# Cohort Study Design: An Underutilized Approach for Advancement of Evidence-Based and Patient-Centered Practice in Athletic Training

Gary B. Wilkerson, EdD, ATC, FNATA\*; Craig R. Denegar, PhD, PT, ATC, FNATA†

\*Graduate Athletic Training Education Program, University of Tennessee–Chattanooga; †Department of Kinesiology, University of Connecticut, Storrs

**Objective:** Providing patient-centered care requires consideration of numerous factors when making decisions that will influence a patient's health status.

**Background:** Clinical decisions should be informed by relevant research evidence, but the literature often lacks pertinent information for problems encountered in routine clinical practice. Although a randomized clinical trial provides the best research design to ensure the internal validity of study findings, ethical considerations and the competitive culture of sport often preclude random assignment of patients or participants to a control condition.

**Clinical Advantages:** A cohort study design and Bayesian approach to data analysis can provide valuable evidence to support clinical decisions. Dichotomous classification of both an outcome and 1 or more predictive factors permits quantification of the likelihood of occurrence of a specified outcome.

technical note

**Conclusions:** Multifactorial prediction models can reduce uncertainty in clinical decision making and facilitate the individualization of treatment, thereby supporting delivery of clinical services that are both evidence based and patient centered.

Key Words: research design, Bayesian analysis, clinical prediction

■ he first sentence of the first chapter of the book Evidence-Based Sports Medicine, authored by MacCauley and Best<sup>1</sup> in 2002, asked, "evidencebased sports medicine-a contradiction in terms?" The use of research evidence to guide clinical decisions has increased over the past decade, but a conceptual change in the approach used to generate research evidence might dramatically advance the rate of improvement in the quality of health care services that are provided to athletes. Evidence-based medicine has been defined as "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients."<sup>2</sup> Another widely advocated paradigm for guiding clinical practice is patient-centered medicine. In contrast to a cognitive-rational application of the best research evidence, a patient-centered approach to medical management carefully considers the unique characteristics of individual patients, including their psychosocial needs and preferences.<sup>3</sup> Despite acknowledging the importance of patient preferences and values as important elements in definitions of evidence-based medicine,<sup>4,5</sup> the 2 clinical practice paradigms have been described as somewhat incompatible with one another.<sup>3</sup> The practice of evidencebased medicine relies heavily on disease-oriented research findings for which the *efficacy* of a treatment (that is, when administered under ideal and highly controlled conditions) has been established.<sup>6</sup> Treatment effectiveness refers to the benefit that a typical patient is likely to derive from administration of the treatment under the usual clinical circumstances,<sup>6</sup> which is well aligned with the patientcentered concept of clinical care.

Progress in advancing clinically meaningful research in athletic health care has been hindered by several factors,

including limited funding, lack of large centralized datacollection networks, and the culture of sport. The rate at which evidence is being developed to inform clinical decisions is also affected by the research methods used. A randomized clinical trial (RCT) is properly viewed as the best method to identify a cause-effect relationship between variables, but a well-designed cohort study may be a more feasible option, and, importantly, may yield evidence that is more informative to daily clinical practice. Furthermore, sole reliance on hypothesis testing for determining a difference between groups, rather than estimating the potential for a desired outcome for an individual patient, can compromise patient-centered decision making. In the quest for evidence to guide practice decisions, reliance on traditional hypothesis testing must be reconsidered as the primary mechanism for advancing evidence-based, athletecentered care. The purposes of this paper are (1) to present the advantages of cohort study design for athletic training research, (2) to review cohort study limitations and strategies for avoiding threats to data validity, and (3) to provide an overview of the Bayesian approach to cohort study data analysis and interpretation of results.

### **COHORT STUDY ADVANTAGES**

Criteria for RCT patient-participant inclusion are often strictly defined to limit the influences of confounding factors. Narrow inclusion criteria increase statistical power and the precision of the estimated treatment effect, but the ability to generalize results to a heterogeneous patient population may be limited. Although the results of RCTs offer the strongest evidence to support the use of interventions intended to prevent, cure, or slow disease processes (ie, treatment efficacy),<sup>4</sup> they do not necessarily provide evidence that administering a given treatment will provide meaningful improvements in the quality of life experienced by an individual patient (ie, treatment effectiveness).

