Impact of Adherence to Quality Measures for Localized Prostate Cancer on Patient-reported Health-related Quality of Life Outcomes, Patient Satisfaction, and Treatment-related Complications

William Sohn, MD,* Matthew J. Resnick, MD, MPH,* Sheldon Greenfield, MD,** Sherrie H. Kaplan, PhD, MS, MPH,§ Sharon Phillips, MSPH,‖ Tatsuki Koyama, PhD,‖ Michael Goodman, MD, MPH,¶ Ann S. Hamilton, PhD, MA,§ Mia Hashibe, PhD, MPH,** Karen E. Hoffman, MD, MHS.c, MPH,†† Lisa E. Paddock, PhD, MPH,†† Antoinette M. Stroup, PhD, §§ Xiao-Cheng Wu, MD, MPH,** David F. Penson, MD, MPH,*††† and Daniel A. Barocas, MD, MPH*

Background and Objective: Quality measures used in pay for performance systems are intended to address specific quality goals, such as safety, efficiency, effectiveness, timeliness, equity, and patient centeredness. Given the small number of narrowly focused measures in prostate cancer care, we sought to determine whether adherence to any of the available payer driven quality measures influences patient centered outcomes, including health related quality of life (HRQOL), patient satisfaction, and treatment related complications.

Methods: The Comparative Effectiveness Analysis of Surgery and Radiation study is a population based, prospective cohort study that enrolled 3708 men with clinically localized prostate cancer during 2011 and 2012, of whom 2601 completed the 1 year survey and underwent complete chart abstraction. Compliance with 6 quality indicators endorsed by national consortia was assessed. Multivariable regression was used to determine the relationship between indicator compliance and Expanded Prostate Cancer Index Composite (EPIC-26) instrument summary scores, satisfaction scale scores (service satisfaction scale for cancer care), and treatment related complications.
Results: Overall rates of compliance with these quality measures ranged between 64% and 88%. Three of the 6 measures were weakly associated with 1 year sexual function and bowel function scores (β 4.6, 1.69, and 2.93, respectively; P ≤ 0.05), whereas the remaining measures had no significant relationship with patient reported HRQOL outcomes. Satisfaction scores and treatment related complications were not associated with quality measure compliance.

Conclusions: Compliance with available nationally endorsed quality indicators, which were designed to incentivize effective and efficient care, was not associated with clinically important changes in patient centered outcomes (HRQOL, satisfaction, or complications) within 1 year.

Key Words: prostate cancer, quality, health related quality of life, patient centered, PQRS

Procedure: The Comparative Effectiveness Analysis of Surgery and Radiation (CEASAR) study is a population-based, prospective cohort study that enrolled 3708 men with clinically localized prostate cancer from January 2011 to February 2012, of whom 2601 completed the 12-month survey and

TABLE 1. Compliance With Quality Measures

<table>
<thead>
<tr>
<th>Measures</th>
<th>Source</th>
<th># Compliant</th>
<th>Total</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Avoidance of overuse of bone scan in men with low-risk tumors</td>
<td>PQRS #102, PCPI #3, NQF 0389</td>
<td>881</td>
<td>1155</td>
<td>76.3</td>
</tr>
<tr>
<td>2. ADT for high-risk patients undergoing XRT</td>
<td>PQRS #104, PCPI #5, NQF 0390</td>
<td>160</td>
<td>210</td>
<td>76.2</td>
</tr>
<tr>
<td>3. Documentation cT stage, biopsy Gleason in newly diagnosed</td>
<td>PCPI #2</td>
<td>1663</td>
<td>2310</td>
<td>72.0</td>
</tr>
<tr>
<td>4. Documentation of DRE, cT stage, biopsy Gleason before primary therapy</td>
<td>PCPI #1</td>
<td>1228</td>
<td>1924</td>
<td>63.8</td>
</tr>
<tr>
<td>5. Documentation of discussion of treatment options</td>
<td>PCPI #4</td>
<td>1338</td>
<td>1897</td>
<td>70.5</td>
</tr>
<tr>
<td>6. Documentation of pathologic T and N stage, Gleason score, and margin status on pathology report in men undergoing RP</td>
<td>PQRS #250, NQF 1853</td>
<td>1096</td>
<td>1252</td>
<td>87.5</td>
</tr>
</tbody>
</table>

