulled from the EMR. These included detecting and deleting records with invalid values (eg, the text “unavailable” for a patients' EMR HGBA1C value or 000/00 for a blood pressure reading) and removing double slashes in blood pressure values (eg, 140//90 was revised to read 140/90). For blood pressure reading, both the systolic and diastolic pressure were required, otherwise the record was deleted. Of the 41,243 blood pressure readings on the patients in

this sample during the observation period, only 168 records (0.4%) were unusable given these conditions. The EMR record for each patient that was closest to the EPRP review date was then retained.

## Analysis

The proportion of missing values in the EMR and EPRP measures were compared using  $\chi^2$  tests of independence overall and by individual sites of care. Among cases in which data were available from both sources, agreement between the EMR and EPRP values were assessed using ordinary least squares regression and kappa. Differences in mean values were assessed using matched *t* tests. Sensitivity and specificity were estimated for the thresholds established within the performance measures. Thresholds for blood pressure are stated above, whereas for LDL, high was defined as  $\geq 120$  mg/dL, and for HgbA1c, a high reading was considered as 9% or above. All analyses were done using SAS version 9.1.

## RESULTS

### Description of Sample

EPRP selected 2095 patients with at least one of the disorders of interest (hypertension [HTN], diabetes [DM], or ischemic heart disease [IHD]) from those eligible patients in the VA New England Healthcare System. Eighty-seven percent ( $n = 1827$ ) of the sample was male, and the mean age was 69 years (median = 71 years, standard deviation = 12.1). Eighty-three percent of patients were receiving treatment of HTN, 42% for DM, and 16% for IHD. Thirty-one percent of patients had 2 of these conditions; 5% had all 3.

### Completeness of Data

EPRP data were more complete than EMR data. Averaging over all measures, EPRP contained 95% of information on eligible patients, whereas EMR contained 83%. The EPRP contained a significantly higher proportion data on LDL cholesterol values (88%), blood pressure values (>99%), and HgbA1c (94%) compared with EMR data (Table 1). The proportion of missing EMR values on the patients assessed by EPRP varied significantly by clinical site ( $P < 0.01$ , Fig. 1). Patterns of missing EMR data were distinct for HgbA1c and for LDL cholesterol. One site in particular, site "H," had the highest proportion of missing values for blood pressure and HgbA1c and the second highest for LDL cholesterol.

To assess whether patients with EMR values were systematically different from those without EMR data, we compared the mean values on EPRP measures by patients' missing status in the EMR. We found that patients missing EMR values did not differ significantly on EPRP measures compared with patients with EMR values (Table 2).

### Agreement Among Nonmissing Matches

The correlation between patients' EMR and EPRP LDL cholesterol values was nearly perfect ( $r = 0.98$ ,  $P < 0.01$ ) (Table 3). Both the systolic blood pressure and diastolic blood pressure readings ( $r = 0.89$ ,  $P < 0.01$ ) from the EMR and EPRP were highly correlated. Among diabetic patients, EMR and EPRP HgbA1c levels were also highly correlated ( $r = 0.98$ ,  $P < 0.01$ ). Although EMR and EPRP measures were highly correlated, the EMR diastolic and systolic blood pressure measures were higher than EPRP measures (1.2 and 0.3 mm Hg, respectively,  $P < 0.01$ ). In Figure 2, the line of identity reveals that the agreement between the sources is consistent over the range of values. The correlation between EMR and EPRP LDL, systolic, diastolic, or HgbA1c levels did not differ substantially across sites (eg, range 0.97–0.99 for LDL) or by patients' gender (eg, 0.94–0.95 for systolic).

Concordance on “high” LDL cholesterol, in which LDL values were dichotomized at 120 mg/dL, was in the excellent range ( $\kappa = 0.98$ ).<sup>16</sup> The EMR-based measure showed a high degree of sensitivity (99.6), specificity (98.5), positive predictive value (99.5), and negative predictive value (98.8) when compared with the EPRP-based measure as the gold standard (Table 4).

