REVIEW ARTICLE

How To Build and Interpret a Nomogram for Cancer Prognosis

Alexia Iasonos, Deborah Schrag, Ganesh V. Raj, and Katherine S. Panageas

From the Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY; and Department of Urology, University of Texas Southwestern Medical Center at Dallas, Dallas, TX.

Corresponding author: Alexia Iasonos, PhD, Assistant Attending Biostatistician, Epidemiology-Biostatistics, Memorial Sloan-Kettering Cancer Center, 307 E 63rd St, 3rd Floor, New York, NY 10021; e-mail: iasonosa@mskcc.org.

ABSTRACT

Nomograms are widely used for cancer prognosis, primarily because of their ability to reduce statistical predictive models into a single numerical estimate of the probability of an event, such as death or recurrence, that is tailored to the profile of an individual patient. User-friendly graphical interfaces for generating these estimates facilitate the use of nomograms during clinical encounters to inform clinical decision making. However, the statistical underpinnings of these models require careful scrutiny, and the degree of uncertainty surrounding the point estimates requires attention. This guide provides a nonstatistical audience with a methodological approach for building, interpreting, and using nomograms to estimate cancer prognosis or other health outcomes.

J Clin Oncol 26:1364-1370. © 2008 by American Society of Clinical Oncology

INTRODUCTION

Oncologists and patients alike desire reliable prognostic information tailored to the individual patient. In recent years, statistical prediction models have been developed across the majority of cancer types.[1-5] One such predictive tool is the nomogram, which creates a simple graphical representation of a statistical predictive model that generates a numerical probability of a clinical event. For many cancers, nomograms compare favorably to the traditional TNM staging systems[6] and thus have been proposed as an alternative or even as a new standard.[7-11] The ability of nomograms to generate individualized predictions enables their use in the identification and stratification of patients for participation in clinical trials. The combination of user-friendly interfaces and widespread availability via the web have contributed to their popularity among oncologists and patients themselves.[12-18]

We intend to clarify the process of nomogram development so that clinicians gain understanding of the statistical underpinnings. In addition, we present guidelines for reporting standards to facilitate proper use of published nomograms. In this article, we outline a series of discrete steps involved in nomogram construction and recommend a systematic approach for their evaluation. For purposes of illustration of these steps, we use the data set for a published nomogram[19] designed to predict the probability of a malignant renal clear-cell carcinoma for patients undergoing surgery for a renal mass (Fig 1). The initial steps in nomogram development include definition of the patient population and outcome, identification of important covariates, specification of the statistical model, and validation of its performance (Table 1).

STEP 1. IDENTIFY THE PATIENT POPULATION

The first step in nomogram construction is to identify the source population. The eligibility criteria for patients to be included should be decided a priori. The source population may be derived from a single institution, multiple centers, or a population-based cohort. Models derived from multicenter or population-based cohorts are more likely to be generalizable; however, they may be hampered by a lack of consistent availability of detailed data elements, such as specific tumor markers that may improve prognostic accuracy. The inherent tradeoffs between generalizability and detailed data elements require careful consideration from the outset. Importantly, consideration must be given to whether the source population resembles the population for whom the nomogram estimates will be applied. Sample questions in evaluating the source population may include: Is the source population unique? Does it represent the entirety of the age spectrum? Are treatment patterns representative? These factors must be considered before nomogram construction, because the choices may have important implications for how useful the model will be when applied to other populations.

In our case study, the derivation cohort included patients who underwent surgery for a renal
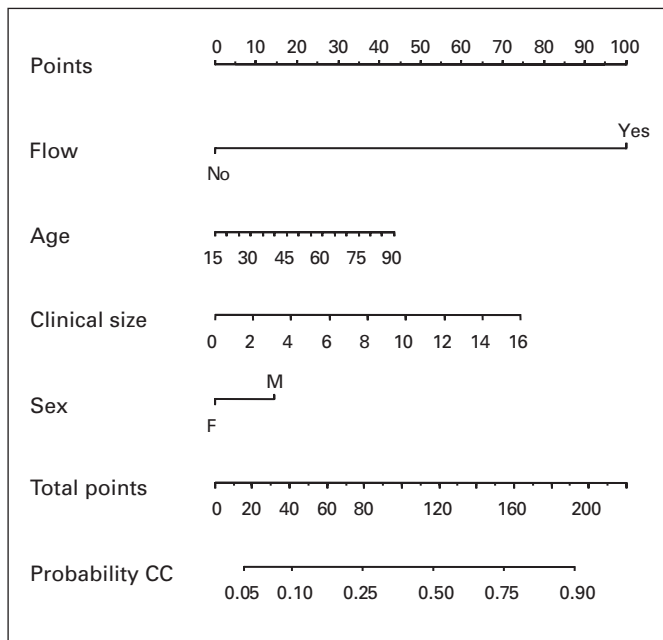
Fig 1. The case study nomogram as published in the *Journal of Urology*[19] predicting the probability of clear-cell (CC) histology in renal cell cancer.

