Systems Pathology and Predicting Outcome After Radical Prostatectomy

Eric A. Klein, Andrew J. Stephenson, Derek Raghavan, and Robert Dreicer, Glickman Urologic and Kidney Institute and Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH

A major challenge in clinical oncology is to accurately answer every newly diagnosed patient’s plaintive question, “What are my chances, Doc?” Despite recent advances in understanding clinical, pathologic, and genomic-based prognostic factors, our responses remain imprecise, or—as statisticians might observe—our estimates are surrounded by wide confidence intervals. Although physicians are accustomed to practicing clinical medicine with this degree of uncertainty and have adapted to working within its bounds, most patients are not so accustomed and desire a higher degree of certainty. More accurate estimates of response rates, clinical benefit, and the potential for cure should help patients deal with their diagnosis realistically and would provide better insight into the selection of the most appropriate primary therapy and adjunctive therapies.

Attempts at predicting individual outcomes after therapy for localized prostate cancer have a long history, dating to early efforts to correlate preoperative grade, stage and prostate-specific antigen (PSA) with biochemical failure. This was followed by the establishment of formal definitions of low-, intermediate-, and high-risk groups using these parameters. These models were based on easily assessed and widely available clinical data and estimated prognosis and the potential need for more aggressive therapy for patients diagnosed early in the PSA era. Contemporaneously, these attempts were systematized and linked to pathologic outcomes by the development of tables predicting the likelihood of organ-confined disease. These suggested that patients presenting with features that yielded a small chance of cure by surgery might best be treated using another modality. Although these studies were validated informally by daily clinical use and formally by statistical comparisons in additional populations, their predictions were imprecise because of wide variations in outcomes for individual patients within each defined risk category. The next advance in outcome prediction was the development of nomograms, based on Cox regression modeling that combined known predictors and linked to pathologic outcomes by the development of tables predicting the likelihood of organ-confined disease. These suggested that patients presenting with features that yielded a small chance of cure by surgery might best be treated using another modality. Although these studies were validated informally by daily clinical use and formally by statistical comparisons in additional populations, their predictions were imprecise because of wide variations in outcomes for individual patients within each defined risk category. The next advance in outcome prediction was the development of nomograms, based on Cox regression modeling that combined known predictors and linked to pathologic outcomes by the development of tables predicting the likelihood of organ-confined disease. These suggested that patients presenting with features that yielded a small chance of cure by surgery might best be treated using another modality. Although these studies were validated informally by daily clinical use and formally by statistical comparisons in additional populations, their predictions were imprecise because of wide variations in outcomes for individual patients within each defined risk category. The next advance in outcome prediction was the development of nomograms, based on Cox regression modeling that combined known predictors and linked to pathologic outcomes by the development of tables predicting the likelihood of organ-confined disease. These suggested that patients presenting with features that yielded a small chance of cure by surgery might best be treated using another modality.

Currently there are more than 65 published, externally validated prostate cancer nomograms and other tools that use standard clinical parameters such age, clinical or pathologic stage, grade, percent of cancer on biopsy cores, and PSA or its derivatives to predict various clinical and pathologic outcomes. Although these have achieved clinical currency, it is clear that additional improvements in accuracy are needed, as few nomograms based on these parameters exceed a predictive accuracy of 80%. For example, the two most widely used preoperative nomograms that predict for PSA recurrence at 5 or 10 years after radical prostatectomy achieve accuracies of only 75% to 79%, despite more than 3,500 patients in the latter nomogram in the test and validation cohorts combined. Next-generation predictive tools incorporate additional biologic variables that may add more objective, tumor-specific information than can be achieved by clinical information and white-light histology alone. A recent study of a preoperative nomogram predicting likelihood of cure after radical prostatectomy demonstrated an improved accuracy of 88%, compared with 71% for clinical variables alone, by the addition of serum transforming growth factor β1 and interleukin-6-SR levels to the nomogram, validating a prior report. For breast cancer, the biologic approach has already yielded three commercially licensed products that rely on gene expression profiles (MammaPrint, Agenda BV, Amsterdam, the Netherlands; Oncotype DX, Genomic Health, Redwood City, CA; and H/I AvariaDX, Carlsbad, CA).

