



Making Recommendations in the Face of Imperfect Evidence

David Atkins, MD, MPH

Center for Outcomes and Evidence

Agency for Healthcare Research and Quality

datkins@ahrq.gov

Outline of Talk

- Overview of systematic, evidence-based methods to evaluate screening
- Illustration of specific issues
- Newborn hearing screening as an example

Observations

- Policy makers often lack ideal evidence at the time they must make a decision
- Debates over scientific evidence confusing to clinicians and public
- Most debates reflect differences in perspective and values rather than disagreement over evidence
- An explicit and systematic approach to evidence can help separate issue of evidence from those of values

From Atkins D et al. *Health Affairs*, 2005

Requirements of an effective screening test

- Condition has important health consequences
- Condition can be detected in pre-symptomatic period
- Acceptable screening test with adequate sensitivity and specificity
- Early intervention more effective than treatment at time of symptoms
- Benefits of early detection outweigh any harms

Misperceptions about Evidence-based Methods

- Overly reliant on RCTs
 - sets unattainable standard for evidence
- Tool to limit health services, save money
- Ignores realities of practice – reimbursement, liability concerns, patient expectations
- Not useful when evidence is poor

Evidence-based Health Policy

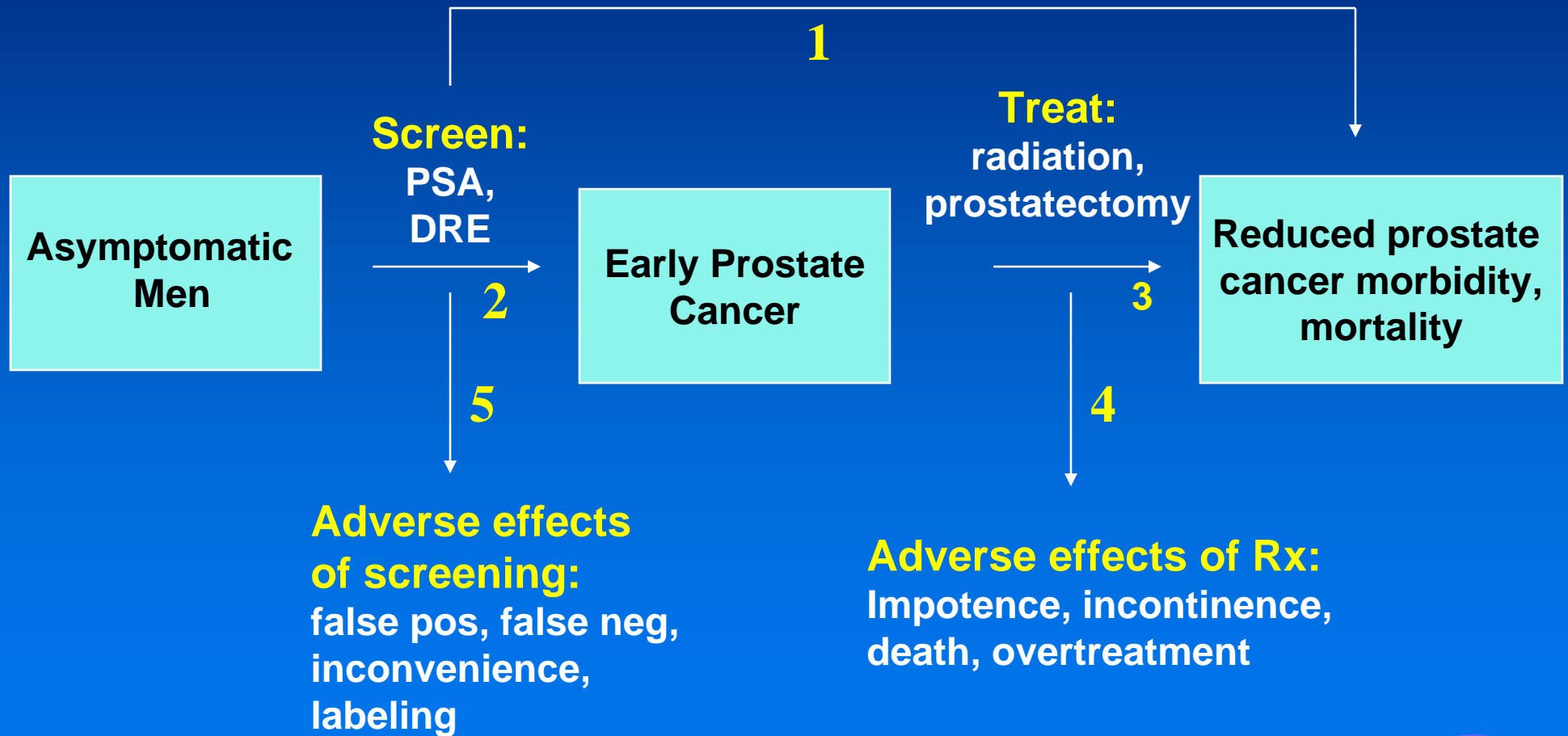


From Muir Gray – *Evidence-based Health Care*

Questions for Setting Policy: A Systematic Process

1. What is the outcome I care most about?
