0022-5347/99/1622-0439/0
THE JOURNAL OF UROLOGY
Vol. 162, 439–444, August 1999
Printed in U.S.A.
Copyright © 1999 by AMERICAN UROLOGICAL ASSOCIATION, INC.

STATISTICAL CONSIDERATIONS WHEN ASSESSING OUTCOMES FOLLOWING TREATMENT FOR PROSTATE CANCER

PETER C. ALBERTSEN, JAMES A. HANLEY, AND MARLENE MURPHY-SETZKO

From the Division of Urology, University of Connecticut Health Center, Farmington, Connecticut, and Department of Epidemiology and Biostatistics, McGill University, Montreal, PQ, Canada

ABSTRACT

Purpose: We explore the impact of study designs, biases, outcome variables and statistical techniques when interpreting studies concerning prostate cancer management.

Materials and Methods: Examples from the current literature and a recently assembled population based sample of patients 55 to 75 years old at diagnosis identified by the Connecticut Tumor Registry as having newly diagnosed localized prostate cancer between 1971 and 1984 are provided to assist the reader to understand the principles discussed.

Results: Most reports concerning prostate cancer outcomes suffer from obvious and subtle biases that confound the reader's understanding of the impact of different treatment alternatives.

Conclusions: By remaining vigilant to these confounding issues, clinicians and patients can gain greater insights into the medical literature and can make individual interpretations concerning the potential impact of treatment interventions on men with prostate cancer.

KEY WORDS: prostate, prostatic neoplasms, outcome assessment (health care), survival analysis

In 1998 it was estimated that approximately 200,000 Americans would be diagnosed with prostate cancer.[1] Many of these men would be offered treatments designed to cure and/or control disease progression. When selecting among currently available treatment options, patients must weigh the potential benefit of increased longevity and improved quality of life against the potential risk of complications associated with treatment or the absence of treatment. Ideally patients need information concerning the natural history of the disease, increased longevity or symptom improvement provided by specific therapies, frequency and severity of complications associated with these therapies, and factors that predict varying outcomes for specific subgroups of patients. Unfortunately the data needed to perform these assessments are often either unavailable or lack precision.

Many patients select primary treatment following a review of information available in the medical literature or interpreted by the media. However, most available studies suffer from obvious and subtle biases that serve to confound the reader's understanding of the impact of different treatment alternatives. Problems readers encounter while interpreting the medical literature are often related to issues surrounding study designs that offer varying abilities to make valid comparisons, biases that impact the construct of the patient population being studied, and statistical analyses and outcome metrics that describe results in ways that are often difficult for patients and/or clinicians to interpret.

We explore the impact of these factors on the presentation and interpretation of available data on prostate cancer management. Examples from the current literature and a recently assembled population based sample of patients 55 to 75 years old at diagnosis identified by the Connecticut Tumor Registry as having newly diagnosed localized prostate cancer between 1971 and 1984 are provided to assist the reader.[2] This review should provide clinicians and patients with greater insights concerning how to interpret studies reporting outcomes associated with various treatments offered to men with prostate cancer.

STUDY DESIGN

Patients and clinicians must pay careful attention to the study design used in research reports. While randomized, controlled, experimental protocols are frequently the ideal way to evaluate new and existing diagnostic tests and procedures, this interventional approach is not always practical, especially for chronic diseases such as prostate cancer that require many years of followup. The choice of study design depends on the question that is being addressed and usually entails a selection between an experimental design in which patients are assigned to a treatment or a nonexperimental design in which data are assembled after patients select their own treatment.

Nonexperimental study designs. A typical example of a nonexperimental design is a case series report that describes and inventories practice patterns or outcomes. Many physicians use case series reports to document clinical outcomes following a specific medical intervention. While these reports provide some insight concerning treatment efficacy, data from case series frequently suffer from numerous confounding factors and biases that limit their usefulness when generalizing results to a community practice population. Since data from case series lack a comparison group, patients and clinicians are unable to differentiate between the impact of treatment and the natural progression of the disease or the impact of competing medical hazards. Researchers rarely provide information concerning the likely outcome of patients who have not received treatment. Readers often assume that a majority of patients receiving treatment would otherwise be destined to die of disease in the absence of treatment. Without a comparison population, however, it is impossible to assess the relative efficacy of the treatment alternative being evaluated as demonstrated in the report of Gerber et al.[3]

