Published in final edited form as:

Med Care. 2012 May ; 50(5): 374–380. doi:10.1097/MLR.0b013e318245a086.

Reducing bias in the assessment of treatment effectiveness: androgen deprivation therapy for prostate cancer

Yong-Fang Kuo, PhD¹, James E. Montie, MD², and Vahakn B. Shahinian, MD, MS³

¹Department of Internal Medicine and Sealy Center on Aging, University of Texas Medical Branch, Galveston, TX

²Department of Urology, University of Michigan, Ann Arbor, MI

³Department of Internal Medicine, University of Michigan, Ann Arbor, MI

Abstract

BACKGROUND—Indication bias is the major challenge in assessing treatment effectiveness in observational studies. We explored the potential advantages of using an instrumental variable approach in the context of primary androgen deprivation therapy (ADT) for prostate cancer.

METHODS—We identified 31,930 men in the linked Surveillance, Epidemiology, and End Results (SEER)-Medicare database with a diagnosis of prostate cancer who were not treated definitively with radical prostatectomy or radiation in the years 1992 through 2002, with follow-up through 2005. The association between use of primary ADT and overall, prostate cancer-specific, and non-prostate cancer survival was assessed using multivariable regression and instrumental variable methods. Two instrumental variables, based on region and urologist prescribing preference, were constructed and analyzed by exogenous probit models. Pre-specified subgroup analyses in patients with lower risk and higher risk prostate tumors were also performed.

RESULTS—In the overall cohort, standard multivariable regression analyses showed a significantly increased risk of prostate cancer death whereas the instrumental variable approaches showed a protective effect of primary ADT, which was significant for the urologist prescribing preference instrument (HR: 0.74; 95% CI: 0.60–0.93). In the high risk subgroup, using urologist preference for primary ADT as the instrument, there was a significant reduction in overall mortality (HR: 0.75; 95% CI: 0.57–0.99), driven by a large reduction in prostate cancer-specific mortality.

CONCLUSIONS—Instrumental variable analysis appears to provide better control of bias when assessing the effectiveness of primary ADT for prostate cancer although the results may be more applicable to policy rather than clinical decisions.

INTRODUCTION

There has been much recent interest in comparative effectiveness research, which aims to identify the optimal management for a given condition (1). Although clinical trials represent the gold standard for determining the efficacy of treatments, they often lack generalizability due to highly selected enrolled populations. This issue may be potentially addressed with

Please address all correspondence to: Yong-Fang Kuo, PhD, Department of Internal Medicine, University of Texas Medical Branch, 301 University Boulevard, Galveston, TX 77555-0177, Phone: (409) 266-9665, Fax: (409) 772-8931, yokuo@utmb.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

observational study designs but a major challenge is indication bias (2–4). Standard methods such as stratification, matching, multivariable regression and propensity score approaches often inadequately control for this typically very strong bias (4–5).

Instrumental variable analysis is a technique borrowed from econometrics that shows promise in health care settings (6–7). The idea entails finding a variable, the instrument, that is strongly associated with a patient's likelihood of receiving treatment, but is independent of the outcomes. One of the main challenges is finding a suitable instrument, with geographic region being one of the most commonly used (5, 8, 9). These studies capitalized on the fact that prognostic factors that relate to outcome are usually evenly distributed across regions, despite often dramatic differences in rates of treatment. The most frequently used estimation method for these analyses is a two-stage linear regression to estimate the effect of treatment (5, 10–13). This method has several disadvantages, however. Mortality is treated as a continuous rather than binary variable, potentially resulting in model misspecification and an inconsistent estimate of treatment effect (14). The results from the model are specified as a risk difference rather than a relative measure of treatment effect (i.e. relative risk or hazard ratio), which may be more desirable. Moreover, only mortality at a specific time point can be analyzed such that patients without complete follow-up are excluded. Therefore, without a censoring mechanism, cause-specific mortality cannot be analyzed.

In this paper, we explored the potential advantages of using an instrumental variable approach for the assessment of primary androgen deprivation therapy (ADT) for prostate cancer. Despite its very common use, substantial controversy remains about ADT, particularly for localized disease, because of a dearth of supporting evidence of efficacy and concerns about serious adverse effects (15–18). The uncertainty about ADT resulted in substantial variations in its use across regions and physicians through the 1990s which we leveraged to construct instruments based on region and physician prescribing preference (19–20). Previous observational studies have paradoxically demonstrated an increased risk of prostate cancer death associated with use of ADT (4, 21–23). This may represent inadequately controlled bias since the indication for use of ADT usually relates to high risk characteristics of the prostate tumor. We hypothesized that an instrumental variable approach would provide better control for this type of bias over standard multivariable regression methods. We therefore used instrumental variable analysis, with an alternative estimation method - exogenous probit modeling (14), to specifically examine the effect of ADT on prostate cancer-specific mortality.