mass at a single specialty center and excluded patients known preoperatively to have metastatic disease.[19] Providing details about the derivation of the cohort's characteristics enables nomogram users to determine the applicability of the resulting estimates to their own patient groups.

STEP 2. DEFINE THE OUTCOME

Construction of a nomogram requires precise definition of the primary outcome. The outcome is typically an event, such as diagnosis of a malignancy, or time to event, such as time to recurrence or death. Nomograms are used to predict the probability of a specific event, such as a positive biopsy, or the probability of recurrence, or survival (using fixed-time anchors, such as 3-year probability of recurrence). In our case study, the goal of the nomogram was to predict the probability of a specific malignant histology, conventional clear-cell

Table 1. Initial Steps in Building a Nomogram

1. Define the patient population: Data source
   Single institution
   Multicenter
   Population-based
2. Specify the outcome of interest: What is the nomogram predicting?
3. Specify all predictors under consideration
4. Construct the nomogram: Model selection
   Select covariates
   Assess confounders/interactions
   Consider model fit and the need for transformations
5. Validation: Data source
   Internal data
   External data

carcinoma, based on preoperative clinical and radiographic parameters before nephrectomy.[19]

STEP 3. IDENTIFY POTENTIAL COVARIATES

Before building a nomogram, it is important to identify the spectrum of prognostic factors that may predict the outcome of interest (Table 1). Prognostic parameters must be selected a priori, based on either prior research or sound clinical reasoning, so that excluding variables because of missing data is eliminated and consistent data collection is maintained.[20] In our case study, potential covariates included age, sex, presence of symptoms at diagnosis, vascular flow on color Doppler ultrasound, clinical tumor size, central location within the kidney, necrosis on imaging studies, and multifocality. In published nomograms, the range of variables considered is usually determined based on data availability and clinical evidence rather than on statistical significance.[10]

STEP 4. CONSTRUCT THE NOMOGRAM

(i) Select the Model

Nomograms may convey the results of a variety of statistical models. In our case study, the intention was to predict a binary outcome of a malignant clear-cell histology (presence/absence) using the above-mentioned variables. The underlying logistic model is given by the equation:

$$\text{Probability (Clear Cell)} = \frac{e^{(\beta_0 + \beta_{age}age + \beta_{gender}gender + \beta_{tumorsize}tumorsize + \beta_{flow}vascularflow + \ldots)}}{1 + e^{(\beta_0 + \beta_{age}age + \beta_{gender}gender + \beta_{tumorsize}tumorsize + \beta_{flow}vascularflow + \ldots)}}$$

If the outcome is time to event (censored outcome), such as overall survival, then the Cox proportional hazard model is often selected, as it models the hazard, which is an instantaneous failure rate as a function of time. If our intention was to determine whether the same factors predict if a patient will experience the event (death) at 3 years, then the underlying model is given by: Hazard (experiencing the event at 3 years) =

$$\text{baseline Hazard (time} = 3))^* \, e^{(\beta_{age}age + \beta_{gender}gender + \beta_{tumorsize}tumorsize + \beta_{flow}vascularflow + \ldots)},$$

where the baseline hazard corresponds to the hazard of experiencing the event (dying) when all covariates are zero.

The right-hand side of the above equations specifies the underlying function of the model. The left-hand side of the equations is the predicted probability that is presented in a nomogram and communicated to patients. Beta coefficients must be estimated for each covariate and converted to odds ratios (hazard ratios for time to event outcomes) as a measure of effect, as in any statistical report. To obtain the predicted probability of the event in question, the above equation is calculated using a patient's individual characteristics and the model-derived beta coefficients.

