

Decision Analysis for the Evaluation of Diagnostic Tests, Prediction Models, and Molecular Markers

Andrew J. VICKERS

The traditional statistical approach to the evaluation of diagnostic tests, prediction models, and molecular markers is to assess their accuracy, using metrics such as sensitivity, specificity, and the receiver-operating-characteristic curve. However, there is no obvious association between accuracy and clinical value: it is unclear, for example, just how accurate a test needs to be in order for it to be considered “accurate enough” to warrant its use in patient care. Decision analysis aims to assess the clinical value of a test by assigning weights to each possible consequence. These methods have been historically considered unattractive to the practicing biostatistician because additional data from the literature, or subjective assessments from individual patients or clinicians, are needed in order to assign weights appropriately. Decision analytic methods are available that can reduce these additional requirements. These methods can provide insight into the consequences of using a test, model, or marker in clinical practice.

KEY WORDS: Decision support techniques; Outcome assessment; Prognosis.

1. INTRODUCTION

Much of clinical medicine concerns diagnosis and prediction: patients want to know what they have (“Do I have cancer?”), and what is likely to happen to them (“Will I be cured or will the cancer come back?”); clinicians want to know what to treat (“Should I operate?”) and how aggressive treatment should be (“Should I also give chemotherapy?”). Diagnosis and prediction have traditionally been based on clinical history and physical examination. Recent years has seen an upsurge of interest in molecular markers of disease, based on sophisticated analysis of blood or tissue, particularly with respect to genomic information. For example, predicting a breast cancer patient’s risk of recurrence traditionally depended on determining how far the cancer had spread (cancer stage); it has recently been suggested that genetic mutations in breast cancer cells also predict disease behavior (Marchionni et al. 2008).

From a statistical standpoint, prediction and diagnosis present similar analytic challenges: in both cases our dataset consists of an estimate on the basis of a test T (with values T^+ and T^- , for “positive” and “negative”), a true disease state D

(with values D^+ and D^- , for “diseased” and “nondiseased”). The traditional statistical approach has been to assess *accuracy*, typically defined using a measure of association between T and D . However, accuracy metrics have questionable clinical relevance. As an alternative, decision analytic methods have been proposed that can evaluate diagnostic tests, predictive models or molecular markers in terms of their real clinical consequences. A drawback of such methods is the requirement of additional information about clinical benefits and harms. However, some novel statistical methods are available to reduce these additional requirements.

2. STATISTICAL APPROACHES TO EVALUATION OF TESTS, MODELS, AND MARKERS

Consider the diagnosis of prostate cancer using the molecular marker prostate-specific antigen (PSA). Men with elevated levels of PSA in the blood are typically referred for prostate biopsy. However, only a minority of men with high PSA, around 20–25%, actually have prostate cancer. Some researchers have suggested that the level of unbound PSA (“free” PSA) can distinguish prostate cancer from benign prostate disease; specifically, cancer is more likely if the ratio between total and free PSA is low (Roddam et al. 2005). To investigate the value of free-to-total PSA ratio, I will use a dataset from the Göteborg site of the European Randomized Study of Screening for Prostate Cancer (ERSPC) (Schroder et al. 2003). The dataset consists of 753 Swedish men with elevated PSA (3 ng/ml or higher) who were biopsied, of whom 192 were found to have prostate cancer. The study question is whether free-to-total PSA ratio can help determine which men really have prostate cancer, and hence should undergo biopsy, and which men do not have cancer, and who should therefore avoid what would be an unnecessary biopsy.

The simplest approach is to use free-to-total PSA ratio as a binary test: for example, men with a ratio of 0.18 or less are defined as positive (T^+) and require biopsy, whereas those with free-to-total PSA ratio above 0.18 are defined as negative (T^-) and do not need biopsy. Traditional biostatistical analysis starts by creating a two-by-two table of D by T and then calculates accuracy metrics such as sensitivity $\text{Sens} = P(T^+|D^+)$ and specificity $\text{Spec} = P(T^-|D^-)$. Table 1 gives sensitivity and specificity estimates, using the ERSPC dataset, for various cut-points of free-to-total PSA ratio. If such calculations are repeated for the entire range of free-to-total PSA values, we can plot Sens versus $1 - \text{Spec}$ to obtain the receiver-operating-characteristic (ROC) curve, with the area-under-the-curve (AUC) providing a global metric of test accuracy. Figure 1 gives the ROC curve for free-to-total PSA ratio and positive biopsy in men with elevated PSA, with $\text{AUC} = 0.769$.