METHODS

Data Sources

We used the merged Surveillance, Epidemiology and End Results (SEER)-Medicare database for the study analyses. The SEER program is overseen by the National Cancer Institute, consists of regional and state-based tumor registries located throughout the country, and represented approximately 14% of the U.S. population until 2001 and 26% thereafter. The American Medical Association (AMA) Physician Masterfile was used to identify physicians with urology as their primary specialty.

Study Subjects

All men 66 years old with a diagnosis incident prostate cancer from 1992 through 2002 were initially selected for the study (n=163,613). Follow-up was available through the end of 2005. To ensure complete information, patients not enrolled in both Medicare Part A and Part B for 12 months before and 6 months following their cancer diagnosis, who died within 6 months of diagnosis, were members of an Health Maintenance Organization, or diagnosed

by autopsy or on a death certificate were excluded (n=56,993). Since the focus of the study was use of primary ADT (i.e. use of ADT in the absence of other therapies), patients receiving radiation or radical prostatectomy within 6 months of diagnosis were excluded (n=61,735). In order to perform the instrumental variable analyses, we identified a treating urologist for each patient using previously described algorithms (n=10,994 patients who could not be assigned to an urologist were excluded from analyses) (24).

Variables

Patient demographic and tumor characteristics were derived from SEER records and used to categorize patients by age, ethnicity, SEER region of residence, year of diagnosis, clinical stage (T1 through T4), and grade (low - Gleason 2–4; moderate - Gleason 5–7; high - Gleason 8–10). The socio-economic characteristics of each patient were based on percent of adults with less than 12 years of education and median household income within the zip code of residence using U.S. Census 2000 data. Comorbidity was assessed using an adaptation of the Charlson Comorbidity Index for use with claims data (25–26). Use of ADT was defined based on the receipt of at least one dose of a gonadotropin-releasing hormone (GnRH) agonist or orchiectomy occurring within 6 months of diagnosis, using previously published and validated methods (18, 22).

Two instrumental variables were constructed. The first was based on the geographic region of the patient's residence, using Health service area (HSA). An HSA is defined as one or more counties that are relatively self-contained with respect to hospital care. There were a total of 81 HSAs for the study cohort (n<11 [precise number suppressed for confidentiality] patients who had missing HSA information were excluded). The variable was constructed as the proportion of patients receiving primary ADT over the study period within each HSA. The second instrumental variable was based on urologist preference for use of primary ADT. Similar to a previous conceptualization, a patient is classified as having a urologist with preference for primary ADT if the urologist administered primary ADT to any of his patients in the most recent previous month (10). This variable was focused on urologists as have previously noted that over the study period, the vast majority of ADT was administered by urologists (20). Patients diagnosed under the urologist's first prescription month during the study period were excluded (n=1947).

The outcome variable was overall mortality and cause-specific mortality (divided as prostate cancer-specific and non-prostate cancer mortality), obtained from SEER records.

Statistical Analyses

Demographic and tumor characteristics were stratified by whether or not patients received primary ADT. The same stratification was also performed based on the instrumental variables - whether patients resided in an HSA with a rate of primary ADT use above or below the median (44%) or whether or not patients were assigned a urologist with a preference for use of primary ADT. Standard survival analyses were performed using Cox proportional hazards regression adjusted for stage, grade, year of diagnosis, age, comorbidity, ethnicity, marital status, SEER region, education, and income. Patients were censored at death or at the end of the study (12/31/2005). For comparability to the instrumental variable analyses (see below), patient record data was converted to person-month data, in which each patient had multiple records and was coded as 0 if alive or 1 if dead, for each month. The total number of records for each patient is equal to the number of months of follow-up with the same censoring as the Cox model above. Unlike survival data analyses, analyses on person-month data do not take into account that the outcome at a given month is conditional on the outcome in previous months. In order to compare the results between the two different approaches, we conducted multivariable probit regression models

on the person-month data with the same adjustments as for the Cox model above, with relative risks estimated by multiplying the probit coefficients by a factor of 1.8 to approximate logistic coefficients (a factor of 1.6 is also suggested in the literature but did not alter findings enough to change conclusions, so data was not presented) (14, 27)

