Use of systematic reviews of effects in policy brief development

Femi Oladapo MD, MPH, FWACS, FRCOG(H)
Maternal and Perinatal Health
Department of Sexual and Reproductive Health and Research
WHO/HQ

Capacity Building for Development of Policy Briefs: Advanced Training of Trainers Workshop for Researchers and Policy-Developers

29-30 March 2022
Cairo, Egypt

World Health Organization
REGIONAL OFFICE FOR THE Eastern Mediterranean

Outline

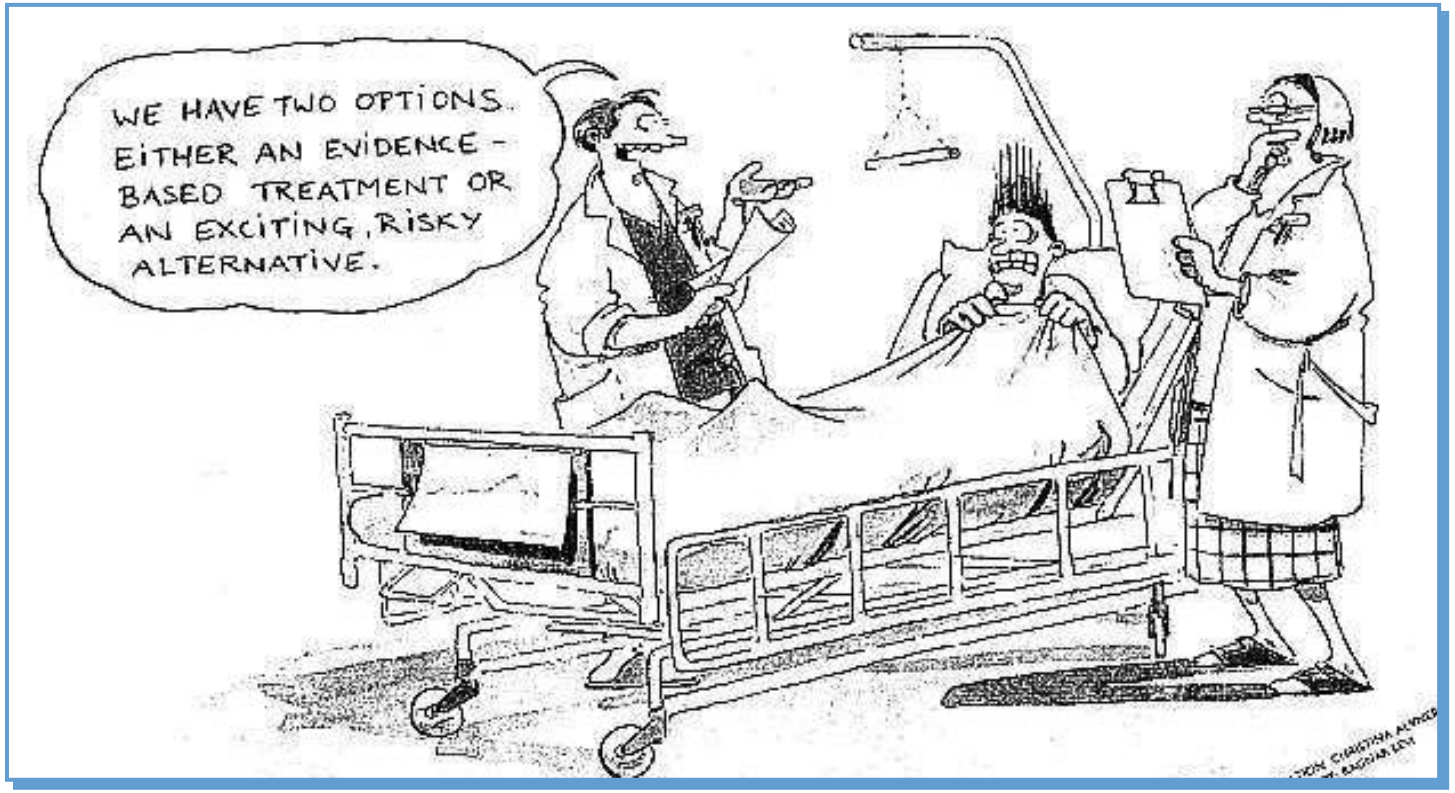
- Appraising findings of systematic review of effect
- How to use systematic reviews findings in policy brief development
- Consolidating SR findings with local evidence
- Leveraging effectiveness evidence for identification of policy options (from evidence to recommendation)

Evidence-based health care policy

- Lack of reliable and valid research evidence or lack of access to such evidence may lead to the use of policies with unknown effectiveness or even those that have been demonstrated to be harmful

What option would you prefer?