Andrew J. Vickers is Associate Attending Research Methodologist, Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, 307 E. 63rd Street, 2nd Floor, New York, NY 10065 (E-mail: vickersa@mskcc.org). Dr Vickers’ work on this research was funded by a P50-CA92629 SPORE from the National Cancer Institute

Table 1. Sensitivity and specificity for various cut-points of free-to-total PSA ratio for predicting prostate cancer in men with elevated PSA

Cut-point for free-to-total PSA ratio	Sensitivity	Specificity
0.18	88%	53%
0.15	71%	68%
0.20	91%	39%

3. LIMITATIONS OF ACCURACY METRICS

There are two general problems with accuracy metrics such as sensitivity, specificity, and AUC. In general, an accurate test, model, or marker is more likely to be useful than one less accurate, but it is difficult to know for any specific situation whether the accuracy of a test is high enough to warrant implementation in the clinic. Does an AUC of 0.769 mean that free-to-total PSA ratio should be used to determine who does or does not get biopsy, or would some higher value, say, 0.850, be required?

The second problem with accuracy metrics concerns the choice of cut-point. Assuming that we did decide to use free-to-total PSA ratio to determine which men with elevated PSA were referred to biopsy, what value should we use as the criterion for biopsy? In the case of cancer, sensitivity is valued over specificity, but it is difficult to say which combination of sensitivity and specificity in Table 1 is optimal.

4. INCORPORATING CLINICAL CONSEQUENCES

In decision analysis, one identifies possible actions and consequences, and selects the action with the best expected conse-

quence. Often this process is aided by constructing a “decision tree,” such as that shown in Figure 2 for prostate biopsy in men with elevated PSA. The principle of the decision tree is first to identify every possible decision, then identify every possible consequence of each decision, and finally to assign a probability and a benefit to each consequence (Hunink et al. 2001). We denote probabilities as p_{xy} and benefits as b_{xy} where x is an indicator for the test result and y is the indicator for disease.

When faced with an elevated PSA, a patient has to decide among three options: undergo biopsy without further testing; refuse biopsy; undergo further testing and decide whether or not to have biopsy dependent on the results of those tests. A patient either has cancer or does not, and so the four possible consequences are finding cancer (true positive, p_{11} and b_{11}); unnecessary biopsy (false positive, p_{10} and b_{10}); missing cancer (false negative, p_{01} and b_{01}); and avoiding unnecessary biopsy (true negative; p_{00} and b_{00}). In the case of decisions for or against biopsy without further testing, the probability of each outcome depends on the prevalence $\pi = P(D^+)$ of prostate cancer. In the case of additional testing, these probabilities are also dependent on the sensitivity and specificity of the additional test as follows:

$$p_{11} = \text{Sens} \times \pi$$

$$p_{10} = (1 - \text{Spec}) \times (1 - \pi)$$

$$p_{01} = (1 - \text{Sens}) \times \pi$$

$$p_{00} = \text{Spec} \times (1 - \pi).$$

The values of each outcome b_{11} , b_{10} , b_{01} , and b_{00} are difficult to specify. One approach is to use published estimates. For

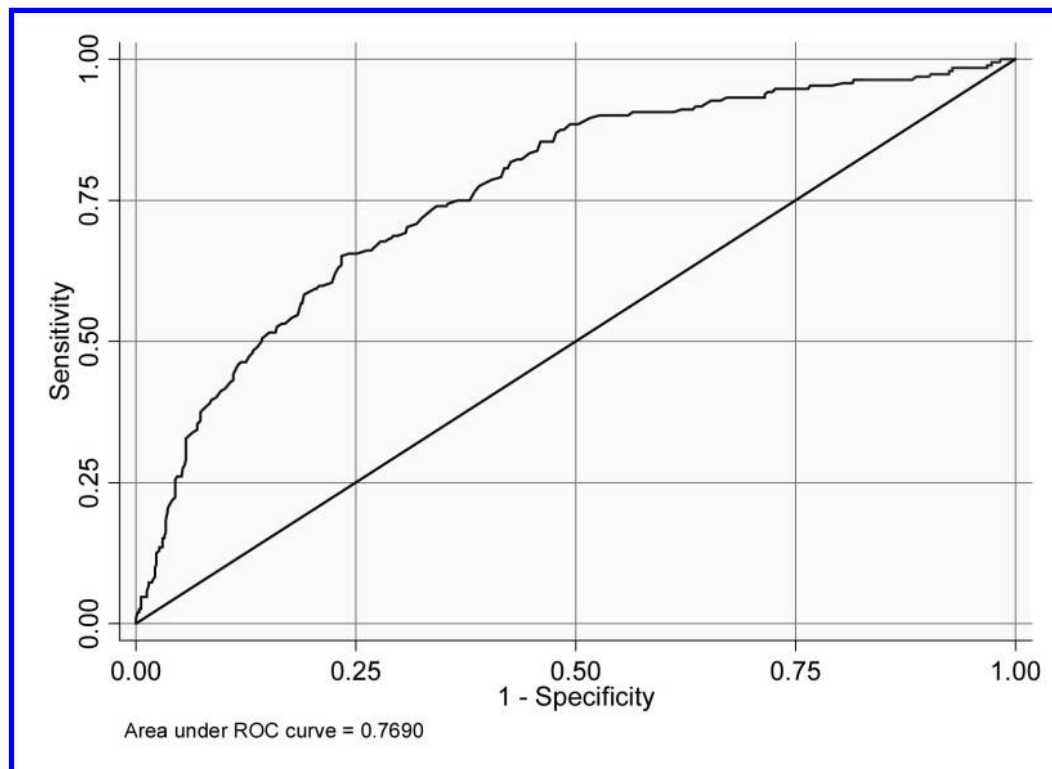


Figure 1. ROC curve for free-to-total PSA ratio in men with elevated PSA.

