

# GRADE

## Evidence-Based Medicine Methodology

### Health Evidence Review Commission



**Wiley Chan, MD**

**Physician, Internal Medicine**

**Methodologist, KP National Guideline Program**

*care management* | institute

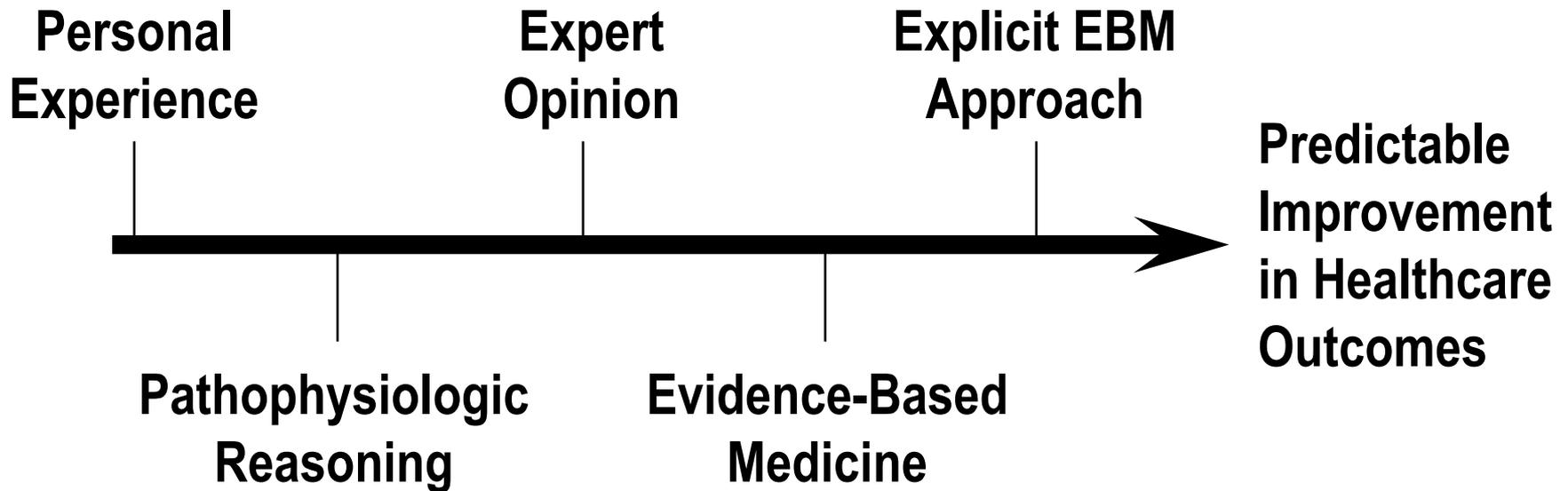


# Statement of Disclosure

## Wiley Chan, MD

- I have no commercial or academic conflicts of interest
  - **Employment: Northwest Permanente Medical Group**
    - Physician, Internal Medicine
    - Director: Guidelines & Evidence-Based medicine
    - Methodologist: KP National Guideline Program
  - **Member: US National Heart, Lung and Blood Institute (NHLBI) Expert Panel on Integrated Cardiovascular Risk Reduction**
    - Co-Chair: NHLBI Implementation Science Work Group
  - **Member: State of Oregon Health Evidence Review Commission (HERC)**
    - Chair: HERC Evidence-Based Guidelines Subcommittee
  - **Member: GRADE Working Group**

# Continuum of Clinical Decision Making



- As one moves to the right there is:
  - Decreasing bias in information
  - Increasing ability to predict health outcomes with confidence
- There will still be a need for pathophysiologic reasoning and expert opinion

# Evidence-Based Medicine In HERC

## ➤ GRADE Methodology

- Grading of Recommendations, Assessment, Development and Evaluation
- International collaboration
- Rigorous, systematic, transparent approach to grading quality of evidence and strength of recommendations

## ➤ Clinical efficacy is paramount

- Patient Oriented Evidence That Matters (POEMs)
- Cost-effectiveness can be explicitly addressed

# What Outcomes & Evidence Matter?

## ➤ POEMs: Patient Oriented Evidence that Matters

- Addresses a question that patients & doctors encounter
- Measures outcomes that patients care about
  - Symptoms
  - Morbidity
  - Quality of Life
  - Mortality
- Has the potential to change patient & doctor behavior

Slawson DC, Shaughnessy AF, Bennett JH. *J Fam Pract.* 1994;38(5):505-513.