(ii) Select the Predictors

Investigators start with the covariates that they anticipate, a priori, may have an effect on the outcome. Statistical testing can identify

whether the data support these initial beliefs. However, it is important to consider both clinical and statistical significance when selecting covariates for inclusion. The statistical significance (typically expressed as *P* value) depends on the magnitude of the effect, the sample size, and the spread in the data (variance). As a result, large studies can detect small differences, whereas small retrospective studies can fail to detect important clinical findings. Many nomograms have been developed using retrospective or single-institution databases, and as such, may not have an adequate sample size to identify a significant effect estimate. For this reason, sample size considerations are important. On the basis of Harrell's guidelines,[20] when the outcome is binary, the minimum value of the frequencies of the two response levels should be greater than 10 times the number of predictors. In our case, there were 169 clear-cell and 130 non–clear-cell renal tumors. Thus the limiting sample size is 130, and based on Harrell's guidelines, no more than 13 predictors can be accommodated. When the outcome is overall survival, the number of deaths should be greater than 10 times the number of predictors, so that the expected error in the predicted probabilities from the Cox model is less than 10%.[20]

Other factors may also affect statistical significance, can change the magnitude of the effect on outcome, or provide false effect estimates. From a statistical perspective, these underlying complexities are essential, and we summarize a few examples below.

(a) Confounders/multicollinearity. Confounders are factors that are related to the outcome and to other independent predictors in the model, but note that there is no causal pathway in this relationship. In the presence of a confounder, the statistical significance of a covariate may be affected. In our illustrative example, a predictor such as evidence of necrosis on imaging may be related to the outcome (malignancy), as well as another predictor (multifocality). A model developed with these predictors may show that the presence of necrosis on imaging is a prominent factor in predicting malignant tumors in renal masses, but other predictors that relate to necrosis, such as multifocality, might not seem to be significant. In this example, the effect of multifocality is confounded by necrosis. Relationships among the predictors, also known as multicollinearity, can influence the beta coefficients in the model, resulting in spurious associations and possibly unreliable effect estimates. Unexpected (counterintuitive) magnitudes or signs of parameter estimates, and opposite results from univariate to multivariate regression are some ways to identify multicollinearity.[21] A common approach is to assess the correlation of these confounders and decide whether some covariates can be omitted, given the presence of others. However, omitting one or more predictors on the basis that they are redundant is not always a viable solution. Despite the fact that a mathematical correlation exists, omitting predictors may be erroneous, because each provides important and specific information. A statistician may assess correlations, variance inflation factors and eigenvalues, and use a number of statistical tools to try to resolve this problem. Ridge regression, principal component analysis, and other methods[20,22] for variable selection, which do not solely depend on the criterion of *P* values less than .05, have been suggested to handle multicollinearity,[23,24] but these methods have not yet been used in the context of nomograms and perhaps should be considered.

(b) Interactions. Assessment of interaction effects should be considered when building a model. An interaction is a synergistic effect (ie, the way two or more factors act together). For example, each pair-wise combination of age ($\leq 65\ v > 65$ years) and sex (male *v* female) may have its own effect on the outcome. That is, the effects of age and sex alone may not be sufficient to explain the outcome, while the interaction term would provide different predictions for younger versus older male patients and younger versus older female patients. In a hypothetical example of 200 patients, with 15% of patients aged $\leq$ 65% and 50% women, the subset of younger women consists of only 15 patients. Failing to show an effect of age or an age/sex interaction does not mean this effect is nonexistent, but rather is a result of small sample size. As a consequence, simultaneous effects on outcome are rarely examined, and this is one reason why some models, even with good sensitivity, lack specificity. We do not advocate that all possible pair-wise interactions be routinely assessed in model building, but rather that assessment of potential interactions be guided by clinical expertise. In the illustrative example, a trend toward an interaction of sex and vascular flow was not clinically meaningful and thus was omitted from the model. Moreover, 39% of the patients were female and only 21% (62 of 299 patients) of the entire cohort had the outcome of interest, which implied that this trend might not be reproducible in a larger cohort.

(c) Transformations. Models discussed thus far assume that the relationship between a given factor and the outcome follows a linear function (the logistic model assumes a linear relationship on the logit scale and the Cox model on the log hazard scale). Such assumptions may not always be appropriate. When a relationship is not linear, a transformation of the predictor may be required. For example, a nonlinear relationship can be a U-shaped relationship between hemoglobin levels and age. Transformations of the predictors can be incorporated in any multivariate model and thus are not unique advantages of nomograms. In the illustrative nomogram, the probability of conventional clear-cell histology increased linearly as clinical tumor size and age increased, thus a transformation was not needed. The need for transformations should be assessed graphically and justified in the process of nomogram building.