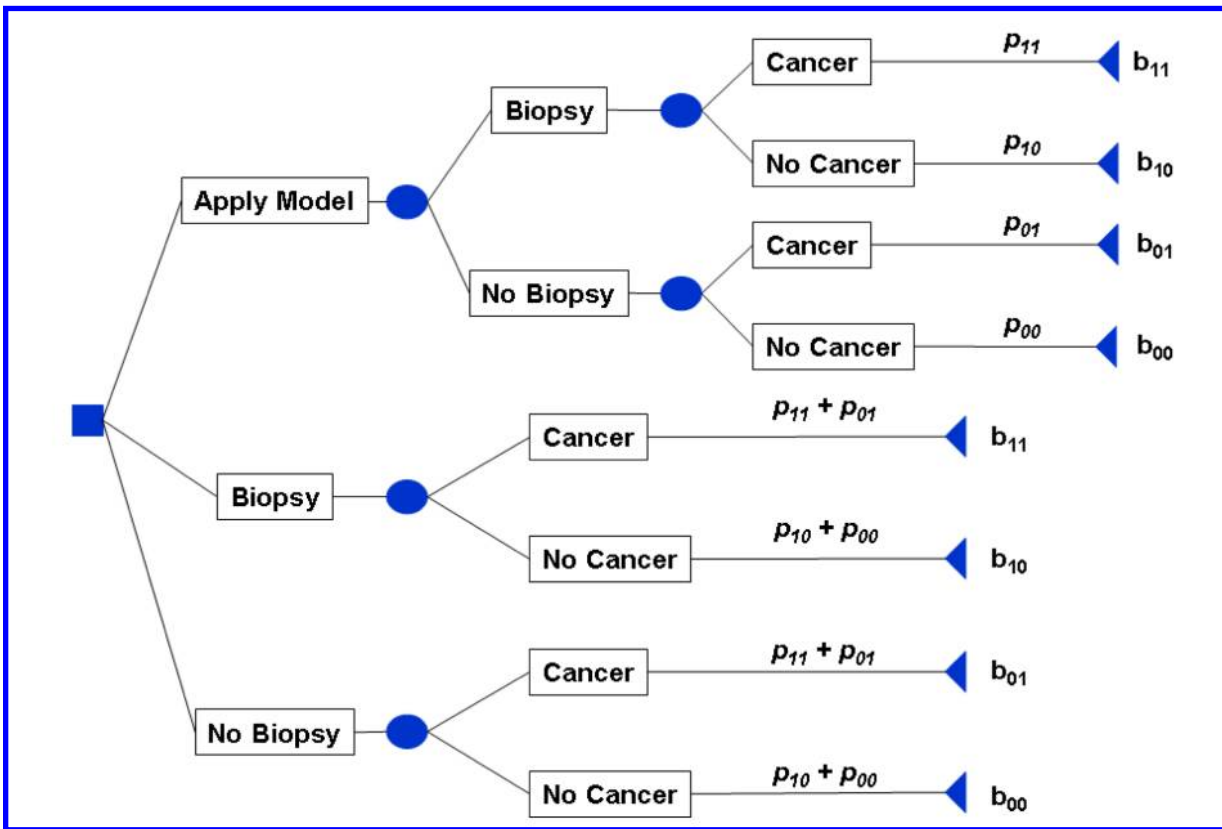


Figure 2. Decision tree for prostate cancer biopsy in men with elevated PSA.

example, Berry and Parmigiani (1998) used published values related to quality of life in a decision analysis of breast cancer screening. The problem for the practicing biostatistician is that obtaining such values from the published literature can be time-consuming, that estimates can vary substantially between different papers and that converting published estimates (such as the probability of a localized cancer progressing in 10 years) into a benefit parameter on a 0–1 scale can be complex.

A simple alternative is to fix the best possible outcome at 1 (in this case, no biopsy and no disease, true negative, $b_{00} = 1$) and

the worst at 0 (no biopsy and disease, false negative, $b_{01} = 0$), hence only two remaining values need be specified. For now assume that, following a discussion with a clinician, we obtain b_{11} and b_{10} of 0.6 and 0.85, respectively.

