Assessing the Quality of Surgical Care for Clinically Localized Prostate Cancer: Results from the CEASAR Study


From the Department of Urology (PAR, AAL, DFP, DAB), Vanderbilt University Medical Center, Nashville, Tennessee, Department of Biostatistics (ZZ, LCH, TK), Vanderbilt University Medical Center, Nashville, Tennessee, Office of Patient and Community Education (RC), Patient Advocacy Program, Vanderbilt Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, Tennessee, Department of Radiation Oncology (KH), University of Texas M. D. Anderson Center, Houston, Texas, Department of Epidemiology (MG), Emory University Rollins School of Public Health, Atlanta, Georgia, Department of Preventative Medicine (ASH), Keck School of Medicine at the University of Southern California, Los Angeles, California, Department of Epidemiology (KCA), Louisiana State University New Orleans School of Public Health, New Orleans, Louisiana, Department of Epidemiology (LEP, AS), Cancer Institute of New Jersey, Rutgers Health, New Brunswick, New Jersey, Department of Urology (MRC), University of California, San Francisco, San Francisco, California, Department of Family and Preventative Medicine (MH), University of Utah School of Medicine, Salt Lake City, Utah, Department of Urology (BBO), University of Utah Health, Salt Lake City, Utah, and Department of Medicine (SHK, SG), University of California Irvine, Irvine, California

Purpose:
Prior studies suggest that nationally endorsed quality measures for prostate cancer care are not linked closely with outcomes. Using a prospective, population based cohort we measured clinically relevant variation in structure, process and outcome measures in men undergoing radical prostatectomy.

Materials and Methods: The Comparative Effectiveness Analysis of Surgery and Radiation (CEASAR) Study enrolled men with clinically localized prostate cancer diagnosed from 2011 to 2012 with 1,069 meeting the final inclusion criteria. Quality of life was assessed using the Expanded Prostate Index Composite (EPIC-26) and clinical data by chart review. Six quality measures were assessed, including pelvic lymphadenectomy with risk of lymph node involvement 2% or greater, appropriate nerve sparing, negative surgical margins, urinary and sexual function, treatment by high volume surgeon, and 30-day and 1-year complications. Receipt of high quality care was compared across categories of race, age, surgeon volume and surgical approach via multivariable analysis.

Results:
There were no significant differences in quality across race, age or surgeon volume strata, except for worse urinary incontinence in Black men. However, robotic surgery patients experienced fewer complications (3% vs 9.3% short-term and 11% vs 16% long-term), were more likely to be treated by a high volume surgeon (47% vs 25%) and demonstrated better sexual function.

Conclusions:
In this cohort we did not identify meaningful variation in quality of care across racial groups, age groups and surgeon volume strata, suggesting that men are receiving comparable quality of care across these strata. However, we

Abbreviations and Acronyms
CEASAR = Comparative Effectiveness Analysis of Surgery and Radiation
EPIC-26 = Expanded Prostate Index Composite
HRQoL = health related quality of life
MCID = minimal clinically important difference
RALP = robotic assisted laparoscopic prostatectomy
RP = radical prostatectomy

Editor’s Note: This article is the second of 5 published in this issue for which category 1 CME credits can be earned. Instructions for obtaining credits are given with the questions on pages 1391 and 1392
disease, a prostate specific antigen less than 50 ng/ml, and between February 2011 and March 2012 with cT1 or cT2 within 6 months of a diagnosis of localized prostate cancer accrued 3,269 men younger than 80 years who were

Prostate cancer is the most commonly diagnosed noncutaneous malignancy in men in the United States, and accurately measuring and optimizing the quality of care has significant implications for longevity, patient health related quality of life and cost to the health care system.1 Identifying clinically relevant, consensus based quality indicators enables comparison, drives quality improvement and enables value based reimbursement.

Although quality measurement reporting in prostate cancer has existed since the 1980s, it is not well understood how closely quality measures correlate with actual outcomes.2–4 Some studies have identified variation in adherence to previously published prostate cancer quality measures, describing disparities in treatment modality, time to treatment and all cause mortality.5,6 Racial, ethnic, geographical and socioeconomic disparities in the delivery of prostate cancer care are well documented and recognized as significant quality problems.7–9 Equally concerning is that prior work by our group found a weak and likely clinically insignificant correlation between compliance with the previously mentioned nationally endorsed quality measures and patient reported HRQoL outcomes.10

Taken together, these data would suggest that the current quality measures for prostate cancer care are not linked closely with outcomes, such as cancer control and treatment related morbidity. Using a prospective, population based cohort, we sought to measure clinically relevant variation and compliance in 6 quality measures that are evidence-based, correlate with outcomes and have obvious clinical relevance. We assessed these measures across racial groups, age groups, surgical approach and surgeon volume to quantify variation in the quality of care for patients who underwent RP for clinically localized prostate cancer.

