Update

Development and Application of Clinical Prediction Rules to Improve Decision Making in Physical Therapist Practice

Clinical prediction rules (CPRs) are tools designed to improve decision making in clinical practice by assisting practitioners in making a particular diagnosis, establishing a prognosis, or matching patients to optimal interventions based on a parsimonious subset of predictor variables from the history and physical examination.[1,2] Clinical prediction rules have been developed to improve decision making for many conditions in medical practice, including the diagnosis of proximal deep vein thrombosis (DVT),[3] strep throat,[4] coronary artery disease,[5] and pulmonary embolism.[6] Clinical prediction rules also have been developed to assist in establishing a prognosis such as determining when to discontinue resuscitative efforts after cardiac arrest in the hospital,[7] determining the likelihood of death within 4 years for people with coronary artery disease,[7] identifying children who are at risk for developing urinary tract infections,[8] and identifying the characteristics of patients who are likely to develop postoperative nausea and vomiting after anesthesia.[9]

Clinical prediction rules have recently been developed that can improve decision making in physical therapist practice. Examples include prediction rules to improve the accuracy of diagnosing ankle fractures (ie, "the Ottawa Ankle Rules")[10] and knee fractures (ie, "the Ottawa Knee Rules")[11] in people with acute injuries and to determine when to order radiographs in patients with neck trauma.[12] Other prediction rules have been developed to diagnose patients with cervical radiculopathy[13] and carpal tunnel syndrome.[14] A CPR also has been developed to establish the prognosis of patients with neck pain following a rear-end motor vehicle accident.[15]

John D Childs, Joshua A Cleland

With increasing attention focused on the rising costs of health care, CPRs provide practitioners with powerful diagnostic information from the history and physical examination that may serve as an accurate decision-making surrogate for more expensive diagnostic tests. For example, the Ottawa Ankle Rules identify only those patients in which the probability of having a fracture is sufficiently large to warrant radiographic imaging, thus reducing costs and avoiding exposing patients to unnecessary radiation.[16] Similarly, if the CPRs used to diagnose patients with cervical radiculopathy[13] and carpal tunnel syndrome[14] are eventually validated, the demand for electrodiagnostic testing may be reduced, potentially saving costs and avoiding the discomfort and anxiety associated with these procedures.

In addition to their diagnostic utility, CPRs pertinent to physical therapist practice have recently been developed to assist with subgrouping patients into specific classifications that are useful in guiding management strategies. For example, CPRs have been developed to help practitioners match patients to optimal treatment approaches such as spinal manipulation[17,18] and a lumbar stabilization exercise program.[19] An advantage of CPRs is that they use the diagnostic properties of sensitivity, specificity, and positive and negative likelihood ratios (LR); thus, their interpretation can be readily applied to individual patients.[1] Although helpful for guiding the early stages of treatment and assigning patients to a particular classification, they are not always useful for prescribing the exact treatment techniques to be used within the context of the patient's assigned classification.

Because CPRs are designed to improve decision making, it is important that they be developed and validated according to rigorous methodological standards. McGinn et al[1] have suggested a 3-step process for developing and testing a CPR prior to widespread implementation of the rule in clinical practice. The purpose of this update is to describe the different steps involved in developing and validating CPRs and illustrate how CPRs can be used to improve decision making in physical therapist practice.

The First Step: Creating the Clinical Prediction Rule

The initial step in the development of a CPR involves creation of the rule (Fig. 1). Researchers and practitioners may initially brainstorm to develop a list of all possible factors that they believe have some predictive value for identifying the condition of interest. Ultimately, a reasonable list of predictors are selected for consideration based on clinical experience and previous research, which demonstrates that the factor or set of factors has some diagnostic or prognostic accuracy. Although it may be ideal to include every possible factor from the clinical examination to ensure that no possible predictor variables are overlooked, the researcher must weigh the benefits of including a complete set of potential predictor variables against the increase in sample size required for each additional variable under consideration. Some authors[20,21] have recommended that 10 to 15 subjects should be enrolled into the study to identify one predictor variable.

The sample size also must be judged in the context of the risks and benefits of decision making based on the rule and the prevalence of a particular phenomenon. For example, there may be significant consequences associated with the failure to identify a clinically relevant cervical spine injury in a patient who has sustained neck trauma or with the failure to identify the presence of an ankle fracture. These studies, therefore, tend to enroll thousands of patients to achieve sufficiently narrow

JD Childs, PT, PhD, MBA, OCS, FAAOMPT, is Assistant Professor and Director of Research, US Army-Baylor University Doctoral Program in Physical Therapy, Fort Sam Houston, San Antonio, Tex. Address all correspondence to Dr Childs at 508 Thurber Dr, Schertz, TX 78154 (USA) (childsjd@sbcglobal.net).