## ➤ GRADE Categories of Outcomes

- **Critical:** 1<sup>0</sup> factors influencing a recommendation
- **Important, but not Critical:** 2<sup>0</sup> factors influencing a recommendation
- **Limited Importance:** May or may not influence a recommendation

# GRADE: Strength of Recommendation

## Implications

	Strong	Weak
For Patients	<b>Most individuals in this situation would want the recommended course of action</b> and only a small proportion would not.	The majority of individuals in this situation would want the suggested course of action, but <b>many would not.</b>
For Clinicians	<b>Most individuals should receive the recommended course of action.</b> Formal decision aids are not likely to be needed.	Recognize that <b>different choices will be appropriate for different patients. Decision aids may well be useful.</b>
For Policy Makers	<b>The recommendation can be adapted as policy in most situations,</b> including for use as <b>performance indicators.</b>	Policy making will require substantial debates and involvement of many stakeholders.

# GRADE: Strength of Recommendation

## Factors that Inform the Decision

Factor	Comment
Balance between desirable and undesirable effects	The larger the difference, the higher the likelihood that a strong recommendation is warranted.
Quality of evidence	The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted
Values and preferences	The more they vary, or the greater the uncertainty, the higher the likelihood that a weak recommendation is warranted
Resource use (Costs)	The higher the costs of an intervention, the lower the likelihood that a strong recommendation is warranted

# GRADE: Balance

## Between Desirable and Undesirable Effects

- Desirable effects include beneficial health outcomes and less burden
- Undesirable effects include harms and more burden
- Burdens are the demands of adhering to a recommendation that patients or caregivers (e.g. family) may dislike, such as having to take medication or the inconvenience of going to the doctor's office

# GRADE: Balance

## Factors that Inform the Decision

- **Importance of Outcomes (to Patients)**
  - **Stronger recommendations for interventions that increase the probability of beneficial outcomes or decrease the risk of adverse outcomes with high patient importance**
- **Baseline Risk**
  - **The higher the baseline risk, the greater the magnitude of benefit and the more likely the recommendation will be strong**
- **Magnitude of Effect**
  - **If large relative effects of an intervention consistently point in the same direction (towards benefits or towards harms and burdens) they are likely to lead to a strong recommendation**
- **Precision of Estimate**
  - **The more precise are the estimates of the effect of an intervention, the more likely the recommendation will be strong**

# GRADE: Quality of Evidence

- Reflects confidence in estimates of efficacy
- A systematic review of the evidence is done for each critical/important outcome
- The overall quality of evidence is the combined grade of the quality of evidence across all critical outcomes
  - **If the quality differs across critical outcomes, and:**
    - Outcomes point in different directions
      - The lowest quality of evidence for any of the critical outcomes determines the overall quality of evidence
    - All outcomes point in the same direction
      - The highest quality of evidence for a critical outcome that by itself would suffice to recommend an intervention determines the overall quality of evidence

# GRADE: Quality of Evidence

## Putting it All Together

Step 1 Starting grade based on study design	Step 2 Reduce grade	Step 3 Raise grade High-quality observational studies	Step 4 Final grade
RCT - High	<p><b>Study quality (risk of bias)</b> Serious (-1) or very serious (-2) limitations</p>	<p><b>Large magnitude of effect</b> <u>Large effect (+1)</u> RR &gt; 2 or &lt; 0.5, based on consistent evidence from two or more observational studies with no plausible confounders</p>	<p><b>High</b> Further research unlikely to change confidence in the estimate of effect</p>
<p><b>Observational – Low</b> Quasi-RCT Cohort Case-control</p>	<p><b>Inconsistency</b> Important inconsistency (-1)</p> <p><b>Indirectness</b> Some (-1) or major (-2) uncertainty about directness</p>	<p><u>Very large effect (+2)</u> RR &gt; 5 or &lt; 0.2, based on direct evidence with no major threats to validity</p>	<p><b>Moderate</b> Further research likely to have an important impact on confidence in the estimate of effect and may change the estimate</p>
<p><b>All others – Very Low</b> Case reports Case series</p>	<p><b>Imprecision</b> Imprecise or sparse data (-1)</p> <p><b>Publication bias</b> High suspicion (-1)</p>	<p><b>Dose response gradient (+1)</b></p> <p><b>All plausible confounders would have reduced the effect (+1)</b></p>	<p><b>Low</b> Further research very likely to have an important impact on confidence in the estimate and may change the estimate</p> <p><b>Very Low</b> Any estimate of effect is very uncertain</p>

