Bridging the Chasm Between Research & Clinical Practice in Athletic Training:
A Discussion of Methods and Analyses

Craig Denegar, PhD, ATC, FNATA
Gary Wilkerson, EdD, ATC, FNATA
Overview

- Perspective on EBP and outcome measures
- The Frequentist perspective
- Considering cohort designs
- The Bayesian perspective
- Other models and open discussion

Clinical Practice

- A problem-solving process that requires decisions –
  - Patient history questions focused on acquisition of the information needed to establish an accurate diagnosis
  - Clinical test selection guided by the need for more information to support or refute differential diagnoses
  - Information synthesized to establish prognosis and to select treatment for realization of a favorable outcome

- Each step in the decision-making process requires estimation of the likelihood for correct categorization
  - Prediction models more accurate than clinical judgment alone

Outcomes – Primary Focus of the Evening

- Preventive measures
  - How does the clinician decide if an exercise or device intended to reduce the risk of injury should be recommended to the athlete in front of her?

- Treatment effectiveness
  - How does the clinician decide which treatments to recommend and what can she tell the patient about their chances of success?

Model for Evidence-Based Clinical Decision Making

Model for EBP

- The model moved healthcare toward a more patient-centered structure by considering the patients' values and preferences.
- Clinical research, particularly in athletic training, is behind in advancing patient-centered care.
- Can a broader paradigm of methods, measures and analyses bridge the gap?

Clinical Expertise Manifestation –

- Accuracy of the “clinical impression”
  - Prerequisite for every decision
    - Diagnosis – Prognosis – Treatment

- Comparison of the individual patient to other patients
  - Each individual patient’s experience will differ to some extent
    - Research reports the collective experience of patient groups

- Key factor influencing the patient’s outcome –
  - Synthesis of information (history, attributes, presentation)
    - Skill in making comparison to experiences of similar patients
    - Identification of “unique” characteristics of patient’s case
Research Evidence – Gold Standard

- **Randomized Clinical Trial (RCT)**
  - Random selection (sampling) of subjects from a defined population
  - Random assignment of subjects to experimental vs. control groups
  - Rigorous control of extraneous variables (strong internal validity)

- **Data analysis**
  - Parametric statistical procedures
    - Sample statistics estimate population parameters
    - “Frequentist” interpretation of results
    - Arbitrary standard for statistical significance

“Frequentist” Hypothesis Testing

Based on theoretical “normal” distribution

$\alpha = .05$
Parametric Comparison of Central Tendency Values for Randomly Formed Groups

\[ t = \frac{\overline{X}_E - \overline{X}_C}{\sqrt{\frac{s_E^2}{n_E} + \frac{s_C^2}{n_C}}} \]

\( H_0: \overline{X}_E = \overline{X}_C \)

\( \alpha = .05 \)

Randomization – Sampling Error

ANOVA Estimation of Parameters \( \mu \) and \( \sigma \)

A frequency distribution of observations (\( X \)'s) in the parent population when \( H_0 \) is true

\[ \text{Large sample size increases the precision of the parameter estimates} \]
“Frequentist” Interpretation of Results

- Rejection of null hypothesis of no difference ($\alpha = .05$)

For a given sample size (number of subjects) and 100 replications of the random sampling process –

- The expected “frequency” that a mean difference as large as that observed will result from random variation:

  No more than 5 times with 100 replications of the study

- Difference between groups attributable to a treatment effect, rather than random sampling error

Questionable Evidence

Why most published research findings are false

- The majority of published research may be reporting false findings

- Findings are NOT most appropriately represented by P-values

  - Large effects: Relative Risk values $\geq 3$
Questionable Evidence

Sifting the evidence – What’s wrong with significance tests?

Sterne JAC, Davey Smith G. BMJ. 2001;322:226-231

- Use of statistics in medicine dominated by division of results into significant or non-significant
- Type II error rate receives little or no consideration
- Confidence intervals should be reported to move away from a mechanistic accept-reject dichotomy

Misinterpretation of P-value:
Proportion of Incorrect Null Hypothesis Rejections

- Assumptions (Sterne and Davey Smith, BMJ, 2001):
  - 1000 small studies of comparative effectiveness of treatments
  - Null hypothesis: Experimental = Standard (no sig. difference)
  - Relatively low power level (1 – β) = 0.50 for each study
  - 10% of experimental treatments truly superior to standard treatment
  - Type I error rate (α) = 0.05

<table>
<thead>
<tr>
<th></th>
<th>Exp Tx Superior (Null False)</th>
<th>Exp Tx Similar (Null True)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null Rejected</td>
<td>50 (correct)</td>
<td>45 (type I error)</td>
<td>95</td>
</tr>
<tr>
<td>Null Accepted</td>
<td>50 (type II error)</td>
<td>855 (correct)</td>
<td>905</td>
</tr>
</tbody>
</table>

- True null hypothesis rejected in only 5% (45/855) of cases, **BUT, 47% (45/95) of null hypothesis rejections are INCORRECT**
Solutions and Limitations

- Confidence intervals
  - Address magnitude and probability and should be required in results reporting