In 1996 Gerber et al published a multi-institutional pooled analysis of men with clinically localized disease treated with radical prostatectomy between 1970 and 1993. They reported excellent 10-year disease specific survival estimates of 94, 80 and 77% for men with well (Gleason score 2 to 4), moderately (Gleason score 5 to 7) and poorly (Gleason score 8 to 10) differentiated disease. Initial review of these data suggests that radical prostatectomy is most efficacious among men with well differentiated disease and least efficacious among those with poorly differentiated disease. Unfortunately no population is available to compare clinical outcomes in the absence of treatment. While not an ideal comparison popu-

lation, we recently analyzed the long-term outcome of 767 men followed conservatively for newly diagnosed localized prostate cancer. The 10-year disease specific survival for this sample was 94, 71 and 30% for men with well, moderately and poorly differentiated disease, respectively.[2] These results are identical to those reported by Gerber et al for men with well differentiated disease, suggesting that radical prostatectomy may not provide any survival advantage among these patients. Conversely, results were much worse for men with poorly differentiated disease, suggesting a potentially significant advantage following surgery for them. For men with Gleason 5 to 7 tumors, the group most frequently targeted for aggressive intervention, disease specific survival outcomes do not appear to be dramatically different. Gerber et al reported a 10-year disease specific survival of 80% (95% confidence interval 74 to 85), while our data suggest 72% (67 to 76). Because of the significant selection biases inherent to the construct of both study cohorts and the inadequate staging of many cases managed conservatively in our series, it is impossible to determine the relative efficacy of surgery for this subset of patients. Similar problems are seen when analyzing data from case series of men receiving external beam radiation therapy.[4]

Historical and contemporary controls. Researchers seeking to estimate the relative efficacy of a treatment frequently compare data from a contemporary case series with data from a group diagnosed and treated in a different era. Unfortunately cancers diagnosed in patients from 1 era are rarely comparable to those in another. A good example was recently reported by Helgesen et al in 1996.[5] They analyzed a population based cohort comprising all 80,901 men diagnosed with prostate cancer in Sweden from 1960 through 1988, and demonstrated a significant survival improvement despite the absence of any effective therapeutic innovations. These findings are most likely consistent with an increase in the detection of nonlethal tumors coupled with the significant effect of lead time bias associated with tumor identification following transurethral resection of the prostate (fig. 1). A similar effect can be attributed to testing for serum prostate specific antigen (PSA). Patients diagnosed before the advent of PSA testing were much more likely to have advanced disease at presentation compared with contemporary patients.[6] By advancing the date of diagnosis, clinicians and researchers guarantee a survival advantage to contemporary patients that is independent of treatment compared with patients diagnosed in the pre-PSA era.

Contemporary general populations are also inappropriate comparison groups because men who choose surgery are frequently healthier compared to the general population. A good example is a recent analysis by Barry et al which documents long-term outcomes of patients who underwent radical prostatectomy at either the Mayo Clinic or the University of Utah.[7] In this case series report men undergoing radical prostatectomy had a survival that was superior to an age matched series of contemporary controls who had not been diagnosed with prostate cancer. Even considering its many merits, it is unlikely that radical prostatectomy will improve longevity beyond that which would occur in the absence of prostate cancer.

