Evidence-Based Medicine:
Application to Athletic Training

Gary Wilkerson, EdD, ATC

Basis for clinical decisions?

- Anecdotes
  - Personal experiences or those of colleagues
  - Efficient vehicle for information conveyance and modification of behavior
  - Generally includes an emotional component!

Basis for clinical decisions?

- Heuristics (Greek: method or process)
  - Simple "rules of thumb" derived from anecdotal information
  - Reduces complex tasks to much simpler operations
  - Often useful in everyday life to produce positive outcomes
  - Easily distorted by personal needs and/or preferences
  - Can lead to potentially harmful incorrect conclusions

Basis for clinical decisions?

- Often based on unsubstantiated theoretical rationale that has questionable validity
  - Widespread variability in practice patterns
  - Sub-optimal outcomes often apparent
  - Escalating health care costs

Basis for clinical decisions?

- Research literature
  - Overwhelming volume
  - Relevant information not readily accessible
  - Conflicting results
  - Variation in research methodology

A systematic approach is necessary to apply the best current research evidence!

Evidence-Based Medicine (EBM)

- The conscientious, explicit and judicious use of current best evidence in making clinical decisions about the care of individual patients
  - Integrated with individual clinical expertise/experience
  - Includes thoughtful consideration of patient preferences and values

Evidence-Based Medicine (EBM)

- An approach to clinical practice:
  - Integration of clinical expertise with the best available evidence from **"systematic research"**
    - Especially "patient-centered" clinical research
  - **Not** based primarily on observational studies, logical intuition, and expert opinions

Athletic Training & EBM

- Relatively few studies have addressed:
  - Accuracy & precision of injury evaluation methods
  - Power of exam findings for prediction of outcome
  - Effectiveness of therapeutic, rehabilitative, & preventive procedures

The Practice of EBM

1) Clearly define a clinical question
   - Population, evaluation/treatment, & outcome of interest
2) Gather the best relevant evidence
   - Define search criteria (key terms)
3) Critically evaluate the quality of the evidence
   - Statistical significance vs. clinical meaningfulness

The Practice of EBM

4) Clinical application of the evidence
   - Utilization of specific clinical procedures
5) Evaluation of result produced by EBM process
   - Functional outcome assessment in patients
   - Clinician behaviors (practice patterns)

Basic EBM Concepts

- Statistical significance is not the same thing as clinical meaningfulness
- Be skeptical of results from small trials
  - Type II errors (low power to reject null hypothesis)
- A review article is not the same thing as a "systematic review"

Basic EBM Concepts

- Consider publication bias
  - Statistical significance necessary for publication
- Weak evidence should not be given greater weight on the basis of numerous weak studies
  - One RCT is better than numerous quasi-experimental studies

Sports Medicine Research

- No single source of clear & simple definitions or standard for classification of research types
  - Wide range of methodology utilized
- Hierarchy of evidence quality:
  - Oxford Centre for Evidence-Based Medicine
  - American Journal of Sports Medicine
    - Strongest evidence: Double-blind RCT
    - Weakest evidence: Case study

Sports Medicine Research

- **"Clinical Trial"**
  - Any research project that prospectively assigns human subjects to intervention or comparison groups to study the cause-effect relationship between a medical intervention and a health outcome.

International Committee of Medical Journal Editors

Clinical Trials

- FDA requirement for marketing clearance
  - Phase I: Small number of subjects; protocol safety
  - Phase II: Therapeutic exploratory trials (n>100)
  - Phase III: Therapeutic confirmatory trials (n>1000)
  - Phase IV: Therapeutic use trials

Classification of Research Design: Experimental

- **Randomized Clinical Trial**
  - Random assignment of patients (treatment-control)
  - Prospective manipulation of variables (intervention)
  - Precise measurement of physiologic variables
  - Rigorous control of extraneous variables
  - Strong internal validity
  - Provides evidence of treatment "efficacy" (beneficial therapeutic effect under ideal conditions)

Classification of Research Design: Experimental

- **Controlled Laboratory Study**
  - "in vivo"
    - Inside the living body
    - Real-life situation
  - "in vitro"
    - Outside the living body
    - Artificial environment

Experimental Research Designs

- Parallel (between-group) design
  - Each group of patient-subjects exposed to only one condition
  - Random assignment of subjects to groups (experimental & control)
- Cross-over (within-group) design
  - Each subject receives every condition in succession
  - Random order of treatment administration
- Factorial design
  - 2 or more conditions evaluated simultaneously
  - Mixed model design includes between-group & within-group factors

Classification of Research Design: **Quasi-Experimental**

- Non-randomized assignment of subjects to groups (experimental & control)
  - RCTs can be impractical or unethical for evaluation of a medical treatment
  - Evidence can be biased toward treatments that are easier to evaluate with RCTs
    - Lack of evidence of treatment effectiveness is not the same thing as evidence of lack of effectiveness!