Agreement beyond chance on blood pressure control, dichotomized at pressure of 140/90 mm Hg or below for nondiabetic patients and at 130/80 mm Hg or below for diabetic patients, was also in the excellent range ( $\kappa = 0.94$ ). It too showed a high degree of sensitivity (95%) and specificity (>99%). Agreement beyond chance on high blood pressure, dichotomized at systolic pressure of 160 mm Hg or a diastolic pressure of 100 mm Hg, was also excellent ( $\kappa = 0.86$ ) with a sensitivity of 100% (indicating that all patients identified as not having high blood pressure by the EPRP were so identified in the EMR). Concordance on high HGBA1C among diabetic patients, dichotomized at >9%, was also high ( $\kappa = 0.99$ ) (Table 4). In addition, kappa scores did not differ significantly over time for any of the measures assessed (eg, agreement on high HgbA1c was 0.96, 1.00, 0.97, and 1.00 for the third quarter of 2002 and first through third quarters of 2003, respectively).

## DISCUSSION

We sought to determine the extent to which selected EPRP chart review-based measures could be accurately replicated using readily available and directly analyzable EMR data. We performed a minimum amount of data cleaning on the EMR data, and such routines could easily be automated. Even if some need to be run manually, they take much less time and resources than formal chart review. We then compared the degree of missing data, the correlation of the data elements, evidence of bias, and the degree of concordance in identifying patients for whom the guidelines were met.

We found that, although the EMR data available for each of the measures were substantial (over 80% of the data for this analysis was available), it was more limited than EPRP and the extent of missing varied by site. Among those for whom both EPRP and EMR data were available, we found a high degree of correlation between the 2 sources in the actual values of the measures assessed ( $\rho = 0.89$ – $0.98$ ) and in the concordance between the measures using the performance cut points ( $\kappa = 0.86$ – $0.99$ ). The low positive predictive value for the EMR high blood pressure measure is a reflection of the prevalence of this condition. This is a well-described issue in low prevalence conditions.<sup>17</sup> We also found the correlation between EMR and EPRP values was consistent across sites despite variation in the amount of missing data. Finally, the differences in mean values between EPRP and EMR were very small, suggesting little bias in the performance estimates derived using EMR compared with EPRP.

We believe that the variance in the proportion of missing seen for the 3 measures we used (blood pressure, hemoglobin A1C, and LDL cholesterol) has to do with differences in the clinical management related to these data. The differences in missing by site may reflect differences in clinical care and clinical documentation at these sites, differences in provider expertise and training with the EMR, as well as other competing demands that vary between sites even in an integrated system such as the VA.<sup>18</sup> The VA is currently implementing alternative methods of capturing data such as clinical reminders when patients are tested outside for LDL, blood pressure, and so on. Thus, we anticipate that in the near future, missing data will be less of an issue.

Systolic and diastolic blood pressure is routinely collected by the nurse or health technician before the physician visit and recorded in a directly analyzable field. If the reading is high,

the physician will recheck the measurement before responding to the value. Physicians may or may not enter the new reading in the vital signs package as an analyzable field. Many instead record the new reading in the progress note (an open text field). This would suggest that the EMR value for blood pressure should be higher on average than the EPRP value. We observed a statistically significant higher value in the EMR data, but the difference was small (1.2 mm Hg for systolic, 0.34 mm Hg diastolic). Clinical sites differ in the degree to which they are actively working with physicians to encourage them to record the readings in the vital signs package (a directly analyzable field). This may explain why the degree of missing varied by site.

We suspect that the reason so many LDL cholesterol tests were missing from the EMR data is because they are associated with veterans who receive care both inside and outside the VA.<sup>19</sup> Because the VA offers a pharmacy benefit, some veterans getting health care outside the VA seek to fill prescriptions within the VA system. This practice is allowed so long as the veteran is enrolled in and has at least one annual visit with a VA primary care provider, the veteran provides written proof of test results relevant to the medications requested, and the provider has the expertise to manage the medication and feels that it is indicated. When this occurs, VA providers are required to document these indications, including relevant outside laboratory test results provided by the patient in the VA progress note. Cholesterol-lowering medications are often provided under this arrangement. As a result, LDL cholesterol test at outside laboratories may be recorded in the progress note (and available for EPRP analyses) but not available for EMR data analyses.