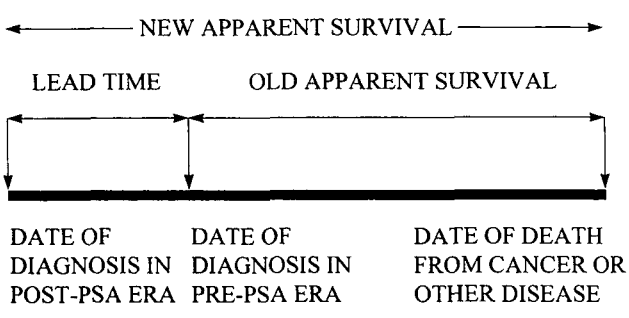

FIG. 1. Lead time bias

BIAS

Classification bias. Comparisons between data obtained from multiple contemporary and/or historical case series also pose problems because of different biases that influence the construct of the study cohorts. An obvious example is a comparison between a surgical series stratified by Gleason scores determined from evaluation of a surgical specimen and a radiation therapy series stratified by Gleason scores determined from evaluation of a biopsy specimen only. Because thorough pathological review of a surgical specimen frequently results in an upgrade of the Gleason score, a case series based on surgical pathology is impacted by the "Will Rogers" phenomenon described by Feinstein et al in an article on survival outcomes of lung cancer patients (fig. 2).[8] By reclassifying some cases into higher categories, average survival in all categories will appear to improve although none has improved individually. A similar phenomenon occurs when prostate cancer cases are re-staged from local disease to either regional or distant disease following surgical exploration. After re-staging these surgical cohorts will demonstrate a survival improvement that will not occur among those receiving external beam radiation, brachytherapy or cryosurgery. The descriptive title of this effect is based on a famous joke by Will Rogers who commented that when all of the "Oklahomans" moved to California during the depression of the 1930s the average IQ of both states went up!

Another good example of classification bias is the report by Aus et al concerning a cohort of 301 Swedish men identified by the Swedish Cancer Registry who were originally diagnosed with localized prostate cancer and who died in Goteborg during 1988 and 1990.[9] The study design selected by the authors is sensitive to several critical issues.[10] Unlike traditional population based studies that accrue patients based on the date of diagnosis, Aus et al accrued patients based on date of death. As a consequence, they selected patients with different characteristics during different times. Men with aggressive tumors were identified from a more contemporary population, while those with more indolent tumors were identified from an earlier era. Because the population at risk changed in size and age distribution and the incidence of the disease increased during the accrual period, results are biased in favor of selecting men with more aggressive disease. As a result their estimates of the 15-year mortality rate from prostate cancer are much higher than those reported by others.[11–13]

Selection bias. Patients choose different treatments at dif-

FIG. 2. Will Rogers' effect. Based on biopsy score 5 patients have Gleason 6 tumors and 3 patients have Gleason 7 tumors. Postoperatively 2 cases are reclassified from Gleason 6 to Gleason 7 according to the surgical pathology. Classifying cases according to surgical scores will yield better survival curves than classifying according to biopsy scores although actual survival of all patients is identical. Illusion occurs because cases with most aggressive tumors in Gleason 6 category are removed, leading to apparent survival improvement of remaining cases classified as Gleason 6. Similarly cases previously classified as Gleason 6 are added to cases with Gleason 7 tumors, leading to apparent survival improvement of these patients because less aggressive lesions are added to group.

ferent institutions for myriad reasons. Patients receiving conservative treatment are clearly different from men who undergo radiation therapy or radical prostatectomy. How these differences impact long-term outcomes is unknown. When readers compare data from different case series they must be alert to potential selection biases that can have a profound effect on the outcomes described.

A way to decrease selection biases in nonexperimental observational studies is to enroll all patients in a defined geographic region, which is known as a population based cohort. A good example is the 1997 report by Johansson et al concerning 15-year outcomes of 642 consecutive men diagnosed with prostate cancer at Orebro Medical Centre, a hospital with a strictly defined catchment area, between 1977 and 1984.[11] They noted that prostate cancer accounted for 210 of all 541 deaths (37%) in the cohort. However, restricting the analysis to the 300 men who presented with localized disease resulted in a death rate from prostate cancer of only 11%. Careful review of this subset of patients revealed that approximately half had well differentiated disease and only 85 were less than 70 years old. Clearly these men are different from the 342 who presented with more advanced disease. To determine whether the 342 men diagnosed with regional or advanced disease would have benefited from modern screening and treatment efforts requires an experimental study design and a comparison population.

Other potential selection biases result from referral patterns. Patients seeking care at tertiary medical centers are often more educated, wealthier and more health conscious than those who remain in their local communities. These factors can favorably impact clinical outcomes independent of the treatment being evaluated. Referral patterns within institutions can also impact outcomes. Good examples are reports by Smith[14] and Walsh[15] et al from the same tertiary medical center. Subtle selection biases beyond surgical technique alone are introduced when analyses are restricted to patients undergoing surgery by specific surgeons.