Epidemiology

- The study of a society's collective health:

1) Descriptive epidemiology: patterns & trends

2) Analytic epidemiology: causation (etiology)

Epidemiologic Research

- **Cohort**
  - A group of people with a common experience over a defined period of time
    - Subject selection:
      Availability and/or characteristics

Classification of Research Design: **Observational (Quasi-Experimental)**

- **Cohort Study (Analytic Epidemiology)**
  - Prospective (cause-to-effect hypothesis)
  - Variability in subject exposure to risk factor(s)
  - Subjects who are not exposed to risk factor(s) are considered "controls"
  - Large number of subjects & long duration required

Classification of Research Design: **Observational (Quasi-Experimental)**

- **Cohort Study**
  - Observation to determine differences in incidence of injury in relation to exposure to risk factor(s)
  - Risk Ratio: Incidence rate for those exposed divided by incidence rate for those unexposed
    - Incidence rate: Number of new cases over defined period divided by person-time
      - Garrick & Requa: *Med Sci Sports, 1973*
      - 12 ankle sprains in 670 intramural basketball games
      - .0179 = 18 ankle sprains per 1000 player-games

Classification of Research Design: **Observational (Quasi-Experimental)**

- **Case-Control Study (Analytic Epidemiology)**
  - Retrospective (effect-to-cause hypothesis)
  - Subjects with condition (*cases*) compared to similar subjects without the condition (*controls*)
    - Closely matched on as many factors as possible
  - Variability in past exposure to risk factor(s)

Classification of Research Design: Observational (Quasi-Experimental)

- **Case-Control Study**
  - Determination of whether cases and controls differ in terms of past exposure to risk factor(s)
    - Examples: low H/Q ratio, foot pronation, high BMI
  - Odds Ratio (estimate of Risk Ratio): Prevalence rate for exposed divided by prevalence rate for unexposed
    - Prevalence rate: Number of cases with condition in relation to those without condition in population of interest
      - Example: prevalence of ACL injury among female Division I basketball players

Injury Prevention: Numbers Needed to Treat (NNT)

- Number of treatments necessary to prevent one injury occurrence
  - Absolute risk reduction: Incidence rate for control group minus incidence rate for treatment group
  - NNT = 1 / absolute risk reduction
    - Olmstead et al: *J Athl Train*, 2004
      - Garrick & Requa: *Med Sci Sports*, 1973
      - Hx of ankle sprain untaped: 24/434 = .0553
      - Hx of ankle sprain taped: 13/794 = .0164
      - NNT = 1 / (.0553 - .0164) = 26

Injury Prevention: Relative Risk Reduction

- Relative Risk: Injury rate for those treated divided by injury rate for those not treated
- Relative Risk Reduction = (1 – RR) X 100
  - Example: Ankle re-injury rates with wobble-board training vs. standard rehab program
    - RR = 1 – .25 / .54 = .46
    - Relative Risk Reduction = (1 – .46) X 100 = 54%
    - Wobble-board training makes re-injury 54% less likely

Olmsted-Kramer & Hertel: ATT, 2004 (Wester et al: JOSPT, 1996)

Classification of Research Design: Observational (or Experimental)

- **Outcomes Research**
  - Typically equivalent to a cohort study (prospective):
    - Cohort: Patients treated in the same clinical facility
  - May be conducted as a true experimental study:
    - Random assignment of subjects to different protocols

Classification of Research Design: Observational (or Experimental)

- Emphasis on patient perception of benefit
  - Dependent variable: Rating of functional capabilities
- Provides evidence of treatment "effectiveness"
  - Beneficial therapeutic effect from procedures performed by clinicians in a typical clinical setting

Classification of Research Design: Observational

- **Case Series**
  - Information on a series of patients having the same injury/illness
  - No control or comparison group
- **Case Report**
  - Information about a single patient

Improving Quality of Research Reports

- CONSORT Statement
  - Consolidated Standards of Reporting Trials
  - www.consort-statement.org
    - International Committee of Medical Journal Editors
    - Council of Science Editors
    - World Association of Medical Editors
- TREND Statement
  - Transparent Reporting of Evaluations with Non-randomized Designs
  - www.trend-statement.org
    - Prevention Research Synthesis Team of the Centers for Disease Control and Prevention