Transformations of the predictors or complex functions, such as splines, are frequently featured in models underlying nomograms.[1] There are times where a more complex function is chosen compared with a simple nonlinear transformation as a result of an overfit.[25] Overfitting of a model occurs when the model is trying to fit every detail present in a particular data set, even if this reduces down to modeling a small number of cases. The resulting model may be highly specific to the particular data set and therefore less generally useful. Thus an overfit occurs when the model performs well in the derivation data set, but is not generalizable to other data sets.

STEP 5. FINALIZE THE MODEL: VALIDATION

The goal of an individualized risk prediction model is to predict the outcome as accurately as possible. The ability of a model to separate patients with different outcomes is known as discrimination. How far the predictions are from the actual outcomes is referred to as calibration. Calibration is typically assessed by reviewing the plot of predicted probabilities from the nomogram versus the actual probabilities. A perfectly accurate nomogram prediction model would result in a plot where the observed and predicted probabilities for given groups fall along the 45-degree line. The distance between the pairs and the 45-degree line is a measure of the absolute error of the nomogram's prediction. The calibration plot for the renal cell nomogram with the CIs added is shown in Fig 2. An evaluation of both the error (how far
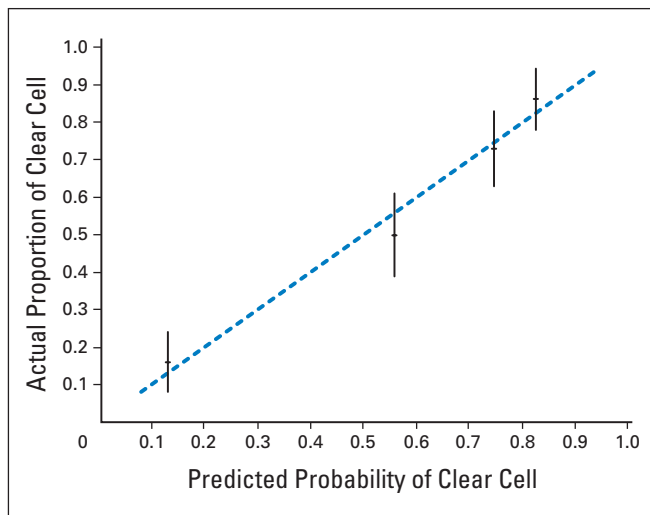
Fig 2. Calibration plot for the case study nomogram. Dashed line indicates ideal reference line where predicted probabilities would match the observed proportions. Dashes represent nomogram-predicted probabilities grouped for each of the four quartile groups, along with the respective CIs.

from the diagonal line the points fall) along with the width of the CI should be assessed when examining calibration plots. Note that the width of the CI depends on the number of patients included in each group, and it will be wider with smaller group sizes.

A nomogram's predictive accuracy (discrimination) is measured via a concordance index (c-index) which quantifies the level of concordance between predicted probabilities and the actual chance of having the event of interest. The c-index denotes the proportion of pairs, with the responder having a higher predicted probability of response than the nonresponder. A c-index with its respective CI provides a more comprehensive measure of discrimination. A CI for a c-index can be obtained either by bootstrap resampling or by the method proposed by Pencina and Agostino.[26] Concordance indices are always higher in the data set used to build the nomogram, compared with the concordance index of the same nomogram used with a new data set. As such, when finalizing the model, cross-validation is required to address model overfit. In this way, an estimate of how well the nomogram will perform when it is used in a new patient cohort is provided. Cross-validation methods include split-sample or bootstrap techniques, which use a separate sample to build the model and a test sample to test the model. Definitions of validation methods are given below.

(a) Cross-Validation

Data splitting and jackknifing are similar approaches for cross-validation, in that data are randomly split into groups. The jackknife approach splits the data into the same number of groups as there are observations, and is also called "leave-one-out cross-validation." Alternatively, they may be split into bigger groups, such as groups with 1/10 of the number of observations, which is called 10-fold cross-validation. One group is removed, and the model is built on the reduced sample set, which is considered fixed. This fixed model derived from the reduced data is used to predict the group of patients that was left out. Repeating this process, by leaving out each group once, provides predictions for all patients in the original cohort and hence a model performance index (concordance index).

To protect against the influence of the random splits, cross-validation is repeated a large number of times (eg, 200 times), and the average of the 200 indices is the bias-corrected index. For example, for the illustrative nomogram, the concordance index of the original cohort was 0.82, which reduced to 0.79 after the bias correction with 10-fold cross-validation.

