Clinical prediction guides

Prediction is central to most of our actions as clinicians. We are faced with the task of predicting on the basis of history, physical examinations, and laboratory results when we diagnose, prognosticate, discuss causes, and choose treatment options. Is the patient with pain radiating down her left arm having a myocardial infarction, and to what extent should my prediction change if the pain radiates only to her right shoulder? What is the risk for an embolic stroke in a hypertensive woman with nonvalvular atrial fibrillation? What degree of benefit can a man with symptomatic 90% carotid stenosis derive from carotid endarterectomy if he also has diabetes and a 60-pack-year history of smoking?

In some instances, a prediction can be made with a single feature that overwhelms all others. For example, the facial features in the Down syndrome are so specific that they clinch the diagnosis and their absence rules it out (1). Single prognostic markers can also occasionally outweigh all others (e.g., ascites as the initial presentation of ovarian cancer), and sometimes individual patient characteristics, such as smoking, can predict huge differences in response to therapy.

We encounter several problems, however, in our attempts to make predictions. First, single prognostic markers are more often the exception than the rule. As a result, we usually are required to integrate more pieces of predictive information than can be easily manipulated in our heads. Second, predictors may be interrelated and, when each is given full weight, an overestimation of our prediction may result. Third, although certain statistical methods can help us identify independent predictors, they rely on mathematical correlation, not human biology, to decide whether a characteristic is important. As a result, statistical methods will identify with equal vigor a truly powerful biological manifestation of rapidly advancing disease and a nonsensical statistical clustering found only in the patients assembled for the report.

Fortunately, some solutions exist for these problems. The first and second problems can be solved with an array of multivariate statistical methods that permits us not only to consider several predictors simultaneously but also to distinguish the powerful predictive features from the hitchhiker findings that do not add information. By using these techniques, investigators can quantify the independent contributions of clinical predictors and combine them into clinical prediction guides (CPGs) for us to use on the clinical front lines. The third problem, of nonsensical statistical correlates, can be solved by testing the combination of predictors, or CPGs, that were developed in 1 group of patients in a second, independent sample of relevant patients to see whether the CPG retains its predictive power (2, 3).

Evidence-Based Medicine and ACP Journal Club have offered commentaries on CPGs in therapy, diagnosis, prognosis, and causation. However, we previously treated CPGs as diagnostic tests (see the Purpose and Procedure section in each issue). That situation changes with this issue. Articles on therapy, diagnosis, prognosis, and causation that test CPGs will now be required to meet the following methodologic criteria in addition to those already applied: CPGs must be generated in 1 set of patients (training set) and validated (i.e., they must have performed in a clinically useful way) in an independent set of patients (test set).

When we publish a CPG, we will ask the commentator to note whether the authors listed patient characteristics (both useful and useless) that were evaluated in the training set, along with their frequency (because potentially powerful predictors may occur too infrequently to emerge in the analysis). The commentator will also be asked to note whether the authors have documented the CPG's effect on clinical behavior and have thus shown that it is useful and practical in clinical settings (4). For example, did the application

Apply the Clinical Prediction Guide

Low PTP

u/s nl

Rules out DVT

Venography

Moderate PTP

u/s abn

Repeat u/s in 1 week

DVT anti-coagulation therapy

High PTP

u/s abn

DVT anti-coagulation therapy

Venography

Figure 1. Application of the clinical prediction guide to determine the probability of deep venous thrombosis. abn = abnormal; DVT = deep venous thrombosis; nl = normal; PTP = pretest probability; u/s = ultrasonography. (Adapted with permission from The Lancet. 1995; 345:1327.)
Checklist for Prediction of DVT

Major Points
- Active cancer (treatment ongoing or within previous 6 mo or palliative)
- Paralysis, paresis, or recent plaster immobilization of the lower extremities
- Recently bedridden > 3 days and/or major surgery within the past 4 weeks
- Localized tenderness along the distribution of the deep venous system
- Thigh and calf swollen (should be measured)
- Calf swelling 3 cm > symptomless side (measured 10 cm below the tibial tuberosity)
- Strong family history of DVT (≥ 2 first-degree relatives with history of DVT)
- Recent trauma (< 60 days) to the symptomatic leg
- Dilated superficial veins (nonvaricose) in symptomatic leg only
- Hospitalization within previous 6 months
- Erythema

Clinical Probability
- High
  - ≥ 3 major points and no alternative diagnosis
  - ≥ 2 major points and ≥ 2 minor points and no alternative diagnosis
- Low
  - 1 major point and ≥ 2 minor points and an alternative diagnosis
  - 1 major point and ≤ 1 minor point and no alternative diagnosis
  - 0 major point and ≤ 3 minor points and an alternative diagnosis
  - 0 major point and ≤ 2 minor points and no alternative diagnosis
- Moderate
  - All other combinations

Minor Points
- History of recent trauma (< 60 days) to the symptomatic leg
- Pitting edema in symptomatic leg only
- Hospitalization within previous 6 months
- Erythema

For example, consider a 72-year-old man who presents to the emergency department with swelling of the left calf and thigh. He has a history of lung cancer and recent trauma to his left leg. Physical examination confirms that the swollen calf is 4 cm larger in circumference (10 cm below the tibial tuberosity) than the other leg and reveals pitting edema and erythema of the affected leg. Bilateral compression ultrasonography is reported as normal, and the ultrasonographer recommends outpatient follow-up. However, the patient’s pretest probability, as determined by 2 major and 3 minor criteria in the CPG, is very high—about 85%. This quantified determination of the pretest probability gives the clinician added confidence to require a more invasive and expensive venogram to confirm or rule out DVT.

The CPG by Wells and colleagues illustrates both the principles of a CPG and our new criteria. First, the study met our criteria for testing clinical diagnostic procedures: 1) an independent blind comparison between the clinical findings plus ultrasonography results compared with the diagnostic standard of venography, 2) clearly defined comparison groups (at least 1 of which was free of the target disorder), and 3) the diagnostic standard (venography) was applied to all patients regardless of the diagnostic test (clinical findings plus results of the ultrasonography) results. In addition, the study meets our new criteria for CPGs: The investigators showed that their CPG remained accurate at 3 centers in 2 countries (test set). Thus, the CPG was transported to a completely different setting with a prevalence and spectrum of disease different from that of the training set. Wells and colleagues used the ideal process of validation of their CPG by testing it prospectively in a different location from where it was derived. However, CPGs checked through retrospective assessment in a different and independent set of patients will also be valid and eligible for abstraction.