METHODS

Study Cohort

Included in this study were men enrolled in the Comparative Effectiveness Analysis of Surgery and Radiation study who underwent radical prostatectomy, which has been described previously.11 Briefly, the study accrued 3,269 men younger than 80 years who were within 6 months of a diagnosis of localized prostate cancer between February 2011 and March 2012 with cT1 or cT2 disease, a prostate specific antigen less than 50 ng/ml, and who were from 5 SEER (Surveillance, Epidemiology, and End Results) registry catchment areas.12 We included 1,069 men who underwent RP as their primary treatment and had the surgical approach (open vs robotic) documented (see figure). Patients completed mailed baseline surveys that captured socioeconomic data and comorbidity. Patient reported functional baseline and outcomes data were collected at enrollment, and at 6, 12 and 36 months after enrollment. Clinical data were obtained from the medical record by chart review at 12 months. Institutional review board approval was obtained from the coordinating center and each study site, and informed consent was obtained from all participating patients.

Outcomes/Exposures

The primary outcomes were compliance with 6 quality measures across 3 categories (structure, process and outcomes measures) while the primary exposures were race (Black vs nonBlack), surgical approach (open vs robotic), age at surgery (40 to 59 years vs 60 to 69 vs 70 to 79) and surgeon volume (stratified into 3 groups based on number of RP cases included in the CEASAR cohort attributed to individual surgeons as low [1 to 3 cases], medium [4 to 10 cases] and high [11 cases or greater]). The cases captured in the CEASAR cohort serve as a proxy for individual surgeon volume. For each quality measure the denominator for computing adherence was the number of patients eligible for that specific measure. We assessed 6 quality measures.

1. Underwent indicated pelvic lymphadenectomy (Process measure)

There are data supporting a prognostic and therapeutic benefit from pelvic lymphadenectomy at the time of radical prostatectomy as up to 15% of patients demonstrate lymph node involvement on pathological evaluation.13 Appropriate lymphadenectomy represents the gold standard to detect lymph node invasion and is critical for cancer staging. The National Comprehensive Cancer Network® (NCCN®) and American Urological Association guidelines recommend the use of nomograms based on preoperative characteristics to calculate the likelihood of lymph node involvement and determine need for lymphadenectomy.14,15 We used preprostatectomy calculated risk of lymph node involvement of 2% or more as our cutoff, calculated using the Memorial Sloan Kettering Cancer Center prediction tool.14

2. Underwent indicated nerve sparing (Process measure)

Preservation of the neurovascular bundles at the time of RP influences the recovery of sexual function after RP for localized prostate cancer.16,17 Recovery of sexual function after RP is an important component of HRQoL
after surgery. In men with poor preoperative erectile function, recovery of function is unlikely even with nerve sparing, and efforts to spare the nerves may increase the risk of positive margins, particularly for men with higher risk cancers. Therefore, we restricted the cohort for this measure to the subcohort of men with low risk disease and good baseline sexual function (EPIC domain score 80 or greater).

3. Positive surgical margins on final pathology (Outcome measure)

Positive surgical margins at the time of RP are reported in 11% to 48% of men and are associated with an increased risk of biochemical recurrence. High volume surgeons demonstrated a decrease in positive margins over time. We chose rate of positive surgical margin as an objective quality measure with direct implications for oncologic control. We investigated this measure separately for pT2 and pT3 disease as the likelihood of positive margins varies by pathological stage.

4. Treatment by a high volume surgeon (Structure measure)

Multiple large meta-analyses have demonstrated that increasing surgeon volume is associated with decreased complications, mortality, length of stay and positive margin rate at the time of RP. Of note, treatment by a high volume surgeon is included as a quality measure outcome. Increasing surgeon volume is also included as an exposure variable.

5. Short and long-term complications (Outcome measure)

Short and long-term complications after RP have a significant impact on the patient and the health care system. We suggest that lower complication rates are associated with higher quality care. For short-term (less than 30-day) complications we measured the rate of perioperative blood transfusion, bowel/rectal injury, myocardial infarction, deep vein thrombosis, pulmonary embolism, obturator nerve injury, pneumonia, anastomotic leak and wound infection. For long-term complications within 1 year we measured the rate of bladder neck contracture, bowel obstruction, deep vein thrombosis, pulmonary embolism, hernia, subsequent impotence procedure, subsequent incontinence procedure, myocardial infarction, urethral stricture and urinary retention.