- Effect size
  - Appealing and useful in meta-analysis, lack context

- Issues of patient-centeredness not fully addressed

Expanding Horizons

- A broader approach in terms of research methods and analyses is needed to grow the clinical research base more rapidly, and advance evidence-based, patient-centered care

- Cohorts, odds, risk and beyond!
Levels of Evidence Quality
Oxford Centre for Evidence-Based Medicine

- **Level 1:**
  - Randomized Clinical Trial or Systematic Review of RCTs

- **Level 2:**
  - Cohort study, low-power RCT, or prospective outcomes study

- **Level 3:**
  - Case-control study or retrospective outcomes study

- **Level 4:**
  - Case series (no control/comparison group)

- **Level 5:**
  - Case report or expert opinion

Contrasting Research Paradigms

- **Parametric statistics**
  - Key assumption: **RANDOMIZATION**
    - Null hypothesis: no difference between group means
      - Statistical test p-value < specified alpha level (typically .05)

- **Bayesian analysis of cohort study data**
  - Association between “exposure(s)” and “outcome”
    - Group comparison: Relative Risk (RR) or Odds Ratio (OR)
      - Lower limit of confidence interval for RR or OR >1.0

http://www.cebm.net
Decision-Making Under Uncertainty

1. **Consequence of the choice** (risk vs. benefit)
   - Severity of loss associated with an adverse outcome
   - Degree of benefit that may be derived

2. **Confidence in the outcome** (desirable vs. undesirable)
   - **Objectively known** probabilities
     - Games of chance (coin flips, dice, cards, roulette)
   - **Subjectively estimated** probabilities
     - Absolute determination of event likelihood impossible, due to the multitude of factors influencing it (treatment outcome)

---

**Subjective Expected Utility (SEU)**

- Degree of belief in the validity and usefulness of an association between a present circumstance and a future outcome
  - Conceptual basis for Bayesian statistical inference
  - Emergence of decision analysis as an applied science in engineering, business, and economics

Objective Probability

- The expected frequency of occurrence of a specified event in relation to the set of possible events
  - Always associated with some degree of uncertainty
  - Expressed as a proportion (or percentage)
    - Coin toss: 2 possible events: Head or Tail
      - .5 probability for either one (1/2)
    - Probability of Heads on consecutive coin flips
      - 2 consecutive tosses: (.5)^2 = .25
      - 3 consecutive tosses: (.5)^3 = .12
      - 4 consecutive tosses: (.5)^4 = .06
      - 5 consecutive tosses: (.5)^5 = .03

Relative Risk (RR) vs. Odds Ratio (OR)

- Relative Risk (or Risk Ratio):
  - Ratio of the probability for a specified outcome (injury) for one group in relation to that for another group
    - Low-risk group vs. High-risk group
    - If the probability is the same for both groups, RR = 1.0

- Odds Ratio:
  - Ratio of the odds for a specified outcome (injury) for one group in relation to the odds for another group
    - If the odds are the same for both groups, OR = 1.0
Probability vs. Odds

- The likelihood for occurrence of a given outcome
  1. Expressed as proportion or percentage (proportion x 100)
  2. Expressed in terms of odds “for” vs. “against” outcome

- Probability vs. Odds of “Heads” on both of 2 coin flips:
  - 4 possibilities:
    - HH
    - HT
    - TH
    - TT
  - Probability = 1/4 = 0.25
  - Odds for 1:3 = 1/3 = 0.33
  - Odds against 3:1

Bayesian Analysis of Observations

- Association between “exposure” and “outcome”
  - Dichotomized Exposure (High-Risk vs. Low-Risk)
    - Trait, status, behavior, event, treatment, etc.
  - Dichotomized Outcome:
    - Injury vs. No Injury
    - Diagnosis Positive vs. Negative
    - Optimal vs. Suboptimal Recovery
  - Null Hypothesis: No exposure-outcome association
    - Incidence of outcome same for both groups
    - Lower limit of confidence interval for RR or OR ≤1.0
Prediction of Core or LE Sprain or Strain  
2009-2011 Football Seasons (N=256)

Relative Risk vs. Odds Ratio

<table>
<thead>
<tr>
<th>Starter Status ≥ 1 Game</th>
<th>Core/LE Sprain-Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Injury</td>
</tr>
<tr>
<td>YES (106)</td>
<td>63</td>
</tr>
<tr>
<td>NO (128)</td>
<td>40</td>
</tr>
<tr>
<td>Total (256)</td>
<td>103</td>
</tr>
</tbody>
</table>

- 103/256 cases sustained Core or LE sprain or strain in 2009-2011 (40%)
- 63/106 who started ≥ 1 game sustained Core or LE sprain or strain (59.4%)
  - High-Risk Group: Odds of injury = .594 / .406 = 1.46
- 40/150 who did not start in any game sustained Core or LE sprain or strain (26.7%)
  - Low-Risk Group: Odds of injury = .267 / .733 = .36

Relative Risk = .594 / .267 = 2.2  Odds Ratio = 1.46 / .36 = 4.0

Sensitivity = 61%  Specificity = 72%

Likelihood for Injury Occurrence (Risk)