ADT indicates androgen-deprivation therapy; cT, clinical T stage; DRE, digital rectal examination; NQF, National Quality Forum; PCPI, Physician Consortium for Performance Improvement; PQRS, Physician Quality Reporting System; RP, radical prostatectomy.
underwent complete medical chart abstraction, and, therefore, were included in the analytic cohort (supplemental Fig., Supplemental Digital Content 1, http://links.lww.com/MLR/B193). The parent study design and patient characteristics have been described previously\(^8\) (supplemental Table 1, Supplemental Digital Content 2, http://links.lww.com/MLR/B194). Patients were accrued from 5 population-based Surveillance, Epidemiology, and End Results (SEER) registry catchment areas (Atlanta, Los Angeles, Louisiana, New Jersey, and Utah), as well as an additional prostate cancer patient registry (Cancer of the Prostate Strategic Urologic Research Endeavor; CaPSURE).\(^9\)

**Data Collection**

Data were collected through manual chart abstraction at 1-year as well as patient surveys at baseline and at 1-year. The Expanded Prostate Cancer Index Composite (EPIC-26), characterizes disease-specific function or health-related quality of life (HRQOL) domains (sexual function, urinary incontinence, urinary irritation/obstruction, and bowel function) scored from 0 to 100 with 100 being better HRQOL.\(^10\) EPIC-26 is widely used in prostate cancer research and practice. There is evidence that supports its high test-retest reliability and internal consistency reliability (each \(r \geq 0.80\) and the Cronbach \(\alpha \geq 0.82\)) for most domain-specific subscales and excellent criterion validity without excessive overlap when compared with instruments that measure related but distinct domains.\(^11\) Clinically meaningful differences in subscale scores have been quantified as 4–6 points in the bowel domain, 6–9 points in the urinary domains and 10–12 points in the sexual domain.\(^12\) We also collected a satisfaction scale based on the service satisfaction scale for cancer care (SSS-CC)\(^13\) \(^14\) at 1-year. The internal consistency for the SSS-CC has been estimated as \(>0.80\), and validation was performed comparing patient and spouse responses, treatment outcomes, and associations with other health measures, and has been used in longitudinal prostate cancer cohort studies.\(^13\) \(^14\) The abbreviated SSS-CC includes 5 questions regarding satisfaction with cancer care, relief of symptoms, and effectiveness. The SSS-CC was scored from 0 to 100, with 100 indicating higher satisfaction. A list of early (< 30 d) and delayed (30–365 d) complications (Fig. 3) was agreed upon by an expert panel and collected through extensive medical chart review. These were evaluated as secondary outcomes. In addition, sociodemographic data, disease characteristics, comorbidity, and psychometric scales measuring participatory decision making\(^16\) (PDM-7), social support\(^17\) (MOS Social Support Survey), and depression\(^18\) (CES-D) were collected and included as covariates.

**Quality Measures**

Six quality measures were chosen based on their endorsement by the National Quality Forum, the Physician Consortium for Performance Improvement, and PQRS. Of these measures, 3 have been adopted for use in PQRS.\(^19\)

For the measures that describe documentation at new diagnosis and before treatment, adherence to PSA documentation was not available, and was either omitted from the measure (documentation in newly diagnosed) or exchanged for digital rectal examination (documentation before primary therapy). Measures containing multiple elements (eg, complete pathology documentation) required compliance with every element to be compliant with the quality measure.

**Statistical Analysis**

We calculated physician compliance with each measure in the relevant patient population. D’Amico risk stratification\(^20\) was used to determine the proper patient groups for guideline concordant imaging use. Pretreatment EPIC scores were calculated for both measure compliant and noncompliant groups, and compared using appropriate parametric statistical tests. Multivariable linear regression analysis was performed to determine the effect of measure compliance on EPIC domain scores, adjusting for baseline EPIC score, treatment type (surgery vs. radiation), age (below 65 vs. 65 y and above), race (white vs. other), household income level (≤ $50 K vs. > $50 K), insurance status (Medicare, Private, or other), comorbidity [total illness burden index for prostate cancer\(^21\) (TIBI-CaP)], D’Amico risk classification, and SEER site. We performed planned subgroup analyses using different age, race, and socioeconomic groups to test the hypothesis that compliance with these measures may be of importance in certain vulnerable subgroups. Similar models were constructed to determine the effect of measure compliance on composite satisfaction scale score, with additional adjustment for social support, CES-D, and PDM-7. Multivariable logistic regression was performed with adjustment for patient and disease characteristics to determine the effect of measure compliance on complication rates.