JA Cleland, DPT, OCS, is Assistant Professor, Department of Physical Therapy, Franklin Pierce College, Concord, NH, and Research Coordinator for Rehabilitation Services of Concord Hospital, Concord, NH.
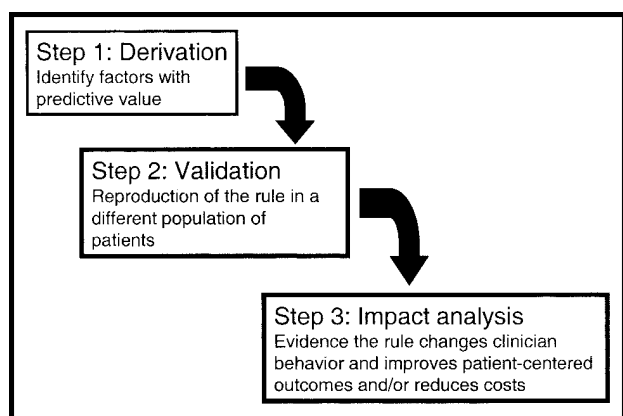
Foundation for
Physical Therapy

Figure 1.
Steps in the development of a clinical prediction rule.

confidence intervals so that practitioners can be virtually certain that application of the rule will not lead to an error in decision making. These injuries also are relatively rare in light of the total number of traumatic injuries; thus, larger sample sizes are necessary to observe a sufficient number of events on which to base the accuracy calculations.

On the other hand, although failing to identify a patient likely to benefit from a specific treatment approach such as spinal manipulation may result in a less-than-optimal outcome or a delay in improvement, the patient is unlikely to have a serious complication. Furthermore, the pretest probability of achieving a successful outcome with spinal manipulation (ie, the probability associated with a successful outcome before considering the patient's status on the rule) was 45%, which is considerably higher than the prevalence of a clinically relevant cervical spine injury or ankle fracture among individual patients with an acute injury. This also permits a smaller sample size because fewer cases are necessary to observe a sufficient number of events (ie, a successful outcome) to characterize the accuracy of decision making within an acceptable level of confidence. The development of CPRs, therefore, requires the researcher to consider the prevalence that a particular event will occur and then balance the benefits of achieving ever more narrower confidence intervals against the additional costs associated with recruiting an increasingly large sample size.

Once the initial set of possible predictor variables is established, patients are examined to determine the presence or absence of each predictor variable at baseline. To minimize bias, it is essential that the examiner be blinded from knowing whether the patient actually has the condition of interest. For example, in the development of the spinal manipulation CPR, a variety of demographic, historical, and physical examination findings were considered.[18] Patients with low back pain first completed several questionnaires consisting of self-report measures of pain and function. Factors such as the mechanism of injury, nature of current symptoms, distal extent of symptoms, and previous episodes of low back pain were considered as possible predictor variables. In addition, psychosocial considerations such as fear-avoidance beliefs and nonorganic signs and symptoms were considered. Physical examination findings that were considered to be potential predictor variables included hip and lumbar spine range of motion, lumbar spine mobility testing, and a variety of traditional landmark symmetry and provocation tests purported to identify dysfunction in the lumbopelvic region.

Applying the Reference Criterion

After patients are examined at baseline for the presence of the possible predictor variables, a second examiner who is blinded to the results of the clinical examination should then establish whether the patient actually has the condition of interest according to a standardized and well-accepted reference criterion (ie, "gold standard" or "reference standard"). Reid et al[22] suggested that an appropriate reference criterion is one that accurately represents the condition the diagnostic test is attempting to identify. For the spinal manipulation CPR, the purpose was to "diagnose" patients with low back pain who were likely to achieve a dramatic improvement from spinal manipulation after 1 week.[18] The reference criterion, therefore, was based on response to a standardized manipulative intervention according to a predetermined clinically relevant cutoff score. In this case, patients who achieved at least 50% improvement on the Oswestry Low Back Pain Disability Questionnaire, a patient's perceived level of disability, were considered to have achieved a successful outcome. Previous research[23–25] has shown that 50% improvement on the Oswestry Low Back Pain Disability Questionnaire distinguishes between patients responding to manipulation versus those simply benefiting from the favorable natural history of low back pain.

During development of the Ottawa Ankle Rules[10] and Ottawa Knee Rules,[11] radiographs were the reference criterion to determine whether an ankle or knee fracture was present. Neural conduction using electrodiagnostic testing based on well-established guidelines for the diagnosis of carpal tunnel syndrome and cervical radiculopathy was used as the reference criterion to determine which patients actually had the condition.[13,14] The credibility and usefulness of the CPR that is eventually developed hinges upon the selection of an appropriate and clinically meaningful reference criterion.

Data Analysis

Once the collection of possible predictor variables and blinded determination of the presence of the condition of interest have been completed, the data can be ana-

lyzed. Although other techniques may be considered, logistic regression is a commonly used statistical approach to determine the most parsimonious set of predictor variables in a multivariate CPR that maximizes the accuracy of diagnosing the condition of interest.[1,2] The details of how to conduct logistic regression are beyond the scope of this article, but a more detailed discussion of this topic is presented in Kleinbaum et al.[26]