The optimal decision is the one with the highest expected benefit. Table 2 shows the results of the decision tree using a cut-point of 0.18 for free-to-total PSA ratio, where the prevalence of prostate cancer is $\hat{\pi} = 25.5\%$. The estimated expected benefit for using the free-to-total PSA ratio is 0.827, which is higher than the values for either biopsying everyone (0.786)

Table 2. Use of a decision tree to make a biopsy decision.

Parameters	Value	Expected outcome			
		Strategy	Expected value	Expected value	Total
Health values for each outcome		Test	$b_{11} \times p_{11}$	$0.6 \times 22.44\%$	0.135
b_{11} (true positive, biopsy cancer)	0.6		$b_{10} \times p_{10}$	$0.85 \times 35.02\%$	0.298
b_{10} (false positive, unnecessary biopsy)	0.85		$b_{01} \times p_{01}$	$0 \times 3.06\%$	0.000
b_{01} (false negative, missed cancer)	0		$b_{00} \times p_{00}$	$1 \times 39.49\%$	0.395
b_{00} (true negative, avoid biopsy)	1				0.827
Prevalence (π)	25.5%	Biopsy	$b_{11} \times (p_{11} + p_{01})$	$0.6 \times 25.5\%$	0.153
Sensitivity (Sens)	88%		$b_{10} \times (p_{10} + p_{00})$	$0.85 \times 74.5\%$	0.633
Specificity (Spec)	53%				0.786
Probability of each outcome		No biopsy	$b_{01} \times (p_{11} + p_{01})$	$0 \times 25.5\%$	0
p_{11} : Sens \times π	22.44%		$b_{00} \times (p_{10} + p_{00})$	$1 \times 74.5\%$	0.745
p_{10} : (1 – Spec) \times (1 – π)	35.02%				0.745
p_{01} : (1 – Sens) \times π	3.06%				
p_{00} : Spec \times (1 – π)	39.49%				

or no one (0.745). Therefore the strategy of undergoing further testing for all men with elevated PSA, and then sending for biopsy those with a ratio of 0.18 or less, is estimated to be optimal.

5. DRAWBACKS OF CONVENTIONAL DECISION ANALYSIS

Determining values for b_{11} , b_{10} used in Table 2 involve judgments about the relative harm of a missed cancer versus an unnecessary biopsy. Patients' and physicians' judgments can vary on this point: for example, some patients do not tolerate invasive procedures such as biopsy particularly well. So while a standard decision analysis such as that in Table 2 may be a good starting point, in theory we should ask each physician or patient to evaluate benefits and harms individually, use their answers in the decision tree, and then work out whether outcome would be improved by using free-to-total PSA. This can be difficult to do, especially as the benefit parameters are not intuitive to specify.

From the analyst's point of view, the determination of b_{11} , b_{10} , whether from the literature or from individual patients or physicians, is problematic. In the next section, I show how an alternative parameter, the threshold probability, can be used in decision analysis.

6. USING THRESHOLD PROBABILITY IN DECISION ANALYSIS

Suppose it is possible to specify a p_t , the threshold probability of disease for taking some action, such as biopsying a man for prostate cancer: if a patient's estimated probability of disease is greater than p_t he will opt for biopsy; if it is less than p_t , he will not opt for biopsy. By definition, when the probability of disease is equal to the threshold probability p_t , the benefits of opting for biopsy or no biopsy are equal. Thus:

$$b_{11} \times p_t + b_{10} \times (1 - p_t) = b_{01} \times p_t + b_{00} \times (1 - p_t).$$

And therefore

$$(b_{00} - b_{10}) / (b_{11} - b_{01}) = p_t / (1 - p_t). \quad (1)$$

Now $b_{00} - b_{10}$ is the benefit of true negative result compared to a false positive result; in clinical terms, the benefit of avoiding unnecessary treatment such as a negative biopsy. Comparably, $b_{11} - b_{01}$ is the benefit of a true positive result compared to a false negative result; in other words, the benefit of treatment where it is indicated. Equation (1) therefore tells us that the threshold probability at which a patient will opt for treatment is informative of how a patient weighs the relative benefit of appropriate treatment compared to the benefit of avoiding unnecessary treatment (Pauker and Kassirer 1980).

