**Potential Pitfalls of Clinical Prediction Rules**

**What Are Clinical Prediction Rules?**

A clinical prediction rule (CPR) is a combination of clinical findings that have statistically demonstrated meaningful predictability in determining a selected condition or prognosis of a patient who has been provided with a specific treatment. CPRs are created using multivariate statistical methods, are designed to examine the predictive ability of selected groupings of clinical variables, and are intended to help clinicians make quick decisions that may normally be subject to underlying biases. The rules are algorithmic in nature and involve condensed information that identifies the smallest number of indicators that are statistically diagnostic to the targeted condition. The number of derived or validated CPRs is increasing, specifically in rehabilitation medicine where prescriptive studies have been developed for musculoskeletal interventions for low back pain, cervical pain, and knee dysfunction.

Clinical prediction rules may best be classified into three distinct groups: 1) diagnostic, 2) prognostic, and 3) prescriptive. Studies that focus on predictive factors related to a specific diagnosis are known as diagnostic CPRs. Clinical prediction rules that are designed to predict an outcome such as success or failure are considered prognostic. Clinical prediction rules designed to target the most effective interventions are identified as prescriptive, and these require prospective, longitudinal, randomized controlled trials that compare outcomes after selected interventions for subjects who meet a similar score on the CPR.

Clinical prediction rules are generally developed using a 3-step method. First, CPRs are derived prospectively using multivariate statistical methods to examine the predictive ability of selected groupings of clinical variables. The second step involves validating the CPR in a randomized controlled trial to reduce the risk that the predictive factors developed during the derivation phase were selected by chance. The third step involves conducting an impact analysis to determine the extent that the CPR improves care, reduces costs, and accurately defines the targeted objective.

Although there is little debate that carefully constructed CPRs can improve clinical practice, to my knowledge, there are no guidelines that specify methodological requirements for CPRs for infusion into all clinical practice environments. Guidelines are created to improve the rigor of study design and reporting. The following editorial outlines potential methodological pitfalls in CPRs that may significantly weaken the transferability of the algorithm. Within the field of rehabilitation, most CPRs have been prescriptive; thus, my comments here are reflective of prescriptive CPRs.

**Methodological Pitfalls**

CPRs are designed to specify a homogeneous set of characteristics from a heterogeneous population of prospectively selected consecutive patients. Typically, the resulting applicable population is a small subset of a larger sample and may only represent a small percentage of the clinician’s actual daily caseload. The setting and location of the larger sample should be generalizable, and subsequent validity studies require assessment of the CPR in different patient groups, in different environments, and with a typical patient group seen by most clinicians. Because many CPRs are developed based on a very distinct group, that may or may not be reflective of a typical population of patients, the spectrum transportability of many current CPR algorithms may be limited.

Clinical prediction rules use outcome measures to determine the effectiveness of the intervention. Outcome measures must have a single operational definition and require enough responsiveness to truly capture appropriate change in the condition; in addition, these measures should have a well-constructed cut-off score and be collected by a blinded administrator. The selection of an appropriate anchor score for measurement of actual change is currently debated. Most outcome measures use a patient recall-based questionnaire such as a global rating of change score (GrRoC), which is appropriate when used in the short term but suffers from recall bias when used in long-term analyses. Other studies may use minimally detectable change scores that were originally validated using the GROC and also may be affected by both recall bias and differences in sample severity or pathology. Lastly, outcome measures that use scores that are influenced by administrative factors (discharge date, length of stay, patient charges), socio-demographic factors, or internal behavioral characteristics (changes in fear avoidance or attitude) are not consistent among populations.

A potential drawback for CPRs is the failure to maintain the quality of the tests and measures used as predictors in the algorithm. The prospective test and measures should be independent of one another during modeling; each should be performed in a meaningful, acceptable manner; and clinicians or data administrators should be blinded to the patient’s outcomes measures and condition. Further development of these recommendations is needed to improve the meaningfulness and validity of the preoperative care literature.
thermore, the tests should demonstrate acceptable reliability (≥ 0.60) and require administration within an acceptable timeframe of the outcome measure; equivocal or indeterminate results necessitate reporting. Recognizing the likelihood of a true positive finding in the absence of any information will avoid the representative heuristic pitfall that may compel us toward identifying a clinical test as positive simply because the result fits the pattern of other findings. CPRs that use tests and measures with reliability or agreement below 0.60 may result in variable findings depending on the clinician who performs the examination and depending on the findings of other tests and measures.

It is my impression that the most frequent current pitfall of CPRs is associated with the failure to meet statistical assumptions during regression modeling. CPRs are typically underpowered falling below the suggested requirements of 10 to 15 subjects for each prospective predictor variable. Validation cohorts require sampling sizes of 100 or greater with use of logistic regression (used as a standard for CPR assessment). Rarely is the statistical significance of the model reported in the rehabilitation-based CPRs, nor is the R² or R²-equivalent of the model identified. An R² or R²-equivalent outlines the strength of association of the predictor variables (both independently and as a group) in explaining the variance of the outcome measure. Low R² or R²-equivalents may suggest that other variables more accurately predict the outcome of the study and generally suggest a low effect size of the independent variables identified and retained in the analyses. Most CPRs do report confidence intervals, and when reported, wide confidence intervals imply poor precision or too small of a sample size.