(b) Bootstrap Validation

An alternate approach is bootstrap validation, in which random samples drawn with replacement from the original data set are the same size as the original cohort. A bootstrap sample for the illustrative nomogram would include 299 patients, but in this new sample, patient A could appear three times, whereas patient B could appear zero times. Although each patient has the same probability of being sampled, random chance can lead to this uneven outcome. In fact, each bootstrap sample will typically include approximately two thirds of the original observations at least once.[27] The same model derived from the original cohort is fit to the bootstrap sample. Repeating this process 200 (or more) times would produce 200 model performance indices. The performance index of the model built on the entire cohort is always higher than the average of these 200 indices. The difference of the two is an estimate of the overfit or optimism, and the average value of the 200 indices is considered the bias-corrected estimate of how well the model would perform in the future.[27]

(c) External Validation

Although cross-validation and bootstrapping are sample reuse methods that prevent against over-interpretation of current data, they cannot ensure external applicability. Whether a nomogram is generalizable to a new patient population is a far bigger concern than over-fitting, and it is a question that requires careful clinical judgment. Published nomograms have tried to assess external applicability by validating the results in an external patient population.[28] In the illustrative nomogram, when the model was evaluated externally on a prospective cohort, the concordance index was 0.76, indicating that measures of predictive accuracy are biased when measured on the same data set that was used to build the nomogram.

Researchers have outlined these limitations in the literature, with remarks such as, "The nomogram needs to be validated externally" or "Whether it can be universally applied is still to be determined." But can we explain to patients, in simple words, the nomogram limitations? If there is no race effect included in the model, and the model was built on an 85% white population (approximate race distribution in clinical trials) with minimal comorbidities and risk factors, it is not clear that outcomes for a nonwhite patient with certain comorbidities and risk factors will be accurately predicted from this specific nomogram. If the nomogram was built on patients who underwent surgery and had large tumor lesions, it will not perform as well on patients with small lesions, because they were underrepresented or absent in the original data. A probability will still be estimated by the nomogram, but this estimate will not be relevant. One must be cautious about extrapolating from regression models built on different populations.[21]

STEP 6. INTERPRET THE FINAL NOMOGRAM

The usefulness of a nomogram is that it maps the predicted probabilities into points on a scale from 0 to 100 in a user-friendly graphical

interface. The total points accumulated by the various covariates correspond to the predicted probability for a patient.[12,29] The point system works by ranking the effect estimates, regardless of statistical significance, and it is influenced by the presence of other covariates. Figure 3 shows a two-variable nomogram as an example. Assume we include two nonsignificant effects in the model, sex ($\beta = 0.12$, $P = .63$) and symptoms ($\beta = 0.51$, $P = .10$). In this example, symptoms has the highest effect, thus it is converted into 100 points. A patient with symptoms is assigned 100 points, whereas a patient without symptoms gets 0 points. Regardless of statistical significance, the effect with the highest beta (absolute value) will be assigned 100 points on the scale, and the remaining variables are assigned a smaller number of points proportional to their effect size. A male patient would be given 23.5 points, which is equal to the ratio of $\beta_{sex}/\beta_{symptoms}$ multiplied by 100. This represents the relative importance of the least significant variable compared with the most significant variable. However, by looking at the nomogram with the assumption that the higher the number of points, the more important the effect, one might falsely interpret that symptomatic presentation is strongly predicting malignant tumors.