2. How good is the evidence that the interventions can improve those outcomes?
3. How sure am I that it will work in “real world”?
4. How do the potential benefits compare to possible harms and costs?
5. What constitutes “good enough” evidence?
6. What other considerations are relevant?

Analytic Framework - 1



2. How good is the evidence that the intervention will improve the outcome?

Systematic review of the evidence:

- Explicit methods, avoid bias
- Distinguish intermediate from clinical outcomes
- Systematic search for relevant studies
- Consistent evaluation of quality of individual studies
- Transparent reasoning, reproducible results

AIM: Distinguish what we know from what we don't

AIM: Facilitate decision making

Misconception about Systematic Reviews

- Distinguished by number of studies examined
- Requires elaborate methods for assessing individual studies
- Relies on quantitative synthesis
- Most useful when large number of RCTs

Assess quality of evidence

■ What do we mean by quality?

“Extent to which a study’s design, conduct, and analysis has minimized selection, measurement, and confounding biases.”

– Lohr, *J Qual Improvement*, 1999

“Extent to which one can be confident that an estimate of effect is correct”

– GRADE, *BMJ* 2004

Assessing Quality of Individual Studies

- GOAL: Identify those studies least likely to be biased (*internal validity*)
- Quality is function of:
 - study design (e.g., RCT or controlled cohort vs. case series)
 - study execution (e.g., loss to follow-up)
- Critical elements vary by topic

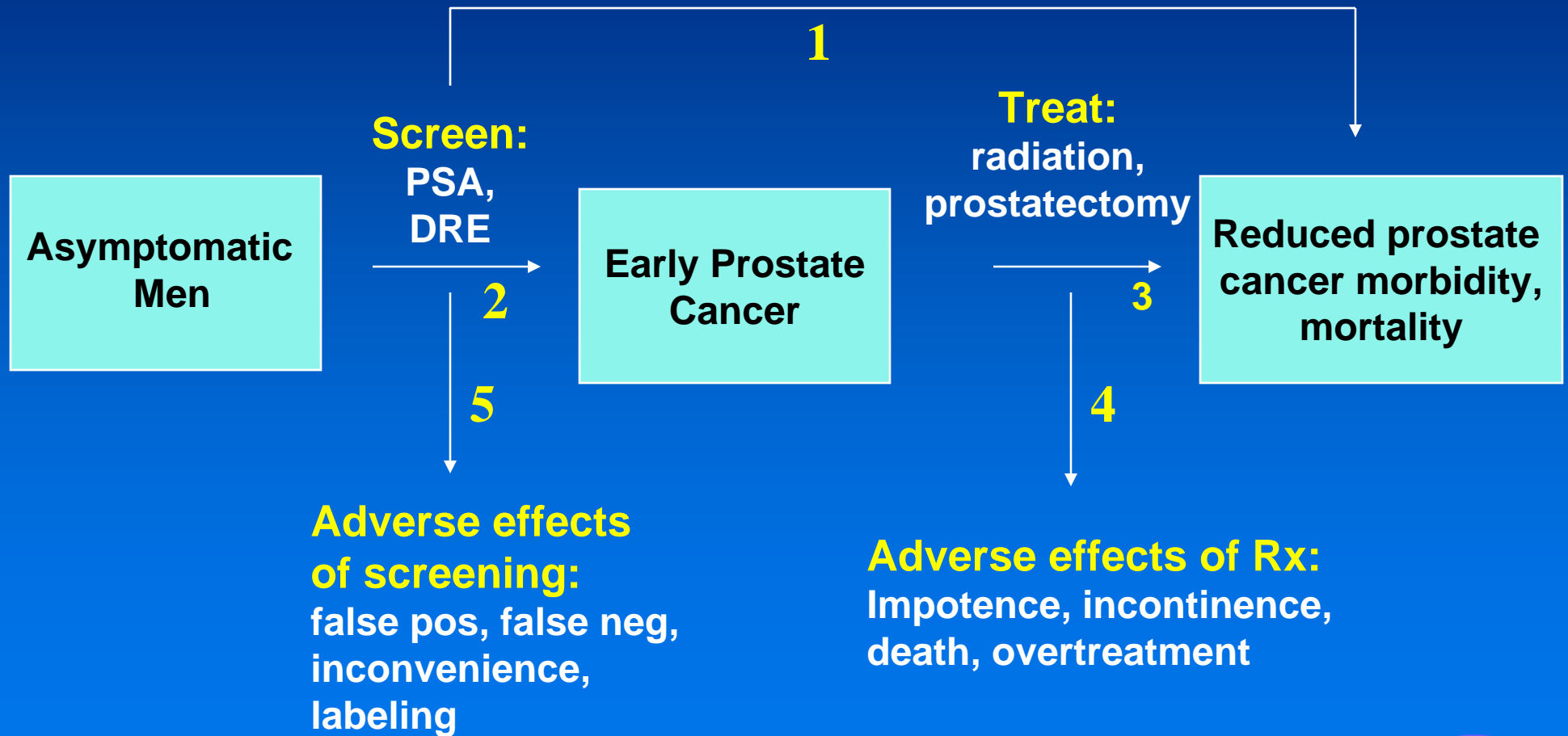
3. Will it work in the real world?

