Evidence-based Medicine and the Pelvic Girdle

Sayers John Miller, PhD, PT, ATC
Penn State University
Evidence-based Medicine

“Evidence-based medicine (EBM) is the integration of best research evidence with clinical experience and patient values”

Sackett et al. (2000)
How to Practice EBM

- 1- Convert the need for information into an answerable question
- 2- Track down the best evidence with which to answer the question
- 3- Critically appraise evidence for validity, impact and applicability
- 4- Integrate the critical appraisal with our clinical expertise and with our patient’s unique biology, values and circumstances
- 5- Evaluate our effectiveness and efficiency in executing steps 1-4 and seek ways to improve them both for next time

(Sackett et al. 2000)
Types of Questions

- **Diagnosis**
  - Reliability
    - Kappa, ICC
  - Validity
    - Sensitivity, specificity, LR
    - Methodology

- **Treatment**
  - Outcomes
    - Systematic reviews, RCT, cohort studies, case-control, case-series, expert opinion
    - LR, effect size, risk reduction, numbers needed to treat
Kappa

- $K = \frac{(Po-Pe)}{(1-Pe)}$
  - $K = \text{generalized Kappa}$
  - $Po = \text{proportion of observed agreement}$
  - $Pe = \text{proportion of agreement expected by chance}$

- **Interpretation**
  - $K = 1$, perfect agreement
  - $K = 0$, chance agreement
  - $K = -N$, less than chance agreement
Reliability

- **Kappa**
  - 0.00-0.20  Slight
  - 0.21-0.40  Fair
  - 0.41-0.60  Moderate
  - 0.61-0.80  Substantial
  - 0.81-1.00  Almost perfect
Intraclass Correlation Coefficients (ICC)

- Rates reliability by comparing variability of different raters of same subject to total variation across all ratings and all subjects
- \[ ICC = \frac{\sigma^2(b)}{\sigma^2(b) + \sigma^2(w)} \]
  
  Where:
  - \( \sigma^2(b) \) is variance of trait between subjects and,
  - \( \sigma^2(w) \) is pooled variance within subjects
Sensitivity/ Specificity

- **Sensitivity**
  - Proportion of people with the disorder who have a positive test

- **Specificity**
  - Proportion of people without the disorder who have a negative test

- **SpPin**
  - When a test has high Specificity, a positive result rules in the diagnosis.

- **SnNout**
  - When a test has high Sensitivity, a negative result rules out the diagnosis.
### Calculation of Sensitivity/ Specificity/ LR

<table>
<thead>
<tr>
<th></th>
<th>Disease positive</th>
<th>Disease negative</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test positive</strong></td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td><strong>Test negative</strong></td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

- **Sensitivity** = $\frac{a}{a+c}$
- **Specificity** = $\frac{d}{b+d}$
- **LR+** = sensitivity/(1-specificity) = $\frac{a}{a+c}\div(b/b+d)$
- **LR-** = (1-sensitivity)/specificity = $\frac{c}{a+c}\div(d/b+d)$
- **PPV** = $\frac{a}{a+b}$
- **NPV** = $\frac{d}{c+d}$
- **Pre-test probability** = $\frac{a+c}{a+b+c+d}$
How Do You Know Condition is Present?

- **Gold standards**
  - **SI joint**
    - Pain mapping: Provocative injection
    - Pain: Fluoroscopically guided anesthetic injection (double block) with specified level of pain reduction (75-90%)
    - Movement: Radiostereophotogrammetric analysis
  - **Disc**
    - Discogram
  - **Facet joint**
    - Pain: Fluoroscopically guided anesthetic injection (double block)
Likelihood Ratio (LR)

- Likelihood ratio
  - The likelihood that a given test result would be expected in a patient with the disorder compared with the likelihood that the same result would be expected in a patient without the disorder

- Pre-test probability
  - The proportion of people with the disorder in the population at risk at a specific time or time interval
Positive LR (LR +)

- Probability of patient with the condition having a positive test/probability of patient without the condition having the result

- $\text{LR}^+ = \frac{\text{Sensitivity}}{1-\text{Specificity}}$

- Score interpretation (Jaeschke et al. 1994)
  - 2.0-5.0 generates small shift in probability
  - 5.0-10.0 generates moderate shift in probability
  - >10.0 generates large and often conclusive shift in probability
Negative LR (LR-)

- Probability of an individual with the condition having a negative test/probability of an individual without the condition having a negative test
- $LR^- = 1 - \text{Sensitivity/Specificity}$
Likelihood Ratio Interpretation