Suppose that instead of sex we add a significant effect in the model, for example, vascular flow on Doppler ultrasound ($\beta_{flow} = 2.85$, $P < .0001$; $\beta_{symptoms} = 0.30$, $P = .42$; Fig 4). The strongest variable in this model, vascular flow, would be converted into 100 points, and a patient with presence of symptoms will now be given 10.5 points, compared with Fig 3, where a patient with symptoms was given 100 points. The 10.5 points reflect the relative importance of $\beta_{symptoms}$ to β_{flow} (0.30/2.85 multiplied by 100). The total points axis in Fig 4 can go up to a maximum of 110.5 points, and the predicted probability ranges from 0.15 to 0.80. The second nomogram would have been chosen over the first because of its higher discrimination ability (c-index of 0.77 v 0.55) and the fact that the effect of flow is statistically significant. Nomograms rank the importance of an effect in predicting the outcome only in the context of the other covariates currently in the model. The number of points does not reflect the
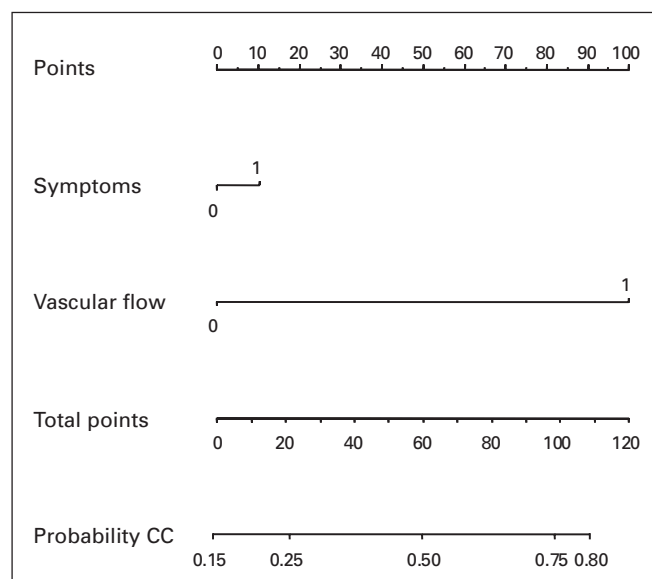

Fig 4. A nomogram with symptoms and vascular flow predicting the probability of clear-cell (CC) histology in renal masses.

association with the outcome, in a broader sense, nor does it represent statistical significance in terms of P values.

STEP 7. APPLY THE NOMOGRAM

When the nomogram is indeed used in clinic to provide a prediction, for example, the probability of 5-year recurrence to a patient, it is important both for the clinician and patient to understand the correct interpretation. Recall that the beta coefficients from the model formula are random, hence there is variability in the predictions. Accordingly, these predictions come with associated CIs around them that should be related to patients. Moreover, clinicians should also use these intervals to assess their own confidence in these predictions. Available statistical packages provide these CIs and they could potentially be programmed in handheld devices (logistic and phreg procedures in SAS).

Patients with equal predicted probability from a nomogram may have a different uncertainty around their prediction. Nomogram predictions were evaluated for the case study nomogram. For example, a 69-year-old male patient with a 3.9-cm renal mass, without vascular flow on Doppler ultrasound (patient A), and a 48-year-old female patient with 9.7-cm renal mass, without vascular flow (patient B), both have an 18% probability of clear-cell histology. The 95% CI for patient A is 10% to 29%, whereas the 95% CI for patient B is 7% to 41%. In the case study, the median width of the 95% CI was 18% (range, 11% to 47%), indicating that 50% (147 of 299 patients) of the patients have an approximately $\pm$ 9% uncertainty regarding their prediction. Ninety-five percent (283 of 299) of patients have less than $\pm$16% uncertainty regarding their prediction. Patients who deviate from the average patient profile (in this case, young female patients with large tumors) were rare and will tend to have wider CIs compared with patients with common characteristics. We recognize that conveying uncertainty regarding estimates of prognosis is challenging for patients, particularly when literacy and numeracy are sometimes limited. However, because nomograms are first and foremost interpreted
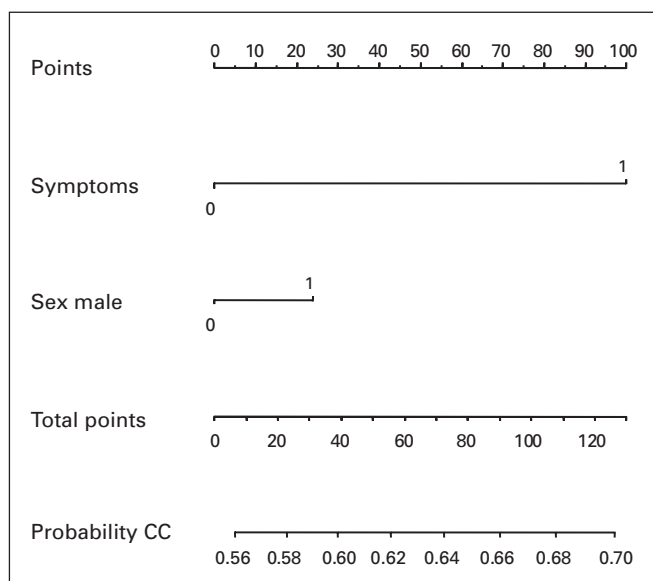

Fig 3. A nomogram with symptoms and sex predicting the probability of clear-cell (CC) histology in renal masses.