# GRADE: Values and preferences

- **Uncertainty concerning values and preferences or their variability among patients may lower the strength of a recommendation.**
  - **While it is ideal for clinicians to elicit patient preferences and values directly from patients or to obtain values and preference estimates from population based studies, such studies are often unavailable.**
  - **There is some systematic research of values and preferences, and guideline panel members' experience with patients provides additional insight.**
- **When value or preference judgments are particularly important for the interpretation of recommendations, authors should describe the key values they have attributed in making a recommendation**

# GRADE: Resource Use

## General Guidance

- May legitimately choose to leave considerations of resource use aside
  - **But be explicit about the decision**
- First decide on the quality of evidence regarding other outcomes, before addressing cost
  - **Resource use usually becomes important when advantages and disadvantages are closely balanced**
- Offer only a single recommendation
  - **Refrain from issue two recommendations, one not taking resource use into account, and a second doing so**

# GRADE: Strength of Recommendation

## Factors that Inform the Decision

Factor	Comment
Balance between desirable and undesirable effects	The larger the difference, the higher the likelihood that a strong recommendation is warranted.
Quality of evidence	The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted
Values and preferences	The more they vary, or the greater the uncertainty, the higher the likelihood that a weak recommendation is warranted
Resource use (Costs)	The higher the costs of an intervention, the lower the likelihood that a strong recommendation is warranted

# GRADE: Strength of Recommendation

## Implications

	Strong	Weak
For Patients	<p><b>Most individuals in this situation would want the recommended course of action</b> and only a small proportion would not.</p>	<p>The majority of individuals in this situation would want the suggested course of action, but <b>many would not.</b></p>
For Clinicians	<p><b>Most individuals should receive the recommended course of action.</b> Formal decision aids are not likely to be needed.</p>	<p>Recognize that <b>different choices will be appropriate for different patients. Decision aids may well be useful.</b></p>
For Policy Makers	<p><b>The recommendation can be adapted as policy in most situations,</b> including for use as <b>performance indicators.</b></p>	<p>Policy making will require substantial debates and involvement of many stakeholders.</p>

# Emerging Construct: Rationale Table

## (For Information Only)

- Concise description of key elements behind a recommendation
- Basis of Recommendation Statement
  - **Describes how the 4 GRADE domains of strength of recommendation were utilized to derive the final recommendation and its strength**
    - Example: “Despite the lack of direct evidence that demonstrates that rescreening is beneficial, we believe that clinicians and patients are likely to place a high value on the potential, but unproven, benefit of rescreening. Therefore, we conclude that rescreening for low Bone Mineral Density with DXA is an option and we have suggested rescreening intervals based on the patient’s initial T-score.”
- Four GRADE Domains of Strength of Recommendation
  - **Brief, high-level overview**
  - **Hyperlinks to underlying Systematic Review(s)**

Slide 16

# Main Modifications of GRADE: HERC Processes

- External SRs & CPGs commonly used
  - Outcomes examined are subject to external sources
  - Evidence-grading is subject to external sources
  - Summary of Findings tables for Critical/Important outcomes are subject to external sources
- Values & Preferences less directly addressed
- Exploration of Resource Allocation subject to external sources
- More explicit consideration of policy context
- Adapted for the development of Coverage Guidance
  - GRADE Framework is designed for clinical recommendations