In general, the accuracy of CPRs is best expressed using diagnostic accuracy statistics such as sensitivity, specificity, and positive and negative LRs. Detailed definitions of these terms are described elsewhere,[27] but a brief review is provided here. With respect to CPRs, *sensitivity* is the proportion of patients with the condition or outcome of interest who are positive on the CPR (ie, true positive rate). It reflects the ability of a test to identify the condition when present. *Specificity* is the proportion of patients who do not have the condition or outcome of interest and are negative on the CPR (ie, true negative rate).[28] It reflects the ability of a test to recognize when the condition or outcome of interest is absent. Likelihood ratios combine the information from sensitivity and specificity. A *positive LR* expresses the change in odds favoring the diagnosis or outcome when the patient satisfies the criteria of the CPR (ie, to rule in the diagnosis), whereas a *negative LR* expresses the change in odds favoring the diagnosis or outcome when the patient does not satisfy the rule's criteria (ie, to rule out the diagnosis).[28] An accurate CPR, therefore, would have a large positive LR to rule in the diagnosis, or a small negative LR to rule out the diagnosis. According to Jaeschke et al,[29] accuracy can be considered moderate when a positive LR is greater than 5.0 or a negative LR is less than 0.20. Accuracy is substantial when a positive LR is greater than 10.0 or a negative LR is less than 0.10.

The Second Step: Validating the Clinical Prediction Rule

Before a CPR can be recommended for use in clinical practice, it is necessary to validate the CPR in a "test set" or "validation set" to ensure that similar results are replicated in a different population of patients or in a different health care setting (Fig. 1).[28] It is possible that some of the predictor variables that emerged in the development phase may have occurred by chance.[31] This is because the strategy of identifying predictive variables may not consider whether the factors identified are biologically plausible. The data could reveal a "biologically nonsensical and random, non-causal quirk" to be predictive of a given outcome, analogous to a type I error in classic hypothesis testing.[28] Examining the criteria for face validity also can shed light on whether the predictors make sense. For example, it seems intuitive that patients who are likely to benefit from spinal manipulation may tend to have more acute symptoms, a

Table 1.
Clinical Prediction Rule for Diagnosing Deep Vein Thrombosis (DVT)[32,a]

Clinical Finding	Score
Active cancer (treatment ongoing, within previous 6 mo, or palliative)	1
Paralysis, paresis, or recent plaster immobilization of the lower extremities	1
Recently bedridden for >3 d or major surgery within 4 wk	1
Localized tenderness along the distribution of the deep venous system[b]	1
Entire lower extremity swelling	1
Calf swelling >3 cm when compared with the asymptomatic lower extremity[c]	1
Pitting edema (greater in the symptomatic lower extremity)	1
Collateral superficial veins (nonvaricose)	1
Alternative diagnosis as likely or greater than that of proximal DVT[d]	−2

[a] Probability of having a proximal DVT: 0=low, 1–2=moderate, and ≥3=high.
[b] Tenderness along the deep venous system is assessed by firm palpation in the center of the posterior calf, the popliteal space, and along the area of the femoral vein in the anterior thigh and groin.
[c] Measured with a tape measure 10 cm below tibial tuberosity.
[d] More common alternative diagnoses are cellulitis, calf strain, Baker cyst, or postoperative swelling.

distribution of symptoms that does not extend distal to the knee, low fear-avoidance scores, and some degree of stiffness in the lumbar spine; however, there is still no guarantee these factors will persist in a different group of patients.

Predictors identified in a CPR's development phase also may be unique to a particular population of patients or the practitioners who participated in the study. If this is the case, the predictor variables may not be generalizable to other patient populations or different practitioners.[30] For example, Wells and colleagues[31–34] have developed a CPR to identify patients suspected of having a proximal DVT (Tab. 1). A patient's overall score on the rule is obtained by adding each item judged to be positive. A score of 0 corresponds to a low probability of having a proximal DVT, a score of 1 or 2 corresponds to a moderate probability, and a score that is greater than or equal to 3 corresponds to a high probability that a proximal DVT exists.[32]

Until recently, studies attempting to validate the rule included a heterogeneous groups of outpatients with a broad range of diagnoses, making it unclear whether the rule was accurate specifically among outpatients with orthopedic conditions, one of the subgroups most likely to experience a proximal DVT.[3] Riddle et al[3] addressed this question in a validation study by considering the accuracy of the rule in a group of patients with exclusively orthopedic conditions, including patients who

were recovering from lower-extremity surgery, who had sustained a traumatic injury or fracture, or who had been diagnosed with one or more soft tissue disorders of the spine or lower extremity.[3] The most common diagnoses included patients who had sustained orthopedic trauma, who were recovering from orthopedic surgery, or who had experienced a ruptured Baker cyst.[3] Importantly, estimates as to whether these patients had a low, moderate, or high risk for proximal DVT were consistent with the estimates derived in the heterogeneous group of patients with a broad range of diagnoses, providing validity for the rule's use among outpatients with orthopedic conditions suspected to have proximal DVT.[3] Specifically, 5.6% (95% confidence interval=3.5%–8.7%) of patients in the low probability group, 14.1% (95% confidence interval=8.6%–22.4%) in the moderate probability group, and 47.4% (95% confidence interval=35.3%–60%) in the high probability group were found to have a proximal DVT. The fact that the confidence intervals in the different probability categories did not overlap also was offered as evidence to support the CPR's validity in this subgroup of patients.[3] The results of this study highlight the need for validation studies and suggest that orthopedic surgeons, physical therapists, and other health care professionals should routinely use the CPR to screen outpatients with orthopedic conditions who are at risk for proximal DVT.