- >10  Large increase in likelihood
- 5-10  Moderate increase in likelihood
- 2-5  Small increase in likelihood
- 1-2  Minimal increase in likelihood
- 1  No change in likelihood
- 0.5-1.0  Minimal decrease in likelihood
- 0.2-0.5  Small decrease in likelihood
- 0.1-0.2  Moderate decrease in likelihood
- <0.1  Large decrease in likelihood
Predictive Value

- Positive predictive value (PPV)
  - Proportion of people with a positive test who have the disorder

- Negative predictive value (NPV)
  - Proportion of people with a negative test who are free of the disorder
Levels of Evidence
Oxford Centre of Evidence-based Medicine Levels of Evidence (May 2001)

- 1a: SR with homogeneity of RCT’s
- 1b: Individual RCT with narrow CI
- 1c: All or none
- 2a: SR with homogeneity of cohort studies
- 2b: Individual cohort study (low quality RCT’s)
- 2c: Outcomes research
- 3a: SR with homogeneity of case-control studies
- 3b: Individual case-control study
- 4: Case-series (poor quality cohort or case studies)
- 5: Expert opinion without explicit critical appraisal, or based on physiology, bench research or first principles
Grades of Recommendation

- **A**: Consistent Level 1 studies
- **B**: Consistent Level 2 or 3 studies or extrapolations from Level 1 studies
- **C**: Level 4 studies or extrapolations from Level 2 or 3 studies
- **D**: Level 5 evidence or troubling inconsistent or inconclusive studies of any level
PEDro Scale

1. Eligibility criteria specified
2. Subjects randomly allocated to groups
3. Allocation concealed
4. Groups similar at baseline regarding most important prognostic indicators
5. Blinding of all subjects
6. Blinding of all therapists performing treatment
7. 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 analyzed by “intention to treat”
10. The results of between-group statistical comparisons are reported for at least one key outcome
11. The study provides both point measure and measures of variability for at least one key outcome

www.pedro.fhs.usyd.edu.au
Prevalence

- 15% in chronic low back pain patients
  - Maigne et al. 1996 (18.5%)
  - Schwarzer et al. 1995

- Chronic low back pain (Bogduk 2005)
  - Internal disc disruption (39%)
  - Sacroiliac joint (15%)
  - Zygapophysial joint (<10%, young, 32% elderly)
  - Iliac crest syndrome (30-50%)
  - Iliolumbar ligament (<10%)
  - Spondylolysis (<6%)
Sacroiliac Joint Examination

- History
- Pain location
- Landmark palpation
- Motion palpation
- Pain provocation tests
- Multiple tests
Diagnosis with Orthodox Clinical Examination

- Poor ability to diagnose SI joint pain
  - Schwarzer et al. 1995
- No component of history, physical exam test or ensemble of them is able to identify SI joint pain
  - Dreyfuss et al. 1996
Pain Location

- **Fortin et al. (1994)**
  - Asymptomatic subjects
  - Provocative fluoroscopically guided injections
  - Common referral
    - Area 10 cm caudal and 3 cm lateral to PSIS
  - Confirmed with injections in patients with LBP
    - 10 of 16 had at least 50% reduction in pain

- **Dreyfuss et al. (1996)**
  - Documented pain referral patterns in LBP patients with positive response to SI joint block and without
  - Broad referral patterns: buttocks, groin, entire LE
  - No referral above L5 in SI joint patients
Fortin Finger Test

- Fortin and Falco (1997)
  - 56 consecutive LBP patients
  - 2 examiners selected the same 12 patients based upon (4 additional after Patrick’s test)
    - Pain localized with one finger
    - Pain inferomedial to PSIS within 1 cm
    - Patient pointed to same area on two trials
  - All 16 had provocation-positive SI joint injections
    - 4 injected bilaterally negative on non-painful side
  - 10 patients had further evaluation with negative lumbar imaging
Pain Location

- **Dreyfuss et al. (1996)**
  - Sacral sulcus tenderness
    - Sensitivity: 0.84-0.93, Specificity: 0.10-0.17
    - LR = 1.00
  - PSIS pointing
    - Sensitivity: 0.71, Specificity: 0.47
    - LR = 1.33

- **Kristiansson & Svärdsudd (1996)**
  - Sacrospinous ligament
    - Sensitivity: 0.64, Specificity: 0.89
    - + LR = 0.72
  - PSIS
    - Sensitivity: 0.35, Specificity: 0.98
    - + LR = 0.36
Landmark Palpation

- Potter & Rothstein (1985)
  - 17 low back pain patients, 8 therapists
  - Inter-examiner agreements
    - In standing
      - ASIS (38%), PSIS (35%), iliac crests (35%)
    - In sitting
      - ASIS (44%), PSIS (35%), iliac crests (41%)
Landmark Palpation