Classification of Research Design: Synthesis of Research Evidence

- **Systematic Review (SR)**
  - Criteria for inclusion of studies established prospectively
  - Objective quality rating of evidence relevant to a specific clinical question
  - May or may not include **meta-analysis** of findings for multiple studies (effect sizes)

Systematic Review Example

- The use of ice in the treatment of acute soft-tissue injury: A systematic review of randomized controlled trials.
  - Bleakley, McDonough, MacAuley: *AJSM*, 2004
    - Search of 8 databases (1981-2002) and 10 key journals (1976-2001) for studies on physiologic effects of ice
    - 22 RCTs met inclusion criteria for review
      - Marginal evidence that ice + exercise is more effective than ice only
      - Little evidence that addition of ice to compression had sig. effect
      - No evidence of an optimal mode or duration of treatment

Systematic Review Example

- Does cryotherapy hasten return to participation? A systematic review.
  - Hubbard, Aronson, & Denegar: *JAT*, 2004
    - 1976-2003 English-language literature searched
    - 4 RCTs identified that assessed return to participation
      - Cryotherapy has a positive effect on return to participation
      - Confounding effects of compression and different rehabilitation techniques; more research needed

Classification of Research Design: Synthesis of Research Evidence

- Meta-Analysis
  - Quantitative comparison of multiple studies that have addressed the same clinical problem with different methods
    - Individual study results (ESs) used as measurement units
    - Standards: Quality of Reporting of Meta-Analysis (QUORUM) statement
      - www.consort-statement.org

Meta-Analysis Example

- Cordova et al: *JOSPT*, 2000
  - Influence of ankle support on joint range of motion before and after exercise: A meta-analysis.
    - 19 studies (repeated measures design: pre-ex & post-ex)
    - Video analysis or goniometer (Inv, Evr, DF, & PF)
    - ES = (Exp Mean – Control Mean) / Control SD
    - Correction factor for studies with small sample size
    - Mixed-model factorial ANOVA
      - Random factor: studies
      - Fixed factors: testing time, support condition, interaction
    - Conclusion: Braces provide greater restriction of frontal plane motion than taping

Evidence Quality (Internal Validity)
Oxford Centre for Evidence-Based Medicine

- **Level 1:**
  - RCT or Systematic Review of Level 1 RCTs
- **Level 2:**
  - Cohort study, low-power RCT, or prospective outcomes study
- **Level 3:**
  - Case-control study or retrospective outcomes study

Evidence Quality (Internal Validity)
Oxford Centre for Evidence-Based Medicine

- **Level 4:**
  - Case series (no control/comparison group)
- **Level 5:**
  - Case report or expert opinion

Evidence Quality (Internal Validity)

- Overall evidence supporting a given clinical procedure
  - **A**: Conclusive
    - Multiple Level 1 studies
  - **B**: Acceptable
    - At least one Level 1 study and/or consistent findings from Level 2-4 studies

Evidence Quality (Internal Validity)

- Overall evidence supporting a given clinical procedure
  - C: Suggestive
    - At least one Level 2 study and/or consistent findings from Level 3-4 studies
  - **D**: Weak
    - Inconsistent findings among Level 1-4 studies

0-10 Internal Validity Scale
Physiotherapy Evidence Database (PEDro)

The Center of Evidence-Based Physiotherapy
University of Sydney http://www.pedro.fhs.usyd.edu

1. Subjects randomly allocated to groups (or treatment order)
2. Allocation concealed (blinded group assignment process)
3. Groups similar at baseline (baseline measure of condition severity)

0-10 Internal Validity Scale
Physiotherapy Evidence Database (PEDro)

4. Blinding of all subjects (unaware of whether or not experimental treatment is received)
5. Blinding of therapists administering treatments (unaware of group assignment for individual subjects)
6. Blinding of all assessors (unaware of group assignments)

0-10 Internal Validity Scale
Physiotherapy Evidence Database (PEDro)

7. Measures of key outcome obtained from >85% of subjects initially assigned
8. Treatment or control condition received as initially allocated

http://www.pedro.fhs.usyd.edu

0-10 Internal Validity Scale
Physiotherapy Evidence Database (PEDro)