Hemoglobin A1c was much less likely to be missing. Because diabetes management requires frequent adjustment of insulin and other medications, those with diabetes and other more serious chronic conditions are more likely to receive more of their care within the VA system. Similarly, there was no significant difference in the values of the hemoglobin A1C gathered by EPRP and those directly collected in the EMR.

Missing data is a concern with either approach to performance measurement. In the case of chart review, the costs of review mean that the measures can only be assessed on a small, but representative, subset. This constraint makes it difficult for administrators to use these data to determine whether individual clinics or providers may be in need of targeted intervention. In the case of directly analyzable text fields, data may be missing in a nonrandom manner making conclusions biased with respect to the target sample, but if this bias is small or of a quantifiable size and direction, it may still provide useful information for quality monitoring and improvement.

This study had several limitations. First, it was limited to measures deemed amenable to direct EMR evaluation and must be considered an initial step toward the use of EMR data for direct performance measurement. There are many approved VA performance measures that would likely be inaccurate using only analyzable fields from the EMR. This is largely because the only consistent record of their application is found in “open text” progress notes. These include the use of aspirin to avoid stroke (the copay is more than the cost of generic so many veterans buy this medication over the counter), counseling for smoking cessation, or screening for depression. Over time, VA facilities are requiring that these be addressed in response to physician prompts, which can create analyzable fields for these measures as well. Thus, we expect the number of measures that might be directly assessed to increase with time and with increasing incentives to facilities and providers to standardize this data. Finally, future use of the EMR text fields may also be facilitated through text-searching engines.<sup>20</sup>

A second limitation is that we evaluated only one point in time to measure agreement between the sources of information. Given the evolving nature of the VA electronic medical records system, and the increasing number of clinical reminders to clinicians to assess and document care, we suspect that more recent measures may exhibit higher agreement. Third, we assessed facilities in only one geographic area of a national system of care. As noted previously, we found substantial variation by site even within the limited geographic area we assessed; therefore, assessing a larger area may reveal even more variation. Lastly, generalizability may be somewhat limited given the long history and experience with electronic medical records in the VA compared with other institutions that may be just beginning to establish EMR systems.

In addition, VA patients may obtain care outside the VA system. To the extent that the veteran reports this care to the physician and the physician documents it in a text field, external care would be reflected as asymmetric missing data for the EMR and not for the EPRP review. However, to the extent that the patient fails to report the care at all, this limitation would apply to both methods. Of note, external care is one argument in favor of using outcome rather than process measures because outcomes are measured more reliably in the setting of outside care.

The use of directly analyzable EMR data for performance measurement has distinct advantages. Specifically, use of EMR fields would allow efficient and standardized determination of performance measures on every patient in the healthcare system rather than the limited, but in-depth, sample collected by EPRP. This dramatic increase in statistical power would allow us to consider whether quality of care varies by: (1) means of identifying the appropriate target population (diagnostic codes, laboratory abnormalities, or vital sign abnormalities); (2) clinical site of care (facility and specific clinic); or (3) important patient characteristics (race/ethnicity, gender, age, distance from facility, multimorbidity, and so on). These analyses will likely help inform more effective interventions to further improve the quality of care delivered. In addition, use of the EMR to assess quality measurement may result in significant cost reductions. An important strength of our approach is that we did not rely on derivative data sources such as “data dumps,” “administrative data,” or “data warehouses.” The data used in our analyses were captured directly from the electronic medical record (VistA) using KB SQL (Knowledge Based Systems, Inc.). This product runs as a server and allows direct ODBC connections from Access or SQL server to VistA tables mapped to a relational schema.