For the instrumental variable analyses, we used an exogenous probit model to analyze person-month data. The standard error of the instrumental variable parameter is estimated robustly to account for intra-patient correlation across months. This method follows the properties of a binary distribution for the outcome of mortality, incorporates a mechanism for censoring, and provides a relative estimate of treatment effect. The equation differs from the two-stage linear regression approach typically used for instrumental variable analysis, and is listed below:

$$T_i^* = \alpha_0 + \alpha_1 Z_i + \alpha_2 \underline{X}_i + U_{1i}$$

$$\Pr(Y_i = 1) = \Phi(\beta_0 + \beta_{1iv} T_i^* + \beta_2 \underline{X}_i)$$

Where T_i is observed treatment (an endogenous factor) which can be characterized by a latent variable of T_i^* and linearly depends on an instrument variable of Z_i (an exogenous factor), other risk factors - a vector of $\underline{X}_i$, and a random error of U_{1i} . Y_i is the study outcome and Φ is the standard normal distribution function with density $\phi = \Phi'$. All parameters in the model are maximum likelihood estimators. Each of the instrumental variables was analyzed in different exogenous probit models using the same adjustments as the Cox proportional hazards model described above.

To test the assumption that the instrumental variables were strongly associated with receipt of ADT we used a partial F-test that has as its null hypothesis that the coefficient for the effect of the instrument in the first stage regression model is zero (12). For both instruments, the null hypothesis was rejected at $p < 0.001$, with F statistics of 363 and 939 for HSA and urologist preference, respectively. Generally an F statistic > 10 suggests that the instrument is not weak (28). To assess the degree to which the instruments were successful in balancing prognostic factors, we developed a prostate cancer severity index using a Cox proportional hazards regression model to predict one year and five year prostate cancer mortality using baseline patient and tumor characteristics (c statistics=0.77 and 0.51, respectively) (29) and compared this across levels of the instruments. We also pre-specified two clinically relevant subgroup analyses. Accordingly, all analyses were repeated for two sub-cohorts. One cohort included patients with lower risk disease (clinical stage T1 or T2 tumors with low to moderate grade histology) who would be expected to derive minimal benefit from ADT. The other cohort included patients with higher risk disease (either patients with high risk clinically localized disease, or locally advanced disease) who would be expected to benefit from primary ADT. The high risk clinically localized disease group included patients with clinical stage T1 or T2 tumors but high grade histology, or T3a tumors. The locally advanced group included patients with T3b tumors, or T4 tumors excluding cases with positive lymph nodes, unknown lymph node status or distant metastases ($n=3,385$ patients excluded). . All analyses were performed using SAS v.9.1 (SAS institute, Cary, NC) and STATA 9.2 (StataCorp, College Station, TX). All statistical tests were two-sided with $p < 0.05$ considered to be statistically significant. The study was approved by the Institutional Review Board of the University of Michigan.

RESULTS

Table 1 presents patient and tumor characteristics, as well as predicted one year prostate cancer mortality, stratified by receipt of primary ADT, and by the instrumental variables.

Patients who received primary ADT were older, and had more advanced stage and higher grade prostate tumors. In contrast, across categories of both instrumental variables, there was a substantially more balanced distribution of age, clinical stage and tumor grade. Crude overall, prostate cancer-specific, and non-prostate cancer survival across various time points were substantially lower, and predicted one year and five year prostate cancer mortality were higher, in primary ADT users versus non-users. In contrast, crude survival and predicted mortality were well balanced across categories of the instrumental variables.

Table 2 presents the association between use of primary ADT and overall, prostate cancer-specific and non-prostate cancer mortality, using standard multivariable regression and instrumental variable approaches in the entire cohort of patients. In the multivariable regression models, there was a substantial and significant increase in the risk of prostate cancer-specific mortality associated with use of primary ADT, with more modest but still significant increases in the risk of overall and non-prostate cancer mortality. In contrast, there was a complete reversal of the pattern with prostate cancer-specific mortality with the instrumental variables approaches, with a lower marginal risk associated with either a higher primary ADT use HSA or being seen by a urologist with a preference for primary ADT. The results were statistically significant only for the urologist preference variable, however.

Table 3 presents pre-specified subgroup analyses using the same analytical approaches in Table 2, in men with lower risk disease, and in men with high risk disease. In both subgroups, the multivariable regression models demonstrated an increased risk of prostate cancer-specific mortality associated with use of primary ADT, which was particularly pronounced in the group with lower risk disease. For both groups using the instrumental variable approaches, there were trends towards reduction in prostate cancer-specific mortality, with a statistically significant reduction only in the high risk subgroup using urologist preference as the instrument. In the lower risk group, the instrumental variable approaches demonstrated no significant effect on overall mortality. In the high risk group, there was a significant reduction in overall mortality using urologist preference for primary ADT as the instrument.

