Teaching Critical Appraisal and Application of Research Findings

Providing a foundation in patient management

Craig Denegar, PhD, PT, ATC
Jay Hertel, PhD, ATC

2015 Athletic Training Educators' Conference

OUTLINE

• Making Evidence-Based Clinical Decisions
• Strategies for Finding Clinical Evidence
• Critical Appraisal Strategies
• Applying Research Findings
  – Prevention
  – Treatment
  – Diagnosis
• Teaching Strategies

Elements of Patient Management

• Examination
• Evaluation
• Diagnosis
• Intervention / Treatment
• Outcomes

Examination

Patient demographics and medical history

History of current concern

Systems Review - Brief examination of each of the 4 systems

Tests and Measures

Model for Evidence-Based Clinical Decision Making

Evaluation and data synthesis

• Synthesis of data gathered to establish a diagnosis*, prognosis, and plan of care.

• Include problem list – impairment, function, participation, quality of life

WHO-ICF

Health Condition

Impairment

Activity Limitations

Participation Restrictions

Contextual Factors

personal / environmental

Diagnoses and syndromes

• Medical (e.g., patellofemoral pain)

• Biomechanical (patellofemoral pain due to excessive hip adduction / internal rotation in landing)

• Important consideration when considering the inclusion/exclusion criteria in clinical research

Prognosis and plan of care

• Consider the problems identified

• Identify all elements in a plan of care

• Link the clinical research and your clinical experience in formulating a prognosis

Interventions

• What is known regarding the impact of an intervention on goal achievement in patients similar to the one in your care?
  – Intervention may include the use of a modality, prescription of an exercise regimen, use of a device of orthosis, a change in behavior, planned patient education individually or in any combination
Outcomes

• Results of patient/client management
  – Impact on:
    • Pathology/pathophysiology
    • Impairments
    • Functional limitations
    • Disability
    • Risk reduction
    • Societal resources
    • Patient satisfaction

Outcomes

• Patient derived
• Clinician derived
• Longitudinal *
  – important in assessing change in risk
  – also important to consider with regard to follow-up (are benefits of care maintained, further improve or diminish over time)

OUTLINE

• Making Evidence-Based Clinical Decisions
• Strategies for Finding Clinical Evidence
• Critical Appraisal Strategies
• Applying Research Findings
  – Prevention
  – Treatment
  – Diagnosis
• Teaching Strategies

WHERE TO FIND CLINICAL EVIDENCE?

• Mailbox
• Cross-referencing
• Search engines
• Resources that specialize in EBP

FINDING EVIDENCE IN YOUR MAILBOX

• Wait for the evidence to come to you

Cross-referencing

• Identifying related articles in the bibliography of an article you’re reading

• Not very efficient, but should always be done
KEYS TO USING SEARCH ENGINES

- Use all appropriate search terms
  - Patellofemoral pain, anterior knee pain,
- Combining search terms
  - “AND” vs. “OR”
  - “NOT”
- Setting search limits
  - Human subjects
  - Language
  - Dates
  - Type of study (RCT, systematic review, …)
  - Patient demographics (age, …)
- Use your librarian to help teach search strategies!

EXAMPLES OF FILTERED RESOURCES THAT SPECIALIZE IN EBP

- sportsmedresearch.blogspot.com/
- Cochrane Reviews
- POEMs
- PEDro

CLINICAL EVIDENCE STRATEGIES

- Make the evidence come to you
- BUT, know how to find evidence when you need to
  - Be systematic in your searches
  - Look for quality meta-analyses, systematic reviews, & critically-appraised topics first
  - Review individual studies when needed
  - Concentrate on the “best evidence”
- Build discussion of “evidence” into your work routine
OUTLINE

• Making Evidence-Based Clinical Decisions
• Strategies for Finding Clinical Evidence
• Critical Appraisal Strategies
• Applying Research Findings
  – Prevention
  – Treatment
  – Diagnosis
• Teaching Strategies

CRITICAL APPRAISAL

• Why do we need to critically appraise research if it is peer-reviewed?

• All Evidence is NOT EQUAL

Evidence_1 \neq \ Evidence_2

Appraising Evidence:

FACE VALIDITY

• Do the characteristics of the study I am appraising match my PICO question?
  – Patients
  – Intervention
  – Comparison
  – Outcomes

• Is the study relevant to your patients?