We can rearrange (1) to obtain

$$-(b_{10} - b_{00}) = (b_{11} - b_{01}) \times p_t / (1 - p_t). \quad (2)$$

This states that the harm of a false positive compared to a true negative, is equal to the benefit of a true positive compared to a false negative, multiplied by the odds at p_t . A "net benefit"

is benefit minus harm, thus the theoretical relationship in (2) allows us to define a net benefit (first described by C.S. Peirce (Baker and Kramer 2007)):

$$\frac{\text{True Positive Count} - \text{False Positive Count} \times \left(\frac{p_t}{1 - p_t} \right)}{\text{Total Sample Size}}. \quad (3)$$

This expression is equivalent to

$$\text{Sens} \times \pi - (1 - \text{Spec}) \times (1 - \pi) \times \left(\frac{p_t}{1 - p_t} \right). \quad (4)$$

There are three advantages to using the threshold probability p_t in place of the benefit parameters b_{xy} . First, only a single parameter needs to be chosen. Second, the units of the parameter are more intuitive: patients and clinicians understand the concept of risk much more easily than the idea of a health state value on a scale of 0 to 1. Indeed, threshold probability is closely related to a widely used statistic, positive predictive value. For example, it has been argued that the positive predictive value of a screening test for ovarian cancer needs to be at least 10%, because clinicians would be unwilling to conduct more than 10 surgeries to find a single case of ovarian cancer (Skates et al. 1995). Accordingly, we might therefore use a p_t of 10% in a decision analysis of ovarian cancer. Third, a threshold probability can be used both for weighting true and false positive test results and for determining the cut-off for a positive test result: instead of arbitrarily choosing a free-to-total PSA ratio cut-off of 0.18, 0.15, or 0.20, we calculate probabilities of cancer by logistic regression and use the threshold probability as the cut-off.

Following Vergouwe et al. (2002), a straightforward decision analytic method for determining the value of a diagnostic test, predictive model or molecular marker is as follows:

1. Obtain a threshold probability (p_t) for treatment.
2. If necessary, use logistic regression to convert the results of the test, marker or model into a predicted probability of disease \hat{p} .
3. Define patients as test positive if $\hat{p} \geq p_t$ and negative otherwise. For a binary diagnostic test, \hat{p} is 1 for positive and 0 for negative.
4. Calculate net benefit of the test, marker, or model using the formula for net benefit in Equation (3) or (4).
5. Calculate clinical net benefit for the strategy of treating all patients. As sensitivity is 100% and specificity 0%, (4) simplifies to

$$\pi - (1 - \pi) \times \left(\frac{p_t}{1 - p_t} \right).$$
6. The net benefit for the strategy of treating no patients is defined as zero.
7. The optimal strategy is that with the highest clinical net benefit.

Table 3. Net benefit at a threshold probability p_t of 20%.

Strategy	Sample size	Negatives	True positives	False positives	Net benefit calculation	Net benefit
Model with free-to-total ratio	753	384	149	220	$[149 - 220 \times (0.2 \div 0.8)] \div 753$	0.1248
Biopsy all men	753	0	192	561	$[192 - 561 \times (0.2 \div 0.8)] \div 753$	0.0687

Note that the unit for net benefit is the number of true cases found per patient and therefore has a maximum value at the prevalence π : all cases found, with no false positives.

To illustrate calculation of a net benefit, we will use a p_t of 20%. To calculate the net benefit for free-to-total ratio, we first have to convert values of the marker into predicted probabilities of cancer by logistic regression. Table 3 shows that of the total of 753 patients, there were 369 who, on the basis of the predictive model using free-to-total PSA ratio, had a predicted probability of cancer of 20% or more. Of these, 149 had cancer and 220 did not. This gives a net benefit of $[149 - 220 \times (0.2 \div 0.8)] \div 753 = 0.1248$. In comparison, the net benefit for a strategy of biopsying all men is 0.0687; the net benefit for biopsying no men is, by definition, zero.

As was for the case for expected value in a traditional decision analysis, we take the strategy with the highest net benefit, irrespective of the size of the difference. Hence for men who would accept a biopsy if their risk of prostate cancer was 20% or more, but not if their risk was less than 20%, the optimal strategy is to calculate their probability of cancer from a logistic model using free-to-total ratio as the predictor and then biopsy those with predicted risk from the model of 20% or more.

7. DECISION CURVE ANALYSIS

As pointed out earlier, different men will weigh differently the relative benefits of finding a prostate cancer compared to an unnecessary biopsy. Accordingly, we can vary p_t , calculate net benefit at each p_t , and then plot net benefit on the y axis against threshold probability on the x axis. This gives what is known as a decision curve (Vickers and Elkin 2006).

The decision curve for free-to-total PSA is shown in Figure 3. To interpret the decision curve, we need an estimate of the range of threshold probabilities in typical patients. We can obtain such an estimate from clinicians: a typical response is that few men would opt for biopsy if they were told they had a risk of prostate cancer less than 10%; on the other hand, it is hard to imagine that a man taking a PSA test would want at least a 50:50 chance of cancer before agreeing to biopsy. Figure 4 shows the decision curve for free-to-total PSA ratio in our reasonable range of 10–40%. Net benefit is superior to biopsying all or no men across the whole range. We can therefore conclude that using free-to-total PSA ratio to determine biopsy in men with elevated PSA will improve clinical outcomes irrespective of any differences in patient and physician preferences.