Observational research, which is performed in a typical clinical setting, assigns patients to groups on the basis of 1 or more defined characteristics.<sup>6</sup> A cohort study defines groups as exposed (eg, positive) versus unexposed (eg, negative) with regard to some factor that is believed to be associated with a specified outcome, whereas a case-control study designates participants as either cases (eg, injured) or controls (eg, uninjured) and compares exposure status (ie, exposed versus unexposed) between the groups. Because a measurement obtained after an injury has occurred may be affected by the injury itself, the retrospective nature of a case-control study severely limits the inference that can be made about the exposure-outcome association. The temporal sequence of a cohort study normally involves exposure classification before outcome event occurrence (ie, documentation of baseline characteristics at the beginning of a defined study period), so its results may have prognostic value. In such cases, the term predictive factor is often used as a synonym for exposure variable. A cohort study is assumed to involve a prospective approach, unless the adjective retrospective designates an exception to the usual procedure (ie, a study that is initiated after the outcome event occurs and involves the analysis of previously collected baseline data).<sup>7</sup>

The magnitude of association between a predictive factor and occurrence of an adverse outcome event within a specified period of time is often expressed as a risk ratio (RR), which is the proportion of exposed group members who experience the event divided by the proportion of unexposed group members who experience the event. The term rate ratio is used to compare incidence rates (ie, events per unit of time) in the exposed and unexposed groups. The term *relative risk* is sometimes used in an indiscriminate manner to refer to either of these ratios. The odds of event occurrence among the members of a given group are calculated by dividing the probability of event occurrence (eg, the proportion injured) by the probability of nonoccurrence of the event (eg, the proportion uninjured). The odds ratio (OR) represents the relative difference in the odds of event occurrence between the exposed and unexposed groups. Alternatively, the OR can be interpreted as the relative difference in the odds for exposure to a given factor between cases (eg, injured) and controls (eg, uninjured). The OR value will always be substantially greater than the RR value, unless the event occurrence is relatively rare (eg, injury incidence <10%).

#### Perspective: Patellofemoral Pain Syndrome

Patellofemoral pain syndrome (PFPS) is a common condition among active individuals and has been studied extensively. Bolgla and Boling<sup>8</sup> recently reported that a PubMed search of the terms *patellofemoral pain syndrome* and *anterior knee pain* resulted in 1230 citations published from 2000 to 2010. The authors concluded that current evidence supports the continued use of quadriceps exercises for conservative management of PFPS, but limited evidence and inconsistent findings precluded conclusive recommendations about any other interventions. Given the extensive body of literature, why is strong evidence to guide treatment planning for PFPS so lacking?

Although a diagnosis of PFPS suggests certainty about the condition's causes, multiple factors may contribute to its development in individual patients, including strength of the hip and quadriceps muscles, patellofemoral joint laxity, foot and ankle biomechanics, and lower extremity structural alignment.<sup>8</sup> Three recently published reports pertaining to PFPS treatment illustrate challenges to the development of recommendations for clinical application of research evidence.9-11 Ferber et al9 investigated the therapeutic benefit of hip-strengthening exercises in patients with PFPS who reported running at least 3 times per week for 30 minutes per session. Dolak et al<sup>10</sup> assessed hip and quadriceps strengthening in female PFPS patients, but physical activity level was not considered. Chui et al<sup>11</sup> investigated quadriceps strengthening in a mixed-sex sample of PFPS patients but excluded those who were participating in a resistance-training program. Each of these studies provided evidence that strengthening exercises reduced pain, but both Ferber et al<sup>9</sup> and Dolak et al<sup>10</sup> ultimately included additional exercises in their treatment regimens. Ferber et al<sup>9</sup> reported that all PFPS patients had returned to pain-free running at a preinjury level after having participated in "a more comprehensive rehabilitation program" for 3 additional weeks.