• When this isn’t possible, still must identify and interpret the “best available evidence” for your question

Appraising Evidence:

LEVEL OF EVIDENCE

CEBM “Levels of Evidence 2”

SORT Levels of Evidence

1—High-quality diagnostic cohort study: cohort design, adequate size, adequate spectrum of patients, blinded, and a consistent, well-defined reference standard.
2—High-quality diagnostic case series (for uncontrolled studies).
3—Other evidence: Consensus guidelines, extrapolated from bench research, usual practice, opinion, diagnosis-oriented evidence (interpreters or physiologic outcomes only), or case series for studies of diagnosis, treatment, prevention, or screening.

Filtered Information

Unfiltered Information

SORT Levels of Evidence

Level 1—good quality, patient-oriented evidence

- Systematic reviews
- Randomized controlled trials (RCTs)
- Critical appraisal

Level 2—limited-quality, patient-oriented evidence

- Meta-analysis of lower-quality studies or studies with inconsistent findings
- Critical appraisal

Level 3—other evidence

- Consensus guidelines
- Extrapolation from bench research
- Usual practice
- Diagnosis-oriented evidence

1. Levels of evidence may be altered based on the quality of evidence.
2. Levels 1 and 2 may be altered based on the quality of evidence.
3. Levels 3 and 4 may be altered based on the quality of evidence.
Appraising Evidence:  
**METHODOLOGICAL QUALITY**

- Is the study you are evaluating internally valid?
  - RAMMbo

- Quantifying the quality of individual studies
  - PEDro
  - CONSORT

**RAMMbo**

- **Representative**
  - Who did the subjects represent?
- **Allocation**
  - Was the assignment to treatments randomised?
  - Were the groups similar at the trial’s start?
- **Maintenance**
  - Were the groups treated equally?
  - Were outcomes ascertained & analyzed for most patients?
- **Measurements blinded OR objective**
  - Were patients and clinicians “blinded” to treatment?
  - Were measurements objective & standardised?

**PEDro Scale**

- Physiotherapy Evidence Database  

- 10-point scale that allows quantification of the quality of a research study
  - Yes = 1 point, No= 0 point
  - Higher total score = Higher quality of study

- Designed specifically for clinical trials of treatment and prevention interventions, but may also be used for other types of human studies

**PEDro Scale Item 1**

1. eligibility criteria were specified.

- This criterion is satisfied if the report describes the source of subjects and a list of criteria used to determine who was eligible to participate in the study.

*Not used to compute the total Pedro score*

**PEDro Scale**

2. subjects were randomly allocated to groups  
   (in a crossover study, subjects were randomly allocated an order in which treatments were received).

3. allocation was concealed.

4. the groups were similar at baseline regarding the most important prognostic indicators.

**PEDro Scale**

5. there was blinding of all subjects.

6. there was blinding of all therapists who administered the therapy.

7. there was blinding of all assessors who measured at least one key outcome.
8. measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups.

9. all subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by "intention to treat".

CONSORT STATEMENT

- What is it?
  - Consolidated Standards for Reporting Trials
  - Recommendations for reporting clinical trials

- Why is it needed?
  - Inconsistency in reporting in the previous literature

- Many journals now require clinical trial submissions to comply with these guidelines

22 Items on Consort Checklist

<table>
<thead>
<tr>
<th>PAPER SECTION And topic</th>
<th>Issue</th>
<th>Description</th>
<th>Beyond or Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE &amp; ABSTRACT</td>
<td>1</td>
<td>How participants were allocated to interventions (e.g., “random allocation”, “randomized”, or “randomly assigned”)</td>
<td></td>
</tr>
<tr>
<td>INTRODUCTION Background</td>
<td>2</td>
<td>Scientific background and explanation of rationale</td>
<td></td>
</tr>
</tbody>
</table>

METHODS

- Participants
- Interventions
- Objectives
- Outcomes
- Sample size
- Randomization -- Sequence generation
- Randomization -- Allocation concealment
- Randomization -- Implementation
- Blinding (masking)
- Statistical methods

RESULTS

- Participant flow
- Recruitment
- Baseline data
- Numbers analyzed
- Outcomes and estimation
- Ancillary analyses
- Adverse events

PEDro Scale

10. the results of between-group statistical comparisons are reported for at least one key outcome.