- O’Haire & Gibbons (2000)
  - 10 subjects, Ten 5th year osteopathic students
  - PSIS, SS, ILA positional relationships (prone)
  - Intra-examiner
    - PSIS, K = 0.33 (0.07 to 0.58), ILA, K = 0.21 (-0.05 to 0.69), SS, K = 0.24 (0.02 to 0.60)
  - Inter-examiner
    - PSIS, K = 0.04, ILA, K = 0.08, SS, K = 0.07

- Flynn et al. (2002)
  - 55 patients, 2 physical therapist examiners
  - Inter-examiner (kappa)
    - ASIS in standing (0.31), Iliac crests in standing (0.23), PSIS in sitting (0.23), PSIS in standing (0.13), ischial tuberosity in prone (0.03), pubic tubercle in supine (-0.04)
Landmark Palpation

- Albert et al. (2000)
  - 2269 subjects, pelvic topography - 4 pairs of landmarks
  - ASIS, PSIS, IC, Greater trochanters (standing)
  - Inter-examiner reliability
    - Kappa = 0.55
  - Diagnosis of SIJD
    - Sensitivity (0.77), Specificity (0.32), LR+ (1.13)
Landmark Palpation

  - Systematic review
  - “Insufficient evidence to advocate the use of static palpation tests to indicate SIJ dysfunction”
Diagnosis - Mobility

■ Gold standard
  - Radiostereometric (Sturesson 1998)
  - CT scan (Smidt et al. 1997)
  - 3-D video kinematic (Hungerford & Gildeard 1998)
Motion Palpation

- **Smidt et al. (1997)**
  - 5 fresh cadavers (52-68 years old)
  - CT scan, lead markers on sacrum and iliums
  - Angular (sagittal)
    - Bilateral flexion to extension (7° left, 8° right)
    - Reciprocal hip flex/ext (5° left, 8° right)
    - Range (3° to 17°)
  - Linear
    - PSIS to sacrum (4 to 8 mm)

- **Sturesson et al. (1989)**
  - In Vivo (18-45 years old, with/without back pain)
  - Radiostereophotogrammetric analysis
  - Angular (sagittal): Flex to Ext (2.5° ± 0.5°), range (1.6° to 3.9°)
  - Translation: 0.7 mm (0.1-1.6 mm)

- **Sturesson et al. (2000)**
  - In Vivo (28-35 yrs. old, 6 women with pelvic pain after pregnancy)
  - Radiostereophotogrammetric analysis, standing in straddle position
  - Angular movement of innominates: 1.9°, range (1.3° to 2.4°)
Gillet’s Test

- **Validity**
  - Dreyfuss et al. (1994)
    - 16% false positives (4.2% males, 26.4% females)
    - Test positive in asymptomatic subjects (20%)
  - Dreyfuss et al. (1996)
    - Sensitivity (0.43), Specificity (0.68), LR (1.3)
  - Sturesson (1998)
    - Invalid, equal movement both sides
    - Invasive radiostereometric testing
  - Hungerford & Gillear (1998)
    - Valid, 7.01 degrees of posterior rotation
    - Video kinematics
  - Hungerford et al. (2003)
    - Delayed onset of OI, Multifidus & Glut Max in stance leg in SIJP patients versus controls
    - Early onset of biceps femoris
Gillet’s Test

- **Reliability**
  - **Intra-examiner**
    - Meijne et al. (1999)
      - Kappa = -0.005 to 0.073
    - Carmichael (1987)
      - Kappa = 0.02 (85.3% agreement)
    - Herzog et al. (1989)
      - 68% positive finding, 79% negative finding, 72% positive on correct side (p<0.01)
Gillet’s Test

- **Reliability**
  - **Inter-examiner**
    - Flynn et al. (2002)
      - Kappa = 0.59
    - Meijne et al. (1999)
      - Kappa = -0.032 to 0.081
    - Dreyfuss et al. (1996)
      - Kappa = 0.22 (54% agreement)
    - Carmichael (1987)
      - Kappa = 0.18 (89.2% agreement)
    - Herzog et al. (1989)
      - 65% positive finding, 61% correct side (P<0.01)
    - Potter & Rothstein (1985)
      - 47% agreement
Standing Flexion Test

- Validity
  - Colachis et al. (1963)
    - Symptomatic males
    - Radiographs with Kirschner wires in PSISs
    - Superior movement of PSISs with forward flexion
  - Dreyfuss et al. (1994)
    - 101 asymptomatic subjects
    - 13% false positives (10.4% in males, 17% in females)
  - Egan et al. (1996)
    - 128 asymptomatic subjects with or without history of LBP
    - 24% false positive
    - Positive test correlated with taller, heavier, male and asymmetric subjects. Also those with history of PBP
Standing Flexion Test