A significance level of 0.05 was used for all statistical inference. Stata version 11.2 (StataCorp, College Station, TX) and R3.0 (R Core Team, Vienna, Austria) were used for all statistical analyses.

**RESULTS**

The study cohort included 2601 men, with a mean age of 64.4 (median 65). Compliance with the 6 quality measures studied ranged from 63.8% to 87.5% (Table 1).

Baseline EPIC domain summary scores were calculated for measure compliant and noncompliant groups (supplemental Table 2, Supplemental Digital Content 3, http://links.lww.com/MLR/B195). For most measures, baseline scores were similar between groups. However, mean baseline scores in the EPIC sexual domain for measure 2 (ADT for high-risk patients) were markedly lower in the compliant group (mean = 47 vs. 74, \(P = 0.001\)). This may be expected, as patients with erectile dysfunction at baseline may be more likely to be offered or to accept ADT, without concerns about its adverse effects on libido and sexual performance. There were small but statistically significant differences in EPIC bowel for measure 1 (bone scan avoidance in low-risk patients) and in EPIC sexual for measure 5 (discussion of treatment options).

In 21 of the 24 multivariable models predicting 1-year EPIC domain scores, compliance with the quality measure was not associated with any difference in functional outcome.
The remaining 3 models showed small magnitude, but statistically significant effects. EPIC bowel scores were slightly better among patients whose physician complied with measure 5 (documentation of the discussion of treatment options; $\beta = 1.69$, $P = 0.01$). EPIC sexual scores were slightly worse among patients whose physician was compliant with measure 1 (avoidance of bone scan in low risk; $\beta = -4.6$, $P = 0.04$) and slightly better for patients whose physician was compliant with measure 4 (documentation of disease characteristics before treatment; $\beta = 2.93$, $P = 0.04$). The forest plot (Fig. 1) displays the difference in EPIC domain score by quality measure compliance with 95% confidence intervals.

Subgroup analysis was performed using age, race, and income level, with interaction terms between these variables and measure compliance (data not shown). Of these subgroups, patients with income <$50,000 were found to have a significant reduction in EPIC sexual domain score with compliance with adjuvant ADT with XRT in high-risk patients ($\beta = -23.85$, $P = 0.001$). This may be related to the known impact of ADT usage on sexual function, although it is not clear why the effect is less significant for men with higher income.

Multivariable linear regression analyses were repeated with satisfaction composite scores as the outcome measure (Fig. 2). Compliance with quality measures did not have a significant effect on composite satisfaction scores at 1-year.

Treatment-related complication rates were calculated in the early and delayed periods. Logistic regression was performed, with odds ratios calculated to summarize the likelihood of complication with measure compliance. Documented complication rates were low overall, with 19 occurring within 30 days and 120 occurring within 1-year. There was no significant association between measure compliance and complication rates (Fig. 3). This finding must be tempered due to small-sample bias affecting the maximum likelihood estimation in the model.

**DISCUSSION**

Assessment of patient-centered outcomes is essential in the evaluation of the patient experience after cancer treatment. Indeed, HRQOL and satisfaction represent 2 critically important outcomes that reflect health care quality. However, outcome measures are associated with difficulties in risk-adjustment and large-scale data collection. To this
end, the majority of quality measures in pay-for-performance programs have been structure or process measures, most of which are intended to assess efficiency and/or effective care. In the case of prostate cancer, we found few weak associations between compliance with nationally endorsed quality measures and patient-centered outcomes. Although there were associations that met statistical significance, there was no discernable pattern of association between compliance and improved outcomes, and the magnitude of these differences is not clinically important, according to published thresholds for clinically detectable change.12,23

There are several potential reasons for failing to identify an association between these measures and patient-centered outcomes, and process-outcome links have generally been elusive.24,25 We may have lacked statistical power to identify a true association (a type II error). This is unlikely, at least in the case of the patient-reported outcomes, given our sample size.