11. the study provides both point measures and measures of variability for at least one key outcome.
22 Items on Consort Checklist

Discussion

<table>
<thead>
<tr>
<th>DISCUSSION</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision and the dangers associated with multiplicity of analyses and outcomes.</td>
</tr>
<tr>
<td>Generalizability</td>
<td>Generalizability external validity of the trial findings.</td>
</tr>
<tr>
<td>Overall evidence</td>
<td>General interpretation of the results in the context of current evidence.</td>
</tr>
</tbody>
</table>

OUTLINE

- Making Evidence-Based Clinical Decisions
- Strategies for Finding Clinical Evidence
- Critical Appraisal Strategies
- Applying Research Findings
  - Prevention
  - Treatment
  - Diagnosis
- Teaching Strategies

Appraising and applying research results of intervention studies

- How large is the treatment effect?
- Is the magnitude of the treatment effect clinically meaningful?
- Would the size of the treatment effect help my patients enough to change my clinical decision-making?

Appraising and applying research results of intervention studies

- How precise is the treatment effect?
  - How wide is the confidence interval?
  - Does it cross into zones of uncertainty or irrelevance?

Interpreting width of CI in relation to zero
Appraising and applying research results of intervention studies

- Measures of treatment effect:
  - Unit of measure (days to RTP, “ROM,...)
  - Effect size (Cohen’s d, Hedges g, ...)
  - Unitless measure of effect
  - Dichotomous outcomes
  - Odds ratio
  - Numbers needed to treat

### Measures of Treatment Effect

- **Numbers needed to treat**
- **Unit of measure** (days to RTP, Unitless)
- **Odds ratio**

### Studies

- **Gilchrist et al.**
- **Pfeiffer et al.**
- **Olsen et al.**
- **Myklebust et al.**
- **Mandelbaum et al.**
- **Heidt et al.**
- **Soderman et al.**
- **Hewett et al.**
- **Steffen et al.**
- **LaBella et al.**
- **Kiani et al.**

### Odds Ratio

<table>
<thead>
<tr>
<th>Study name</th>
<th>Odds ratio and 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gilchrist et al.</td>
<td>0.35 (0.23 to 0.54)</td>
</tr>
<tr>
<td>Pfeiffer et al.</td>
<td></td>
</tr>
<tr>
<td>Olsen et al.</td>
<td></td>
</tr>
<tr>
<td>Myklebust et al.</td>
<td></td>
</tr>
<tr>
<td>Mandelbaum et al.</td>
<td></td>
</tr>
<tr>
<td>Heidt et al.</td>
<td></td>
</tr>
<tr>
<td>Soderman et al.</td>
<td></td>
</tr>
<tr>
<td>Hewett et al.</td>
<td></td>
</tr>
<tr>
<td>Steffen et al.</td>
<td></td>
</tr>
</tbody>
</table>

### Relative Risk Reduction

- **ACL Prevention Programs**: RRR
- **Non-Contact ACL Injury Prevention**

### Pooled Results

- **RRR = 67.5% (55 - 77)**

### Main Findings:

**ACL Prevention Programs**

**Relative Risk Reduction**

A program of neuromuscular training incorporating plyometrics & balance exercises decreases the incidence of ACL injuries in female athletes by 67.5% (95% CI: 55-77%)

**Numbers Needed to Treat**

In order to prevent one ACL injury, 121 (95% CI: 95-166) female athletes would need to participate in a neuromuscular training program over the course of one season.
OUTLINE

• Making Evidence-Based Clinical Decisions
• Strategies for Finding Clinical Evidence
• Critical Appraisal Strategies
• Applying Research Findings
  – Prevention
  – Treatment
  – Diagnosis
• Teaching Strategies

Identifying Effective Diagnostic Tests:
Likelihood ratios and ROC curves

How is effectiveness of diagnostic tests measured?

• How can good diagnostic procedures (or clusters of tests) be differentiated from poor ones.
• Sensitivity, specificity, and likelihood ratios

Gold Standard Result

<table>
<thead>
<tr>
<th></th>
<th>Condition Present</th>
<th>Condition Absent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Cell A True positive</td>
<td>Cell B False positive</td>
</tr>
<tr>
<td>Negative</td>
<td>Cell C False negative</td>
<td>Cell D True negative</td>
</tr>
</tbody>
</table>

Sensitivity

• Sensitivity - the number of injuries (or illness) that are diagnosed by a test divided by the true number of injuries based upon gold standard

\[
\text{Sensitivity} = \frac{\text{# of injuries dx}}{\text{# of true injuries}} = \frac{A}{A+C}
\]

Specificity

• Specificity - number of negative exams based upon a test divided by number of negative cases based on gold standard.

\[
\text{Specificity} = \frac{\text{# dx without the injury}}{\text{# true negatives}} = \frac{D}{B + D}
\]
Interpretation

- Values for Sp and Sn range from 0.0-1.00 with higher values representing better tests

SpP_in and SnN_out

- Sacket et al - Tests with high specificity are good at ruling in a condition, those with high sensitivity good at ruling out a condition

- Counterintuitive? Remember where the false negative and false positives lie!