9. Results of between-group statistical comparisons reported for key outcome (typically a group x trial interaction effect)
10. Meaningfulness of statistically significant differences between treatment & control groups reported (treatment effect size)

http://www.pedro.fhs.usyd.edu

Clinical Test Accuracy

- **True Positive Rate (TPR) = Sensitivity**
  - Proportion of True Positives among cases with condition
- **False Negative Rate (FNR) = 1 – Sensitivity**
  - Proportion of False Negatives among cases with condition
- **True Negative Rate (TNR) = Specificity**
  - Proportion of True Negatives among cases without condition
- **False Positive Rate (FPR) = 1 – Sensitivity**
  - Proportion of False Positives among cases without condition

Clinical Test Result Accuracy:
Positive Identification of a Condition

Clinical Test Result	Gold Standard Positive
Positive	True Positive
Negative	False Negative

Sensitivity of a Clinical Test (True Positive Rate)

- **Positive identification** when condition exists
- High degree of agreement with "gold standard"
  - + Lachman test & arthroscopic diagnosis of ACL tear
    - High rate of clinical test true positives
    - Low rate of clinical test false negatives (failure to identify)
    - Proportion of cases correctly identified by clinical test in relation to total cases diagnosed by gold standard (0.0 –1.0)

Sensitivity = # True Positive / (# True Positive + # False Negative)

Clinical Test Result Accuracy:
Exclusion of a Condition

Clinical Test Result	Gold Standard Negative
Positive	False Positive
Negative	True Negative

Specificity of a Clinical Test (True Negative Rate)

- **Exclusion** of a condition when it **does not exist**
- High degree of agreement with "gold standard"
  - Negative Lachman test & arthroscopic confirmation of intact ACL
    - High rate of clinical test true negatives (correct exclusion)
    - Low rate of clinical test false positives
    - Proportion of correct clinical test exclusions in relation to total gold standard exclusions (0.0-1.0)

Specificity = # True Negative / (# True Negative + # False Positive)

Likelihood Ratios: +LR & –LR

- **Positive Likelihood Ratio (+LR)**
  - Correct identification of condition
    - High True Positive Rate (among cases with condition)
    - Low False Positive Rate (among cases without condition)
- **Negative Likelihood Ratio (–LR)**
  - Correct exclusion of condition
    - Low False Negative Rate (among cases with condition)
    - High True Negative Rate (among cases without condition)

Clinical Test Accuracy

- Likelihood Ratios
  - Positive Clinical Test Result:
    - +LR = Sensitivity / (1 – Specificity) OR TPR / FPR
  - Negative Clinical Test Result:
    - –LR = (1 – Sensitivity) / Specificity OR FNR / TNR

Clinical Test Result Likelihood Ratios

- +LR = TPR / FPR
  - e.g., +LR = 5.0
    - Positive clinical test result is 5X more likely for a person with the condition than for people without the condition
- –LR = FNR / TNR
  - e.g., –LR = 0.2
    - Negative clinical test result is 80% more likely for people without the condition compared to people with the condition (20% or 4X less likely to have a negative clinical test result)

Clinical Test Result Interpretation

+LR	Probability Shift	–LR
>10.0	Conclusive	<.01
5.0-10.0	Moderate	0.1-0.2
2.0-5.0	Small	0.2-0.5
1.0-2.0	Insignificant	0.5-1.0

Jaeschke et al: *JAMA* 271(9); 1994

Olmstead & Denegar: Physical Examination of the Shoulder: Considerations of Sensitivity and Specificity. *Athletic Therapy Today*, 2003

TABLE 4. SPECIAL TESTS FOR LABRAL PATHOLOGY

Special Test, Author	Sensitivity	Specificity	Positive LR	Negative LR	LR Strength
Compression-rotation test					
MacFarland et al.[6]	24%	76%	1	1	very small
Anterior-slide test					
Kibler[7]	78.4%	81.5%	4.23	0.27	small
MacFarland et al.[6]	8%	84%	0.5	1.1	very small
Active compression (O'Brien test)					
O'Brien et al.[8]	100%	98%	50	0	high
MacFarland et al.[6]	47%	55%	1.04	0.96	very small
Guanche and Quick[9]	64%	85%	4.26	0.42	small
Stetson and Templin[10]	54%	31%	0.78	1.81	very small
Crank test					
Liu et al.[11]	91%	93%	13	0.1	high
Stetson and Templin[10]	46%	56%	1.05	0.96	very small
Pain-provocation test					
Mimori et al.[12]	100%	90%	10	0	moderate-high
Magnetic resonance imaging					
Bencardino et al.[13]	90%	91%	10	0.11	moderate

Clinical Prediction/Decision Rules

- A decision-making rule for clinicians that includes *3 or more variables* from:
  1) History
  2) Physical examination
  3) Simple clinical tests