It is possible that these measures are related to patient-centered outcomes, but only in specific subgroups. We did not identify any such effects based on age, race, or income level, perhaps because of limited statistical power in these smaller subgroups. However, additional analyses could reveal certain high-risk subpopulations that may have improvements in patient-centered outcomes with measure compliance. Recognizing that many patients achieve favorable outcomes regardless of the measured quality-of-care, identification of vulnerable subgroups at high risk for poor outcomes may facilitate targeted application of quality measures.26

It is also conceivable that a composite quality measure comprising multiple measures would be more predictive of patient-centered outcomes. For example, it has been shown that composite measures, including process and intermediate outcome measures, account for a significant proportion of physician-level variation in diabetes outcomes.27

The most likely explanation for our null findings is that these nationally endorsed process measures were developed to address other quality goals, namely effective clinical care, cost-effectiveness, and efficiency, rather than improved patient-reported outcomes. Conversely, there may be other processes of care that more directly influence patient-reported outcomes, such as the use of nerve sparing in low-risk surgical patients. Yet none of the nationally endorsed measures, upon which the value of prostate cancer care may be judged and upon which reimbursement may be based, include patient-reported outcomes as their aim, despite the fact that HRQOL and satisfaction are recognized as the most salient treatment outcomes in this disease with a nearly 100% 5-year survival. Among the measures endorsed by CMS, 2 target effective clinical care and 1 targets efficiency and cost-reduction. This mirrors the overall trend in quality measurement by CMS; of the 175 measures CMS assessed

FIGURE 2. Relationship between quality measure compliance and service satisfaction scale for cancer care satisfaction scale scores. CI indicates confidence interval.
for PQRS in 2015, 103 (59%) are directed at effective clinical care or cost-reduction/cost-efficiency and 11 (6%) specifically target “Patient-Centered Experience and Outcomes.” Yet PQRS is not intended to merely enhance adherence to selected measures. Rather, its intention is to “obtain meaningful data to improve care” presuming that adherence to specific processes will serve as a proxy for more global quality-of-care. Thus, it is expected that adherence to available measures would result in better patient-centered outcomes. The remaining measures for prostate cancer do not have a clear quality-improvement domain specified. This represents an opportunity to develop measures that influence patient-reported outcomes, so that patient-centered care may be recognized as an important quality aim.

Each of the 6 quality measures analyzed in this study is a process measure. Process measures are attractive as they do not require extensive risk-adjustment, are easy to benchmark, and can be collected during the clinical process. However, they may not be associated with important outcomes. In contrast, outcome measures have face-validity and may reflect the impact of multiple processes of care, but require careful risk-adjustment, which may result in lower standards of care for disadvantaged populations. Moreover, the collection of some outcome measures may be onerous (eg, HRQOL) or take place long after the intervention (eg, mortality), and the opportunity for quality improvement may be lost. Nonetheless, direct measurement of patient-reported outcomes with appropriate risk-adjustment is 1 avenue to explore for improving the assessment of quality-of-care for prostate cancer.

The findings in the study must be interpreted in light of the study design and dataset. One significant limitation of our study is the lack of adjustment for structural measures related to the resources and qualifications of hospitals and providers. The main independent variables were process measures because these are the nationally endorsed measures. We were also limited to specific outcome measures. The CEASAR study was designed and powered to measure differences in EPIC-26 1-year after treatment. It was felt that the short-term (and maybe long-term) oncologic control are similar between groups (referring both to treatment and guideline compliance). Therefore, by design, we set out to determine whether process measures influenced patient-reported outcomes, but we recognize that there may be alternative measures that could demonstrate a process-outcome link. Finally, we did not make adjustments for multiple comparisons, for 2 reasons. First, these were a priori analyses. Second, we were not concerned.

**FIGURE 3.** Relationship between quality measure compliance and treatment-related complications. CI indicates confidence interval; OR, odds ratio.
about type I error, as we were interpreting the clinical importance of the magnitude of difference rather than P-values, and no clinically meaningful associations between guideline compliance and patient-centered outcomes emerged.

CONCLUSIONS

Quality assessment is critical to provide the best care possible to patients, and to inform comparative-effectiveness research. Furthermore, it is the backbone of value-based reimbursement initiatives that comprise the payment structure of our health care system. For these reasons, it is important to establish clear expectations for the intended outcomes and impact of adherence to these quality measures. This study did not find a clinically meaningful improvement in functional outcomes, satisfaction scores, or treatment-related complication rates associated with adherence to available quality measures. This represents an opportunity to identify alternative measures that may influence patient-centered outcomes.

REFERENCES

30. Fiscella K, Burnt HR, Nerenz DR. Quality measures and sociodemographic risk factors: to adjust or not to adjust. JAMA. 2014;312: 2615 2616.