In this issue of Journal of Clinical Oncology, Donovan et al take a different approach by adding cellular and biologic features to standard clinical and histologic parameters, a method called systems pathology. Using image analysis and quantitative immunofluorescence of hematoxylin and eosin–stained tissue microarrays and outcome data on 881 patients treated with radical prostatectomy, they identified 27 features reflecting color, texture, and the area of epithelial and stromal cells. Twelve of these were associated with a meaningful composite end point of clinical failure (defined as radiographic or pathologic evidence of metastasis, an increasing PSA in a castrate state, or death attributed to prostate cancer). Using support vector regression for censored data, a final model containing six variables (androgen receptor levels, prostatectomy Gleason grade, lymph node involvement, and three quantitative characteristics) predicting clinical failure within 5 years of surgery was generated. In the test set, the model had a concordance index of 0.92, sensitivity of 90%, and specificity of 91%, whereas in the validation set, the concordance index was 0.84, with a sensitivity of 84% and specificity of 85%. When compared with the
Excitement over this potential improvement in accuracy must be tempered by limitations in its clinical utility. In an additional analysis, the authors compared the concordance index for the systems pathology–derived model with that of a 10-variable model that contained only the usual clinical parameters (stage, PSA, biopsy Gleason score and sum, prostatectomy Gleason score and sum, and pathologic extent in the prostatectomy specimen) generated by Cox regression. In this analysis, the concordance index for the 10-variable clinical model was 0.80 and improved only slightly to 0.83 for the systems pathology approach. The corresponding hazard ratios for clinical failure were 6.37 and 9.11, respectively. Although the difference in concordance indices was statistically significant, the question is whether there is sufficient clinical relevance to justify the extra effort, expense, and expertise needed for the systems approach. The clinical utility of defining high risk for failure after radical prostatectomy is to decide whether patients require closer follow-up than average or whether adjuvant radiotherapy, hormone therapy, or chemotherapy would be of benefit. In contemporary practice, a patient with a hazard ratio of 6.37 generated by the model using easily derived, routinely reported clinical and pathologic parameters is just as likely to be a candidate for closer monitoring or adjuvant therapy than one with a hazard ratio of 9.11 generated by the systems approach. Either patient is at very high risk of recurrence, and in current practice, a clinician cannot derive a follow-up scheme or therapeutic approach that could be applied differentially between them. Even when restricting the analysis to the subset of patients with intermediate risk disease, the hazard ratios are 2.7 and 5, respectively, perhaps approaching the threshold, but not yet sufficiently different to change clinical practice for an individual patient. Another limitation of the study is the use of a validation set derived from the same institution; the authors do not provide details on how members of each set were chosen. The study would be strengthened if a similar high concordance rate were to be found with a large external validation sample, thus allowing for interinstitutional differences in patient populations and variability of measurement and of these specific parameters in the algorithm.

This work at once highlights the utility of standard clinical variables in predicting prognosis for newly diagnosed patients with prostate cancer as well as the potential for the addition of biologic variables to improve precision. It is clear that the standard variables contain a great deal of predictive power (ie, they already capture a lot of the biologic potential of the tumor, especially when postprostatectomy pathologic variables are included) and that new markers of disease progression should be evaluated based on what they add to standard variables, rather than to replace them. The systems pathology approach as applied to prostate cancer may ultimately improve our ability to accurately prognosticate for individual patients and to make meaningful clinical recommendations based on multiple externally validated models. The continuing development of such models is a positive reflection of the dynamic evolution of medical science, reflecting both innovation in development and use of new technology. Until even better models emerge, we must respond to patients who ask “What are my chances, Doc?” by quoting the quantum physicist Niels Bohr, who observed that “Prediction is very difficult, especially if it’s about the future.”

REFERENCES