55

Clinical Prediction/Decision Rules

1) Diagnosis of a condition
   - Clinical guideline used in lieu of "gold standard"
2) Prognosis for recovery of function
   - Estimation of likelihood for full recovery under specific conditions and a given timeframe
3) Likely response to a particular treatment
   - Patient characteristics linked to choice of therapy

Clinical Prediction/Decision Rules

- 3 stages:
  1) Development (identification of predictors)
  2) Validation (assessment of accuracy)
  3) Impact analysis (utilization)

57

Clinical Prediction/Decision Rules

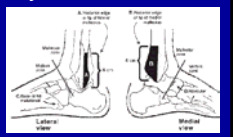
- Derived from systematic clinical observations
- Most useful in complex decision-making circumstances
- Validation of a clinical prediction rule:
  - Sensitivity
  - Specificity
  - Likelihood ratios

58

Ottowa Ankle Rules

Steill et al: *Ann Emerg Med*, 1992

- A means of reducing unnecessary X-rays
  - Only 15% of ankle injuries evaluated in ERs involve Fx
- An ankle X-ray series is required only if:
  1) Pain felt in "malleolar zone"
     AND
  2) Inability to bear weight (4 steps)
     OR
     Point tenderness:
     - **A. Posterior edge or tip of fibular malleolus**
     - **B. Posterior edge or tip of tibial malleolus**
     - **C. Base of the 5th metatarsal**



59

Ottowa Ankle Rules

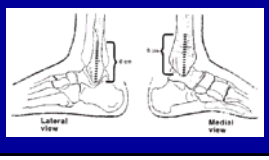
- Steill et al: *Ann Emerg Med*, 1992
  - Published OARs vs. "alternate rules"
    - Sensitivity (TPR): 100% vs. 96%
    - Specificity (TNR): 40% vs. 58%
    - Reduction in X-rays 36% vs. 50%
- Bachmann et al: *Br Med J*, 2003
  - Systematic review of 32 studies of OARs
    - Sensitivity (TPR): 99-100%
    - Specificity (TNR): 26-48%

60

Modified Ottowa Ankle Rules

Leddy et al: *Am J Sports Med*, 1998

- "Buffalo Rules"
  - Point tenderness over midline of the distal 6 cm of either the fibular malleolus or tibial malleolus
    - Sensitivity (TPR): 100%
    - Specificity (TNR): 59%



Clinical Prediction/Decision Rules

- No clinical test or clinical decision rule is both 100% sensitive and 100% specific
- As sensitivity increases (high TPR & low FNR), specificity decreases (low TNR & high FPR)
- How do we maximize the probability of correctly classifying a case as normal vs. abnormal?

Clinical Rule Development

- Event predicted must be clearly defined in terms of dichotomous diagnostic categories or outcomes
  - Presence vs. absence of fracture

    OR
  - Improvement in self-rated function >50% vs. ≤50%
- Regression analysis used to identify the most powerful predictors
  - History – Physical Exam – Clinical Tests

Logistic Regression Analysis

- Used to identify multiple predictors (IVs) of a dichotomous dependent variable (coded 0 or 1)
- Quantifies the relative importance of each IV
- Permits calculation of the likelihood that the DV (outcome) can be predicted from the IVs

Logistic Regression Analysis

- A non-linear transformation of a dichotomous DV
  - Natural logarithm (ln): Exponent to which base 2.71828 must be raised to obtain Y
  - S-shaped distribution of log odds ratio:

    $\ln[p/(1-p)] = a + B_1X_1 + B_2X_2 + B_3X_3$
  - Slope coefficient (B_1) represents rate of change in log odds of event occurrence as X_1 changes
  - Conversion to odds ratio:

    Effect of 1 unit change in X_1 on odds ratio = 2.71828 raised to power of (B_1)

Clinical Rule Accuracy

- Likelihood: Probability (p) of event occurrence when criteria are met (0.0 to 1.0)
  - Example: Probability of positive outcome with 4 out of 5 indicators present = 84 positive / 125 cases = .67
  - Probability of non-event = 1.0 – p OR 1.0 – .67 = .33
- Odds Ratio = p / (1.0 p)
  - Example: 0.67 /.33 = 2.00
    - Positive outcome 2X more likely than negative outcome