Note that decision curve analysis does not require that pa-

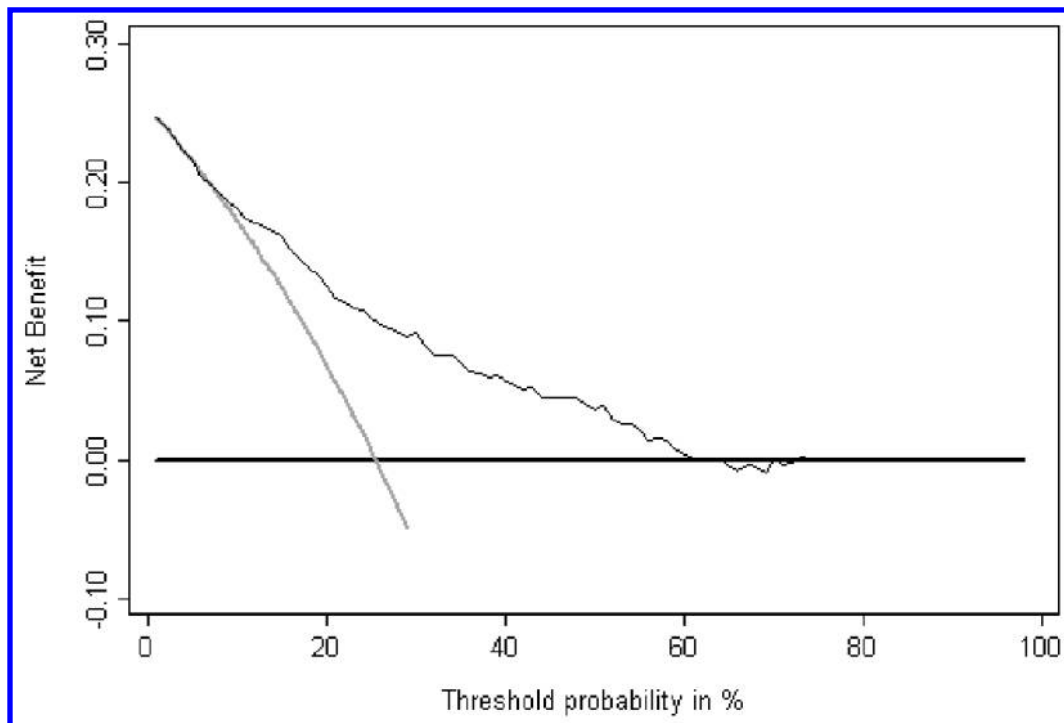


Figure 3. Decision curve analysis for free-to-total PSA ratio in men with elevated PSA. Gray line: biopsy all men. Thin black line: use free-to-total PSA ratio to determine who to biopsy. Thick black line: biopsy no man.

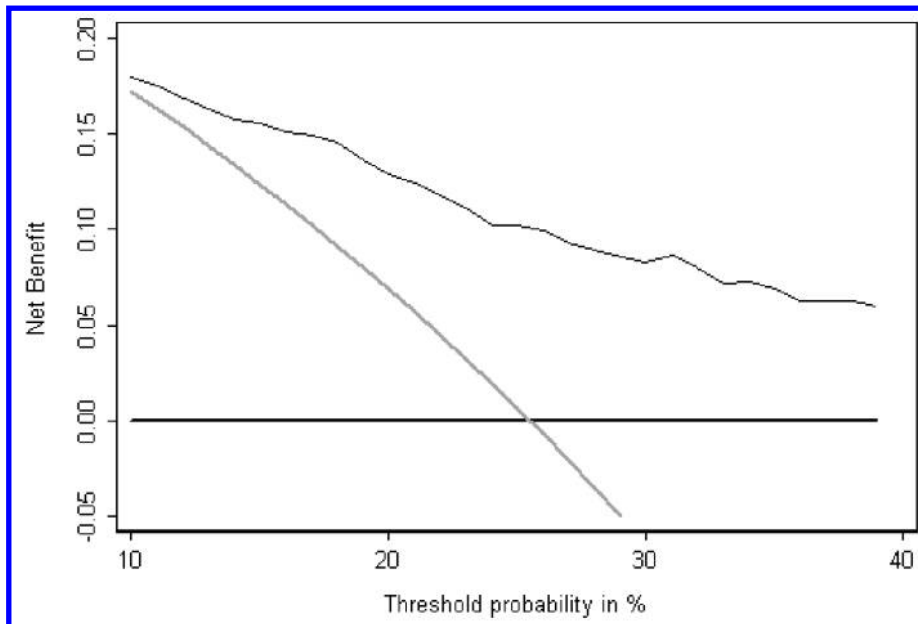


Figure 4. Decision curve analysis for free-to-total PSA ratio in men with elevated PSA, showing the critical range of threshold probabilities, 10–40%. Gray line: biopsy all men. Thin black line: use free-to-total PSA ratio to determine who to biopsy. Thick black line: biopsy no man.

tients be asked about their threshold probability, indeed, our conclusions are independent of the mode of decision making. For example, in the case of prostate cancer biopsy, a clinician has the following options to choosing who to biopsy:

1. Set a threshold and apply to all patients: patients above the threshold are biopsied, patients below the threshold are not.
2. Divide patients into high, low, and intermediate risk. High-risk patients are biopsied; low-risk patients are not. Whether or not to biopsy a patient at intermediate risk is

taken on a case-by-case basis, depending on the patient’s age, co-morbidities, and personal preference.

