EDITORIAL

Counterpoint

Caveats for Modeling Disease Free Survival after Radical Prostatectomy

John T. Wei, M.D.
James E. Montie, M.D.

Section of Urology, The University of Michigan, Ann Arbor, Michigan.

It has been more than a decade since Patrick Walsh described the "anatomic" radical prostatectomy for the treatment of localized prostate carcinoma.[1] Since then, thousands of men with localized prostate carcinoma have been treated and we have come to recognize the tremendous heterogeneity in disease free survival. Early attempts at prognostication after surgery utilized the Whitmore A-D staging system, which has since undergone a number of revisions to become the American Joint Committee on Cancer/International Union Against Cancer TNM staging system. However, the evolution of staging classifications takes years to occur and is based primarily on existing models that do not account adequately for the myriad of other prognostic indicators such as prostate specific antigen (PSA) and tumor specific pathologic markers.

More recently, investigators have incorporated staging into multivariate models.[2-4] These models have elaborated on our understanding of the biology of prostate carcinoma by identifying the more important prognostic elements while simultaneously controlling for confounding variables. The Cox proportional hazards model specifically allows for the prediction of risk according to predetermined variables; however, the growing number of potential predictor variables makes the use of these traditional statistical methods cumbersome. To simplify the clinical application of these models, investigators often will construct classifying tables or nomograms. A major criticism of classifying modeled outcomes has been the arbitrary assignment of risk levels or cutoff values. In this sense, existing modeling techniques for disease free survival after radical prostatectomy have not been very satisfying.

The article by Banerjee et al. in the current issue of *Cancer* addresses this issue by applying a nonparametric modeling technique known generally as recursive partitioning.[5] This modeling technique requires a number of steps including 1) selection of the splits at each node, 2) determination of when to stop splitting, and 3) pruning the tree down to the right size. As a result, the models are only as good as the choices made by the investigator at each of these steps. In contrast to linear regression techniques, which require a predetermined linear model be fitted to the data with the assignment of various weights to each predictor variable, recursive partitioning is less restrictive and permits the data to dictate the structure of the model.

See reply to counterpoint and referenced original article on pages 234–5 and 404–11, this issue.

Dr. Montie is employed as a clinical investigator for Zeneca, Inc., and Centeon, Inc.

Address for reprints: John T. Wei, M.D., Section of Urology, The University of Michigan, 2916 Taubman Center, 1500 East Medical Center Drive, Ann Arbor, MI 48109-0330.

Received January 6, 2000; accepted January 13, 2000.

This allows for the identification of subgroups of patients who are more homogeneous with respect to their risk of disease recurrence and makes readily apparent the "natural cutoff values" in the data. Although recursive partitioning techniques are flexible and the results are more intuitive, they are limited by the requirement that data be classified into a small number of discrete homogeneous groups using cutoff values.

In the analyses by Banerjee et al.,[5] the use of recursive partitioning identified a series of subgroups based on clinicopathologic factors including pathology stage, Gleason score, preoperative PSA, and age. The investigators then examined the median survival for each of these subgroups and combined groups based on their similarity in outcomes. This resulted in the creation of three risk groups that appear to predict disease free survival. Although these risk groups lend themselves to applications such as clinical trials, in which strict inclusion and exclusion criteria are necessary, several caveats should be considered before applying these findings more broadly to clinical practice.

The first caveat relates to the generalizability of this nonrandomly selected study population. As with all statistical models, validation of these results should be performed based on independent datasets because sampling bias often can lead to spurious results. For example, the study population of 485 men represents only 53% of all the men who underwent radical prostatectomy during the study period. Furthermore, the racial composition of the study sample is not typical of the majority of radical prostatectomy series, with 41% of the men identified as African-American. Caution is warranted in generalizing these results to other populations without further study.

Another caveat relates to the use of cutoff values. The findings from this study would suggest that an age cutoff value of 65 years may be an important predictor of disease free survival in those men with extension of the carcinoma to the seminal vesicles or lymph nodes and a PSA level $\leq$ 24.1 ng/mL. Although this may have resulted from sampling bias, the difference in median survival is impressive (13.2 months for the low risk group vs. 53 months for the high risk group). To our knowledge no previous study has found age to have such a remarkable effect on disease free survival. This brings to the fore the limitations of using cutoff values when data unquestionably span a continuous spectrum such as in patient age and serum PSA levels. It is quite likely that a 65-year-old man and a 66-year-old man with carcinoma in the seminal vesicles or lymph nodes and a PSA level $\leq$ 24.1 ng/mL have very similar risk profiles but the use of this model would place them into different risk strata. To our knowledge there currently is no known biologic explanation to support this finding and it certainly is worthy of further investigation.

It is appreciated that there are situations (such as clinical trials) in which artificial cutoff values are necessary and recursive partitioning may be a logical approach; however, without independent validation, the accuracy of this recursive partitioning model cannot be determined (i.e., the correct classification of an individual into a risk subgroup). In other disease processes in which the accuracy of recursive partitioning has been compared with linear techniques, recursive partitioning has not been shown to be superior.[6] Finally, the clinician should be aware that the majority of models for predicting outcomes after radical prostatectomy are reliable for large samples of men, but have faired less well in determining the disease free survival of individual patients.

These caveats notwithstanding, the study by Banerjee et al. represents an important contribution to the body of literature regarding prognostication of outcomes after radical prostatectomy with the promotion of recursive partitioning. In our continuing quest for an understanding of the pathophysiology of prostate carcinoma, it is essential that we not lose the forest for the trees by confining ourselves to only a few modeling techniques but rather consider the limitations of each for modeling clinical data.

REFERENCES

1. Walsh PC. Radical retropubic prostatectomy with reduced morbidity: an anatomic approach. *Monogr Natl Cancer Inst* 1988;7:133–7.
2. Polascik TJ, Pearson JD, Partin AW. Multivariate models as predictors of pathological stage using Gleason score, clinical stage, and serum prostate-specific antigen. *Semin Urol Oncol* 1998;16:160–71.
3. Pound CR, Partin AW, Epstein JI, Walsh PC. Prostate-specific antigen after anatomic radical retropubic prostatectomy. Patterns of recurrence and cancer control. *Urol Clin North Am* 1997;24:395–406.
4. Partin AW, Kattan MW, Subong EN, Walsh PC, Wojno KJ, Oesterling JE, et al. Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update. *JAMA* 1997;277:1445–51.
5. Banerjee M, Biswas D, Sakr W, Wood DP. Recursive partitioning for the prognostic grouping of patients with clinically localized prostate carcinoma. *Cancer* 2000;89:404–11.
6. Hadorn DC, Draper D, Rogers WH, Keeler EB, Brook RH. Cross-validation performance of mortality prediction models. *Stat Med* 1992;11:475–89.