Clinical Predicition/Decision Rule Specificity & Sensitivity

- Relative importance of sensitivity vs. specificity depends on clinical circumstances
- Example: Ottowa Ankle Rules
  - Designed to have 100% sensitivity (TPR)
    - No missed fractures; FNR = 0%
  - Specificity (TNR) = 40%; 1- Specificity (FPR) = 60%
    - Reduction in rate of radiography: 36%
  - Alternate rule: 96% sensitivity & 58% specificity
    - FNR = 4%; FPR = 42%; Reduction in radiology rate: 50%

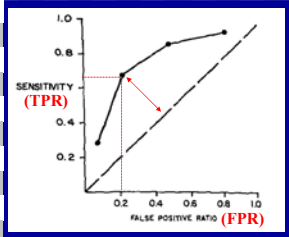
Receiver Operating Characteristic (ROC) Curve

- Illustrates relationship between true-positive ratio and false-positive ratio as definition of "positive" is changed

Factors	Abnormal	Normal
0	2	25
1	3	68
2	8	49
3	16	29
4	12	12
	41	183

# Positive	TPR	FPR
1 or more	39/41=.95	158/183=.86
2 or more	36/41=.88	90/183=.49
3 or more	28/41=.68	41/183=.22
All 4	12/41=.29	12/183=.07

Receiver Operating Characteristic (ROC) Curve



45-degree line represents discriminatory ability that is no better than chance (+LR=1.0)

TPR = Sensitivity

FPR = 1 - Specificity

# Positive	TPR	FPR
1 or more	39/41=.95	158/183=.86
2 or more	36/41=.88	90/183=.49
3 or more	28/41=.68	41/183=.22
All 4	12/41=.29	12/183=.07

Clinical Prediction Rule Example Flynn et al: *Spine*, 2002

- A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation.
  - 71 patients with lumbosacral symptoms (pain/numbness)
  - Baseline Oswestry Disability Questionnaire score ≥30%
  - 61 potential predictor variables
    - **7 patient self-report variables**
    - **14 patient history variables**
    - **22 clinical examination variables**
    - **6 symmetry tests**
    - **5 mobility tests**
    - **7 symptom provocation tests**

Flynn et al: *Spine*, 2002

- Either 2 or 3 treatment sessions separated by 2-4 days
  - Supine/rotated trunk position: posterior/inferior thrust through ASIS
- Definition of treatment success vs. nonsuccess
  - >50% vs. ≤50% improvement in OSW disability score
  - Success: 45% (32 of 71) - Nonsuccess: 55% (39 of 71)
- Data analysis
  - t-tests & χ^2 tests, $p<0.15$: 11 of 61 variables retained
  - Logistic regression analysis: 5 of 11 variables retained
  - Cut-off values established through ROC curve analysis

Flynn et al: *Spine*, 2002

- Identification of patients likely to benefit from spinal manipulation:

1) Duration of symptoms <16 days
2) FABQ work subscale score <19
3) >35° hip internal rotation ROM
4) Hypomobility in lumbar spine
5) No symptoms distal to knee

Flynn et al: *Spine*, 2002

- **A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation.**

Factors	Success	Nonsuccess
0	0	1
1	0	5
2	2	19
3	10	13
4	14	1
5	6	0
	32	39

# Positive	TPR	FPR	+LR
1 or more	32/32=1.00	38/39=.97	1.03
2 or more	32/32=1.00	33/39=.85	1.18
3 or more	30/32=.94	14/39=.36	2.61
4 or more	20/32=.63	1/39=.03	24.38
All 5	6/32=.19	0/39=.00	∞

Flynn et al: *Spine*, 2002

- Likelihood of success in relation to predictors:
  - Without prediction - 71 cases with 32 successes: 45%
  - 1 or more positive in 70 cases with 32 successes: 46%
  - 2 or more positive in 65 cases with 32 successes: 49%
  - 3 or more positive in 44 cases with 30 successes: 68%
  - 4 or more positive in 21 cases with 20 successes: 95%
  - All 5 factors positive in 6 cases with 6 successes: 100%

Clinical Relevance: External Validity

- Few RCTs are designed to address the practical needs of clinicians and patients
  - RCTs typically lack measurement of:
    - Quality of life
    - Patient preferences
    - Costs of service delivery

Clinical Utilization of Research Evidence: Treatment Plan

- **Efficacy:** Level of benefit observed when -
  1) treatment is applied under ideal conditions
  2) to a clearly defined population of subjects
     - Random assignment to experimental and control groups
- **Effectiveness:** Level of benefit expected when -
  1) treatment is provided in customary practice setting
  2) to typical patients