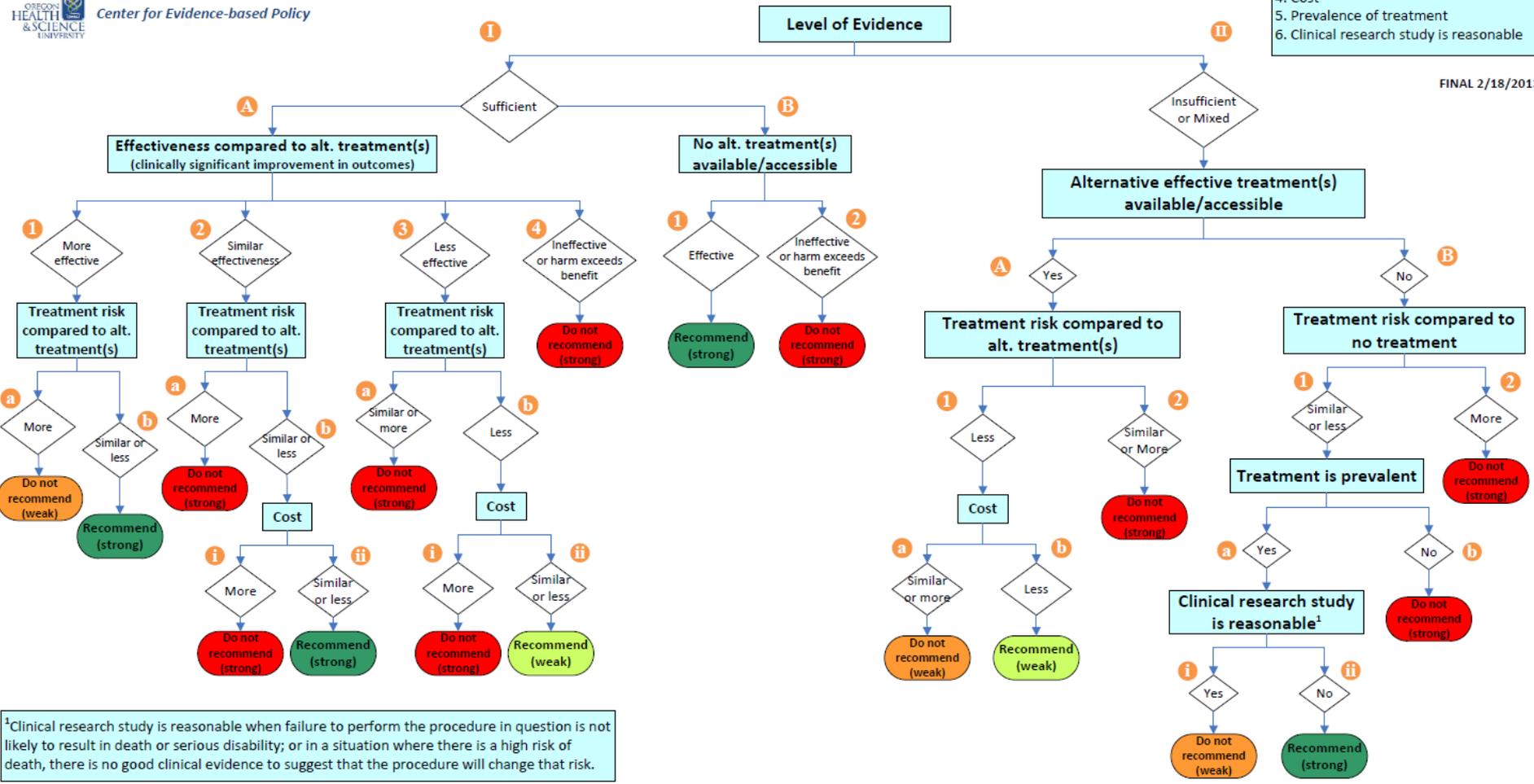
# HERC Guidance Development Framework

## HERC Guidance Development Framework

Refer to *HERC Guidance Development Framework Principles* for additional considerations

- Decision Point Priorities**
1. Level of evidence
  2. Effectiveness & alternative treatments
  3. Harms and risk
  4. Cost
  5. Prevalence of treatment
  6. Clinical research study is reasonable

FINAL 2/18/2013



<sup>1</sup>Clinical research study is reasonable when failure to perform the procedure in question is not likely to result in death or serious disability; or in a situation where there is a high risk of death, there is no good clinical evidence to suggest that the procedure will change that risk.