STUDY VARIABLES AND STATISTICAL TECHNIQUES

In addition to study design and potential biases, readers must be alert to the choice of variables in any study that attempts to determine a cause and effect relationship. The outcome variable (survival) is the factor to be compared, while the intervention variables include those factors that impact the outcome variable and can be measured by the researcher (tumor grade and stage). Other important variables the reader must also consider include confounding variables (co-morbidity) that can influence the outcome variable but that cannot always be fully measured and accounted for by the researcher. Many clinicians frequently select treatment as a critical intervention variable. However, tumor grade and stage may have a much more profound impact on patient outcome. If researchers do not control for these variables, they may make inferences about cause and effect relationships concerning treatment that are not justified.

Multiple regression analysis. The statistical tool often used to search for variables that have the largest impact on outcomes is multiple regression analysis. If the outcome data are binary or categorical (alive/dead at the end of a defined period) logistic regression is used. A good example is the recent report by Partin et al who used logistic regression analysis to identify the key predictors of local tumor extension.[16] They assembled a database containing information on 4,133 men. The key outcome variable of the analysis was pathological stage defined as a categorical variable with the 4 possible values of organ confined disease, isolated capsular penetration, seminal vesicle involvement and pelvic lymph node involvement. Using multiple regression analysis the authors were able to calculate the probability of pathological stage given a clinical stage and preoperative serum PSA level.

Cox's proportional hazards model. If followup times are variable, a life table regression analysis is used usually with Cox's proportional hazards model. With these techniques a statistician systematically searches for variables that impact outcomes. A typical study will evaluate the impact of 3 to 5 variables. The statistician must control for the effect of the other independent variables before drawing cause and effect conclusions concerning the impact of a specific variable on survival or the outcome being assessed. A good example is our recent analysis of data from a cohort of 451 patients followed conservatively for prostate cancer.[13] Gleason score and patient co-morbidities were powerful independent predictors of overall survival. In a more recent analysis men diagnosed with Gleason score 5 tumors had a significantly better outcome than those diagnosed with either Gleason 6 or 7 tumors (fig. 3).[2] Many clinicians reporting results frequently combine data for men with Gleason 5 to 7 tumors.[3] If researchers do not carefully control for the distribution of Gleason 5, 6 and 7 tumors when making comparisons between case series, survival results will be dramatically better or worse depending on the relative number of men with Gleason 5, 6 and 7 tumors that are included in the series. As a result readers will be unable to discern the impact of intervention from the natural history of the disease.

Primary versus secondary data. Readers should also be alert to whether a study uses primary or secondary data. Primary refers to data collected explicitly for the conduct of a particular study, while secondary refers to data gathered for other reasons and used by researchers to address specific research questions. For example, many researchers have used Medicare claims data linked to other large databases, such as the Surveillance, Epidemiology and End Results program sponsored by the National Cancer Institute, to address research questions. Recently Lu-Yao et al evaluated the probability of receiving secondary cancer therapy among patients undergoing radical prostatectomy.[17] By using several administrative databases the authors were able to evaluate 5-year outcomes following radical prostatectomy among patients enrolled in the Medicare program. By identifying claims reflecting the use of luteinizing hormone-releasing hormone agonists or bilateral orchiectomy they were able to estimate the probability of disease progression among men older than 65 years undergoing radical prostatectomy in the general population. The advantage of this approach lies primarily with the large sample size and, consequently, the extraordinary statistical power compared to traditional clinical studies reporting outcomes from tertiary medical centers. The disadvantages stem from the inability to define critical variables such as clinical stage and grade precisely and the absence of associated data such as serum PSA levels. Possible confounding can be introduced by systematic biases that distort the findings of large administrative databases. Examples include changes in coding schemes with time that are used to classify cases by clinical stage or histology criteria. The Surveillance, Epidemiology and End Results program of the National Cancer Institute, for example, classifies tumor histology into the 4 categories of well differentiated, moderately differentiated, poorly differentiated and anaplastic. This coding scheme can lead to misclassifications when tumor pathology is reported according to the Gleason classification system.