- **Reliability**
  - **Intra-examiner**
    - Vincent-Smith & Gibbons (1999)
      - Kappa = 0.46 (68% agreement)
  - **Inter-examiner**
    - Flynn et al. (2002)
      - Kappa = -0.08
    - Riddle et al. (2002)
      - Kappa = 0.32
    - Vincent-Smith & Gibbons (1999)
      - Kappa = 0.052 (42% agreement)
    - Van der Wurff et al. (1996)
      - Kappa = 0.29 (74% agreement)
    - Potter & Rothstein (1985)
      - 44% agreement
Standing Flexion Test

- **Limitations**
  - Positive test may only indicate non-specific change in lumbo-pelvic-hip mechanics
  - False positives
    - Unilateral restriction of motion at hip
    - Unilateral restriction of low lumbar joints
    - Unilateral tightness of myofascia
    - Leg length inequality
Seated Flexion Test

- **Validity**
  - Dreyfuss et al. (1994)
    - 8% false positives (8.3% males, 7.5% females)

- **Reliability (Inter-examiner)**
  - Albert et al. (2000)
    - Kappa = >0 (88% agreement)
  - Flynn et al. (2002)
    - Kappa = 0.25
  - Potter & Rothstein (1985)
    - 50% agreement
Diagnosis - Pain Provocation

- **Gold Standard**
  - Injection with local anesthetic (single)
    - False-positive rate in facet joints of 38%
  - Double blocks advocated in pain literature

- **Problems with gold standard**
  - Only investigates intra-articular pain sources
  - May not affect all parts of joint capsule
  - Needle passage through structures may increase pain
  - VAS cut-off selection
    - Should be as close to 100 as possible
Thigh Thrust

- **Validity**
  - Laslett et al. (2005)
    - Sensitivity (0.88), Specificity (0.69), PPV (0.58), NPV (0.92), LR+ (2.80), LR- (0.18)
  - Ostgaard et al. (1994)
    - Sensitivity (0.80), Specificity (0.81), PPV (0.71), NPV (0.88), LR+ (4.07), LR- (0.23)
  - Dreyfuss et al. (1996)
    - Sensitivity (0.36), Specificity (0.50), LR+ (0.8)
  - Broadhurst & Bond (1998)
    - Sensitivity (0.80), Specificity (1.00)
Thigh Thrust

- **Reliability**
  - Inter-examiner
    - Dreyfuss et al. (1996)
      - Kappa = 0.64 (82% agreement)
    - Flynn et al. (2002)
      - Kappa = 0.70
    - Laslett & Williams (1994)
      - Kappa = 0.88 (94% agreement)
    - Kokmeyer et al. (2002)
      - Kappa = 0.69
    - Albert et al. (2000)
      - Kappa = 0.70
Gaenslen’s Test

- Reliability
  - Inter-examiner
    - Dreyfuss et al. (1996)
      - Kappa = 0.61 (82% agreement)
    - Laslett & Williams (1994)
      - Kappa = 0.72-0.75 (88% agreement)
    - Kokmeyer et al. (2002)
      - Kappa = 0.58
    - Flynn et al. (2002)
      - Kappa = 0.54
Motion/ Pain Provocation

- Potter & Rothstein (1985)
  - Percentage agreement
    - Gillet (47%), seated flexion (50%), standing flexion (44%), prone knee flexion (24%), supine gapping (94%), sidelying compression (76%), supine long sit (40%)
Motion/ Provocation/ Palpation

- Dreyfuss et al. (1996)
- Reference standard
  - Single injection with fluoroscopy, 90% reduction in pain
- Method
  - Tests picked by a panel of experts from various professions
  - 85 patients, 2 examiners (MD, Chiropractor)
- Results
  - Inter-examiner reliability
    - Gillet (0.22), sacral spring (0.15), sacral sulcus tenderness (0.41), and sacral thrust (0.30)
    - Groin pain (0.70), buttock pain (0.71), SI joint pain (0.67), Gaenslen’s test (0.61), thigh thrust (0.64), Patrick’s (0.62), PSIS pointing (0.60)
  - Validity
    - Gillet, thigh thrust, Patrick’s test, Gaenslen’s, sacral thrust (LR 0.1-1.3)
  - No component of history, physical exam test or ensemble of them identified SI joint pain (responders to block)
  - No pain above L5 in patients who responded to SI injection and PSIS pointing or sacral sulcus pointing (PPV 0.60)
Pain Provocation

Laslett & Williams 1994

- Five physical therapist pairs
  - Thigh thrust (0.88), 94% agreement
  - Gaenslen (0.72-0.75), 88% agreement
  - Compression (0.73), 88% agreement
  - Distraction (0.69), 88% agreement
  - Cranial shear (0.61), 84% agreement
  - Sacral thrust (0.52), 78% agreement
Pain Provocation