Once a CPR is developed, it is important to recognize the true benefit of the tool. It has been suggested that for true impact on clinical practice, CPRs should provide a LR+ of 5 or greater. CPR derivations performed on high-risk groups, where failure to provide the appropriate intervention is highly undesirable, should have sensitivity values that are greater than specificity values. This indicates that the final algorithm will accurately provide all of the best treatment(s) possible versus assuring that only those specific to the problem are used.

CPRs should have clinical sensibility. Clinical sensibility implies that the tool makes inherent clinical sense, that it's easy to use, that the tests and measures are truly related to the outcome, and that clinician perception does not overly alter the findings of the tool. Consequently, tests and measures that vary in clinical interpretation (e.g., spring tests of the spine) or that are potentially explained by factors beyond the original scope of the examination (e.g., hip osteoarthritis when addressing hip procedures that affect the knee) may not be as useful as factors that are more explicit during clinical assessment.

Lastly, most rehabilitation-related CPRs are derivation studies, which are the initial steps in the development of clinical decision rules. Derivation studies lack validation and require follow-up studies in diverse centers with different populations of patients and different clinicians. Whether the findings from a derivation study stand up to the scrutiny of further assessment is unknown. In essence, adoption of a derivation-only CPR runs the risk of improper treatment. Careful attention should be made before blindly adopting derivation studies or basing treatment pathways on these tools.

**Summary**

Is this editorial an attack on clinical prediction rules? Actually, it’s quite the contrary. Prescriptive CPRs are useful tools for a select and discrete population of patients. As manually oriented clinicians, we have long realized that sub-sets of the population benefit from manual therapy more so than others. CPRs allow us to isolate a sub-set of desired patient characteristics and to define which techniques are most useful for that population. The current rehabilitation-based CPRs have opened the door for additional research to improve our accuracy as clinicians. Unfortunately, many of the present rehabilitation-based CPRs may have methodological weaknesses that may allow questioning of the utility of the instrument. Although there is no such thing as a “perfect” study, better and more rigorous designs should provide additional, profound and clinically applicable findings. As a clinician and a researcher, I am an advocate of CPRs.

**REFERENCES**


MANUAL THERAPY AWARDS

To encourage research in manual therapy, Cardon Rehabilitation Products, TherEx, OPTP and The Journal of Manual & Manipulative Therapy are sponsoring prizes to be awarded based on scientific merit in three areas:

1. experimental studies,
2. case studies and
3. review articles.

Awards will be selected from all papers published in JMMT in 2008. Award winners will be announced in Volume 17, Number 1, 2009.

The Cardon Award For Excellence in a Published Research Article

To encourage research in manual therapy, Cardon Rehabilitation Products and JMMT are sponsoring a prize in the amount of $1000 for an outstanding experimental study. The goal of the experimental study is to evaluate a manual therapy assessment or treatment technique. The concept of validity is central to the experimental study. Thus within the confines of the study, the results are accurate, the method and analysis used can bear up under scrutiny and the interpretation of the findings is supported by the data collected. In addition, the conclusions reported can be generalized to practice settings and subjects outside those described in the study. The format for the experimental study also requires that the researchers make their assumptions clear, their methods repeatable and their interpretations clearly separate from the methods and results. Experimental papers provide a forum for presenting one’s own findings and conclusions and for arguing for or against competing hypotheses in manual therapy.

The TherEx Award For Excellence in a Published Case Study

To encourage case studies in manual therapy, TherEx and JMMT are sponsoring a prize of $750 for an outstanding case presentation. The goal of the case study is to report on a patient or a small group of patients that a manual therapist feels should be brought to the attention of colleagues. Specifically, the case study method integrates basic science knowledge with patient assessment and treatment techniques. Within each case, emphasis is placed on the most clinically relevant aspects of each musculoskeletal condition. In addition, a selective review of the literature is also included for each case study. In all, the cases provide the reader with information about clinical decision making. The intent is to facilitate critical thinking and to promote professional growth.

The OPTP Award For Excellence in a Published Review of the Literature

To encourage reviews of the literature in manual therapy, OPTP and JMMT are sponsoring a prize of $750 for an outstanding review article. The goal of the review article is to present a large amount of information on a subject comprehensively and efficiently. In addition to a command of the literature in a specific area of manual therapy, the writer must also apply critical appraisal skills to material that is being reviewed. Specifically, it is not enough for a review to summarize the findings of research studies and case reports. Some comment should be made on the research design and methodological quality of the work being reviewed. By combining content review with methodological critique, review articles are intended to bring clinicians and researchers up to date on the state of the art in manual therapy.