This study is focused on a subset of quality measures in use in a system caring for more than 5 million veterans a year (The VA Healthcare System) with a well-integrated electronic medical record system that has been functional for nearly a decade.<sup>21</sup> The VA system has proven to be an outstanding laboratory in which to study issues relevant to the use of the electronic medical record in improving quality of health care.<sup>6,22,23</sup> It has also been recognized as 1 of 4 institutions producing the highest quality research concerning electronic medical records, all 4 of which evolved their own electronic medical record system.<sup>24</sup> As such, the VA offers an invaluable glimpse into the future as other healthcare systems adopt electronic medical record systems. Furthermore, the VA system is available, free of charge, and has been adopted by a number of other institutions.<sup>25</sup> A simplified version of VistA has also been released by Centers for Medicare & Medicaid Services for private office use.<sup>26</sup> Although the particulars how personnel enter data in the EMR may be unique to the VA, the underlying issues of how providers enter information (unstandardized text fields, ie, progress notes), how laboratory data and administrative codes are entered (directly analyzable fields), and the resulting missing data that results if only directly analyzable fields are used to estimate performance are completely generalizable.

In conclusion, this study suggests that directly analyzable fields in the EMR can effectively and accurately reproduce selected EPRP measures on most patients. We found no evidence of systematic differences in performance values among these with and without directly analyzable data in the EMR. Future research should address the applicability of EMR data to other EPRP measures. Studies are needed to determine the degree of site variation in coding practices and to assess whether the EMR data fields are useful for performance measures consistent across the entire system.

## Acknowledgments

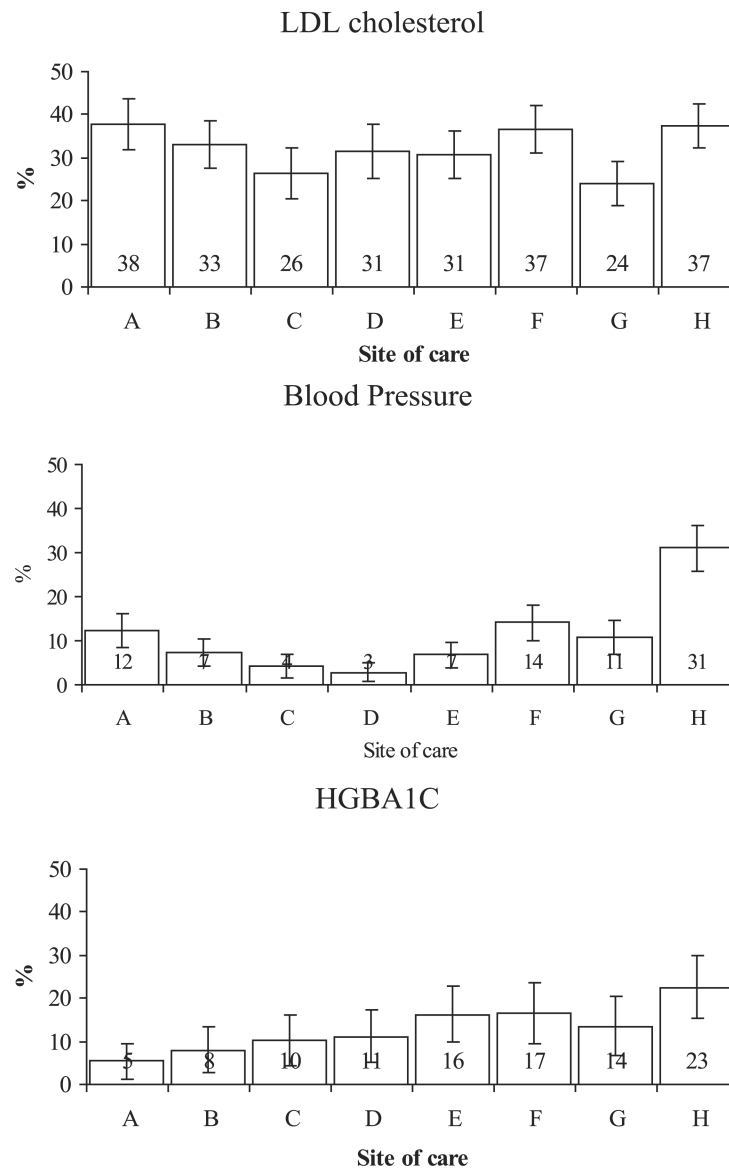
Funded by the National Institute on Alcohol Abuse and Alcoholism (UO1 AA 13566) and the Veterans Health Administration Office of Research and Development.