A final reason why it is important to perform validation studies is that a different group of practitioners may fail to accurately apply the CPR or may perform the tests and measures used to determine the presence of predictor variables in the CPR differently than in the initial study.[30] Therefore, training practitioners in how to perform the examination and treatment procedures in validation studies is essential to eliminate the possibility that a useful CPR will fail in a validation study. For example, the recent validation study for the spinal manipulation CPR[17] involved 14 physical therapists from a variety of different settings. All treating clinicians underwent formal training in study procedures and performance of the manipulation technique, further enhancing the study's internal validity.

The Third Step: Conducting an Impact Analysis

Ultimately, a CPR is useful only to the extent that it can improve clinically relevant outcomes, increase patient satisfaction, and decrease costs once it is implemented into the realities of busy clinical practice. The final step in the development of a CPR, therefore, involves assessing the impact of its implementation on practice patterns, outcomes of care, and costs (Fig. 1). Impact analysis studies are performed in primarily 1 of 3 ways. Ideally, individual patients would be randomly assigned to either receive care based on the CPR or have decisions made that are based on standard practice. However, this

requires that practitioners alternate back and forth between decision making based on the CPR and decision making based on usual care, which may not be clinically feasible in busy practice settings. An alternative approach is to randomly assign clinical sites to either apply the CPR or not for all patients. A third alternative is to use a nonrandomized before-and-after design in which similar outcomes are assessed within the same practice setting both before and after the CPR's implementation. Although the latter is a reasonable approach, the inference of the findings is clearly stronger with the randomized design.

The Ottawa Ankle Rules provide a good example for how successful impact analysis studies can be performed.[16,35,36] Auleley et al[35] randomly assigned 6 emergency departments to either apply the Ottawa Ankle Rules or use conventional models of decision making. Using a traditional approach, ankle radiographs were ordered in 99.6% of patients compared with only 78.9% of patients when decision making was based on the CPR. Although the emergency departments using the Ottawa Ankle Rules missed 3 ankle fractures, none were associated with an adverse outcome.

In a nonrandomized before-and-after design, Stiell et al[16] demonstrated a 28% reduction in the utilization of ankle radiographs and a 14% reduction in foot radiographs upon implementation of the Ottawa Ankle Rules compared with a control hospital not trained in the implementation of the rule ($P<.001$). Compared with patients who received radiographs but did not have a fracture, patients discharged without radiography also spent significantly less time in the emergency department (80 minutes versus 116 minutes) ($P<.0001$) and had lower estimated total medical costs ($62 versus $173) ($P<.001$). The percentage of patients who were satisfied with their care was similar between the groups (95% versus 96%). Importantly, these results were achieved without compromising the quality of care and were maintained over a 12-month period after the formal trial.[36] Similar reductions in utilization, costs of care, and waiting times without compromising patient satisfaction or quality of care have been found with the implementation of the Ottawa Knee Rules.[37,38] One should also assess the validity of data obtained with CPRs in different health care settings (eg, academic medical center versus health maintenance organization setting versus rural setting) to determine whether the rule performs similarly between different health care settings.

Using Clinical Prediction Rules to Improve Decision Making

The application of CPRs in physical therapist practice can have important implications for decision making and can assist practitioners in making informed deci-

sions about potential treatment approaches. It is important to recognize that the primary diagnostic accuracy statistic of interest will vary depending on the CPR's purpose. The calculation of LRs affords equal weight to false positive and false negative findings, both of which may lead to an error in decision making. Maximizing the LR, therefore, is not the best choice when one finding may have more severe consequences than the other. For example, with the Ottawa Ankle Rules, practitioners do not want to overlook patients with an ankle fracture; therefore, it was essential to minimize the number of false negative findings. These false negative findings would be patients for whom the rule suggests radiographs are unnecessary, when in fact they have an ankle fracture. Maximizing the positive LR in this case would maximize the efficiency of ordering radiographs at the expense of missing several patients in which a fracture in fact exists. Clearly, the nominal cost associated with ordering a few additional radiographs judged to be normal (ie, a false positive finding) is well worth the potential risks associated with missing a fracture (ie, a false negative finding). In these instances, the primary statistic of interest is to maximize the CPR's sensitivity, so that being negative on the CPR virtually eliminates the possibility that an ankle fracture exists. Practitioners, therefore, can be confident in a decision not to order radiographs for patients who are negative on the CPR, because they know that they are not exposing their patients to increased risk of an adverse complication from a failure to detect a fracture.