6. Sexual and urinary function at 6 months, 1 year and 3 years after surgery (Outcome measure)

Functional outcomes including urinary continence and sexual function after RP have been positively correlated with patient reported HRQoL. We compared postoperative function to patient baseline function as assessed
by EPIC-26, which is a 26-item validated questionnaire that assesses urinary, sexual and bowel function after treatment for prostate cancer. Skolarus et al describe the concept of MCIDs as the change in EPIC-26 domain score thresholds corresponding to clinically relevant change.

**Statistical Analyses**

Patient demographics, baseline characteristics and outcome measures were summarized with median and quartile statistics for continuous variables or frequency and percentage for categorical variables. To evaluate the associations of the 4 main exposures of race, surgical approach, surgeon volume and age at surgery with the 5 categorical quality measures, adherence to these quality measures was summarized by exposure variable groups and compared using the Pearson chi-square test. Multivariable logistic regression models were fit to investigate the association between the main exposures and outcomes adjusting for other covariates (supplementary Appendix, www.jurology.com). Odds ratios and associated 95% CIs were reported.

We fit multivariable linear models to evaluate the associations between the main exposures with the EPIC sexual and urinary function domain scores. Robust variance-covariance matrices were estimated using the Huber-White method to account for the correlation due to repeated measurement. In order to allow exposures to differently associate with the outcomes at different times since surgery, we fit a model for each exposure that included the interaction term between the exposures and time since surgery, while adjusting for the other 3 main exposures and potential confounders (supplementary Appendix, www.jurology.com). Mean differences between exposure groups and the associated 95% CIs were reported. Statistical significance was considered for all 2-sided p values 5% or less. All analyses were conducted using R version 3.4.

**RESULTS**

**Clinical and Patient Characteristics**

Clinical and patient characteristics at the time of diagnosis are shown in supplementary table 1 (https://www.jurology.com). Notably, the median age of the cohort was 62 years, 11% of the cohort identified as Black and 80% of the cohort underwent robotic assisted laparoscopic prostatectomy. Supplementary table 2 (https://www.jurology.com) shows the proportion of patients meeting each dichotomous quality measure across exposure groups. Supplementary table 3 (https://www.jurology.com) summarizes the number of men eligible and compliant with each individual quality measure. The supplementary figure (https://www.jurology.com) displays the results of multivariable modeling of functional outcomes (our sixth quality metric) by exposure group. Supplementary table 4 (https://www.jurology.com) provides the absolute values used in the analysis for the sexual and urinary functional outcomes.

**Race**

Black and nonBlack men had similar rates of indicated pelvic lymphadenectomy, indicated nerve sparing, positive surgical margins, treatment by high volume surgeons, and rates of short and long-term complications (supplementary table 2, https://www.jurology.com). Sexual function domain scores were similar at 6 months, 1 year and 3 years after surgery. However, Black men reported significantly worse urinary incontinence at all 3 time points (supplementary figure part A, https://www.jurology.com).

**Age at Surgery**

There were no significant differences in rates of indicated lymphadenectomy, indicated nerve sparing, treatment by high volume surgeons or complications between age groups. However, men 60 to 69 years old with pT3 final pathology were more likely to have positive margins compared to men 70 to 79 years old.

**Surgeon Volume**

There were no significant differences in rates of indicated pelvic lymphadenectomy, indicated nerve sparing, positive surgical margins, and short-term and long-term complications between low, medium and high volume surgeons on multivariable analysis. Sexual and urinary function scores were similar at 6 months, 1 year and 3 years after surgery. However, patients who underwent prostatectomy by low volume surgeons actually demonstrated significantly better urinary continence at 6 months (p=0.01) when compared to those treated by high volume surgeons. This difference was not significant at later time points. Conversely, high volume surgeons demonstrated better sexual functional outcomes at 1 year (p=0.03) when compared to medium volume surgeons, but did not meet the threshold for MCID (supplementary figure part B, https://www.jurology.com).

**Surgical Approach**

There was no significant difference between open and robotic surgery in rates of indicated lymphadenectomy, nerve sparing or rate of positive surgical margins stratified by pT2 and pT3 disease. However, on univariable and multivariable analyses significantly more patients who underwent RALP were treated by high volume surgeons (multivariable p <0.001, OR 0.38, 95% CI 0.25–0.55). Also, on univariable and multivariable analyses patients who underwent open prostatectomy experienced a higher rate of short-term (multivariable p <0.001, OR 2.92, 95% CI 1.51–5.62) and long-term complications (multivariable p=0.02, OR 1.75, 95% CI 1.10–2.79). Patients who underwent open
prostatectomy demonstrated significantly worse sexual function at 1 year and 3 years (p <0.001 and p=0.034, respectively) compared with those who underwent RALP (supplementary figure part D, https://www.jurology.com).