Outcome metrics. Physicians reporting results from research studies, such as randomized trials, population based analyses and case series, can select from several potential metrics when reporting outcomes. Historically, overall survival has been the primary outcome metric used to evaluate treatment outcomes in oncology. This metric has the distinct advantage of being quite precise. The date of death of research subjects is rarely open to interpretation, and this

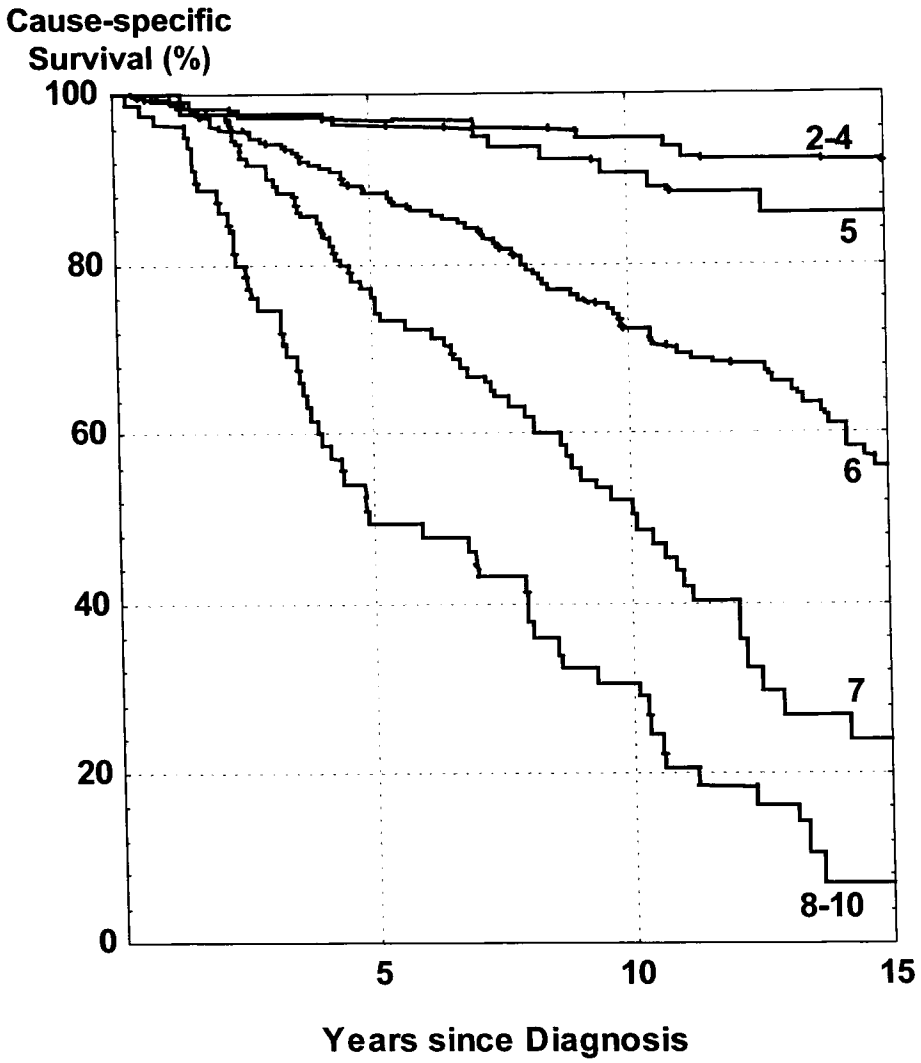

FIG. 3. Cause specific survival of 767 men diagnosed with localized prostate cancer and treated conservatively stratified by Gleason score.

metric allows patients to estimate probability of survival when choosing among different treatment alternatives.

Unfortunately there are 3 major disadvantages to using survival as an outcome metric. 1) Long-term survival does not separate the impact of prostate cancer from that of competing medical problems. As a result, patients and clinicians have difficulty separating the impact of treatment from the impact of co-morbidities. 2) Since prostate cancer is a chronic disease extending for many years death is relatively rare when followup is less than 5 to 10 years. Because modern screening efforts identify patients early in the course of disease, survival is an inadequate measure for a short time horizon. Researchers need large numbers of patients to achieve sufficient statistical power to detect survival differences. 3) Survival is a relatively crude metric and does not incorporate other significant outcomes such as quality of life. Consequently, researchers frequently turn to alternative metrics such as PSA progression and statistical techniques such as cause specific survival.