## REFERENCES

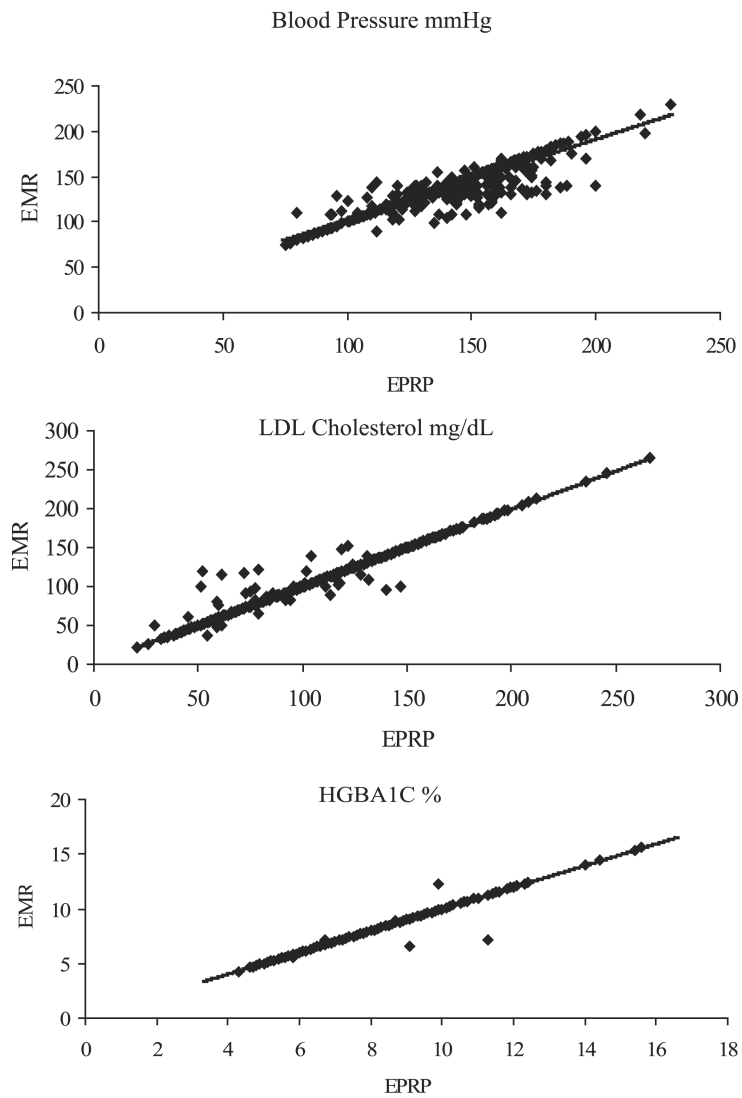
1. McQueen L, Mittman BS, Demakis JG. Overview of the Veterans Health Administration (VHA) Quality Enhancement Research Initiative (QUERI). *J Am Med Inform Assoc.* 2004; 11:339–343. [PubMed: 15187071]
2. Corrigan, JM.; Eden, J.; Smith, BM., editors. *Leadership by Example: Coordinating Government Roles in Improving Healthcare Quality* Committee on Enhancing Federal Healthcare Quality Programs. National Academy Press; Washington, DC: 2002.
3. Perlin JB, Kolodner RM, Roswell RH. The Veterans Health Administration: quality, value, accountability, and information as transforming strategies for patient-centered care. *Am J Manag Care.* 2004; 10:828–836. [PubMed: 15609736]
4. Kizer, KW. *Vision for Change: A Plan to Restructure the Veterans Health Administration.* Department of Veterans Affairs; Washington, DC: 1995.
5. Clancy C, Scully T. A call to excellence: how the federal government's health agencies are responding to the call for improved patient safety and accountability in medicine. *Health Aff (Millwood).* 2003; 22:113–115. [PubMed: 12674413]
6. Kizer KW. The `new VA': a national laboratory for health care quality management. *Am J Med Qual.* 1999; 14:3–20. [PubMed: 10446659]
7. Institute of Medicine. *The Computer-based Patient Record: An Essential Technology for Health Care.* National Academy Press; Washington, DC: 1991.
8. Sawin CT, Walder DJ, Bross DS, et al. Diabetes process and outcome measures in the Department of Veterans Affairs. *Diabetes Care.* 2004; 27(suppl 2):B90–94. [PubMed: 15113789]
9. Reed JF 3rd, Baumann M, Petzel R, et al. Data management of a pneumonia screening algorithm in Veterans Affairs Medical Centers. *J Med Syst.* 1996; 20:395–401. [PubMed: 9087884]
10. Hynes DM, Perrin RA, Rappaport S, et al. Informatics resources to support health care quality improvement in the Veterans Health Administration. *J Am Med Inform Assoc.* 2004; 11:344–350. [PubMed: 15187063]
11. [Accessed May 17, 2006] [www.va.gov/vista\\_monograph/docs/vista\\_monograph2005\\_06.pdf](http://www.va.gov/vista_monograph/docs/vista_monograph2005_06.pdf).
12. Benin AL, Vitkauskas G, Thornquist E, et al. Validity of using an electronic medical record for assessing quality of care in an outpatient setting. *Med Care.* 2005; 43:691–698. [PubMed: 15970784]
13. McDonald CJ. Quality measures and electronic medical systems. *JAMA.* 1999; 282:1181–1182. [PubMed: 10501125]
14. Doebbeling BN, Vaughn TE, Woolson RF, et al. Benchmarking Veterans Affairs Medical Centers in the delivery of preventive health services: comparison of methods. *Med Care.* 2002; 40:540–554. [PubMed: 12021680]
15. [Accessed July 6, 2006] [www.oqp.med.va.gov/cpg/HTN04/htn\\_cpg/frameset.htm](http://www.oqp.med.va.gov/cpg/HTN04/htn_cpg/frameset.htm).
16. Cichetti D. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess.* 1994; 6:284–290.
17. Grimes DA, Schulz KF. Uses and abuses of screening tests. *Lancet.* Mar 9.2002 359:881–884. [PubMed: 11897304]