Table 2. Reviewing a Published Nomogram
Assessment Guidelines
1. Data population: Eligibility criteria
Broad versus stringent
Previous therapies
2. Clinical outcome
3. Assess all variables under consideration
Was a known clinically important variable not included? Why?
Have new variables reached significance that were not known to be related to the outcome?
Are predictor variables confounded?
Have interactions been assessed?
Are nonlinear relationships justified?
4. Reporting
Have they reported estimates of effects, confidence intervals for effect measures?
Have they reported CIs around these predictions?
Have they reported how far predictions are in comparison with the actual values?
5. Is the underlying final model clear and is it clinically meaningful?
6. Are odds ratios/hazard ratios clinically interpretable?
7. How did they validate the model?
Internal versus external data
8. Do you feel that this model is better than existing models?

Table 3. Using and Applying the Nomogram (generalizability)
Pertinent Queries
1. Is the patient population in your clinic similar to that in the nomogram, ie:
Patient characteristics
Clinical characteristics (tumor type)
Previous treatment
2. If your patient population differs in patient characteristics or tumor types, do you feel comfortable using this nomogram?
3. If an important predictor that is useful in characterizing your patient population, for example, race/ethnicity, previous therapy, or clinical size, is not included in the nomogram, is it reasonable to use the published nomogram in your clinic?
4. Do you feel comfortable with the degree of uncertainty regarding the point estimates derived from the underlying regression models?
5. Can you provide predictions to patients regarding their probability of survival by providing a CI, knowing that this uncertainty varies from patient to patient?

by physicians as a tool to assist in conveying prognostic information, inclusion of estimates of uncertainty are important so as to reinforce the degree of uncertainty around the point estimates derived from the underlying regression models. Although this strategy adds complexity to the graphical representation of results, it enhances the integrity of nomograms. Further research to explore strategies for communicating prognostic estimates and their respective uncertainty to patients is a priority and requires collaboration between clinicians, statisticians, and behavioral researchers.

In practice, a clinician can apply a published nomogram that was generated using data on patients that might be different than one's own population of interest. A comparison of the distribution of patient characteristics in the original cohort with the patients in the external cohort would provide a guideline on whether this is a valid representation. In Table 2, we outline points clinicians should review to perform their own assessment of the applicability of an existing nomogram to their own circumstances. Table 3 identifies a set of basic questions for clinicians to consider before applying a nomogram in the clinical setting.

In conclusion, the methodology underlying the construction of nomograms should be understood by clinical users so that prognostic estimates are appropriately communicated. In 2004, the National Cancer Institute sponsored a workshop on methodological issues relevant to prediction models that included development, evaluation, and validation of proposed models.[30] We extend this previous work by providing a step-by-step guide to help clinicians either construct a new nomogram or evaluate and apply published nomograms. The statisti-

cal concepts are often not obvious. It is the statistician's responsibility to explain the importance of these steps during nomogram development as well as to ensure that the prognostic estimates derived from these tools are reliably interpreted. The validity of nomograms lies in the assumptions of the statistical model, which should be carefully reviewed. The relationship between the prognostic variables and outcome measures should be correctly analyzed, the relationships among covariates should be discussed, and synergistic effects should be assessed. Building a nomogram is a trade-off between use of complex mathematical formulas and elegant simplicity: like any other estimation technique, bias and uncertainty, or variability, are inherent in this process. Meaningful relationships should be the foundation of the underlying models in nomograms. Greater reliance on a set of standard criteria for model development, validation, and communication will enhance the value of nomograms for clinical decision making and patient-clinician communication.

REFERENCES

1. Bochner BH, Kattan MW, Vora KC: Postoperative nomogram predicting risk of recurrence after radical cystectomy for bladder cancer. J Clin Oncol 24:3967-3972, 2006

2. Alran S, De Rycke Y, Fourchotte V, et al: Validation and limitations of use of a breast cancer nomogram predicting the likelihood of non-sentinel node involvement after positive sentinel node biopsy. Ann Surg Oncol 14:2195-2201, 2007

3. Karakiewicz PI, Briganti A, Chun FK, et al: Multi-institutional validation of a new renal cancer-specific survival nomogram. J Clin Oncol 25:1316-1322, 2007

4. Steyerberg EW, Roobol MJ, Kattan MW, et al: Prediction of indolent prostate cancer: Validation and updating of a prognostic nomogram. J Urol 177:107-112, 2007