DISCUSSION

This study examined the association between use of primary ADT and survival in men with prostate cancer using both standard multivariable regression and instrumental variable approaches. Most notably, there was a sharp contrast in results between the different approaches, particularly with regards to prostate cancer-specific mortality. The standard multivariable regression approaches uniformly showed an increased risk of prostate cancer death. Both instrumental variable approaches demonstrated trends towards a protective effect of primary ADT, though results were statistically significant only for urologist prescribing preference. In the high risk disease subgroup, using urologist preference for primary ADT as the instrument, there was a significant reduction in overall mortality, driven by the large reduction in prostate cancer-specific mortality.

Several points support the validity of the results from the instrumental variable approaches. The results are consistent with those from clinical trials of primary ADT use in men with locally advanced or metastatic prostate cancer, which demonstrated a clear reduction in prostate cancer-specific mortality, with initial results from the Medical Research Council study showing improvement in overall survival as well (30–31). The results are biologically plausible, with the beneficial effect of primary ADT being mediated through a reduction in prostate cancer-specific mortality. Multivariable regression analyses in this study and in other published work examining use of primary ADT have uniformly shown an increased risk of prostate cancer-specific mortality in localized disease (21, 23), almost certainly

representing insufficient control of indication bias. One likely candidate for an unmeasured confounder in previous analyses is PSA level at diagnosis (which until recently was unavailable in the SEER-Medicare dataset), which is both a strong driver of decisions to use ADT and an important prognostic factor. Important assumptions about the instruments used in this study are likely to be valid. The assumption that the instruments are not independently related to the outcome cannot be tested directly. However, by theoretical considerations, geographic region as defined by HSA is unlikely to have an independent effect on prostate cancer mortality. There is potentially more concern with use of an instrument based on the treating urologist. However, since we limited our analysis to patients who did not undergo definitive treatment such as surgery or radiation, and since dosing regimens for ADT are generally standard, any effect of the urologist should only be mediated through their decision to use ADT, which would not violate the assumption. In addition, stratification by categories of the instruments yielded a good balance of relevant prognostic factors (Table 1). We have also demonstrated in this study and previous work that both our instruments were very strongly correlated with use of primary ADT, an important criterion for an instrumental variable (19–20). A final point is that the pattern of results was consistent across the two different instruments, providing further credibility for the validity of our findings.

There are a number of methodological issues of note in this study. The choice of instrument can be important. Although results were similar across the instruments, urologist preference has potential advantages over HSA. Statistical power is often very limited in instrumental variable analyses, with low precision of point estimates rendering meaningful interpretation of results difficult, as demonstrated in a recent paper examining the impact of radical prostatectomy for localized prostate cancer (32). The precision of an instrumental variable analysis appears to depend on the number of units for the instrument, rather than on the patient sample size, and also on the strength of the instrument in terms of its association with the exposure of interest. In this study, there were 81 HSAs versus 1,321 urologists, and in previous work we have shown that through the period of the study, the urologist of the patient was the strongest predictor of whether the patient received ADT (20). As such, there was more statistical power using urologist preference as the instrument as can be noted by the narrower confidence intervals. This resulted in some of the findings being statistically significant using urologist preference as the instrument, which were not significant using HSA, despite similar point estimates (Table 3). Another methodological issue is the choice of estimation method for the instrumental variable analysis. A previous study using SEER-Medicare data and an instrumental variable approach also examined primary ADT and survival (9). That study also used the proportion of patients receiving primary ADT within the HSA as the instrument, but employed a different analytical approach, based on the Rubin Causal model (33). Although the results were similar to our findings, there was an increased risk of prostate cancer-specific mortality associated with the instrument in the low risk patients, and in the high risk group, the protective effect was more modest such that there was no effect on overall mortality. The authors explained the adverse effect on prostate cancer death in the lower risk group as potentially being related to the selection of more androgen resistant clones in the setting of ADT. However, as pointed out by others (34), a more plausible explanation is simply residual confounding by indication. In the lower risk group, it is more likely that unmeasured factors accounted for the decision to use ADT, making the indication bias much stronger. Since we used a similar instrument and the same database, it is evident that the difference in results must be related to our use of a different estimation method, which appears to have allowed for better control of bias.