Each of the cited studies used sound research methods, and none of the investigators' conclusions were overstated. The clinician can conclude that strengthening exercises for the hip and quadriceps muscles are likely lead to pain relief for patients with PFPS. However, the clinician is not able to provide a patient-athlete with an estimate of the likelihood that he or she will be able to return to sport participation at the preinjury level by complying with a specific treatment plan. The clinician is forced to rely on evidence derived from studies of fairly homogeneous samples of patients, whose treatment responses have been removed from the context of the multitude of factors affecting functional status, to make treatment decisions about an individual patient-athlete who may be a member of a very different population.<sup>3</sup> Cohort studies of the effectiveness of comprehensive PFPS treatment programs could greatly add to our understanding of the prognostic factors that contribute most to restoring pain-free function and may identify subgroups of patient-athletes who are more or less likely to experience a successful outcome. The cited studies certainly contribute to the body of knowledge pertaining to treatment of PFPS, but consistent with the conclusions of the systematic review performed by Bolgla and Boling,<sup>8</sup> the nature of the evidence is insufficient to define best practices for clinical management of the condition. The availability of RR and OR estimates for various treatment options and different patient populations would provide clinicians with a better means to develop an optimal treatment plan for a given patient.

### **Clinical Decision Support**

An updated model for evidence-based clinical decision making emphasizes consideration of the "clinical state and circumstances of the patient" as a key element.<sup>5</sup> Such an individualized approach to injury prevention and treatment decisions needs to be supported by research evidence that

provides some basis for predicting the outcome that will ultimately be realized by a patient who possesses a given set of personal attributes, pathophysiologic indicators, and routine physical demands. Experimental study data that are analyzed by parametric hypothesis testing may have limited relevance to an individual patient. Results are typically reported as "average" values for dependent variables that were measured on a continuous scale. What does an average improvement of 20% in quadriceps strength mean to the patient with PFPS? First, the average may be derived from an exceptionally wide range of values, which would not convey any meaningful information about the proportion of patients who see improvement as a result of a training regimen. More important, an average improvement might not be sufficient to attain the individual patient's goal of returning to sport participation.

Clinical decisions made by athletic trainers to prevent, diagnose, or treat individual patients' injuries relate to an ultimate outcome for each patient that can often be classified in a binary manner (eg, injured versus not injured, diagnosis positive versus diagnosis negative, or optimal versus suboptimal recovery of function).<sup>12</sup> Welldesigned cohort studies have not been extensively used by athletic training researchers but are ideal for assessing associations among multiple characteristics (ie, predictive factors) and clinical outcomes that are important to athletes and athletic trainers (eg, a defined magnitude of global rating of change in functional status or return to unrestricted participation in sport within a defined amount of time).<sup>13,14</sup>

Observational studies and RCTs are viewed by some experts as opposing methods of clinical research, but more empirical evidence is needed to establish the relative merits of each approach for guiding clinical decisions.<sup>15</sup> No observational study of a heterogeneous cohort of participants can match the internal validity of a well-designed RCT, but the strong external validity of a well-designed cohort study can yield results that are highly relevant to clinical decision making.<sup>16</sup> Appreciating the value of both RCT evidence of treatment efficacy and cohort study evidence of treatment effectiveness is key to delivering clinical services that are both evidence based and patient centered.

## COHORT STUDY LIMITATIONS

The primary limitation of a cohort study is *selection bias*, which refers to the potential existence of important differences between the cohort's exposed and unexposed groups that may independently affect outcome.<sup>16,17</sup> Ideally, the exposed and unexposed groups are comparable in all other respects, but this is rarely the case. Confounding, which is distortion of an apparent association between an exposure and an outcome that is due to the influence of another factor, and selection bias often overlap.<sup>16</sup> Thus, anticipating possible confounding factors is necessary to ensure that sufficient information is obtained to statistically control for their effects through data stratification or multivariable regression analysis. Unfortunately, an unknown confounding factor cannot be anticipated. A major advantage of the RCT is neutralization of the effect of an unknown confounder through its random distribution among the groups.