Cause specific survival. An analytic method frequently used by researchers is cause specific survival. This approach is designed to eliminate the impact of competing medical hazards and to focus only on the impact of treatment on prostate cancer mortality. The technique provides researchers with better estimates of treatment efficacy and is frequently used when there is no comparison population. Unfortunately it is not a substitute for an appropriate control. By effectively eliminating deaths from competing medical hazards, researchers inflate estimates of the potential bad outcomes following treatment especially among older patients. As an example, figure 3 shows the cause specific survival of the cohort of 767 men diagnosed with localized prostate cancer in Connecticut from 1971 to 1984. These data are stratified by Gleason grade and show relatively good outcomes for men with Gleason 2 to 5 disease and progressively poorer outcomes for men with Gleason 6 to 10 disease. Patients viewing these data might assume that their 15-year survival with a low grade prostate cancer is greater than 80%.

Figure 4 displays these same data using a competing risk analysis. The cumulative mortality from prostate cancer for the entire cohort is presented as a dark band, cumulative mortality from competing disease hazards is presented as a lighter band and the percentage of men still alive is shown in the white band. The corresponding cause specific survival curve is superimposed for comparison. For men with relatively lethal tumors (Gleason scores 7 to 10) the results of a cause specific analysis and competing risk analysis are com-

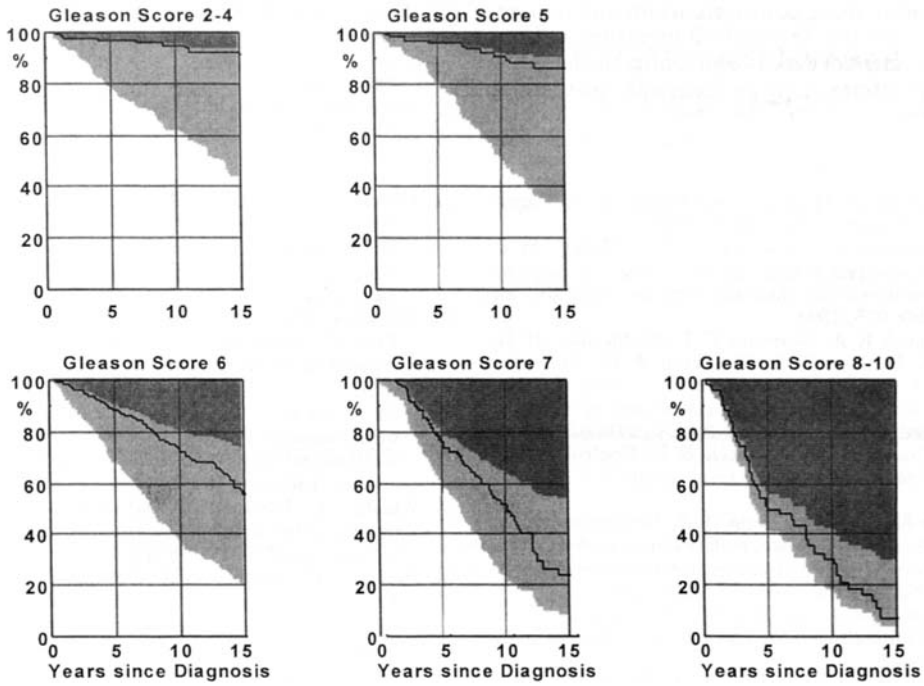

FIG. 4. Competing risk survival compared with cause specific survival for 767 men diagnosed with localized prostate cancer and treated conservatively stratified by Gleason score. Cumulative mortality from prostate cancer (dark area), cumulative mortality from other causes (lighter area), percentage of men surviving (white area) and cause specific survival (staircase solid line) are shown. Note that because competing mortality is substantial, complement of cause specific survival (portion above solid line) overestimates percentage of men who actually die of prostate cancer.

parable, especially during the first 5 years after diagnosis. For men with lower grade disease competing medical hazards pose a significantly greater threat than prostate cancer. Disease specific survival analysis overstates the percentage of patients who actually die of disease, a concept known as the case-fatality rate. Readers can visualize this effect by referring to figure 4. The single black line documenting the disease specific survival analysis is significantly below the break between the black and gray areas, which documents the actual percentage of men who died of prostate cancer. This distortion increases with each year of followup and is more pronounced in older men.