This study has important limitations. With an observational design, it is not possible to determine with certainty what the “true” effect of ADT is on survival and therefore know whether the findings from our instrumental variable analyses provide a more correct

assessment than those from the standard multivariable regression methods. However, we have argued that the instrumental variable analyses are more likely to be correct since they produced findings that are more in line with biological plausibility. In addition, known prognostic variables were better balanced across categories of the instruments, making it more likely that important unmeasured factors would also be better balanced. An important limitation of the instrumental variable approach in general is the difficulty in finding suitable instruments. This requires a clinical context where decisions to treat are highly dependent on non-clinical factors. In this study example, we used an era in which there was extreme variation in use of ADT (20), likely due to both uncertainty about its efficacy but also tremendous financial incentives for its use (to which physicians respond variably). Major cuts in reimbursement since 2004 may have substantially reduced physician variation in use of ADT, making physician preference a weaker instrument in more recent years (35). Finally, the interpretation of results from instrumental variable analyses is not straightforward, since they estimate the average effect in the so-called “marginal population” – those patients in whom treatment assignment depends entirely on the instrument. In our example, the findings would not apply to patients who would “always” or “never” receive ADT, regardless of the instrument. Given the substantial clinical uncertainty about use of ADT during the study period, however, a large proportion of patients would be expected to fall within the marginal population. Nevertheless, it may be difficult for a practicing clinician to identify whether an individual patient under their care is a member of the marginal population, and therefore whether the study results would apply to them. Instrumental variable analyses are therefore often most useful for determining population level effects of policy changes (36). In this context, for example, the study results may help inform the potential impact of policies that reduce reimbursement for ADT.

The study findings have a number of implications. From a clinical standpoint, with the caveats in interpretation discussed above, our results suggest there may be no overall survival benefit to use of primary ADT for lower risk prostate tumors, despite a modest reduction in prostate cancer-specific mortality. In contrast, in the high risk subgroup, our results provide support for use of primary ADT as compared with observation. However, recent clinical trial results suggest that for those men who can tolerate it, the optimal approach for high risk disease may be radiation combined with ADT (37). From a methodological standpoint, instrumental variable analysis appears to be a powerful tool for reducing bias in the assessment of treatment effectiveness using observational data. Even when clinical trial data is already available, the results of a confirmatory instrumental variable analysis can be helpful. Results based on observational data, which usually include less selected populations, may in fact provide a truer assessment of treatment effectiveness as compared with clinical trials (5). Our results also demonstrate that careful consideration should be paid to the choice of instrument and the estimation method as these can have important effects on the findings.

REFERENCES

1. Lohr K. Comparative effectiveness research methods: Symposium overview and summary. *Med Care*. 2010; 48:S3–S6. [PubMed: 20473188]
2. Grimes DA, Schulz KF. Bias and causal association in observational research. *Lancet*. 2002; 359:248–252. [PubMed: 11812579]
3. Brookhart MA, Sturmer T, Glynn RJ, Rassen J, Scheeweiss S. Confounding control in healthcare database research – challenges and potential approaches. *Med Care*. 2010; 48:S114–S120. [PubMed: 20473199]
4. Giordano SH, Kuo YF, Duan Z, Hortobagyi GN, Freeman JL, Goodwin JS. Limits of observational data in determining outcomes from cancer therapy. *Cancer*. 2008; 112:2456–2466. [PubMed: 18428196]