- Based on injury incidence in a particular population
  - For an individual, the injury occurs or does not occur within a given period of time
    - 0% or 100% incidence

- Key consideration:
  - Comparability of the individual patient's characteristics to those of a high-risk versus low-risk population
Pre-season assessment of college athletes at 18 universities
- 80 non-contact ACL tear cases (45 female, 35 male)
- 80 matched controls (gender, height, weight, age, sport, position)

Non-contact ACL tear cases compared to controls –
- Lower verbal memory and visual memory scores
- Slower processing speed and reaction time

Reaction Time
- Cases: 570 ms
- Controls: 530 ms

Uninjured-Injured Comparison
<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uninjured</td>
<td>53</td>
<td>568</td>
<td>67</td>
</tr>
<tr>
<td>Injured</td>
<td>23</td>
<td>583</td>
<td>70</td>
</tr>
</tbody>
</table>

Reaction Time Mean Difference = 15 ms
Independent t-test result:
t(74) = 0.842; P = .403
Clinical Prediction Guides  
(Predictive Modeling)

- **Purpose:**
  - Use of patient-specific information to predict outcome

- **Potential benefit:**
  - The patient can make an educated choice among treatment options on the basis of the outcome he or she is most likely to experience (given personal characteristics)

- **Ultimate value:**
  - Accuracy in discriminating between patients who will or will not experience a specified outcome

**2009 – 2011 Combined Analysis**

**Wall-Sit Hold N=256**

- **2009 N = 83**
  - Mean = 79 s
  - SD = 34 s

- **2010 N = 88**
  - Mean = 61 s
  - SD = 27 s

- **2011 N = 85**
  - Mean = 28 s
  - SD = 14 s

---

**2009 – 2011 Combined Analysis**

N=256

1) Starter   2) Hi ODI   3) Lo WSH

<table>
<thead>
<tr>
<th>Risk Factors</th>
<th>Injury</th>
<th>No Injury</th>
<th>Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9</td>
<td>47</td>
<td>16.1%</td>
</tr>
<tr>
<td>1</td>
<td>36</td>
<td>76</td>
<td>32.1%</td>
</tr>
<tr>
<td>2</td>
<td>45</td>
<td>25</td>
<td>64.3%</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>5</td>
<td>72.2%</td>
</tr>
<tr>
<td>Total</td>
<td>103</td>
<td>153</td>
<td>40.2%</td>
</tr>
</tbody>
</table>
2009 – 2011 Combined Analysis
N=256 Starter Status Stratification

Logistic Regression Result

<table>
<thead>
<tr>
<th>Factor</th>
<th>Cut-Point</th>
<th>p</th>
<th>Adj. OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starter</td>
<td>≥ 1 game</td>
<td>&lt;.001</td>
<td>4.22</td>
</tr>
<tr>
<td>Hi ODI</td>
<td>≥ 4 points</td>
<td>.006</td>
<td>2.26</td>
</tr>
<tr>
<td>Lo WSH</td>
<td>≤ 88-41-30 s</td>
<td>.005</td>
<td>2.22</td>
</tr>
</tbody>
</table>

Model $\chi^2 = 43.64; p < .001$
Nagelkerke $R^2 = .212$

WSH & ODI Risk Factors

2009 – 2011 Combined Analysis
3-Factor Prediction Model N=256

1) Starter (≥ 1 game)  2) Hi ODI (≥ 4)  3) Lo WSH (≤ 88-41-30 s)

Core + LE Strains & Sprains (103)

| 1) ≥ 2 Factors | 58 | 30 |
| 2) 0 or 1 Factor | 45 | 123 |
| Total           | 103 | 153 |

Fisher’s Exact One-Sided p < .001
Sensitivity: = .56  Specificity = .80
Odds Ratio 2.87 / .54 = 5.28
90% CI: 3.31 – 8.44
Relative Risk = .659 / .268 = 2.46
90% CI: 1.93 – 3.14

AUC = .72
Confidence Interval Function

- Graphic representation of magnitude and precision of estimate of exposure-outcome association

Relative Risk = **2.46** (90% CI: 1.93 – 3.14)

Odds Ratio = **5.28** (90% CI: 3.31 – 8.44)

2009 – 2011 Combined Football Data (N=256)

≥ 2 of 3 Risk Factors (Starter, WSH, ODI)

Different Research Paradigms

**Frequentist Approach**

- Randomized Assignment
- IV: Group membership
- DV: Continuous measure
- Error: Random variation
- Focus: Statistical significance
  - Difference between groups
  - Mean values (central tendency)

**Bayesian Approach**

- Observation of Cohort
- IV: Exposure status
- DV: Binary outcome
- Error: Misclassification
- Focus: Precision of estimate
  - Exposure-Outcome association
  - Relative Risk and Odds Ratio
Evidence-Based Practice

- Bayesian approach offers major advantages
  - No need for randomized assignment to groups
  - Data acquired during routine clinical activities
  - Development of multi-factor prediction models
  - Individualization of risk reduction programs

- Multi-center cohort studies are needed
  - Different operational definitions of outcome
  - Different populations (e.g. sports, age groups)

Other Models and Open Discussion