When a therapeutic intervention is defined as an exposure in a cohort study, confounding by indication is a possible source of bias that threatens internal validity. In such a case, the clinical presentation (ie, indication) that results in administration of the therapeutic intervention also has an independent effect on the outcome. For example, an analysis of the association between the occurrence of lateral ankle sprain (ie, outcome) with the use of an ankle support (ie, therapeutic exposure) could be confounded by the influences of characteristics that differentiate athletes who choose to regularly wear ankle supports during sportrelated activities from those who do not. Characteristics that are unevenly distributed between the exposed and unexposed groups may increase the likelihood for ankle sprain occurrence (eg, ligament laxity), but some could also decrease its likelihood (eg, regularly performing strengthening and postural-balancing exercises). Multivariable regression analysis that includes each of the possible confounding variables provides a means to statistically control for their influences on the outcome.

Weak exposure-outcome associations are often due to bias, but a large degree of bias would be necessary to produce a strong invalid association. Thus, the magnitudes of RR and OR values indicate the likelihood that a meaningful exposure-outcome association exists, assuming that the influences of all major confounding factors have been addressed in the data analysis. The calculated point estimates of RR and OR magnitude should be reported with associated confidence intervals that define the precision of the estimated values for the cohort of heterogeneous participants.<sup>16</sup>

# Randomized Assignment Versus Observational Categorization

An experimental research design randomizes assignment of participants to experimental and control groups for the purpose of maximizing the validity of a cause-effect finding.<sup>14</sup> However, assignment to the control group often conflicts with the perceived needs of individual participants. The ethical concept of equipoise dictates that alternative treatments must be equally acceptable on the basis of current knowledge.<sup>14,16</sup> An athletic trainer may consider alternative treatments equally acceptable, but denying access of the control group to an intervention that coaches or athletes perceive to be beneficial can be seen as an obstacle to attaining individual and team goals. Thus, few athletes are likely to agree to voluntarily participate in research that may result in assignment to a treatment perceived to be inferior. Admittedly, factors that influence a patient's decision to choose a given therapeutic intervention may also have an independent influence on the outcome that is ultimately manifested (ie, confounding by indication). An observational cohort study design certainly does not solve the problem posed by extraneous influences that might contribute to an observed outcome, but it can provide a means to gain clinically important insights about predictive exposure-outcome associations that might otherwise remain poorly understood.

The use of an experimental design for injury-prevention research presents another ethical dilemma. If evidence is sufficient to suggest that a preventive intervention can reduce injury incidence, then athletes assigned to a control group may be exposed to a greater level of injury susceptibility than those assigned to an experimental group. Furthermore, injury-prevention research must address potential interactions among biomechanical, physiologic, behavioral, and medical factors that may influence the manner in which forces of different magnitudes, rates, and frequencies affect energy transfer to the body tissues.<sup>18,19</sup> A cohort study design is ideal for assessing multiple predictive factors before injury-risk exposure, which can subsequently be analyzed separately and collectively for their value in predictig the outcome.<sup>13,14,20,21</sup> Such a multifactorial approach to assessing sport-injury causation has not been extensively used.<sup>22,23</sup>

### BAYESIAN VERSUS FREQUENTIST PHILOSOPHICAL APPROACH

Randomization controls for selection bias and enhances the interpretation of parametric statistical tests that evaluate a null hypothesis of no difference between groups of participants at a specified level of probability. In theory, randomized participant selection and group assignment promotes even distribution of unmeasured factors between the study groups, thereby limiting extraneous influences on the observed posttreatment difference between groups. Thus, evidence derived from any research design that does not incorporate randomization is widely viewed as inferior to that derived from an RCT.<sup>24</sup>