5. Stukel TA, Fisher ES, Wennberg DE, Aler DA, Gottlieb DJ, Vermeulen MJ. Analysis of observational studies in the presence of treatment selection bias – effect of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *JAMA*. 2007; 297:278–285. [PubMed: 17227979]
6. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. 2000; 29:722–729. [PubMed: 10922351]
7. Newhouse JP, McClellan M. Econometrics in outcomes research: the use of instrumental variables. *Annu Rev Public Health*. 1998; 19:17–34. [PubMed: 9611610]
8. Landrum MB, Ayanian JZ. Causal effect of ambulatory specialty care on mortality following myocardial infarction: A comparison of propensity score and instrumental variable analyses. *Health Serv Outcomes Res Methodol*. 2001; 2:221–245.
9. Lu-Yao GL, Albertsen PC, Moore DF, Shih W, Lin Y, DiPaola RS, Yao SL. Survival following primary androgen deprivation therapy among men with localized prostate cancer. *JAMA*. 2008; 300:173–181. [PubMed: 18612114]
10. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology*. 2006; 17:268–275. [PubMed: 16617275]
11. Schneeweiss S, Seeger JD, Landon J, Walker AM. Aprotinin during coronary artery bypass grafting and risk of death. *NEJM*. 2008; 358:771–783. [PubMed: 18287600]
12. Hardley J, Polsky D, Mandelblatt JD, Mitchell JM, Weeks JC, Wang Q, Hwang YT, and the OPTIONS research Team. An exploratory instrumental variable analysis of the outcomes of localized breast cancer treatments in a Medicare population. *Health Econ*. 2003; 12:171–186. [PubMed: 12605463]
13. Kahn JM, Ten Have TR, Iwashyna TJ. The relationship between hospital volume and mortality in mechanical ventilation: An instrumental variable analysis. *Health Serv Res*. 2009; 44:862–879. [PubMed: 19674428]
14. Rassen JA, Schneeweiss S, Glynn RJ, Mittleman MA, Brookhart MA. Instrumental variable analysis for estimation of treatment effects with dichotomous outcomes. *Am J Epidemiol*. 2009; 169:273–284.
15. Welch HG, Albertsen PC, Nease RF, Bubolz TA, Wasson JH. Estimating treatment benefits for elderly: the effect of competing risk. *Ann Intern Med*. 1996; 124:577–584. [PubMed: 8597322]
16. Chodak GW, Keane T, Klotz L. Critical evaluation of hormonal therapy for carcinoma of the prostate. *Urology*. 2002; 60:201–208. [PubMed: 12137810]
17. Higano CS. Side effects of androgen deprivation therapy: monitoring and minimizing toxicity. *Urology*. 2003; 61:32–38. [PubMed: 12667885]
18. Shahinian VB, Kuo YF, Freeman JL, Goodwin JS. Risk of fracture after androgen deprivation for prostate cancer. *N Engl J Med*. 2005; 352:154–164. [PubMed: 15647578]
19. Shahinian VB, Kuo YF, Freeman JL, Orihuela E, Goodwin JS. Increasing use of gonadotropin-releasing hormone agonists for the treatment of localized prostate carcinoma. *Cancer*. 2005; 103:1615–1624. [PubMed: 15742331]
20. Shahinian VB, Kuo YF, Freeman JL, Goodwin JS. Determinants of androgen deprivation therapy use for prostate cancer: role of the urologist. *J Natl Cancer Inst*. 2006; 98:839–845. [PubMed: 16788157]
21. Wong YN, Freedland SJ, Egleston B, Vapiwala N, Uzzo R, Armstrong K. The role of primary androgen deprivation therapy in localized prostate cancer. *Eur Urol*. 2009; 56:609–616. [PubMed: 19368995]
22. Lu-Yao G, Moore DF, Oleynick J, DiPaola RS, Yao SL. Use of hormonal therapy in men with metastatic prostate cancer. *J Urol*. 2006; 176:526–531. [PubMed: 16813882]
23. Holmes L, Chan W, Ziang Z, Du XL. Effectiveness of androgen deprivation therapy in prolonging survival of older men treated with locoregional prostate cancer. *Prostate Cancer Prostatic Dis*. 2007; 10:388–395. [PubMed: 17486111]
24. Shahinian VB, Kuo YF, Freeman JL, Orihuela E, Goodwin JS. Characteristics of urologists predict the use of androgen deprivation therapy for prostate cancer. *J Clin Oncol*. 2007; 25:5359–5365. [PubMed: 18048816]

25. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A method of classifying prognostic comorbidity in longitudinal studies : development and validation. *J Chronic Dis.* 1998; 40:373–383. [PubMed: 3558716]
26. Klabunde C, Potosky A, Legler J, Warren J. Development of a comorbidity index using physician claims data. *J Clin Epidemiol.* 2000; 53:1258–1267. [PubMed: 11146273]
27. Agresti, A. *Categorical Data Analysis* 2nd Edition. New York, NY: John Wiley & Sons, Inc; 2002.
28. Cameron, AC.; Trivedi, PA. *Microeconometric using STATA*. College Station, TX: StataCorp; 2009.
29. Pencina MJ, D’Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Statistics in Medicine.* 2004; 23:2109–2123. [PubMed: 15211606]
30. Studer UE, Hauri D, Hanselmann S, Chollet D, Leisinger HJ, Gasser T, Senn E, Trinkler FB, Tscholl RM, Thalmann GN, Dietrich D. Immediate versus deferred hormonal treatment for patients with prostate cancer who are not suitable for curative local treatment : results of the randomized trial SAKK 08/88. *J Clin Oncol.* 2004; 22:4109–4118. [PubMed: 15483020]
31. The Medical Research Council Prostate Cancer Working Party Investigators Group. Immediate versus deferred treatment for advanced prostatic cancer: initial results of the Medical Research Council trial. *Br J Urol.* 1997; 79:235–246. [PubMed: 9052476]
32. Hadley J, Yabroff KR, Barrett MJ, Penson DF, Saigal CS, Potosky AL. Comparative effectiveness of prostate cancer treatments : evaluating statistical adjustment for confounding in observational data. *J Natl Cancer Inst.* 2010; 102:1780–1793. [PubMed: 20944078]
33. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc.* 1996; 91:444–455.
34. Laufman L. Primary androgen deprivation therapy in men with prostate cancer. *JAMA.* 2009; 301:35. letter. [PubMed: 19126809]
35. Shahinian VB, Kuo YF, Gilbert SM. Reimbursement policy and androgen deprivation therapy for prostate cancer. *N Engl J Med.* 2010; 363:1822–1832. [PubMed: 21047226]
36. Harris KM, Remler DK. Who is the marginal patient? Understanding instrumental variables estimates of treatment effects. *Health Serv Res.* 1998; 33:1337–1360. [PubMed: 9865223]
37. D’Amico AV, Manola J, Loffredo M, Renshaw AA, DellaCroce A, Kantoff PW. 6-month androgen suppression plus radiation therapy vs. radiation therapy alone for patients with clinically localized prostate cancer – A randomized controlled trial. *JAMA.* 2004; 292:821–827. [PubMed: 15315996]