A similar rationale has been used in the development of CPRs to determine when radiographs are needed for patients who have a serious knee injury[11,37,39] or trauma to the neck.[12] Readers are referred to the text by Sackett et al[28] for a detailed discussion on the use of LRs to determine changes in the probability that a patient has the condition or outcome of interest. However, a nomogram published by Fagan[40] is a useful tool that can be used to easily make the conversions, and thus is suitable for use in a busy clinical practice. A description of how to use this method has been published elsewhere.[27]

The use of CPRs to select treatment approaches for individual patients can be illustrated by the spinal manipulation and lumbar stabilization CPRs. For the spinal manipulation CPR, 45% of patients achieved at least a 50% improvement in their Oswestry Low Back Pain Disability Questionnaire score regardless of whether they satisfied the criteria on the CPR.[18] That is, if practitioners were to randomly provide manipulation to patients with nonradicular low back pain, approximately 45% of patients will have at least a 50% improvement in disability within 1 week. The study sought to identify patients who would likely benefit from spinal manipulation; thus, the statistic of interest was the

Table 2.
Criteria in the Spinal Manipulation Clinical Prediction Rule[17,18]

Criterion	Definition of Positive
Duration of current episode of low back pain	<16 d
Extent of distal symptoms	Not having symptoms distal to the knee
Fear-Avoidance Beliefs Questionnaire work subscale score	<19 points
Segmental mobility testing	At least one hypomobile segment in the lumbar spine
Hip internal rotation range of motion	At least one hip with >35° of internal rotation range of motion

positive LR. The positive LR was 24.4 for patients who met at least 4 of the 5 criteria (Tab. 2). To put this result in perspective, the probability of achieving a successful outcome increases from 45% to 95%; therefore, practitioners can be increasingly certain that patients who meet at least 4 out of the 5 criteria in the CPR will achieve at least 50% improvement in disability by the end of 1 week.

With 3 criteria present, the positive LR was 2.6, which translates into a 68% probability of success. Given the ease with which this manipulation technique can be performed and in light of the extremely low risks,[41] however, an attempt at spinal manipulation may still be warranted. The validation study resulted in similar findings to the derivation study.[18] Based on a pretest probability of success of 44% and a positive LR of 13.2, a patient who is positive on the rule and treated with manipulation had a 92% chance of achieving a successful outcome by the end of 1 week.[17] A patient's status on the rule was of little relevance in determining the outcome of patients treated with an exercise intervention, supporting the notion that the rule is specifically predicting a response to spinal manipulation.[17]

A similar CPR has been developed to identify patients who are likely to benefit from a lumbar stabilization exercise approach (Tab. 3).[19] Approximately 33% of the subjects had at least 50% improvement on the Oswestry Low Back Pain Disability Questionnaire after 8 weeks of a standard exercise regimen. Therefore, if practitioners were to randomly prescribe lumbar stabilization exercises for patients with LBP, approximately 33% of patients will achieve at least 50% improvement in disability by the end of 8 weeks. However, the positive LR was 4.0 among patients with at least 3 out of 4 predictors of success; thus, the probability of a achieving at least 50% improvement increases to 67%. The researchers also identified the likelihood that a patient would have

Table 3.
Criteria in the Lumbar Stabilization Clinical Prediction Rule for the Prediction of Successful and Unsuccessful Outcome[19]

Prediction of Success	Prediction of Failure
Positive prone instability test	Negative prone instability test
Aberrant movement present	Aberrant movement absent
Average straight leg raise (>91°)	Fear-Avoidance Beliefs Questionnaire physical activity subscale (<9)
Age (<40 y)	No hypermobility with lumbar spring testing

improvement (defined as patients who exhibited less than 50% improvement on the Oswestry Low Back Pain Disability Questionnaire, but achieved an improvement that exceeded the minimal clinically important difference of 6 points with the lumbar stabilization program). The researchers determined that if a patient satisfied 3 or more of the criteria, the likelihood that the patient would exhibit clinically meaningful improvement with lumbar stabilization was 97%.[19]

Clinical prediction rules also can be used to help practitioners determine when a particular treatment approach may *not* be beneficial. For example, the probability of a successful outcome among patients with less than 3 out of 5 criteria on the spinal manipulation CPR is essentially no better than if a clinician were to randomly provide manipulation to patients with nonradicular low back pain (Tab. 2). In a recent validation study,[17] the negative LR for patients who met fewer than 3 of the criteria in the CPR was 0.10 (95% confidence interval=0.03–0.41), reducing the posttest probability of success to only 7.4%. With fewer than 2 criteria met, the posttest probability of responding to manipulation approaches 0%, suggesting that practitioners may want to consider other treatment approaches with a higher probability of success.

Decision making can be further enhanced by recent evidence identifying clinical variables associated with the failure of patients with LBP to improve from spinal manipulation.[42] A CPR also has been developed to identify patients who are likely to fail to respond to a lumbar stabilization exercise approach (Tab. 3).[19] In this study,[19] *failure* was defined as less than 6 points of improvement on the Oswestry Low Back Pain Disability Questionnaire, which has been shown to be the minimum clinically important difference for this instrument.[43] Using this criterion, approximately 28% of patients with low back pain will fail to improve at least 6 points on the Oswestry Low Back Pain Disability Questionnaire after 8 weeks. The likelihood of improvement with a lumbar stabilization program decreased to 32%

Answers "yes" to the question "Do your symptoms improve with moving, 'shaking,' or positioning your wrist or hands?"
Has a wrist ratio index >.67
Has a Symptom Severity Scale score >1.9
Has diminished sensation in the median nerve distribution of the thumb
Is >45 y of age

Figure 2.
Criteria in the carpal tunnel syndrome clinical prediction rule.[14]

Positive Spurling test
Positive distraction test
Positive upper-limb tension test
Presence of <60° of cervical rotation range of motion to the involved side

Figure 3.
Criteria in the cervical radiculopathy clinical prediction rule.[13]

among patients with one or none of the variables in the CPR.