DISCUSSION

In this contemporary, multicenter, prospective, longitudinal, population based study assessing compliance with quality measurements, we did not find evidence of clinically meaningful variation in care across racial groups, age groups or surgeon volume strata. We did identify variation in use of high volume surgeons, complications and some functional outcomes across surgical approach (robotic vs open), although the extent to which this represents a difference in quality versus differences in practice patterns or bias in selection of patients is unclear.

While there are significant differences in incidence and mortality of prostate cancer across racial groups, racial variation in our cohort was limited. Black men did experience significantly worse urinary incontinence than nonBlack men at all 3 postoperative time points. This has been demonstrated previously, but the reasons for this difference are not clear. It could represent variation in the quality of the surgery, but could also reflect actual differences in pelvic anatomy, or cultural differences in perception of incontinence. Overall our results may suggest that there has been improvement in mitigating racial variation in prostate cancer care. Conversely, this finding may be the result of either limited statistical power to identify differences because of the relatively low proportion of Black men in the analytic cohort (11%) or self-selection (responding to surveys).

There are several studies documenting variation in quality of care by surgeon experience. Nonetheless, our study did not find significant differences among surgeon volume strata in rates of indicated lymphadenectomy, nerve sparing, positive margins or complications. However, there were several differences noted in functional outcomes. These differences, while statistically significant, are small and likely not clinically significant differences when taken in the context of the MCID concept mentioned previously.

When comparing open versus robotic surgery, those undergoing open RP experienced more short-term and long-term complications, lower use of high volume surgeons, and worse sexual function outcomes compared to the robotic group. We interpret these findings cautiously, since our study accrued at a point in time when most prostatectomies in the U.S. were being performed robotically with selection bias inherent in these analyses. Previous studies comparing open and robotic prostatectomy have demonstrated generally equivalent oncologic and functional outcomes, but with lower rates of blood transfusion and length of stay in the robotic groups. There are ongoing debates regarding open versus robotic prostatectomy, but RALP has become the overwhelmingly dominant surgical approach. While our findings could represent a quality deficit among those undergoing open surgery, they may simply reflect changing practice patterns such as adoption of robotics in high volume centers, the centralization of surgical care, and the diminishing presence of open prostatectomy in practice and training.

When interpreting our findings it is important to consider several limitations in addition to those previously mentioned. The CEASAR cohort itself was initially designed and powered to measure differences in functional outcomes across treatments, although comparing outcomes by quality of care received was one of the primary aims. Selection bias can be present among survey respondents and may affect our results. As with most prostate cancer cohorts, minority enrollment was relatively low. Nerve sparing status was not documented in 28% of the cases and they were omitted from the analysis for that measure. Also, nerve sparing status may not have been documented accurately in the operative report. In reporting complications we acknowledge that complications were not categorized with a classification system and some important complications after prostatectomy are missing. We also acknowledge that long-term complications may not be directly attributable to the surgical procedure. There have been emerging data in recent years to suggest that we should use a higher calculated risk threshold for pelvic lymph node dissection, and the practice patterns of pelvic lymph node dissection may have been in flux when the CEASAR cohort was assembled in 2011 to 2012. Very small numbers of complications were reported overall, which may not reflect the true complication burden. Also, presumably some of the patients who underwent surgery for low risk disease would now be enrolled in active surveillance due to changing national practice patterns. Finally, the cases captured in the CEASAR cohort serve as a proxy for individual surgeon volume and may not accurately capture individual surgeons’ overall caseload. Thus, findings related to surgeon volume should be interpreted with caution.

Despite these limitations, our study used a rich contemporary data set. We chose patient centered quality measures that reflect the structure, process and outcome domains of quality measurement. These measures are evidence-based, correlate with
outcomes and have obvious clinical relevance. Our findings provide tangible areas around which to build a framework for meaningful quality measurement, comparison and improvement as we seek to improve the quality of care delivered to patients with prostate cancer.

CONCLUSION
In this cohort we found no evidence of clinically meaningful variation in surgical quality of care across racial groups, age groups or surgeon volume strata. We did find variation between open and robotic surgery, with fewer complications, improved sexual function and increased use of high volume surgeons in the robotic group, possibly reflecting differences in quality between approaches, differences in practice patterns and/or biases in patient selection. Our study suggests that variation in quality across these subgroups was low, but further studies are needed to confirm whether this represents a trend toward more equitable care. Ideally, increased compliance with relevant quality measures will lead to improved cancer specific survival and HRQoL.

REFERENCES