Patient characteristics stratified by primary ADT use, high vs. low primary ADT use health service area, and urologist preference for primary ADT

Table 1

Characteristics	Category	Use of Primary ADT		High Primary ADT use Health Service Area		Urologist preference for primary ADT	
		Yes	No	Yes	No	Yes	No
Sample Size	N=	14,450	17,480	15,489	16,441	16,080	15,850
Age (years)	65 – 69	8.6%	11.4%	9.3%	10.9%	10.0%	10.2%
	70 – 74	19.1%	23.3%	20.6%	22.2%	21.0%	21.8%
	75 – 79	28.1%	30.5%	29.4%	29.5%	29.3%	29.6%
	80	44.2%	34.8%	40.8%	37.4%	39.7%	38.4%
	Mean ± Std. Dev.	78.5 ± 6.2	77.2 ± 6.1	78.1 ± 6.2	77.5 ± 6.2	77.9 ± 6.2	77.7 ± 6.2
Ethnicity	White	77.0%	78.9%	76.0%	80.0%	78.3%	77.8%
	Black	8.9%	9.9%	4.7%	13.9%	8.1%	10.8%
	Hispanic	4.8%	4.2%	5.7%	3.3%	4.4%	4.5%
	Others	9.3%	7.0%	13.7%	2.8%	9.2%	6.9%
Year of Diagnosis	1992 – 1995	40.4%	46.1%	43.0%	44.0%	41.6%	45.5%
	1996 – 1999	33.6%	33.2%	33.7%	33.2%	33.4%	33.4%
	2000 – 2002	26.0%	20.7%	23.3%	22.8%	25.0%	21.1%
		14.0%	12.1%	17.1%	9.0%	13.8%	12.1%
Regions	Northeast	31.7%	33.9%	21.7%	43.5%	32.5%	33.3%
	Midwest	7.4%	10.4%	2.8%	15.0%	7.4%	10.8%
	South	46.9%	43.6%	58.5%	32.5%	46.3%	43.8%
	West	38.9%	37.9%	36.8%	39.9%	38.6%	38.2%
Marital Status	Not Married	61.1%	62.1%	63.2%	60.1%	61.4%	61.8%
	Married	19.3%	22.4%	21.4%	20.6%	20.5%	21.5%
Zip code education* (% adults less than 12 years education)	< 8.4%	23.8%	23.0%	24.0%	22.8%	23.9%	22.9%
	8.4 – < 13.3%	27.0%	26.5%	26.1%	27.4%	27.9%	25.6%
	13.4 – < 20.3%	29.9%	28.1%	28.5%	29.2%	27.8%	30.0%
	20.4%	14.9%	14.4%	14.3%	15.2%	14.5%	14.8%
Zip code median income*	Median	28.1%	28.0%	25.8%	30.2%	27.5%	28.6%
	<\$ 37,500	26.4%	25.9%	25.6%	26.6%	26.5%	25.7%
	\$37,500 – < \$47,500	24.5%	23.5%	21.8%	26.0%	24.1%	23.8%
	\$47,500 – < \$62,000						