The data derived from an RCT are often continuous, representing magnitudes of differences between groups in terms of the statistical significance of a difference between their respective mean values. Outcomes expressed in this manner do not allow the clinician or the patient to appraise the chances of achieving or avoiding a particular result. There is increasing recognition of the clinical value of estimates of the RR and the OR for a specified outcome; these compare the frequency of outcome occurrence (eg, injury incidence) in 2 groups of participants that differ on the basis of exposure to a possible risk factor (ie, status associated with high probability for adverse event occurrence) or protective factor (ie, status associated with low probability for adverse event occurrence), such as a behavior (eg, dietary intake), personal trait (eg, body mass), treatment (eg, use of an ankle brace), or event (eg, educational session). The exposed group is sometimes designated as the *index* cohort (eg, high-risk group), in which case the unexposed group is designated as the reference cohort (eg, low-risk group).16 Cohort study results typically include an estimation of the likelihood, or odds, of an event's occurrence, but alternative statistical values include incidence rate and time to event, as well as their corresponding indicators of strength of association between exposure and outcome (eg, rate ratio, hazard ratio).<sup>6</sup> Our report is focused on RR and OR as indicators of the strength of exposure-outcome association. Although the results of an RCT are usually reported in terms of the probability that a statistically significant difference exists between groups for a continuous dependent variable (eg, P < .05), the magnitude of change in the dependent variable can be classified in a dichotomous manner to report the results in terms of risk or odds.

The outcome of interest in a prevention study is the occurrence of an injury or illness, whereas a therapeutic

outcome is often measured on a continuous scale (eg, reduction in pain, improvement in strength or range of motion). With the notable exceptions of patient ratings of pain relief and functional capabilities, most treatment outcomes are more clinician centered than patient centered. Dichotomizing outcomes in terms that are meaningful to patients will permit the calculation of the odds of achieving a desirable outcome. A specified magnitude of change in functional capabilities quantified by a joint-specific survey instrument<sup>25</sup> or a global rating of change questionnaire<sup>26</sup> can be used to dichotomously classify an outcome as either optimal or suboptimal at a given point in the treatment process or at its conclusion (eg, a change score that meets or exceeds a minimal clinically important difference).

An important difference in interpretating statistical findings distinguishes traditional hypothesis testing for a difference between randomly created study groups from evaluation of an association between exposure and outcome within a cohort. The test of a null hypothesis for a study that randomly assigns participants to 2 or more different groups determines the probability (specified by a P value) that the magnitude of measured difference could result from the random variability encountered in numerous replications of the experiment. The term *frequentist* has been used to designate this traditional approach to hypothesis testing, because of its reliance on the theorized frequency that a given finding could be expected to result from repeated random sampling from a large population.<sup>16,17</sup> In contrast, Bayesian analysis reflects a philosophical approach that is focused on the magnitude of an observed association (represented by RR and OR values), and its practical implications for a decision-making process that includes consideration of updated knowledge.27,28 Rather than setting an arbitrary P value as an objective standard for statistical significance (eg, an  $\alpha$  level of .05), the precision of a point estimate of the observed association is subjectively interpreted on the basis of its associated confidence interval. A Bayesian approach interprets probability as a person's degree of belief in the validity and usefulness of an association, which influences the subjective expected utility of a decision made in the face of uncertainty.2

## BAYESIAN APPROACH TO OBSERVATIONAL DATA ANALYSIS

A simple  $2 \times 2$  cross-tabulation analysis can be used to quantify the magnitude of association between a binary exposure variable (eg, starter versus nonstarter status) and a binary outcome variable (eg, injured versus not injured). Although such an analysis quantifies the association between 2 variables (ie, an exposure variable and an outcome variable), the term univariable is often used to distinguish an analysis that is limited to a single exposure variable from a multivariable analysis that quantifies the association of a set of 2 or more exposure variables with an outcome variable. Multivariable logistic regression analysis provides a means to statistically adjust for suspected confounding influences by generating output that represents the collective association of multiple exposure variables with a binary outcome variable. A univariable (ie, simple  $2 \times 2$  cross-tabulation) analysis of the association between each exposure variable and a designated outcome variable