- Maigne et al. 1996
  - 54 patients (unilateral pain, x-ray/CT to rule out low back
  - Double block gold standard (75% reduction)
  - 18.5% prevalence
  - No significant correlation between double block response and SI joint pain provocation tests
    - Distraction, compression, Gaenslen, Patrick, resisted hip external rotation, sacral thrust
      - 19/54 decreased 75% with initial block
      - 10/19 decreased 75% with confirmatory block
Motion/ Provocation/ Palpation

- Albert et al. (2000)
  - Two physical therapists, 34 pregnant women, 8 symptomatic
  - Inter-examiner reliability (kappa)
    - Menell’s (0.87), separation (0.84), compression (0.84), thigh thrust (0.70), Patrick’s (0.54), pelvic topography (0.55), palpation of long dorsal ligament (0.34), Piedallu’s test (>0)
Motion/ Provocation/ Palpation

- Flynn et al. 2002
  - (55 patients, 2 physical therapist examiners)

- Provocation tests (kappa)
  - Posterior shear (0.70), sacral sulcus test (0.64), Patrick test (0.60), Gaenslen (0.54), resisted hip abduction (0.41), sacral thrust (0.41), compression-distraction (0.26)

- Motion tests (kappa)
  - Gillet (0.59), seated flexion (0.25), long-sitting (0.21), prone knee bend (0.21), standing flexion (-0.08)

- Symmetry tests (kappa)
  - ASIS in standing (0.31), Iliac crests in standing (0.23), PSIS in sitting (0.23), PSIS in standing (0.13), ischial tuberosity in prone (0.03), pubic tubercle in supine (-0.04)
<table>
<thead>
<tr>
<th>Test</th>
<th>Author (Method score)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distraction</td>
<td>Blower (45)</td>
<td>.21</td>
<td>1.00</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Russell (39)</td>
<td>.11</td>
<td>.90</td>
<td>No</td>
</tr>
<tr>
<td>Compression</td>
<td>Blower (45)</td>
<td>.00</td>
<td>1.00</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Russell (39)</td>
<td>.07</td>
<td>.90</td>
<td>No</td>
</tr>
<tr>
<td>Gaenslen</td>
<td>Dreyfuss (58)</td>
<td>.71</td>
<td>.26</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Russell (39)</td>
<td>.21</td>
<td>.72</td>
<td>No</td>
</tr>
<tr>
<td>Sacral thrust</td>
<td>Dreyfuss (58)</td>
<td>.53</td>
<td>.29</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Blower (45)</td>
<td>.27</td>
<td>1.00</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Russell (39)</td>
<td>.03</td>
<td>.90</td>
<td>No</td>
</tr>
<tr>
<td>Thigh thrust</td>
<td>Östgaard (64)</td>
<td>.80</td>
<td>.81</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Dreyfuss (58)</td>
<td>.36</td>
<td>.50</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Broadhurst (50)</td>
<td>.80</td>
<td>1.00</td>
<td>Yes</td>
</tr>
<tr>
<td>Patrick’s sign</td>
<td>Dreyfuss (58)</td>
<td>.69</td>
<td>.16</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Broadhurst (50)</td>
<td>.77</td>
<td>1.00</td>
<td>Yes</td>
</tr>
<tr>
<td>Resisted abduction</td>
<td>Broadhurst (50)</td>
<td>.77</td>
<td>1.00</td>
<td>Yes</td>
</tr>
<tr>
<td>Gillet’s</td>
<td>Dreyfuss (58)</td>
<td>.43</td>
<td>.68</td>
<td>No</td>
</tr>
<tr>
<td>Long sitting</td>
<td>Bemis (23)</td>
<td>.41</td>
<td>.83</td>
<td>Yes</td>
</tr>
</tbody>
</table>

van der Wurff et al. 2000
Composite Tests

- **Cibulka et al. (1998)**
  - Standing flexion test, PSIS levels in sitting, supine/long sit test, prone knee flexion test
    - Intra-examiner reliability: $\kappa = 0.86$
    - Inter-examiner reliability: $\kappa = 0.88$ (Cibulka, 1988)
  - Hip internal and external rotation
    - ICC (1,3) = 0.97 (IR) and 0.98 (ER)
  - Hip external rotation ROM greater than internal rotation on side of posterior innominate rotation in patients with SI joint regional pain
Composite Tests

- Cibulka et al. (1999)
- Tests
  - Standing flexion test, PSIS levels in sitting, supine/long sit test, prone knee flexion test
  - Reliability (previous studies)
    - Intra-examiner reliability: Kappa = 0.86 (1998)
    - Inter-examiner reliability: Kappa = 0.88 (1988)
- Composite score
  - 3 of 4 tests positive
    - Sensitivity 0.82, Specificity 0.88
    - LR 6.83
    - PPV 0.86, NPV 0.84
Composite Tests