18. Desai MM, Rosenheck RA, Craig TJ. Case-finding for depression among medical outpatients in the Veterans Health Administration. *Med Care*. 2006; 44:175–181. [PubMed: 16434917]
19. [Accessed May 17, 2006] [www1.va.gov/VHAPUBLICATIONS/ViewPublication.asp?pub\\_ID=220](http://www1.va.gov/VHAPUBLICATIONS/ViewPublication.asp?pub_ID=220).
20. Fisk JM, Mutalik P, Levin FW, et al. Integrating query of relational and textual data in clinical databases: a case study. *J Am Med Inform Assoc*. 2003; 10:21–38. [PubMed: 12509355]
21. Justice AC, Erdos J, Brandt C, et al. The Veterans Affairs healthcare system: a unique laboratory for observational and interventional research. *Med Care*. 2006; 44(suppl):S7–S12. [PubMed: 16849970]
22. Jha AA, Perlin JB, Kizer KW, et al. Effect of the transformation of the Veterans Affairs health care system of the quality of care. *N Engl J Med*. 2003; 348:2218–2227. [PubMed: 12773650]
23. Kizer KW, Demakis JG, Feussner JR. Reinventing VA health care: systematizing quality improvement and quality innovation. *Med Care*. 2000; 38(suppl 1):I–7.
24. Chaudhry B, Wang J, Wu S, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Ann Intern Med*. 2006; 144:742–752. [PubMed: 16702590]
25. VISTA Adopters Worldwide. Hardhats Org. [Accessed September 15, 2006] 2005. Available at: [www.hardhats.org/adopters/vista\\_adopters.html](http://www.hardhats.org/adopters/vista_adopters.html).
26. Kolata G. In unexpected medicare benefit, US will offer doctors free electronic records system. *New York Times*. Jul 21.2005



**FIGURE 1.** Proportion of patients missing electronic medical records low-density lipoprotein, blood pressure, and HgBA1c values by site of care.



**FIGURE 2.** Association of electronic medical records and External Peer Review Program measures of Veterans Integrated Service Network 1 patients' low-density lipoprotein, blood pressure, and HgbA1c values.

