CARD 1: EVIDENCE (Diagnosis & Therapy)

DIAGNOSIS

Big 3 **questions for diagnostic tests**

1. **How *good* is the test/sign/symptom?** (See Card 2: Diagnostic Tests.) 
2. **How *strong* is the evidence behind the claim?** 
3. **How *applicable* is it to my/our patient(s)?**

Quality of the Evidence [2]

1. What type of patient (e.g., *severity of case, practice setting*)?
2. Nature of the blinding (*who*)?
3. Nature of the gold/reference standard (*reasonable*)?
4. Was the gold/reference standard applied to *everyone* in the study?

Diagnosis & Therapy

Applicability to my/our patient(s)? [3]

- Similar to patients in study?
- Verified by any other studies (*initial reports of test validity are often inflated*)?
- Practical (e.g., *cost, skill level*)?
- Compatible with patient's needs, wants, etc.?
- Change patient management?

How precise are the test results?

Confidence intervals (range): 2 rules

1. A narrow range is good. Wide ranges mean the reported values are less precise.
2. If one end of the interval is in the negative or insignificant range, the results should be seriously questioned.

CARD 1: EVIDENCE (Diagnosis & Therapy)

THERAPY

Big 3 **questions**

1. **How effective? (How *likely* is it to help? How *much* is it going to help?)** (See Card 3: Therapeutic Effectiveness.)
2. **How strong is the evidence? TYPE/LEVEL of evidence? QUALITY of the evidence (strengths and weaknesses)?**
3. **Does it apply to my/our patient(s)?**

Types/Levels of Evidence (Oxford) [2]

1a: **Systematic review of RCTs** (with homogeneity vs. worrisome heterogeneity)

1b: Individual **RCT** (with narrow vs. wide confidence intervals)

2a: **Systematic review of COHORT studies** (with homogeneity vs. worrisome heterogeneity)

2b: **Individual COHORT study/low quality RCT** (<80% follow-up / wide confidence interval)

3a: **Systematic review of CASE-CONTROL studies** (with homogeneity vs. with worrisome heterogeneity)

3b: **Individual CASE-CONTROL study**

4: **CASE SERIES**/poor quality cohort & case-control studies

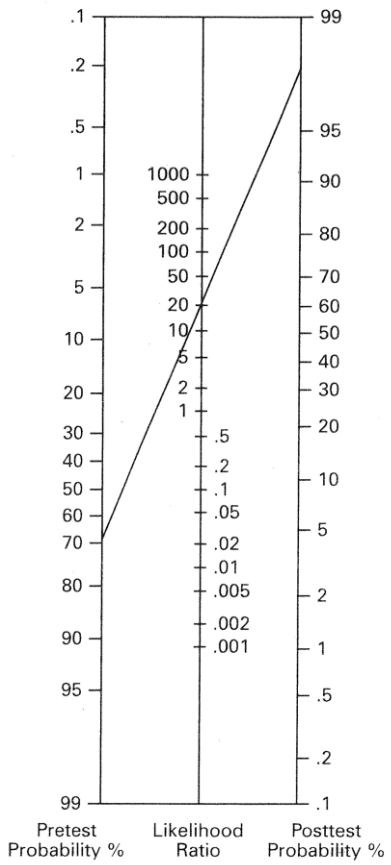
5: **Expert opinion** without explicit critical appraisal, or based on **physiology**, **bench research** or "**first principles**"

Quality of an RCT (***ABCD fix***)

A allocation concealed?
B blinding? (who?)
C comparable groups (beginning to end)?
D drop outs? (5-20%)
F follow up?
I intention to treat analysis?
X x-factors (any other biases)?

CARD 2: DIAGNOSTIC TESTS

Nomogram for Likelihood Ratios (LRs)



Calculating LRs (if not provided)

Positive LR = $\frac{\text{sensitivity}}{1 - \text{specificity}}$ **Negative LR =** $\frac{1 - \text{sensitivity}}{\text{specificity}}$

3/01/10 Ron LeFebvre, DC – University of Western States© 2010

Interpreting LRs: Very useful LRs: > 10 and < .10

Note: + LR reflects the results of a positive test; -LR reflects the results of a negative test.

Estimated % change in probability	+LR Power to RULE IN	SHIFT IN POST-TEST PROBABILITY	−LR Power to RULE OUT	Estimated % change in probability
10 ≈ 45% ↑	10	← LARGE →	< 0.1	0.1 ≈ 45% ↓
5 ≈ 30% ↑	5-10	← MODERATE →	.1-.2	0.2 ≈ 30% ↓
2 ≈ 15% ↑	2-5	← SMALL → (but sometimes important)	.2-.5	0.5 ≈ 15% ↓

Note: These estimates apply best when pre-test probability is 50%; are smaller with higher & lower pre-test probabilities; do not apply when pre-test probability is <10% or >90%.