Cause specific survival also raises issues concerning how cause of death is determined. It is not easy to determine whether a death resulted from prostate cancer or from some competing medical hazard. Several researchers have studied this issue from a population based perspective. In an extensive comparison of death certificates and hospital records in patients with a definite diagnosis of cancer and for whom cancer was noted on the death certificate, prostate cancer was overreported on the death certificate in 4% and underreported in approximately 5% of patients.[18] In another study of death certificates for patients with cancer who were hospitalized within 1 month of death the underreporting of cancer rate on the death certificate was 5 to 10%. For studies involving large numbers of patients, cause of death determination appears to have a relatively small, random error, which may not be true with smaller case series, especially those that do not rely on information supplied by a death certificate. The overall effect for all studies is to add noise to the data, making it more difficult to detect outcome differences among patients treated differently.

PSA progression. Other outcome metrics frequently used by researchers and clinicians include PSA progression, extracapsular extension and quality adjusted life years. All of these metrics have potential advantages and disadvantages. PSA progression often indicates evidence of residual disease but is dependent on the frequency of PSA testing, the assay used and the definition of what constitutes significant elevation. The primary advantage of this metric is its ability to identify tumor recurrence long before disease becomes clinically evident. Whether this metric is a good proxy for long-term survival remains to be determined. Similarly, extracapsular extension is a useful short-term metric that potentially identifies patients in whom radical surgery is destined to fail. Unfortunately this metric cannot be used for nonsurgical therapies.

Health related quality of life. Measurements of patient health related quality of life and specific treatment related complications also raise several issues. How incontinence and impotence are measured can yield potentially different outcome estimates. Patient self-reports are subject to recall biases and the frequent desire to minimize symptoms. Subjective interpretations by surgeons yield different outcomes compared with results derived from well designed, validated survey instruments. The Short Form-36 developed by Ware and Sherbourne at the Rand Corporation[19] and the disease specific measures of bowel, bladder and sexual dysfunction developed by Litwin et al[20] are excellent examples of tools that have recently been used to standardize measurement of clinical outcomes associated with the treatment of prostate cancer.

SUMMARY

All research reports contain biases. Some are obvious, some are more subtle and all conspire to confound patient and researcher understanding of the natural history of prostate cancer and the efficacy of competing medical therapies. When reviewing research reports, readers need to be sensitive to the study design, the biases that impact the construct of the study population, and the research variables and statistical techniques used to quantify the impact of intervention. All of these issues help frame the data and the analysis.

By remaining vigilant to these issues, clinicians and patients can gain greater insights into the medical literature and can, in turn, make independent decisions concerning the potential impact of treatment intervention on men with newly diagnosed prostate cancer.

