Strengthening the Evidence Process: Grading the Evidence

James M. Perrin, MD
Professor of Pediatrics, Harvard Medical School
Director, Center for Child and Adolescent Health Policy
MassGeneral Hospital *for* Children
Director, Evidence Review Group

Alex R. Kemper, MD, MPH, MS
Department of Pediatrics, Duke University

Key Questions affecting AC Decisions

- ***Vary by condition reviewed***
- Test issues
  - Test characteristics
    - Including early vs late, etc.
  - Population testing data
- Value of early identification (screening vs clinical assessment)
- Does treatment help?
- Availability of follow-up diagnosis and treatment

Less Critical Data

- Incidence/prevalence
  - Although important in determining bounds of harms and benefits
- Natural history – alone

Key Questions re CCCHD

- Does adding pulse oximetry improve sensitivity of CCCHD diagnosis (over clinical exam alone)?
- What is the specificity of pulse oximetry?
- What is the effect of early treatment?
- How available is follow up care for test-positive children?

Strength of Evidence for Key CCCHD Questions

Number of studies; subjects	Design	Risk of bias/study quality	Consistency	Directness	Precision	Strength of evidence
Additional sensitivity of pulse oximetry over clinical exam						Moderate
3; 45,754	Prospective Cohort	Good	Inconsistent	Direct	Imprecise	-
Evidence Summary: Pulse oximetry detects most cases of CCCHD. Most studies suggest that pulse oximetry leads to the detection of additional cases over those detected by clinical examination.						
Specificity of pulse oximetry						Moderate
11; 180,773	Prospective Cohort	Good	Inconsistent	Direct	Imprecise	-
Evidence Summary: The specificity of pulse oximetry after 24 hours is high.						
Availability of follow-up care						Poor
0;0	N/A	N/A	N/A	N/A	N/A	-
Evidence Summary: No data identified regarding the availability of follow-up diagnostic care for those with a positive screen.						
Effectiveness of early intervention						Fair
N/A	Case series and reviews	N/A	N/A	N/A	N/A	-
Evidence Summary: Indirect evidence that early intervention is associated with improved outcomes for those with CCCHD.						

Grading the Evidence

- Assessing:
  1. Analytic validity
  2. Quality of data sources
  3. Study quality
  4. Adequacy of the evidence or the strength of linkages in the chain of evidence

- Calonge N, Green NS, Rinaldo P, et al. Committee report: Method for evaluating conditions nominated for population-based screening of newborns and children. *Genet Med*. 2010;12:153-159.

Quality of Data Sources

- Level 1 – usually good quality evidence
- Level 2 – usually fair quality evidence
- Level 3 – usually fair or poor quality evidence
- Level 4 – usually poor quality evidence
- Level 5 – usually poor quality evidence

Assessing Study Quality

1. Clear description of test or disorder/phenotype and outcomes
2. Adequate description of study design and methods
3. Interventions clearly identified, scientifically sound, consistently provided
4. Adequate description of the basis of the “right answer”
5. Avoidance of biases
6. Appropriateness of the data analysis

GRADE

- Grading of Recommendations Assessment, Development and Evaluation Working group: http://www.gradeworkinggroup.org
- Goal: single system to avoid confusion and provide transparency

GRADE

- High – further research is very unlikely to change confidence in the estimate of effect
- Moderate – further research is likely to have an important impact on confidence in the estimate of effect
- Low – further research is very likely to have an important impact on confidence of effect
- Very low – any estimate of effect is very uncertain

Diagnostic Screening and Testing

- Optimal is RCT of screening vs (usually) no screening (or other screening method) – but rarely exists
- PICO
  - **P**atients
  - **I**ntervention (screening)
  - **C**omparison (screening vs no screening
  - **O**utcome (clinical improvement arising from testing)

GRADE

Type of evidence	Randomized trial = high Cross-sectional or cohort studies and comparison with appropriate reference standard = high Any other evidence = very low
Decrease grade if	• Serious or very serious limitation to study quality • Important inconsistency among studies • **Some or major uncertainty about directness** • Imprecise or sparse data • High probability of reporting bias
Increase grade if	• Strong evidence of association—significant relative risk of > 2 (< 0.5) based on consistent evidence from two or more observational studies, with no plausible confounders (+1) • Very strong evidence of association—significant relative risk of > 5 (< 0.2) based on direct evidence with no major threats to validity (+2) • Evidence of a dose response gradient (+1) • All plausible confounders would have reduced the effect (+1)

GRADE

- Challenges for the ERW
  - Almost all studies will be screening vs published comparison – no direct comparison
  - Most evidence will be low or very low
  - Can develop more reliable methods of determining quality

Summary

- Highlighting the questions of most relevance to AC decision-making
- Per earlier presentation, modeling key questions for the AC
- Systematic grading and summarizing the evidence