Treatment Efficacy vs. Effectiveness

Dzewaltowski et al: *Ex Sport Sci Rev*, 2004

- Efficacy established under highly controlled conditions may not realized (or sustainable) under conditions of a typical clinical setting
- RE-AIM (Kansas State U. Community Health Institute)
  - **R**each: % of potential recipients in target population
  - **E**fficacy: Magnitude of benefit (treatment effect size)
  - **A**doption: % of clinical settings (and/or % of clinicians)
  - **I**mplementation: Quantity & quality of treatment delivery
  - **M**aintenance: Long-term (≥ 6 months) effectiveness

www.re-aim.org

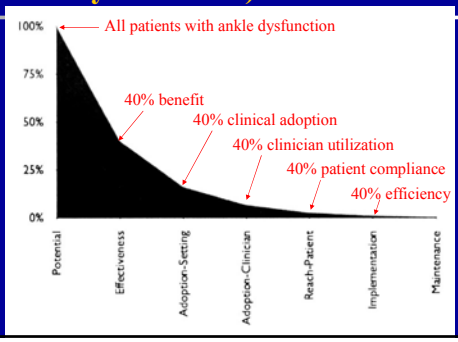
Treatment Efficacy – Clinical Impact

- Reach: % of patients who would derive benefit
  - Election to accept treatment & subsequent compliance may depend on:
    - Gender, age, ethnicity, socioeconomic status, …
      - Often, those in greatest need are the least motivated and most difficult to reach

Treatment Efficacy – Clinical Impact Example: Ankle Strengthening Exercise

- Potential impact: 100% of population with ankle dysfunction
- Efficacious results: Improved function in 40% of research subjects
- Adoption of treatment: 40% of all clinics/institutions
  - Only 16% of population impacted
    - Among 40% of population likely to realize benefit, only 40% have access
- Clinician advocacy: 40% of total clinical staff
  - Only 6.4% of population impacted
- Patient acceptance/compliance: 40% of patients
  - Only 2.6% of population impacted
- Treatment administration: 40% procedural proficiency
  - Only 1% of population impacted
- Maintenance: 40% long-term patient compliance
  - Only 0.5% of population impacted

Theoretical Impact of RCTs (Efficacy Research)



Efficacy vs. Effectiveness Research *RE-AIM: Reach*

- Efficacy Study (Randomized Clinical Trial)
  - Homogeneous sample of highly motivated subjects; exclusion of subjects with complications or co-morbidities
- Effectiveness Study (Outcomes Study)
  - Heterogeneous sample of typical rehabilitation patients; inclusion of subjects with complications or co-morbidities

Efficacy vs. Effectiveness Research *RE-AIM: Efficacy/Effectiveness*

- Efficacy Study (Randomized Clinical Trial)
  - Random assignment to experimental or treatment group; specialized interventions delivered in standardized manner to maximize effect size
- Effectiveness Study (Outcomes Study)
  - Quasi-experimental design; feasible interventions not requiring great expertise or special equipment

Efficacy vs. Effectiveness Research *RE-AIM: Adoption*

- Efficacy Study (Randomized Clinical Trial)
  - Typically one setting to reduce variability; extensive clinical resources and a high level of professional expertise
- Effectiveness Study (Outcomes Study)
  - Data may be collected in multiple clinical settings; equipment design and level of clinician expertise may vary greatly

Efficacy vs. Effectiveness Research *RE-AIM: Implementation*

- Efficacy Study (Randomized Clinical Trial)
  - High degree of coordination among research staff to follow specific treatment protocol
- Effectiveness Study (Outcomes Study)
  - Treatment administered by a independent clinicians with competing daily demands; some degree of treatment protocol modification probable

Efficacy vs. Effectiveness Research
RE-AIM: Maintenance

- Efficacy Study (Randomized Clinical Trial)
  - Cost of treatment often not relevant to research question (typically not an issue or concern)
- Effectiveness Study (Outcomes Study)
  - Cost of treatment becoming increasingly important
  - Sustainability of treatment results is a major issue (work and/or sports functional abilities)

85

Outcomes Research

- True experimental designs with rigorous control rarely feasible
  - Poor internal validity
- Requires utilization of social science research methodology
  - Rating scales

86

Clinical Prediction/Decision Rule Development

- Wasson et al: *N Eng J Med, 1985*
  - Clinical prediction rules: Applications and methodological standards.
    - If possible, the predicted outcome should be biologic rather than sociological or behavioral.
- Wilson & Gansneder: *JOSPT*, 2000
  - Measures of functional limitation as predictors of disablement in athletes with acute ankle sprains.
    - Self-reported measures of functional limitation were stronger predictors of disability duration than measures of impairment (ROM loss & swelling volume).