REFERENCES

1. Landis, S. H., Murray, T., Bolden S. and Wingo, P. A.: Cancer statistics, 1998. Cancer J. Clin., **48:** 6, 1998.
2. Albertsen, P. C., Hanley, J. A., Gleason, D. F. and Barry, M. J.: A competing risk analysis of men age 55–75 years at diagnosis managed conservatively for clinically localized prostate cancer. J.A.M.A., **280:** 975, 1998.
3. Gerber, G. S., Thisted, R. A., Scardino, P. T., Frohmuller, H. G., Schroeder, F. H., Paulson, D. F., Middleton, A. W., Rukstalis, D. B., Smith, J. A., Schellhammer, P. F., Ohori, M. and Chodak, G. W.: Results of radical prostatectomy in men with clinically localized prostate cancer. J.A.M.A., **276:** 615, 1996.
4. Bagshaw, M. A., Cox, R. S. and Hancock, S. L.: Control of prostate cancer with radiotherapy: long-term results. J. Urol., **152:** 1781, 1994.
5. Helgesen, F., Holmberg, L., Johansson, J. E., Bergstrom, R. and Adami, H. O.: Trends in prostate cancer survival in Sweden, 1960 through 1988: evidence of increasing diagnosis of nonlethal tumors. J. Natl. Cancer Inst., **88:** 1216, 1996.
6. Merrill, R. M., Potosky, A. L. and Feuer, E. J.: Changing trends in U.S. prostate cancer incidence rates. J. Natl. Cancer Inst., **88:** 1683, 1996.
7. Barry, M. J., Albertsen, P. C., Bagshaw, M. A., Blute, M. L., Cox, R., Middleton, R. G., Gleason, D. F., Zincke, H., Bergstrahl, E. J. and Jacobsen, S. J.: Outcomes for men with clinically nonmetastatic prostate cancer managed with radical prostatectomy, external beam radiotherapy, or expectant management: a retrospective, intention-to-treat analysis with standardized assessments of state, grade and comorbidity. Submitted for publication, 1998.
8. Feinstein, A. R., Sosin, D. M. and Wells, C. K.: The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. New Engl. J. Med., **312:** 1604, 1985.
9. Aus, G., Hugosson, J. and Norlen, L.: Long-term survival and mortality in prostate cancer treated with noncurative intent. J. Urol., **154:** 460, 1995.
10. Abrahamsson, P. A., Adami, H. O., Taube, A., Kim, K., Zelen, M. and Kulldorff, M.: Re: Long-term survival and mortality in prostate cancer treated with noncurative intent. J. Urol., **155:** 296, 1996.
11. Johansson, J. E., Holmberg, L., Johansson, S., Bergstrom, R. and Adami, H. O.: Fifteen-year survival in prostate cancer. A prospective, population-based study in Sweden. J.A.M.A., **277:** 467, 1977.
12. Chodak, G. W., Thisted, R. A., Gerber, G. S., Johansson, J. E., Adolfsson, J., Jones, G. W., Chisholm, G., Moskovitz, B., Livne, P. and Warner, J.: Results of conservative management of clinically localized prostate cancer. New Engl. J. Med., **330:** 242, 1994.
13. Albertsen, P. C., Fryback, D. G., Storer, B. E., Kolon, T. F. and Fine, J.: Long-term survival among men with conservatively treated localized prostate cancer. J.A.M.A., **274:** 626, 1995.
14. Smith, R. C., Partin, A. W., Epstein, J. I., Brendler, C. B.: Extended followup of the influence of wide excision of the neurovascular bundles on prognosis in men with clinically localized prostate cancer and extensive capsular perforation. J. Urol. **156:** 454, 1996.
15. Walsh, P. C., Partin, A. W. and Epstein, J. I.: Cancer control and quality of life following anatomical radical retropubic prostatectomy: results at 10 years. J. Urol., **152:** 1831, 1994.
16. Partin, A. W., Kattan, M. W., Subong, E. N., Walsh, P. C., Wojno, K. J., Oesterling, J. E., Scardino, P. T. and Pearson, J. D.: Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update. J.A.M.A., **277:** 1445, 1997.
17. Lu-Yao, G. L., Potosky, A. L., Albertsen, P. C., Wasson, J. H., Barry, M. J. and Wennberg, J. E.: Follow up prostate cancer treatments after radical prostatectomy—a population based study. J. Natl. Cancer Inst. **88:** 166, 1996.
18. Percy, C., Stanek, E. and Gloeckler, L.: Accuracy of cancer death certificates and its effects on cancer mortality statistics. Amer. J. Public Health **71:** 242, 1981.
19. Ware, J. E. and Sherbourne, C. D.: The MOS 36-item short form health survey (SF-36): I. Conceptual framework and item selection. Med. Care **30:** 473, 1992.
20. Litwin, M. S., Hays, R. D., Fink, A., Ganz, P. A., Leake, B., Leach, G. E. and Brook, R. H.: Quality of life outcomes in men treated for localized prostate cancer. J.A.M.A., **273:** 129, 1995.

QUESTION AND RESPONSE

Dr. C. A. Olsson. Does disease specific survival overstate true survival? Isn't it really the other way around? As an example, a patient who dies of azotemia as a consequence of ureteral obstruction or a patient who has thromboembolic complications of estrogen treatment is really a prostate cancer related death?

Dr. P. C. Albertsen. No, disease specific survival analysis will overstate true survival. This artifact is small during short periods but becomes more pronounced with longer followup. It appears because the denominator, the number of patients at risk, decreases with a disease specific survival analysis because patients dying of other competing hazards are censored. The denominator remains constant in a standard survival analysis.

The issue raised concerns cause of death attribution, which is also a significant problem with disease specific survival analysis. A standard survival analysis simply requires the investigator to determine date of death. A disease specific analysis demands that the investigator attribute the cause of death to a specific reason such as prostate cancer or some other competing hazard, which is more difficult than it would initially appear as indicated by your question. Different methods of attributing cause of death can have a significant impact on reported outcomes.