5. Wierda WG, O'Brien S, Wang X, et al: Prognostic nomogram and index for overall survival in previously untreated patients with chronic lymphocytic leukemia. Blood 109:4679-4685, 2007

6. Sternberg CN: Are nomograms better than currently available stage groupings for bladder cancer? J Clin Oncol 24:3819-3820, 2006

7. Kattan MW: Nomograms are superior to staging and risk grouping systems for identifying high-risk patients: Preoperative application in prostate cancer. Curr Opin Urol 13:111-116, 2003

8. Mariani L, Miceli R, Kattan MW, et al: Validation and adaptation of a nomogram for predicting the survival of patients with extremity soft tissue sarcoma using a three-grade system. Cancer 103:402-408, 2005

9. Wang L, Hricak H, Kattan MW, et al: Prediction of organ-confined prostate cancer: Incremental value of MR imaging and MR spectroscopic imaging to staging nomograms. Radiology 238:597-603, 2006

10. Wong SL, Kattan MW, McMasters KM, et al: A nomogram that predicts the presence of sentinel node metastasis in melanoma with better discrimination than the American Joint Committee on Cancer staging system. Ann Surg Oncol 12:282-288, 2005

11. Ohori M, Kattan MW, Koh H, et al: Predicting the presence and side of extracapsular extension: A nomogram for staging prostate cancer. J Urol 171: 1844-1849; discussion 1849, 2004

12. Stephenson AJ, Scardino PT, Eastham JA, et al: Postoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. J Clin Oncol 23:7005-7012, 2005

13. Cowen ME, Halasyamani LK, Kattan MW: Predicting life expectancy in men with clinically localized prostate cancer. J Urol 175:99-103, 2006

14. Kattan MW, Ficarra V, Artibani W, et al: Nomogram predictive of cancer specific survival in patients undergoing partial or total amputation for squamous cell carcinoma of the penis. J Urol 175: 2103-2108, 2006

15. Sorbellini M, Kattan MW, Snyder ME, et al: Prognostic nomogram for renal insufficiency after radical or partial nephrectomy. J Urol 176:472-476, 2006

16. White RR, Kattan MW, Haney JC, et al: Evaluation of preoperative therapy for pancreatic cancer using a prognostic nomogram. Ann Surg Oncol 13:1485-1492, 2006

17. Roach M 3rd, Weinberg V, Nash M, et al: Defining high risk prostate cancer with risk groups and nomograms: Implications for designing clinical trials. J Urol 176:S16-20, 2006

18. Eastham JA, Kattan MW, Scardino PT: Nomograms as predictive models. Semin Urol Oncol 20: 108-115, 2002

19. Raj GV, Bach AM, Iasonos A, et al: Predicting the histology of renal masses using preoperative Doppler ultrasonography. J Urol 177:53-58, 2007

20. Harrell FE Jr: Regression Modeling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York, NY, Springer Verlag, 2001

21. Neter JKMH, Nachtsheim JC, Wasserman W: Applied Linear Statistical Models (ed 4). Chicago, IL, Irwin, 1990

22. Tibshirani R: Regression shrinkage and selection via the lasso. J R Stat Soc Ser B (Methodol) 58:267-288, 1996

23. Barker L, Brown C: Logistic regression when binary predictor variables are highly correlated. Stat Med 20:1431-1442, 2001

24. Van Steen K, Curran D, Kramer J, et al: Multicollinearity in prognostic factor analyses using the EORTC QLQ-C30: Identification and impact on model selection. Stat Med 21:3865-3884, 2002

25. Hand DJ: Classifier technology and the illusion of progress. Stat Sci 21:1-14, 2006

26. Pencina MJ, D'Agostino RB: Overall C as a measure of discrimination in survival analysis: Model specific population value and confidence interval estimation. Stat Med 23:2109-2123, 2004

27. Efron B: TJR: An Introduction to the Bootstrap. New York, NY, Chapman & Hall

28. Ferrone CR, Kattan MW, Tomlinson JS, et al: Validation of a postresection pancreatic adenocarcinoma nomogram for disease-specific survival. J Clin Oncol 23:7529-7535, 2005

29. Diblasio CJ, Kattan MW: Use of nomograms to predict the risk of disease recurrence after definitive local therapy for prostate cancer. Urology 62:9-18, 2003 (suppl)

30. Freedman AN, Seminara D, Gail MH, et al: Cancer risk prediction models: A workshop on development, evaluation, and application. J Natl Cancer Inst 97:715-723, 2005

■ ■ ■