- Riddle et al. (2002)
- 65 patients, 34 therapists
- Inter-examiner reliability (kappa)
  - Standing flexion test (0.32)
  - Prone knee flexion (0.26)
  - Supine long sit test (0.19)
  - Sitting PSIS levels (0.37)
- Composite testing reliability
  - Kappa = 0.11-0.23
Composite Tests

- Kokmeyer et al. 2002
  - 2 final year PT students, 78 subjects, 59 symptomatic
  - 5 SIJ tests (Distraction, compression, Gaenslen, Patrick, thigh thrust)
    - Individual tests (kappa): Distraction (0.46), compression (0.58), Gaenslen (0.58), Patrick (0.62), thigh thrust (0.69)
    - Composite tests (kappa): 3/5 (0.70), 1/5 (0.63), 2/5 (0.74), 4/5 (0.71), 5/5 (0.66)
Pain Provocation: Composite Tests

- Laslett et al. 2003
  - Double block gold standard (80% reduction)
  - 3 or more SI tests positive
    - Distraction, compression, thigh thrust, Gaenslen’s, sacral thrust
    - With centralizers in
      - Sensitivity 0.91, Specificity 0.78
      - LR+ = 4.16 (95% CI, 2.16 to 8.39), LR- = 0.12
    - With centralizers removed
      - Sensitivity 0.91, Specificity 0.87
      - LR+ = 6.97 (95% CI, 2.70 to 20.27), LR- = 0.11
Pain Provocation: Composite Tests

- Laslett et al. 2005
  - 3 or more positive tests
    - Distraction, compression thigh thrust, sacral thrust, Gaenslen
      - Sensitivity = 0.94, Specificity = 0.78
      - LR+ = 4.29, LR- = 0.80
      - AUC = 0.842
  - 2 of 4 tests positive
    - Distraction, thigh thrust, compression, sacral thrust
      - Sensitivity 0.88, Specificity 0.78
      - LR+ = 4.00 (95% CI, 2.13 to 8.08), LR- = 0.10
      - AUC = 0.819
Pain Provocation: Composite Tests

- Laslett et al. (2005)
  - Individual tests
    - Distraction: Sensitivity (0.60), Specificity (0.81)
    - Compression: Sensitivity (0.69), Specificity (0.69)
    - Thigh thrust: Sensitivity (0.88), Specificity (0.69)
    - Gaenslen: Sensitivity (0.50-0.53), Specificity (0.71-0.77)
    - Sacral thrust: Sensitivity (0.63), Specificity (0.75)
  - Optimal rule for use in diagnosis
    - Distraction and thigh thrust most sensitive and specific
    - Order of performance
      - Distraction, thigh thrust, compression, sacral thrust
      - Stop when two positive
Systematic Reviews

  - SI joint motion palpation tests slightly reliable but not shown to be valid
  - Only palpation for pain has acceptable results

- French et al. (2000)
  - Common clinical methods of detecting manipulable lesion not reliable

- van der Wurff et al. (2000)
  - Mobility and pain provocation tests unreliable except for Gaenslen test and thigh thrust
  - Need further research emphasizing multiple tests and pain provocation tests
van der Wurff et al. (2000)

- Poor methodological quality in studies of validity of mobility and pain provocation studies
- Only Ostgaard study (thigh thrust) showed potentially acceptable methodology
- Questionable if any SI joint test are of value in clinical practice
Systematic Reviews

- **Waters (2003)**
  - “Insufficient evidence to advocate the use of static palpation tests to indicate SIJ dysfunction”

- **Freburger & Riddle (2001)**
  - “Some evidence to support use of pain provocation tests and patient’s pain description in the identification of dysfunction in the SI joint region”
    - Patrick’s, thigh thrust, compression/distraction, resisted hip abduction, sacral sulcus tenderness
    - No lumbar pain, pain below L5, pain in region of PSIS, pain in groin
Problems with Research

- Low methodological quality
- Execution of tests in non-standard or inexperienced manner
- Multiple tests may change musculoskeletal system and previous test may influence subsequent tests
- Questionable gold standard
- Use of asymptomatic subjects can influence statistical measures
  - Kappa takes into account frequency of (+) and (-), producing higher kappa values for agreement on positive findings
    - Use of asymptomatic subjects make it difficult to demonstrate agreement above chance
  - Sensitivity difficult to demonstrate when there is disproportionate number of negative findings
  - Good specificity can be show
Active Straight Leg Raise Test