Figure 1. Receiver operating characteristic curve for discriminating injured college football players (ie, core or lower extremity sprain or strain) from uninjured players on the basis of ImPACT composite reaction time. Reprinted with permission.<sup>31</sup>

is normally performed as a preliminary procedure for selecting the exposure variables that will be included in a multivariable logistic regression analysis. Exposure variables that have the strongest adjusted association with a dichotomized outcome variable may be designated as predictive factors when they are combined to create a clinical prediction guide.<sup>21</sup>

When an exposure is defined by a continuous variable or is represented on an ordinal scale that includes more than 2 levels, a cut point must be selected to assign a binary exposure classification to each cohort member (eg, high risk versus low risk).<sup>30,31</sup> A receiver operating characteristic (ROC) curve offers a means to identify a cut point that optimizes the balance between sensitivity (eg, correct highrisk classification of injured cases) and 1 - specificity (eg, incorrect high-risk classification of uninjured cases) for a continuous exposure variable.<sup>30,31</sup>

The association between ImPACT neurocognitive test composite reaction time and noncontact ACL injury reported by Swanik et al<sup>32</sup> provided the impetus for an analysis that will be used as an example. An ROC curve for discriminating injured football players (ie, those who ultimately sustained a core or lower extremity strain or sprain) from uninjured players on the basis of neurocognitive reaction time is presented in Figure 1.<sup>33</sup> Although the area under the curve is relatively small, a clearly definable point provided the basis for a binary classification of exposure status that produced evidence of a meaningful association with injury occurrence (Table). An inverse

relationship (ie, fewer injuries among players with prolonged reaction time) would have produced an inverted ROC curve. The expectation that prolonged reaction time will have either no association or a positive association with injury occurrence justifies the use of a 1-sided *P* value for assessing the observed frequency of injury in the 2 groups, relative to that expected in the absence of an exposure-outcome association.<sup>16,17</sup> When performing an exploratory univariable analysis for identifying exposures associated with an adverse outcome (ie, injury risk factors), a *P* value as large as .25 may be interpreted as sufficient evidence to confirm the existence of a meaningful relationship.<sup>21</sup>

The strength of the association between exposure and outcome is represented by RR and OR values derived from

Table. Results of the Cross-Tabulation Analysis for Discriminating Injured College Football Players (ie, Core or Lower Extremity Sprain or Strain) From Uninjured Players on the Basis of ImPACT Composite Reaction Time 0.545-s Cut Point<sup>a</sup>

ImPACT Composite Reaction Time, s	Injury	No. Injury
≥0.545	17	26
<0.545	6	27
Total	23	53

<sup>a</sup> Fisher exact 1-sided P = .038; sensitivity = 0.74; specificity = 0.51; +likelihood ratio = 1.51; -likelihood ratio = 0.51; odds ratio = 1.507/ 0.512 = 2.94; 90% confidence interval = 1.19, 7.25; risk ratio = 0.395/0.182 = 2.17; 90% confidence interval = 1.10, 4.30. Reprinted with permission.<sup>31</sup>



Figure 2. Confidence interval function for point estimate of the risk ratio for core or lower extremity sprain or strain among collegiate football players (n = 76) with ImPACT composite reaction time  $\geq$ 0.545 s versus <0.545 s (risk ratio = 2.17, 90% confidence interval = 1.10, 4.30). The horizontal dashed line corresponds to the 90% confidence interval, and the vertical dashed line identifies the critical value that the lower limit must exceed for a positive association to exist between exposure and outcome. Reprinted with permission.<sup>31</sup>

the  $2 \times 2$  contingency table, both of which exceed 1.0 when exposure is positively associated with the defined outcome. Because a 90% confidence interval for the RR or OR point estimate provides both upper and lower limits, the possible values outside the specified range are equally divided between extremely high and extremely low values. Thus, the lower limit of a 90% confidence interval that exceeds a value of 1.0 for RR or OR can be viewed as the equivalent of a .05  $\alpha$  level for a 1-tailed test of statistical significance. Although calculating a confidence interval involves a parametric statistical procedure, a Bayesian interpretation of analysis results does not involve a hypothesis test at a specified level of statistical significance.<sup>16</sup> A graphical method that depicts all possible confidence intervals around a point estimate is referred to as a confidence interval *function*, which provides a visual representation of both the magnitude of the observed association and the precision of the estimate (Figure 2).<sup>16,34</sup>