Biopsychosocial Model of Dysfunction (Nagi, 1965)

- Injury
  - ligament sprain
- Impairment
  - pain, swelling, ROM limitation
- Functional Limitation
  - inability to walk
- Disability
  - inability to participate in sport/work

88

Outcomes Research

- Health status can be viewed as continuum
  - Biological & physiological measures at one end
    - Cellular processes, organ function, joint stability, etc.
  - Perception of health-related quality of life at the other
    - Ability to function in society

Wilson & Cleary: *JAMA*, 1995

Quantification of Change in Functional Capabilities

- The association between physical impairments and functional capabilities is very weak!
  - Wilson & Gansneder, *JOSPT*, 2000
  - Man & Morrisey, *Med Sci Sports Ex*, 2005
- Patient self-report is the only mechanism available for acquisition of meaningful data
  - The patient is the best judge of change in status

Categories of Health/Function Survey Instruments

- Generic
  - SF-36
  - Permits comparisons across diagnostic categories & demographic subgroups
- Disease/Condition-Specific or Population-Specific
  - SPADI, Oswestry Disability Index, IKDC, FADI
- Patient-Specific
  - Most responsive (most sensitive to change)
  - Not designed for between-group comparisons

Survey Psychometric Properties: Validity

- Construct Validity
  - Criterion-Related: established by comparison to a recognized gold standard
  - Content-Related: proper representation of the various contexts associated with the construct
    - Functional ability contexts:
      - Activities of Daily Living
      - Work-Related Activities
      - Sports-Related Activities
- Responsiveness (sensitivity to change in status)
  - Standardized Response Mean (effect size)

Standardized Response Mean

- Representation of improvement magnitude in standard deviation units (effect size)
  - Basis for quantitative comparison of results from studies that used different methods (meta-analysis)
- Permits comparison of improvement between different clinical samples

SRM = post mean - pre mean / std dev of change scores

Evidence-Based Sports Medicine

- Downhall MacAuley (Queen's University, Belfast, Ireland) & Thomas Best (University of Wisconsin Medical School)

What recommendations should be made concerning exercising with a fever and/or acute infection?

Does stretching help prevent injuries?

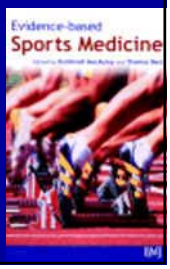
Does exercise help or harm in osteoarthritis of the knee?

How should athletes with chronic low back pain be managed in primary care?

How reliable is the physical examination in the diagnosis of sports related knee injuries?

Can we prevent ankle sprains?

www.evidencebasedsportsmedicine.com



Athletic Training Education

- Denegar & Hertel: *JAT*, 2002 Editorial
  - Clinical education reform and evidence-based clinical practice guidelines.
    - Review of therapeutic ultrasound effectiveness studies: Robinson & Baker, *Phys Ther*, 2001
      - Little evidence that ultrasound is more effective than placebo for treatment of musculoskeletal conditions.
    - Clinical proficiencies: Which skills?

Medical Treatment Guidelines

- McGlynn et al: The quality of health care delivered to adults in the United States. *N Eng J Med* 348:2635-2645, 2003
  - Review of medical records for 6,712 people in 12 cities
  - 439 indicators of quality of care for 30 conditions
  - 55% of cases reviewed received recommended care

"The deficits we have identified in adherence to recommended processes for basic care pose serious threats to the health of the American public."

Improving Health Care: A Dose of Competition

- Report of the Federal Trade Commission & Department of Justice – July 2004
  - Tens of billions of dollars are spent annually on services whose value is questionable or non-existent.
  - At present, most payments to providers have no connection with the quality of care provided.
  - Vigorous competition promotes the delivery of high quality, cost-effective health care, and vigorous antitrust enforcement helps protect competition.

Athletic Training & EBM

- NATA Pronouncements Committee
  - Guidelines for special circumstances:
    - Fluid replacement
    - Lightning
    - Concussion management
    - Exercise-induced asthma
  - Clinical evaluation & treatment methods
    - Needed to guide management of conditions
- National Guidelines Clearinghouse
  - Agency for Healthcare Quality and Research
    - www.guideline.gov

Thank you

Gary Wilkerson, EdD, ATC
University of Tennessee at Chattanooga
423-425-5394
Gary-Wilkerson@utc.edu

Presentation available as PDF file download:
www.utc.edu/gatp → "Program News"