# HERC Guidance Development Framework

## Principles

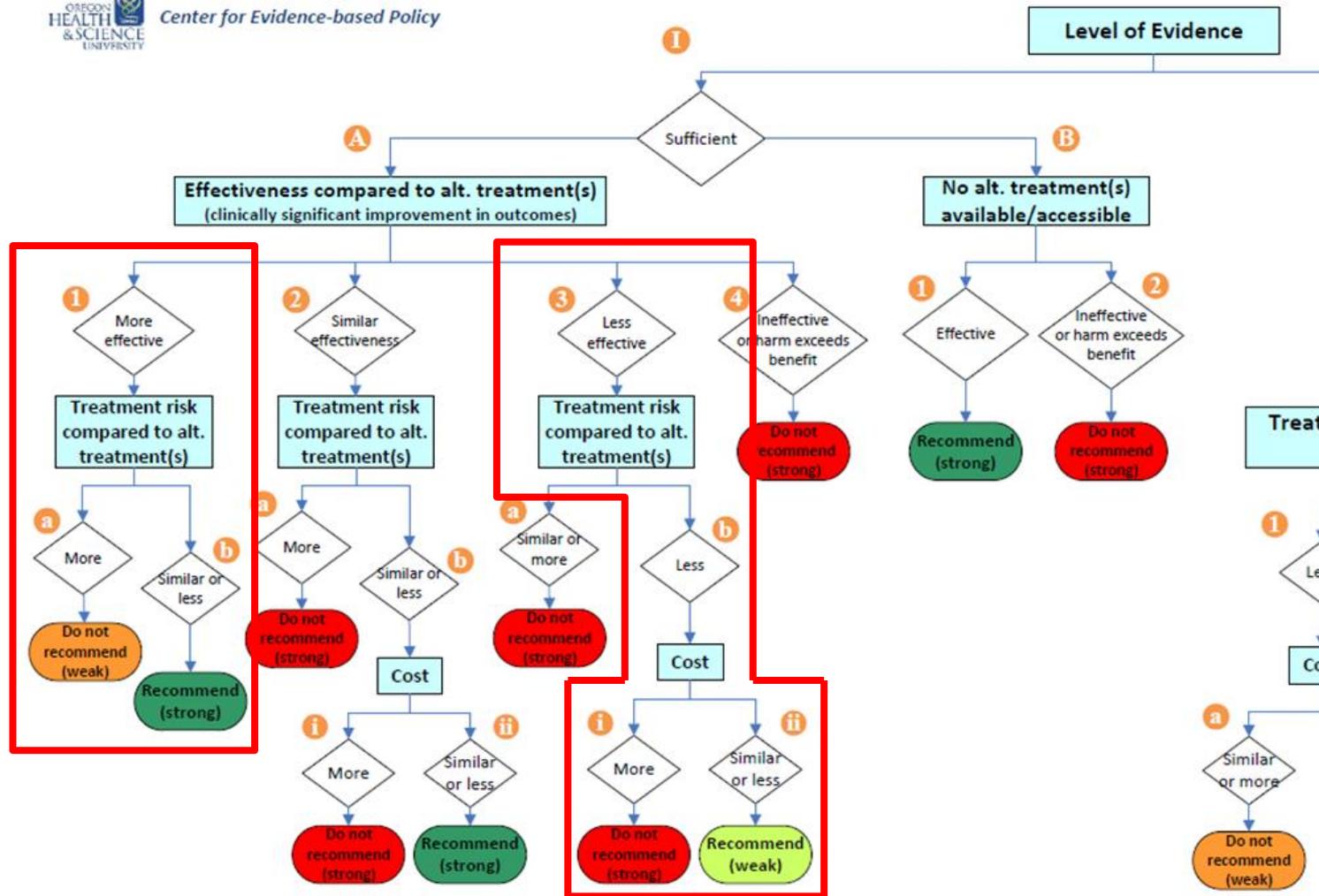
- This framework is a general guide, and must be used in the context of clinical judgment. It is not possible to include all possible scenarios and factors that may influence a policy decision in a graphic format.
- While this framework provides a general structure, factors that may influence decisions that are not captured on the framework include but are not limited to the following:
  - Estimate of the level of risk associated with the treatment, or any alternatives;
  - Which alternatives the treatment should most appropriately be compared to;
  - Whether there is a discrete and clear diagnosis;
  - The definition of clinical significance for a particular treatment, and the expected margin of benefit compared to alternatives;
  - The relative balance of benefit compared to harm;
  - The degree of benefit compared to cost; e.g., if the benefit is small and the cost is large, the committee may make a decision different than the algorithm suggests;
  - Specific indications and contraindications that may determine appropriateness;
  - Expected values and preferences of patients.

# HERC Guidance Development Framework

## Potential Variations From Algorithm Due to Nuances



Center for Evidence-based Policy



# Emerging GRADE Framework: Coverage Guidance (For Information Only)

## ➤ Additional elements:

- **Is the cost small, relative to the net benefits?**
- **Is the total cost (impact on budget) low?**
- **What would be the impact on health inequities?**
- **Is inappropriate use likely to be an important problem?**
- **GRADE Recommendations for coverage:**
  - **Do not cover**
    - Undesirable consequences clearly outweigh desirable consequences
  - **Coverage, with evidence development**
    - Undesirable consequences probably outweigh desirable consequences
  - **Restricted coverage**
    - Desirable/undesirable consequences closely balanced or uncertain
  - **Cover, with price reduction**
    - Desirable consequences probably outweigh undesirable consequences
  - **Cover**
    - Desirable consequences clearly outweigh undesirable consequences

# GRADE EBM Methodology



**Wiley.Chan@kp.org**