3. Discuss biopsy with each patient and obtain a quantitative estimate of their personal preferences. Compare this estimate with their risk from the model and act accordingly.

Irrespective of how the physician decides which patients should be biopsied, decision curve analysis shows that using free-to-total ratio will improve decision making, as long as any thresholds used are in the reasonable range of 10–40%. Hence decision curve analysis can be applied to a dataset without the

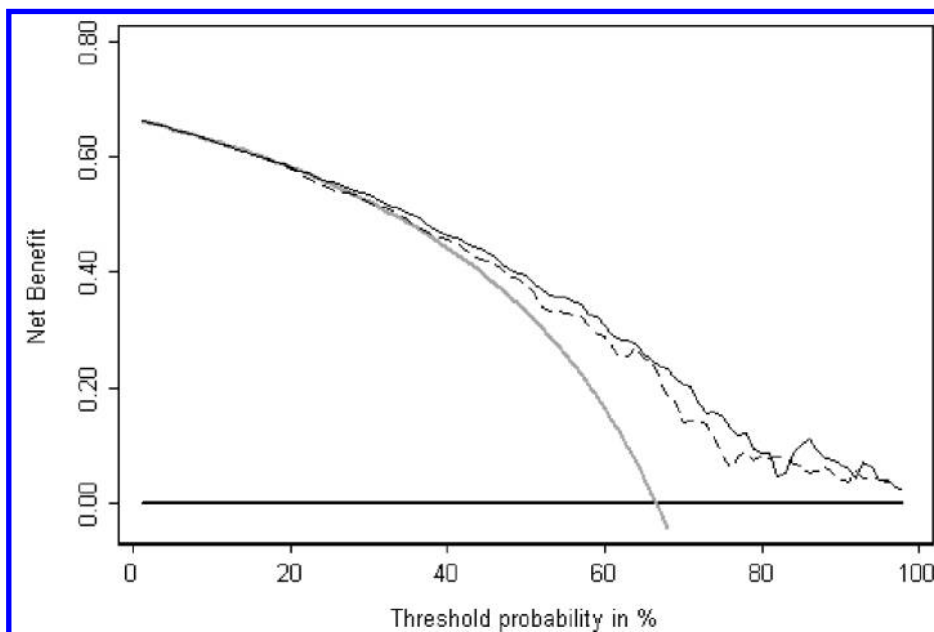


Figure 5. Decision curve analysis for urokinase. Gray line: biopsy all men. Thick black line: biopsy no man. Dashed line: biopsy men on the basis of a statistical model incorporating age and PSA level. Thin black line: biopsy men on the basis of a statistical model incorporating age, PSA level and urokinase.

need to obtain the sort of additional information—such as on the benefits of treatment, or subjective patient preferences—typically required by traditional decision analysis; all that is required is a general estimate for a reasonable range of threshold probabilities.

Figure 5 shows a decision curve for a different marker, evaluated on a different set of patients. This demonstrates two points about decision curve analysis. First, the method can be used to evaluate the marginal value of a molecular marker, by estimating the net benefit of a statistical model including both standard predictors and the new marker with that of the standard predictors alone. Second, decision curve analysis can determine that a model is not of clinical value despite good accuracy. The AUC of the model including urokinase was excellent (0.751) and yet it clearly has minimal clinical value: in the critical range of threshold probabilities of 10–40%, net benefit is no higher than the strategy of biopsying all patients. This is no doubt related to the extremely high prevalence of prostate cancer in this dataset (~65%).

8. CONCLUSIONS

The biostatistical literature has almost exclusively been concerned with methods for evaluating the accuracy of predictive models, diagnostic tests and molecular markers. While novel approaches assessing ROC curves, classification tables, or calibration continue to proliferate, methods that incorporate clinical consequences are almost entirely absent from the literature. The clinical literature is similarly marked by a near-exclusive focus on accuracy. We recently reviewed 129 molecular marker studies in cancer, and although we found that many studies did evaluate whether or not a marker was accurate, not a single study used decision analytic methods to determine whether the marker would improve clinical outcome (Vickers et al. 2008).

Accuracy metrics clearly have their place. In the early phases of research, assessment of accuracy can help determine whether a test, model, or marker is sufficiently promising to warrant further testing, and can help refine techniques before a definitive study. As a trivial example, evidence of miscalibration might prompt an analyst to explore the use of nonlinear terms or a Bayesian correction factor. Moreover, it would surely be overly pragmatic to claim that a decision-analytic approach is all that is required, and that evaluation of accuracy does not aid our understanding of a test, model, or marker that we wish to bring to clinical practice.