- Carefully controlled research studies may overstate benefits of intervention in practice – “external validity”
- Harms minimized, benefits maximized
- Considerations with newborn screening:
 - Loss to follow-up
 - Accuracy of diagnosis
 - Compliance with interventions

4. Are benefits sufficient to justify possible harms and costs?

- How big are the benefits?
- What are the possible harms?
- How to present tradeoffs:
 - Number needed to screen
 - Number needed to treat
- Opportunity costs, resource implications

Analytic Framework - 1



5. What constitutes “good enough” evidence?

- Depends on perspective
- Depends on what values you place on different outcomes
- Risks of acting “too soon” or “too late”

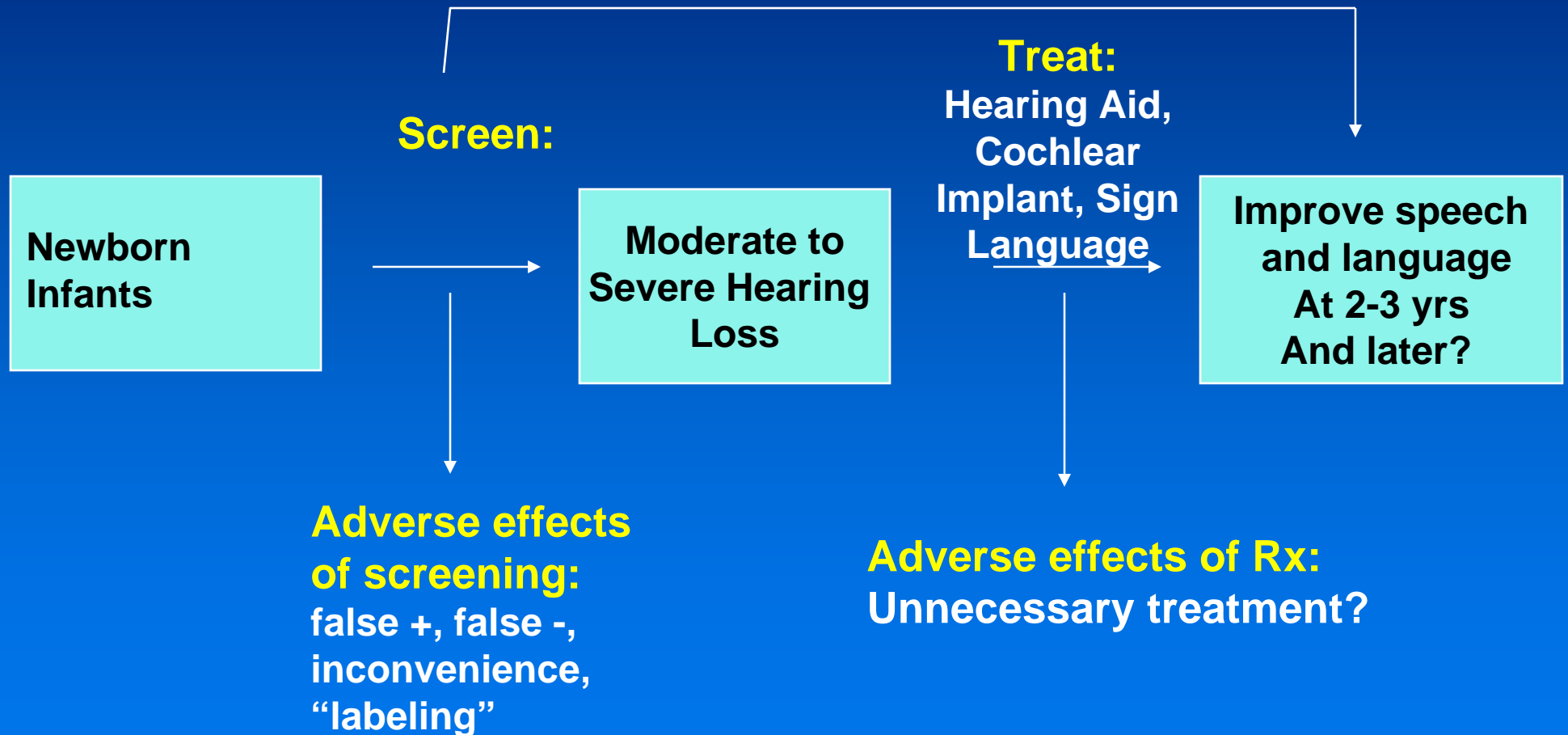
6. What other considerations are relevant?