- Mens et al. (2001)
  - Subjects (200 PPPP, 50 healthy controls)
  - Reliability
    - Patient (ICC = 0.83)
    - Assessor (ICC = 0.82)
    - Patient to assessor (ICC = 0.77)
  - Validity
    - Overall (Sensitivity = 0.87, Specificity = 0.94, LR 13.3)
    - QBPDS score > 45 (Sensitivity = 1.00)
    - QBPDS score < 45 (Sensitivity = 0.73)
    - Thigh thrust (Sensitivity = 0.69)
    - Correlation with thigh thrust (0.27)
Differential Diagnosis

- Young et al. (2003)
  - 81 chronic low back pain patients
  - Discograms, injections (single)
  - Pain generators
    - Discogenic pain
      - Midline lumbar pain, pain with rising from sitting,
      - Centralization (sensitivity = 0.47, specificity = 1.00)
    - Sacroiliac pain
      - Unilateral pain below L5, pain with rising from sitting,
      - 3 or more pain provocation test positive
      - Centralization (sensitivity = 0.09, specificity = 0.79)
    - Facet joint pain
      - Midline lumbar pain?, no pain with rising from sitting
      - Centralization (sensitivity = 0.00, specificity = 0.89)
Centralization

- Validity
  - Young et al. (2003)
    - Sensitivity = 0.47, Specificity = 1.00
    - 47% of patients with (+) discogram centralized
    - 100% of patients who centralized had (+) discogram
  - Donelson et al. (1997)
    - Sensitivity = 0.94, Specificity = 0.52
    - 64% of patients with (+) discogram centralized
    - 74% of patients who centralized had (+) discogram
    - 69% of patients who peripheralized had (+) discogram
Centralization

- Kilpikoski et al. (2002)
- 2 MDT trained therapists
- Reliability
  - Centralization
    - Presence: kappa = 0.7 (95% agreement)
    - Direction: kappa = 0.9 (90% agreement)
  - Patient classification
    - Main syndrome: kappa = 0.6 (95% agreement)
    - Specific subgroup: kappa = 0.7 (74% agreement)
Directional Preference

- Long et al. (2004)
- Prospective, cohort study with 312 patients
- Patient groups
  - 230 (74%) had directional preference and used as subjects
  - Directional preference matched, DP opposite, evidence-based group (multi-directional, mid-range lumbar exercises)
- Results
  - 74% of those meeting inclusion criteria had DP
  - 83% extension, 7% flexion, 10% lateral responders
  - Matched group showed significantly greater improvement than other two groups in all outcome measures
    - Back pain, leg pain, disability, medication, activity interference, depression inventory
Facet Joint Pain

- Revel et al. (1992)
- Low back pain patients
- 7 variables
  - Age greater than 65
  - Pain not worsened with forward flexion
  - Pain not worsened with rising from forward flexion
  - Pain well relieved by recumbency
  - Pain not exacerbated by coughing
  - Pain not worsened by extension rotation
  - Pain not worsened by hyperextension
- Responders to facet injection
  - When 5 variables present
  - Sensitivity = 80%, Specificity = 78%
  - LR = 3.64
Facet Joint Pain

- Revel et al. (1998)
- 80 chronic low back pain patients
  - 43 (+) for 5 or more variables, 37 (+) for less than 5 variables
- Saline or single facet joint injection (Lidocaine)
  - 75% reduction in VAS score
- 5 significant predictors of (+) response to Lidocaine
  - 100% of responders, 66% of non-responders
  - Pain not worsened with rising from forward flexion
  - Pain well relieved by recumbency
  - Pain not exacerbated by coughing
  - Pain not worsened by extension rotation
  - Pain not worsened by hyperextension
  - (Age over 65, pain not worsened with forward flexion)
- 5 variables with pain relieved with recumbency included
  - 92% of responders, 80% of non-responders
Facet Joint Pain

- Laslett et al. (2004)
- 116 patients
- Revel’s criteria
- Double injection with 75% pain reduction
  - Sen (0.11), Spec (0.91), LR+ (1.2)
  - PPV (0.273), NPV (0.771)
- Potentially helpful criteria
  - No increase in pain rising from flexion
  - No pain with cough & sneeze
  - Age over 65
Laslett et al. (2005) studied 216 Chronic LBP patients and compared clinical diagnosis with a reference standard. They found that:

- 67% received pathoanatomic diagnosis
- 10% had more than one tissue involved
- Diagnosis
  - Exact match 32%, partial match 51%, chance 13%
- Patient categorization
  - Disc, facet, SI, hip, nerve root, spinal stenosis
  - PT vs. radiologist: 56% agreement, chance 33%, kappa = 0.31
Diagnostic Categories

- **Disc**
  - Centralization, peripheralisation, directional preference with repeated end-range movements
  - Lumbar pain in exact midline