Specificity (+SPpin)

A test with a specificity of 95% means that if tested on 100 patients without a particular condition, it would only cross react with 5/100 condition-free subjects (very few false positives).

A *positive* test with *high* specificity is most helpful at ruling in a condition (+SPpin), but only ***IF…***

1) The specificity and sensitivity ratings do *not* add up to 100% ***AND***
2) Prior to the test, you estimate at least a 50% chance that the patient *could* have the particular condition ***AND***
3) The subjects that the test is based on are similar to your patient (e.g., in terms of differential diagnoses and severity).

Warning: A specificity of 95% does NOT mean that a positive test reflects a 95% chance that the patient has a particular condition. The actual ***positive predictive value*** is always lower than the isolated specificity number.

Sensitivity (−SNnout)

A test with a sensitivity of 95% means that if tested on 100 patients with a particular condition, it would only *miss* 5/100 people with that condition (very few false negative tests). A *negative* test with *high* sensitivity is most helpful at ruling out a condition (−SNnout) unless coupled with very low specificity.

Test Reliability

Consistent repeatability of a test performed by the same person or a different person (intra vs. inter-examiner). ***Kappa values**** represent how much *better* than chance agreement (e.g., .50 = 50% better than guessing).

.25	Poor agreement	.51-.75	Moderate to good agreement
.26-.50	Fair agreement	.76-1.00	Good to excellent agreement

* ICC values: >.75 = excellent; .40-.75 = fair to good; <.40 =fair.

Ron LeFebvre, DC – University of Western States © 2010 and 2012,

CARD 3: THERAPEUTIC EFFECTIVENESS

Treatment effectiveness is reported in the following ways:

	Good	Poor
NNT (number needed to treat)	A low number	a large number
ES (effect size)	**> 0.8**	<0.2
OR (odds ratio)*	**> 2**	**1**
RR, AR (relative risk, absolute risk)	Judgment	Judgment
Mean difference (e.g., pain score)	MCID	< MCID

When a treatment prevents a bad outcome, good ORs are < 1.0. The lower, the better (e.g., 0.4 represents a 60% decrease in odds of the bad outcome; 0.7 represents a 30% decrease).

NNT

- The number needed to treat for one *additional* patient to achieve a beneficial outcome (compared to no treatment, a placebo, or an alternate therapy).
- NNTs run from 1 to infinity. Effective treatments are usualy single digits, prevention interventions can be double digits.
- It is useful to know how a successful result was defined, what it was compared to, and anything unique about the patient group (e.g., severity or phase of the condition). For example, NNT of 2 *for ROM exercise and joint mobilization improving wrist extension following Colles fracture compared to home exercise.*

Effect Size

Measure of standard mean difference between treatment & placebo.

0.0	no difference
0.2	small treatment effect
0.5	moderate treatment effect
> 0.8	large treatment effect

NOTE: A negative number (e.g., -0.07) = treatment was worse than the comparison.

Relative Risk

Relative risk is the ratio of risk comparing a treatment group to a control. Treatment outcomes are often reported as

- **Relative risk reduction (RRR)—**the percentage change (e.g., patients improved 30% more with therapy A compared to B).
- **Absolute risk reduction (ARR)—**the actual raw difference in numbers (a more useful assessment – can be directly converted to NNT by dividing into 1.0).

ODDS RATIO (OR)

- Odds ratios represent a comparison of odds or change of odds (therapy vs. control).
- Used often in retrospective studies (e.g., case control study).
- *Sometimes* reported in cohort studies (prospective) because ORs are easier to use for other calculations—but are less accurate and may inflate the effect (unless the condition is rare).
- OR of 1.0 = odds of improvement in the treatment group no better than in the control.
- The higher the OR, the better when measuring treatment for a condition.
  - An OR of 1.5 usually calculates to an NNT of 10-44.
  - An OR of 2.5 usually calculates to an NNT of 5-9 (unless the control improvement rate is very high, 90%, or low , 10%).
- When a treatment prevents a bad outcome, good ORs are < 1.0. The smaller the better (e.g., 0.5 is more effective than 0.7).

Minimal Clinically Important Difference (MCID)

An estimate of the minimal amount of improvement above baseline to be considered worthwhile by a typical patient/physician. Smaller differences may still be significant when comparing 2 effective treatments.

Examples

Outcome Measure	Suggested MCID	Comments
Pain scale (11-point)	**2-3**	Patients with chronic pain
Neck Disability Index	**7**	Patients with cervical radiculopathy
Neck Disability Index	**5**	Physical therapy outpatients with musculoskeletal neck pain
Oswestry Disability Index	**4-6**	Patients with low back pain
Roland Morris	**2-5**	Patients with low back pain (1-2 with mild disability, 7-8 with high disability)
PSFS (Patient-Specific Functional Scale)	**2**	For average of 3 activities
PSFS	**3**	For an individual activity
VAS for back pain (100 point)	**18**	Patients with chronic low back pain