Characteristics	Category	Use of Primary ADT		High Primary ADT use Health Service Area		Urologist preference for primary ADT	
		Yes	No	Yes	No	Yes	No
	\$62,000	21.0%	22.6%	26.8%	17.2%	21.9%	21.9%
	Median	\$45,361	\$45,521	\$46,368	\$44,582	\$45,521	\$45,283
Comorbidity	0	69.4%	69.5%	68.9%	70.0%	69.3%	69.7%
	1	19.9%	19.6%	20.3%	19.1%	20.0%	19.4%
	2	7.0%	6.9%	7.0%	6.9%	7.0%	6.9%
	3	3.7%	4.0%	3.8%	4.0%	3.7%	4.0%
Clinical stage	T1	17.2%	39.1%	28.7%	29.7%	27.7%	30.7%
	T2	46.3%	43.0%	43.4%	45.4%	45.6%	43.3%
	T3	4.3%	1.5%	2.8%	2.8%	2.7%	2.8%
	T4	19.0%	4.1%	11.2%	10.5%	10.8%	10.9%
Grade	Unknown	13.2%	12.3%	13.9%	11.7%	13.2%	12.3%
	Well 2–4	5.6%	21.9%	14.7%	14.3%	13.8%	15.3%
	Moderate 5–7	47.5%	53.8%	50.2%	51.7%	51.3%	50.6%
	Poor/Un 8–10	37.6%	13.1%	25.9%	22.3%	25.0%	23.4%
Overall survival	Unknown	9.3%	11.2%	9.2%	11.4%	9.9%	10.8%
	5 years	50.3%	65.7%	59.0%	58.6%	58.9%	58.7%
	8 years	30.3%	45.6%	38.5%	39.1%	38.9%	38.7%
	10 years	20.3%	34.2%	27.7%	28.5%	27.9%	28.3%
Prostate cancer-specific survival	12 years	13.0%	24.6%	19.0%	20.1%	19.5%	19.6%
	5 years	77.1%	92.7%	86.1%	85.6%	86.3%	85.4%
	8 years	67.3%	87.9%	78.9%	79.2%	79.6%	78.4%
	10 years	61.0%	84.4%	74.3%	74.7%	74.7%	74.4%
Non prostate cancer-specific survival	12 years	55.0%	80.2%	68.0%	71.2%	69.5%	69.9%
	5 years	65.4%	70.9%	68.6%	68.5%	68.3%	68.8%
	8 years	45.2%	51.9%	48.9%	49.4%	49.0%	49.4%
	10 years	33.4%	40.6%	37.4%	38.2%	37.5%	38.2%
Predicted one year prostate cancer mortality	12 years	23.8%	30.7%	28.1%	28.2%	28.2%	28.1%
		6.0%	4.1%	4.9%	5.0%	5.0%	5.0%
Predicted five year prostate cancer mortality		32.4%	27.9%	30.0%	30.0%	29.9%	29.9%

Table 2

Mortality associated with use of primary ADT using multivariable regression and instrumental variable approaches

Model	Overall Mortality		Prostate Cancer-Specific Mortality		Non Prostate Cancer-Specific Mortality	
	RR	95% CI	RR	95% CI	RR	95% CI
Multivariable Cox Model	1.15	(1.11,1.18)	1.51	(1.42,1.61)	1.06	(1.02,1.10)
Multivariable Probit Model	1.11	(1.08,.13)	1.29	(1.24,1.34)	1.04	(1.02,1.07)
Instrumental Variable Probit Model with HSA as the instrument	1.08	(0.89,1.31)	0.88	(0.63,1.21)	1.18	(0.96,1.46)
Instrumental Variable Probit Model with Urologist Preference as the instrument	0.94	(0.83,1.06)	0.74	(0.60,0.93)	1.04	(0.90,1.19)

Mortality associated with use of primary ADT using multivariable regression and instrumental variable approaches in subgroups of men with lower risk and high risk disease

Table 3

Model	Overall Mortality		Prostate Cancer-Specific Mortality		Non-Prostate Cancer Mortality	
	RR	95% CI	RR	95% CI	RR	95% CI
Lower Risk Cohort (T1 or T2, low to moderate grade tumors):						
Multivariable Cox Model	1.19	(1.13,1.25)	1.90	(1.68,2.16)	1.10	(1.05,1.16)
Multivariable Probit Model	1.13	(1.09,1.16)	1.44	(1.34,1.54)	1.07	(1.04,1.11)
Instrumental Variable Probit Model with HSA as the instrument	1.22	(0.97,1.55)	0.79	(0.46,1.33)	1.31	(1.02,1.69)
Instrumental Variable Probit Model with Urologist Preference as the instrument	1.06	(0.89,1.26)	0.78	(0.52,1.17)	1.11	(0.92,1.33)
High Risk Cohort (T1 or T2 with high grade tumors, T3 or T4 tumors without regional or distant metastasis):						
Multivariable Cox Model	1.08	(1.01,1.15)	1.38	(1.22,1.56)	0.96	(0.89,1.04)
Multivariable Probit Model	1.05	(1.01,1.11)	1.22	(1.13,1.31)	0.97	(0.92,1.02)
Instrumental Variable Probit Model with HSA as the instrument	1.00	(0.63,1.57)	0.90	(0.44,1.82)	1.07	(0.62,1.84)
Instrumental Variable Probit Model with Urologist Preference as the instrument	0.75	(0.57,0.99)	0.62	(0.40,0.94)	0.87	(0.63,1.20)