- Equity
- Costs and resources
- Feasibility

Particular Challenges for Newborn Screening

- Variety of factors make RCTs impossible
 - Rare disorders
 - Involve children
 - Technology and interventions evolving
- Emotionally charged issue
 - Missed cases provide compelling evidence
- Individual decision making difficult
 - Policy decisions affect large populations

Newborn Hearing Screening - Analytic Framework



Universal Newborn Hearing

- Screening detects one case severe hearing loss per 600 infants screened
- Good evidence that universal screening leads to earlier diagnosis and referral for treatment
- Compared to screening only high-risk infants:
 - Screen 2400 infants to get 1 into early treatment
- Initial false-positive rate 2% to 6%
 - Only one in 50 referred children has severe hearing loss
- Effects on speech and language
 - Poor evidence from observational studies
 - Plausible benefit but magnitude of benefit unclear

What is the comparison?

- Benefits of universal screening smaller if compared to current strategy of screening high-risk infants
- Real world benefits diminished by problems in follow-up testing and referral
- States often lack resources for organized tracking and follow-up

What Benefits Are Important?

- Isn't early detection itself a valuable outcome?
 - Valued by parents of affected children
 - Prevents regret over missed diagnosis
 - Allows for other social interventions
 - Value of information

What Other Considerations are Relevant?

- Screening may help improve resources, effectiveness of early interventions
- Equity concerns from uneven policies
- Individualized policies inefficient with newborn screening
- Resource decisions made at state level
 - Limited resources to address varied child health issues
 - Downstream costs of new screening tests
 - Does state have system in place to screen and follow-up effectively?

What evidence is “good enough”?

- Risks of waiting for better evidence
 - Missed opportunities to help affected infants
- Risks of acting too soon
 - Divert resources to ineffective intervention
 - Possible harm to unaffected infants?
- What is probability of having better information in near future?

Where do we need better information on new screening tests?

- How accurate are the tests in real world of state labs?
- How safe and effective are interventions for specific disorders?
- Are all identified infants at equal risk for developing clinical consequences from their disorder?

Dealing with uncertainty

- Who bears burden of proof?
 - When is evidence “good enough”?
 - E.g. High dose chemo/ABMT for breast ca
- Most controversies involve differences in values and perspective
 - Affected family
 - General pediatrician
 - Public health practitioner
 - State policy maker

Potential solutions to uncertainty

- Shared decision making
 - E.g. prostate cancer screening
 - Difficult in newborn screening
- Conditional coverage
 - lung volume reduction surgery for emphysema
- Individualized state policies
 - Challenge to notions of equity
- Staged implementation
 - Conditioned on specific parameters

Conclusions

- Explicit approaches useful even when evidence is imperfect
- Clarify what we know at present, what we need to know, and what we'd like to know
- Useful to separate issues of evidence from issues of values and resources
- Disputes often reflect legitimate differences in the perspectives of the different parties
- Consider risks of acting “too soon” and acting “too late”