**TABLE 1**

The Proportion of Information Available on Each Measure by the Source of Information

| Measure                         | Total No. | EMR   | EPRP  | P     |
|---------------------------------|-----------|-------|-------|-------|
| Low-density lipoprotein (mg/dL) | 2095      | 67.4% | 88.4% | <0.01 |
| Blood pressure (mm Hg)          | 2095      | 87.8% | 99.8% | <0.01 |
| HgbA1c (%) *                    | 877       | 87.3% | 93.5% | <0.01 |

\* Among diabetic patients.

TABLE 2

Agreement Between EMR- and EPRP-Based Measures Using Clinical Cut Points

| Measure                                     | Threshold | Prevalence* | Percent Agreement | Kappa | Sensitivity | Specificity | Positive Predictive Value | Negative Predictive Value |
|---|-----------|-------------|-------------------|-------|-------------|-------------|---------------------------|---------------------------|
| Low-density lipoprotein (mg/dL)             | >120      | 24.1        | 99.4              | 0.98  | 99.6        | 98.5        | 99.5                      | 98.8                      |
| Blood pressure control (mm Hg) <sup>†</sup> | <140/90   | 56.8        | 97.1              | 0.94  | 95.4        | 99.3        | 99.4                      | 94.3                      |
| High blood pressure                         | >160/100  | 7.2         | 97.8              | 0.86  | 100.0       | 97.7        | 76.9                      | 100.0                     |
| HgbA1c (%) <sup>‡</sup>                     | >9        | 11.4        | 99.7              | 0.99  | 97.7        | 100.0       | 100.0                     | 99.7                      |

\* Percent with the condition in the EPRP.

<sup>†</sup> For diabetic patients, blood pressure control threshold was 130/80 mm Hg.

TABLE 3

Comparison of mean EPRP Values by Availability of EMR Measures

| Measure                          | No.Observations | EMR Available | EMR Missing   | Difference | P *  |
|----------------------------------|-----------------|---------------|---------------|------------|------|
| Low-density lipoprotein (mg/dL)  | 1848            | 100.5 (29.92) | 98.1 (31.94)  | 2.37       | 0.15 |
| Systolic blood pressure (mm Hg)  | 2090            | 130.9 (17.94) | 130.5 (17.74) | 0.48       | 0.69 |
| Diastolic blood pressure (mm Hg) | 2090            | 72.3 (11.58)  | 72.9 (10.66)  | -0.65      | 0.40 |
| HgbA1c (%) <sup>‡</sup>          | 820             | 7.2 (1.56)    | 7.1 (1.11)    | 0.18       | 0.36 |

Standard deviations shown in parentheses.

\* P value on difference in means from paired *t* test.<sup>‡</sup> Among diabetic patients only.

**TABLE 4**  
 Agreement Between EMR and EPRP Based Measures for VISN1 Patient Sample (n = 2095)

| Measure  | No. Pairs of Observations | r    | EMR              | EPRP             | Difference | P *   |
|--|---------------------------|------|------------------|------------------|------------|-------|
| Low-density lipoprotein (mg/dL), mean $\pm$ 1 standard deviation | 1404                      | 0.98 | 100.8 $\pm$ 29.9 | 99.9 $\pm$ 30.4  | 0.9        | 0.08  |
| Systolic blood pressure (mm Hg)                                  | 1840                      | 0.89 | 132.1 $\pm$ 19.9 | 130.9 $\pm$ 17.9 | 1.2        | <0.00 |
| Diastolic blood pressure (mm Hg)                                 | 1840                      | 0.89 | 72.7 $\pm$ 11.9  | 72.4 $\pm$ 11.5  | 0.3        | <0.01 |
| HgbA1c (%) <sup>†</sup>  | 754                       | 0.98 | 7.2 $\pm$ 1.5    | 7.2 $\pm$ 1.5    | 0.0        | 0.47  |

r = Pearson correlation coefficient.

VISN1, Veterans Integrated Service Network 1.

\* P-value on difference in means from paired t test.

<sup>†</sup> Among diabetic patients only.