Using CPRs to improve diagnostic decision making can be illustrated by the prediction rules designed to improve the accuracy of diagnosing cervical radiculopathy[13] and carpal tunnel syndrome. In these cases, the primary aim was to rule in the diagnosis of these conditions; thus, the primary statistic of interest was the positive LR. The most powerful combination of factors for ruling in the diagnosis of carpal tunnel syndrome is illustrated in Figure 2. Based on a pretest probability of 34% that the patient may have carpal tunnel syndrome and a positive LR of 4.6 when at least 4 out of 5 findings are present, the posttest probability of having carpal tunnel syndrome increases to 70%. However, given a positive LR of 18.3 when all 5 findings are present, the posttest probability jumps to 90%. An understanding of the patient's status on the rule, therefore, can help inform the diagnostic process for determining whether the patient has carpal tunnel syndrome.

Similar decision-making principles can be applied to the cervical radiculopathy CPR (Fig. 3). Given a pretest probability of 23% and a positive LR of 6.1 for patients with at least 3 out of 4 findings present, the probability of having a cervical radiculopathy is increased to 65%. When all 4 findings are present, a positive LR of 30.3 increases the posttest probability to 90%, increasing the level of confidence that the patient has a cervical radiculopathy.

Another advantage of CPRs is that, in addition to decision making based on the patient's overall status on the rule, the individual variables that make up the rule continue to be useful. For example, the wrist ratio index, which is calculated by dividing the anteroposterior wrist

Table 4.
Levels of Evidence for Clinical Prediction Rules[45]

Level	Requirements	Extent of Use
I	At least one prospective validation in a different population plus one impact analysis demonstrating change in clinician behavior with beneficial consequences	Rule can be used in a wide variety of settings with confidence that it change clinician behavior and improve outcomes
II	Validation in one large prospective study, including a broad spectrum of patients and clinicians, or in several similar settings that differ in geographical location and experience levels of clinicians	Rule can be used in various settings with confidence
III	Validated in only one narrow prospective sample	Rule can be used only with caution and only if patients in the study are similar to those in the clinician's setting
IV	Derived but not validated, or validated only in split samples or large retrospective databases or by statistical techniques	Further evaluation required before rule can be applied clinically

width by the mediolateral wrist width based on sliding caliper measurements and is thought to be an indicator of carpal canal volume, was the most useful test for ruling out carpal tunnel syndrome when the value exceeded 0.67 (negative LR=0.29).[14] The upper-limb tension test described by Elvey[44] was the best screening test for establishing a diagnosis of cervical radiculopathy.[13] With a negative LR equal to 0.12, a negative upper-limb tension test essentially rules out the presence of cervical radiculopathy.

Hierarchy of Evidence for Clinical Prediction Rules

McGinn et al[45] have established an evidence hierarchy to assist practitioners in determining whether the CPR is appropriate for use in the decision-making process. Clinical prediction rules in which the rule has been derived but not yet validated are classified as level IV (Tab. 4). For example, the carpal tunnel syndrome and cervical radiculopathy prediction rules currently correspond to a level IV CPR, representing the first step in the development of the rule. Validation studies are necessary before these rules can be recommended for widespread use in clinical practice. A level III CPR is one that has only been validated in one narrow prospective sample, and thus should be used with caution among patients in a similar practice setting (Tab. 4).[45]

Clinical prediction rules cannot be recommended for widespread implementation until at least one large prospective validation study in a broad spectrum of patients and practitioners has been carried out in a variety of practice settings. For example, the spinal manipulation CPR validation study incorporated numerous clinicians with a variety of experience levels working in different health care settings; thus, the rule's increased generalizability qualifies it as level II on the evidence hierarchy (Tab. 4).[45] This increases the practitioner's confidence that the spinal manipulation CPR can be used in a broad spectrum of patients with low back pain to improve decision making and patient outcomes. The validation study of the CPR to identify outpatients with orthopedic conditions who have a proximal DVT also is consistent with a level II CPR,[3] increasing a practitioner's confidence in the rule's accuracy among patients with outpatient orthopedic conditions.

A level I CPR corresponds to the highest level of evidence, and requires at least one prospective validation study in a different population plus results from an impact analysis study showing improvements in practice patterns, outcomes of care, and costs (Tab. 4).[45] For example, practitioners can be quite confident that decision making based on the Ottawa Ankle Rules will not only be accurate in a wide variety of health care settings, but will also reduce the rate at which ankle radiographs need to be ordered, thus reducing health care costs.