- **Facet**
  - Revel’s criteria and no centralization

- **SI**
  - 3 or more (+) pain provocation tests with no centralization

- **Nerve root**
  - Referred pain with nerve tension tests

- **Spinal stenosis**
  - Intermittent claudication relieved by sitting or flexed postures

- **Hip**
  - Movements of hip provoked pain more than lumbar or SI tests

Laslett et al. 2005
Manipulation

- Tullberg et al. (1998)
  - Radiostereophotogrammetric analysis
  - No change in sacrum/ilium alignment after manipulation
  - Positive clinical tests prior became negative after
  - ? soft tissue/neuromuscular reflex

- Herzog et al. (1999)
  - SI joint manipulation produced immediate increase in EMG activity in gluteals, erector spinae, latissimus dorsi, trapezius deltoid and splenius muscles
  - Neurologic effect rather than biomechanical effect
Manipulation

- **Piva et al. 2003**
  - Iliac crest level measurement device
  - ICC (1,1) for standing 0.80 (95% CI = 0.7-0.9) and for seated 0.73 (95% CI = 0.6-0.8) with SEM 0.91 and 0.86 degrees, respectively

- **Childs et al. 2003**
  - SI manipulation
  - Improved IC and WB symmetry after manipulation
  - Decreased pain and improved function 3-4 days after study
  - Results due to neuromuscular system effects
Muscle Energy Technique

- **Lenehan et al. (2003)**
  - Increased trunk rotation ROM in asymptomatic but restricted subjects
    - 10.66° improvement on restricted side (1.09 effect size)

- **Schenk et al. (1997)**
  - Increased lumbar extension ROM in asymptomatic but restricted subjects
    - 7° improvement in restricted subjects

- **Schenk et al. (1994)**
  - Increase cervical rotation ROM but not flexion, extension or sidebending in asymptomatic but restricted subjects
    - 7-8° ROT, 5-8° SB, 2° F, 2° E
Muscle Energy in LBP

- Wilson et al. (2003)
- Pilot study (8 men, 8 women)
  - All had lumbar ERS dysfunction

Groups
- Neuromuscular re-ed and strength training (Control)
- Neuromuscular re-ed and strength plus MET

Results
- 83% reduction in Oswestry in MET group (mean score 7%)
- 65% reduction in Oswestry in control group (mean score 15%)
- 38 point drop in Oswestry, 6 considered clinically significant
- Average of 3 MET treatments in MET group
- Patients can return to pre-injury occupation with Oswestry scores as high as 10-12%
Clinical Prediction Rule

- Flynn et al. (2002)
- Prospective, cohort study of 71 LBP patients
- Reference standard
  - 50% improvement in Oswestry over a 2-4 day period
- Identified 5 variables
  - Duration of symptoms < 16 days
  - At least one hip with >35° of internal rotation
  - Lumbar hypomobility
  - FABQ work score <19
  - No symptoms distal to the knee
Clinical Prediction Rule

- If 4 of 5 variables present, there is a 95% chance of decreasing Oswestry score by 50% over a 2-4 day period by manipulating the lumbopelvic region with a technique reported to effect the SI joints.

- CPR
  - Sensitivity 0.63, Specificity 0.97, PLR 24.38
  - Probability of success with manipulation 95%
  - Pre-test probability of success 45%
Clinical Prediction Rule

- Individual variables
  - Duration of symptoms <16 days
    - Sensitivity 0.56, Specificity 0.87, PLR 4.39
  - FABQ Work score <19
    - Sensitivity 0.84, Specificity 0.49, PLR 1.65
  - Symptoms not distal to knee
    - Sensitivity 0.88, Specificity 0.36, PLR 1.36
  - At least one hip IR > 35°
    - Sensitivity 0.50, Specificity 0.85, PLR 3.25
  - Hypomobility at one or more lumbar levels
    - Sensitivity 0.97, Specificity 0.23, PLR 1.26
Clinical Prediction Rule

- Childs et al. (2004)
- PEDro score = 8/10
- Groups (131 consecutive patients)
  - Manipulation vs. stabilization exercise
  - Positive on CPR, negative on CPR
- Results
  - Manipulation + CPR (+) showed greater improvement in 1wk, 4 wk and 6 mo disability than:
    - Manipulation + CPR (-)
    - Exercise + CPR (+)
  - No difference between exercise groups
  - LR+ 13.2 for success at one week
  - NNT for success at 1 wk = 1.3, 4 wk = 1.9
  - 92% chance of success (+CPR), 7% (-CPR) at 1 wk
Where Do We Go From Here?

- Patient subgroup classification
- Diagnostic tests
- Validity and reliability
- Improved methodology
- Better reference standards
  - Clinical outcomes
- Outcome studies
  - Larger clinical trials
  - Meta-analysis
- Think outside the box
Thank You