Nonetheless, I have argued here that thinking only in terms

of accuracy is limited. We need to know not only whether a diagnostic test, predictive model, or molecular marker is accurate, but whether it is helpful clinically. As such, I would argue for increased attention to decision-analytic techniques in the methodologic and clinical literature. In particular, analysts should consider methods that can be based on general clinical estimates, but which can provide insight into the consequences of using a test, model, or marker in the clinic. These methods involve only the most trivial of computations and are thus straightforward to implement in biostatistical practice.

[Received January 2008. Revised June 2008.]

REFERENCES

- Baker S. G., and Kramer, B. (2007), "Peirce, Youden, and Receiver Operating Characteristic Curves," *The American Statistician*, 6, 1–4.
- Berry, D. A., and Parmigiani, G. (1998), "Assessing the Benefits of Testing for Breast Cancer Susceptibility Genes: A Decision Analysis," *Breast Disease*, 10, 115–125.
- Hunink, M., Glasziou, P., and Siegel, J. (2001), *Decision-Making in Health and Medicine: Integrating Evidence and Values*, New York: Cambridge University Press.
- Marchionni, L., Wilson, R.F., Wolff, A.C., Marinopoulos, S., Parmigiani, G., Bass, E.B., and Goodman, S.N. (2008), "Systematic Review: Gene Expression Profiling Assays in Early-Stage Breast Cancer," *The Annals of Internal Medicine*, 148, 358–369.
- Pauker, S. G., and Kassirer, J. P. (1980), "The Threshold Approach to Clinical Decision Making," *New England Journal of Medicine*, 302, 1109–1117.
- Roddam, A.W., Duffy, M.J., Hamdy, F.C., Ward, A.M., Patnick, J., Price, C.P., Rimmer, J., Sturgeon, C., White, P., and Allen, N.E. (2005), "Use of Prostate-Specific Antigen (PSA) Isoforms for the Detection of Prostate Cancer in Men with a PSA level of 2–10 ng/ml: Systematic Review and Meta-analysis," *European Urology*, 48, 386–399.
- Schroder, F. H., Denis, L. J., Roobol, M., Nelen, V., Auvinen, A., Tammela, T., Villers, A., Rebillard, X., Ciatto, S., Zappa, M., Berenguer, A., Paez, A., Hugosson, J., Lodding, P., Recker, F., Kwiatkowski, M., and Kirkels, W. J. (2003), "The Story of the European Randomized Study of Screening for Prostate Cancer," *BJU International*, 92, Suppl 2, 1–13.
- Skates, S.J., Xu, F.J., Yu, Y.H., Sjövall, K., Einhorn, N., Chang, Y., Bast, R.C. Jr., Knapp R.C. (1995), "Toward an Optimal Algorithm for Ovarian Cancer Screening with Longitudinal Tumor Markers," *Cancer*, 76(10 Suppl), 2004–2010.
- Vergouwe, Y., Steyerberg, E. W., Eijkemans, M. J., and Habbema, J. D. (2002), "Validity of Prognostic Models: When is a Model Clinically Useful?" *Seminars in Urologic Oncology*, 20, 96–107.
- Vickers, A., Jang, K., Sargent, D., Lilja, H., and Kattan, M. (2008), "A Systematic Review of Statistical Methods used in Molecular Marker Studies in Cancer," *Cancer*, 112(8), 1862–1868.
- Vickers, A. J., and Elkin, E. B. (2006), "Decision Curve Analysis: A Novel Method for Evaluating Prediction Models," *Medical Decision Making*, 26, 565–574.

This article has been cited by:

1. Shahrokh F. Shariat, Caroline Savage, Thomas F. Chromecki, Maxine Sun, Douglas S. Scherr, Richard K. Lee, Giovanni Lughezzani, Mesut Remzi, Michael J. Marberger, Pierre I. Karakiewicz, Andrew J. Vickers. 2011. Assessing the clinical benefit of nuclear matrix protein 22 in the surveillance of patients with nonmuscle-invasive bladder cancer and negative cytology. *Cancer* n/a-n/a. [[CrossRef](#)]
2. Q. Yang, T. Liu, R. Valdez, R. Moonesinghe, M. J. Khoury. 2010. Improvements in Ability to Detect Undiagnosed Diabetes by Using Information on Family History Among Adults in the United States. *American Journal of Epidemiology* 171:10, 1079-1089. [[CrossRef](#)]
3. Shahrokh F Shariat, Michael W Kattan, Andrew J Vickers, Pierre I Karakiewicz, Peter T Scardino. 2009. Critical review of prostate cancer predictive tools. *Future Oncology* 5:10, 1555-1584. [[CrossRef](#)]