The Ultimate Goal: Changing Practitioner Behavior to Improve Outcomes of Care

It seems reasonable to believe that publishing evidence for a level I or level II CPR should be sufficient to change practice patterns; however, despite their intuitive attraction, changing a practitioner's behavior to be consistent with the evidence is a difficult task.[46] Even having a level I CPR such as the Ottawa Ankle Rules does not guarantee that it can be easily incorporated into clinical practice. Cameron and Naylor[47] found no change in the use of ankle radiography among emergency department physicians who had been trained in the use of the Ottawa Ankle Rules. The challenge for practitioners is to find an effective means to implement CPRs in a busy clinical setting. Practitioners are required to recall the individual factors in the CPR and how to examine patients with respect to each criterion, and they must remember them in the overall context of the decision-making process to maximize the accuracy of their use. Therefore, CPRs that have too many predictor variables may be burdensome for the practitioner to remember and apply in clinical practice. Unless practitioners are confident that the CPR

is easy to use and will improve costs or outcomes of care, systematic implementation may be difficult to achieve. "Magic bullet" strategies to change practitioner behavior do not appear to exist[46]; thus, efforts should be made to utilize effective implementation strategies that encourage standardized practice patterns that are consistent with the evidence.[48–50] Clearly, customized strategies tailored to the unique circumstances of each practice setting are needed.

Conclusion

Clinical prediction rules have the potential to improve outcomes, increase patient satisfaction, and decrease costs of care in physical therapist practice. They can be useful tools to save practitioners valuable time and to better inform patients about their diagnosis or prognosis. Their development has helped to reverse common misperceptions among health care professionals that diagnostic tests such as radiographs or laboratory tests provide "hard" data useful for decision making, whereas information from the patient history is more "soft" and not as useful. Importantly, they also may be useful to increase the power of clinical research by permitting researchers to study more homogenous groups of patients.

References

1 McGinn TG, Guyatt GH, Wyer PC, et al; Evidence-Based Medicine Working Group. Users' guides to the medical literature, XXII: how to use articles about clinical decision rules. *JAMA.* 2000;284:79–84.

2 Laupacis A, Sekar N, Stiell IG. Clinical prediction rules: a review and suggested modifications of methodological standards. *JAMA.* 1997;277: 488–494.

3 Riddle DL, Hoppener MR, Kraaijenhagen RA, et al. Preliminary validation of clinical assessment for deep vein thrombosis in orthopaedic outpatients. *Clin Orthop.* 2005;432:252–257.

4 Ebell MH, Smith MA, Barry HC, et al. The rational clinical examination: does this patient have strep throat? *JAMA.* 2000;284:2912–2918.

5 Morise AP, Haddad WJ, Beckner D. Development and validation of a clinical score to estimate the probability of coronary artery disease in men and women presenting with suspected coronary disease. *Am J Med.* 1997;102:350–356.

6 Ferreira G, Carson JL. Clinical prediction rules for the diagnosis of pulmonary embolism. *Am J Med.* 2002;113:337–338.

7 van Walraven C, Forster AJ, Stiell IG. Derivation of a clinical decision rule for the discontinuation of in-hospital cardiac arrest resuscitations. *Arch Intern Med.* 1999;159:129–134.

8 Gorelick MH, Shaw KN. Clinical decision rule to identify febrile young girls at risk for urinary tract infection. *Arch Pediatr Adolesc Med.* 2000;154:386–390.

9 Sinclair DR, Chung F, Mezei G. Can postoperative nausea and vomiting be predicted? *Anesthesiology.* 1999;91:109–118.

10 Stiell IG, Greenberg GH, McKnight RD, et al. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. *Ann Emerg Med.* 1992;21:384–390.

11 Stiell IG, Greenberg GH, Wells GA, et al. Derivation of a decision rule for the use of radiography in acute knee injuries. *Ann Emerg Med.* 1995;26:405–413.

12 Stiell IG, Wells GA, Vandemheen KL, et al. The Canadian C-spine rule for radiography in alert and stable trauma patients. *JAMA.* 2001;286:1841–1848.

13 Wainner RS, Fritz JM, Irrgang JJ, et al. Reliability and diagnostic accuracy of the clinical examination and patient self-report measures for cervical radiculopathy. *Spine.* 2003;28:52–62.

14 Wainner RS, Fritz JM, Irrgang JJ, et al. Development of a clinical prediction rule for the diagnosis of carpal tunnel syndrome. *Arch Phys Med Rehabil.* 2005;86:609–618.

15 Hartling L, Pickett W, Brison RJ. Derivation of a clinical decision rule for whiplash associated disorders among individuals involved in rear-end collisions. *Accid Anal Prev.* 2002;34:531–539.

16 Stiell IG, McKnight RD, Greenberg GH, et al. Implementation of the Ottawa ankle rules. *JAMA.* 1994;271:827–832.

17 Childs JD, Fritz JM, Flynn TW, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med.* 2004;141:920–928.

18 Flynn T, Fritz J, Whitman J, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine.* 2002;27:2835–2843.

19 Hicks GE, Fritz JM, Delitto A, McGill SM. Predictive validity of clinical variables used in the determination of patient prognosis following a lumbar stabilization program. *Arch Phys Med Rehabil.* In press.

20 Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. *Ann Intern Med.* 1993;118:201–210.

21 Stevens J. *Applied Multivariate Statistics for the Social Sciences.* 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates; 1996.