SpP_in – few false positives so a positive test is conclusive

SnN_out – few false negatives so a negative test is conclusive

Interpretation

- Sp and Sn are useful but ...
- Tests may have high Sp and low Sn or vice versa
- Difficult to apply directly in clinical practice

Likelihood Ratios

- A logical extension of Sn and Sp
- A positive likelihood (+LR) ratio is indicative of the impact of a positive examination finding on the probability that the condition in question exists.
  - $+ LR = \frac{\text{Sensitivity}}{(1 - \text{Specificity})}$
Likelihood Ratios

- A negative likelihood ratio addresses the impact of a negative examination on the probability that the condition in question is present.
  - \( LR = \frac{1 - \text{Sensitivity}}{\text{Specificity}} \)

Interpretation

- \((+LR > 10 \text{ or } -LR < .1)\) – large, often conclusive shift in probability a condition is present
- \((+LR \text{ of } 5-10 \text{ or } -LR \text{ 0.1-0.2})\) – moderate, but usually important shifts in probability a condition is present

Interpretation

- \((+LR \text{ of } 2-5 \text{ or } -LR \text{ 0.2-0.5})\) – small, sometimes important shifts in probability a condition is present,
- \((+LR \text{ of } 1-2 \text{ or } -LR \text{ 0.5-1})\) – very small, usually unimportant shifts in probability a condition is present.

Connecting the Dots in Evaluation

- Pre-test probability? (remember your interview and review of the medical record)
- Shifts in probability with testing results?
- Certainty is fleeting – and uncertainty uncomfortable!

Key points
- Before you begin a physical exam what you think is wrong? The probability of 1 or more differential diagnoses must be weighed
- Remember: Diagnostics is an uncertain business

Application of LRs

Consider a 38 year old teacher and avid tennis player who presents complaining of intermittent medial knee pain of insidious onset, with occasional catching and giving way (primarily while playing tennis) and intermittent swelling. History of 1 prior knee injury (believes an MCL sprain) in high school that prevented participation in football for 3 weeks. What do you think is wrong?

Application of LRs

- Pre-test probability - e.g. 80%
- Pre-test odds - probability / (1 - probability) \( .8/(1-.8) = 4 \)
- Pre-test odds x LR = post-test odds
- If LR = 3 then post-test odds would be 12:1
- Convert post-test odds to probability by post-test odds / post-test odds + 1 (92%)
Certainty

• What level of certainty is sufficient to recommend a course of treatment?
• How serious is the condition? What are the consequences of being wrong? Greater certainty is warranted as the stakes rise!

Continuous rather than dichotomous data?

• Receiver Operator Characteristic (ROC) curves provide information regarding the discriminating effects of specific values
• A very important concept with applications to risk, prevention and diagnostics

Notes on methodological issues in studies of diagnostic tests.

• How “gold” is the standard
• Are investigators blinded to the results of other testing
• What is the patient sample (selection bias)
• Do all subjects receive the same work-up (work-up bias)

Methodological issues in studies of diagnostic tests.

• Reliability of testing
• Qualifications of those performing the diagnostic tests - Can we generalize between, for example orthopaedic surgeons and physical therapists?

OUTLINE

• Making Evidence-Based Clinical Decisions
• Strategies for Finding Clinical Evidence
• Critical Appraisal Strategies
• Applying Research Findings
  — Prevention
  — Treatment
  — Diagnosis
• Teaching Strategies

Teaching Strategies

• Entry-level
• Post-professional
• Continuing education (in the workplace)
RESOURCES

• WWW.CEBM.NET
• WWW.NNT.COM

CONCLUSIONS

• Keep patient care decisions at the center of evidence-based practice
• Have strategies to make the evidence find you
• Not all evidence is equal
• Confidence (intervals) to appraise the clinical applicability of research results

THANK YOU

Craig.denegar@uconn.edu
Jhertel@virginia.edu