### CONCLUSIONS

Virtually every clinical scenario encountered by an athletic trainer demands 1 or more decisions, and relatively few of them can be made with certainty that a given choice will yield optimal results. The evolution of the evidencebased medicine paradigm has elevated RCT results to the pinnacle of the evidence hierarchy, but cohort studies have much to offer in advancing the practice of athletic training. Increased use of the cohort study design can provide important and clinically meaningful evidence of effectiveness that is well suited to the delivery of patient-centered care. Although a cohort study design cannot achieve a level of internal validity to match that of a well-designed RCT, its greater feasibility in many settings could yield valuable evidence about a variety of clinical problems that would otherwise remain poorly understood.

The Bayesian approach to interpreting research findings offers a powerful means to support clinical decision making

with a quantifiable likelihood that a positive or negative outcome will result from a given patient characteristic, circumstance, or treatment option. Well-designed cohort studies can provide RR and OR values for either a single exposure variable or a combination of multiple exposure variables (ie, a clinical prediction guide) that can reduce guesswork and individualize treatment, thereby optimizing patient outcomes. Despite the limitations of observational research, the aggregation of a large volume of standardized clinical data as part of multisite trials could provide estimates of exposure-outcome associations that have not previously been addressed by more traditional research methods. Widespread familiarity with the advantages of the cohort study design and the relevance of RR and OR values to clinical decision making among researchers, educators, and clinicians could greatly advance evidence-based and patient-centered practice in athletic training.

## REFERENCES

- 1. MacAuley D. Best T. *Evidence Based Sports Medicine*. London, United Kingdom: BMJ Books; 2002:3.
- 2. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996; 312(7023):71–72.
- Bensing J. Bridging the gap. The separate worlds of evidence-based medicine and patient-centered medicine. *Patient Educ Couns*. 2000; 39(1):17–25.
- Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. Evidence-Based Medicine: How to Practice and Teach EBM. 2nd ed. Edinburgh, Scotland: Churchill Livingstone; 2000.
- Haynes RB, Devereauz PJ, Guyatt GH. Clinical expertise in the era of evidence-based medicine and patient choice. *Evid Based Med*. 2002;7(2):36–38.
- Verhagen E, van Mechelen W. Sports Injury Research. New York, NY: Oxford University Press; 2010.
- 7. Samet JM, Munoz A. Evolution of the cohort study. *Epidemiol Rev.* 1998;20(1):1–14.