22 Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research: getting better but still not good. *JAMA.* 1995;274:645–651.

23 Delitto A, Cibulka MT, Erhard RE, et al. Evidence for use of an extension-mobilization category in acute low back syndrome: a prescriptive validation pilot study. *Phys Ther.* 1993;73:216–222.

24 Erhard RE, Delitto A, Cibulka MT. Relative effectiveness of an extension program and a combined program of manipulation and flexion and extension exercises in patients with acute low back syndrome. *Phys Ther.* 1994;74:1093–1100.

25 Fritz JM, George S. The use of a classification approach to identify subgroups of patients with acute low back pain: interrater reliability and short-term treatment outcomes. *Spine.* 2000;25:106–114.

26 Kleinbaum DG, Kupper LL, Muller KE, Nizam A. Logistic regression analysis. In: *Applied Regression Analysis and Other Multivariable Methods.* Pacific Grove, Calif: Duxbury Press; 1998:656–686.

27 Fritz JM, Wainner RS. Examining diagnostic tests: an evidence-based perspective. *Phys Ther.* 2001;81:1546–1564.

28 Sackett DL, Straus SE, Richardson WS, et al. *Evidence-Based Medicine: How to Practice and Teach EBM.* 2nd ed. New York, NY: Churchill Livingstone Inc; 2000.

29 Jaeschke R, Guyatt GH, Sackett DL; The Evidence-Based Medicine Working Group. Users' guides to the medical literature, III: how to use an article about a diagnostic test, B: What are the results and will they help me in caring for my patients? *JAMA.* 1994;271:703–707.

30 Guyatt G, Rennie D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice.* Chicago, Ill: AMA Press; 2002.

31 Wells PS, Hirsh J, Anderson DR, et al. Accuracy of clinical assessment of deep-vein thrombosis. *Lancet.* 1995;345:1326–1330.

32 Wells PS, Anderson DR, Bormanis J, et al. Value of assessment of pretest probability of deep-vein thrombosis in clinical management. *Lancet.* 1997;350:1795–1798.

33 Wells PS, Hirsh J, Anderson DR, et al. A simple clinical model for the diagnosis of deep-vein thrombosis combined with impedance plethysmography: potential for an improvement in the diagnostic process. *J Intern Med.* 1998;243:15–23.

34 Wells PS, Anderson DR, Ginsberg J. Assessment of deep vein thrombosis or pulmonary embolism by the combined use of clinical model and noninvasive diagnostic tests. *Semin Thromb Hemost.* 2000;26: 643–656.

35 Auleley GR, Ravaud P, Giraudeau B, et al. Implementation of the Ottawa ankle rules in France: a multicenter randomized controlled trial. *JAMA.* 1997;277:1935–1939.

36 Verbeek PR, Stiell IG, Hebert G, Sellens C. Ankle radiograph utilization after learning a decision rule: a 12-month follow-up. *Acad Emerg Med.* 1997;4:776–779.

37 Stiell IG, Wells GA, Hoag RH, et al. Implementation of the Ottawa Knee Rule for the use of radiography in acute knee injuries. *JAMA.* 1997;278:2075–2079.

38 Nichol G, Stiell IG, Wells GA, et al. An economic analysis of the Ottawa knee rule. *Ann Emerg Med.* 1999;34(4 pt 1):438–447.

39 Stiell IG, Greenberg GH, Wells GA, et al. Prospective validation of a decision rule for the use of radiography in acute knee injuries. *JAMA.* 1996;275:611–615.

40 Fagan TJ. Nomogram for Bayes' theorem [letter]. *N Engl J Med.* 1975;293:257.

41 Haldeman S, Rubinstein SM. Cauda equina syndrome in patients undergoing manipulation of the lumbar spine. *Spine.* 1992;17: 1469–1473.

42 Fritz JM, Whitman JM, Flynn TW, et al. Factors related to the inability of individuals with low back pain to improve with a spinal manipulation. *Phys Ther.* 2004;84:173–190.

43 Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther.* 2001;81:776–788.

44 Elvey RL. The investigation of arm pain: signs of adverse responses to the physical examination of the brachial plexus and related tissues. In: Boyling JD, Palastanga N, eds. *Grieve's Modern Manual Therapy.* 2nd ed. New York, NY: Churchill Livingstone Inc; 1994:577–585.

45 McGinn T, Guyatt G, Wyer P, et al. Diagnosis: clinical prediction rule. In: Guyatt G, Rennie D, eds. *Users' Guide to the Medical Literature: A Manual for Evidence-Based Clinical Practice.* Chicago, Ill: AMA Press; 2002:471–483.

46 Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ.* 1995;153:1423–1431.

47 Cameron C, Naylor CD. No impact from active dissemination of the Ottawa Ankle Rules: further evidence of the need for local implementation of practice guidelines. *CMAJ.* 1999;160:1165–1168.

48 Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance: a systematic review of the effect of continuing medical education strategies. *JAMA.* 1995;274:700–705.

49 Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA.* 1999;282:1458–1465.

50 Davis D, O'Brien MA, Freemantle N, et al. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *JAMA.* 1999;282:867–874.