- 8. Bolgla LA, Boling MC. An update for the conservative management of patellofemoral pain syndrome: a systematic review of the literature from 2000 to 2010. Int J Sports Phys Ther. 2011;6(2):112-125.
- 9. Ferber R, Kendall KD, Farr L. Changes in knee biomechanics after a hip-abductor strengthening protocol for runners with patellofemoral pain syndrome. J Athl Train. 2011;46(2):142-149.
- 10. Dolak KL, Silkman C, Medina McKeon J, Hosev RG, Lattermann C, Uhl TL. Hip strengthening prior to functional exercises reduces pain sooner than quadriceps strengthening in females with patellofemoral pain syndrome: a randomized clinical trial. J Orthop Sports Phys Ther. 2011;41(8):560-570.
- 11. Chiu JK, Wong YM, Yung PS, Ng GY. The effects of quadriceps strengthening on pain, function, and patellofemoral joint contact area in persons with patellofemoral pain. Am J Phys Med Rehabil. 2012; 91(2):98-106.
- 12. Denegar CR, Cordova ML. Application of statistics in establishing diagnostic certainty. J Athl Train. 2012;47(2):233-236.
- 13. Bahr R, Holme I. Risk factors for sports injuries: a methodological approach. Br J Sports Med. 2003;37(5):384-392.
- 14. Kocher MS, Zurakowski D. Clinical epidemiology and biostatistics: a primer for orthopaedic surgeons. J Bone Joint Surg Am. 2004;86-A(3):607-620.
- 15. Ioannidis JP, Haidich AB, Lau J. Any casualties in the clash of randomised and observational evidence? BMJ. 2001;322(7291):879-880
- 16. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.
- 17. Grimes DA, Schultz KF. Bias and causal associations in observational research. Lancet. 2002;359(9302):248-252.
- 18. McIntosh AS. Risk compensation, motivation, injuries, and biomechanics in competitive sport. Br J Sports Med. 2005;39(1):2-3.
- 19. van Mechelen W, Hlobil H, Kemper HC. Incidence, severity, aetiology and prevention of sports injuries: a review of concepts. Sports Med. 1992;14(2):82-99.
- 20. Finch C. A new framework for research leading to sports injury prevention. J Sci Med Sport. 2006;9(1-2):3-9.
- 21. Guyatt GH. Determining prognosis and creating clinical decision rules. In: Haynes RB, Sackett DL, Guyatt GH, Tugwell P, eds.

edu.

Clinical Epidemiology: How to Do Clinical Practice Research. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2006:323-355.

- 22. Meeuwisse WH. Assessing causation in sport injury: a multifactorial model. Clin J Sport Med. 1994;4(3):166-170.
- 23. Van Tiggelen D, Wickes S, Stevens V, Roosen P, Witvrouw E. Effective prevention of sports injuries: a model integrating efficacy, efficiency, compliance and risk-taking behavior. Br J Sports Med. 2008;42(8):648-652.
- 24. Sterne JA, Davey Smith G. Sifting the evidence: what's wrong with significance tests? BMJ. 2001;322(7280):226-231.
- 25. 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(2):776-788.
- 26. Lesher JD, Sutlive TG, Miller GA, Chine NJ, Garber MB, Wainner RS. Development of a clinical prediction rule for classifying patients with patellofemoral pain syndrome who respond to patellar taping. J Orthop Sports Phys Ther. 2006;36(11):854-866.
- 27. Freedman L. Bayesian statistical methods. BMJ. 1996;313(7057): 569-570.
- 28. Lilford RJ, Braunholtz D. The statistical basis of public policy: a paradigm shift is overdue. BMJ. 1996;313(7057):603-607.
- 29. Vallverdu J. The false dilemma: Bayesian vs. frequentist. E-Logos. 2008. http://nb.vse.cz/kfil/elogos/science/vallverdu08.pdf. Published 2008. Accessed January 7, 2014.
- 30. Bewick V, Cheek L, Ball J. Statistics review 13: receiver operating characteristic curves. Crit Care. 2004;8(6):508-512.
- 31. Lasko TA, Bhagwat JG, Zou KH, Ohno-Machado L. The use of receiver operating characteristic curves in biomedical informatics. J Biomed Inform. 2005;38(5):404-415.
- 32. Swanik CB, Covassin T, Stearne DJ, Schatz P. The relationship between neurocognitive function and noncontact anterior cruciate ligament injuries. Am J Sports Med. 2007;35(6):943-948.
- 33. Wilkerson GB. Neurocognitive reaction time predicts lower extremity sprains and strains. Int J Athl Ther Train. 2012;17(6):4-9.
- 34. Sullivan KM, Foster DA. Use of the confidence interval function. Epidemiol. 1990;1(1):39-42.

Address correspondence to Gary B. Wilkerson, EdD, ATC, FNATA, Graduate Athletic Training Education Program, University of Tennessee–Chattanooga, 615 McCallie Avenue, Department 6606, Chattanooga, TN 37403-2598. Send e-mail to Gary-Wilkerson@utc.