Faith versus Facts

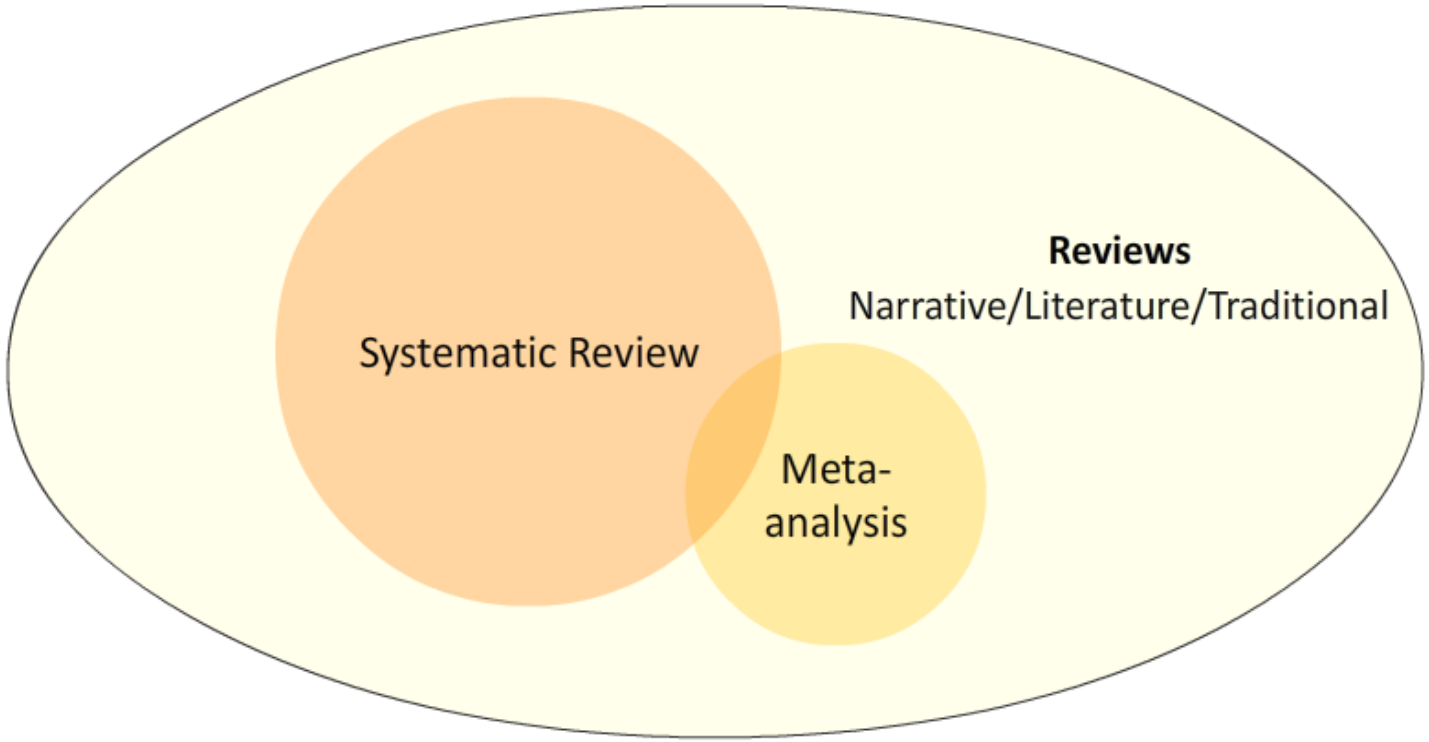


"How do we know if one policy option will work better than another, or if it will do more harm than good?"

Systematic reviews

- A review of clearly formulated question that uses explicit and replicable methods to identify, select and critically appraise relevant research, and to collect and analyse data from the studies that address a particular question
- Statistical methods (meta-analysis) may or may not be used to analyse and summarise the results of the included studies
- The goal is to limit bias in the identification, evaluation and synthesis of the body of relevant studies that address a specific research question

Systematic review, meta-analysis, and narrative review



Systematic reviews of effects of interventions

- Published systematic reviews of studies of health care interventions has increased rapidly and extensively used for clinical and policy decisions
- Subject to a wide range of biases, particularly with inclusion of non-randomised studies
- Critical for policy developers to be able to differentiate high from poor quality reviews
- Several instruments have been designed to evaluate different aspects of reviews

"How much confidence do we have in the estimates of effects published in a systematic review?"

Appraisal of findings of systematic reviews of effect

- Making a choice between policy options requires evaluation of comparative advantages and disadvantages of different options
- Level of confidence in the findings of a systematic review depends of methodological rigor of the review
- For the particular subject in question, is this the best available systematic review and are the findings trustworthy?
- Appraisal of reviews of effects can be at two levels
  - Quality rating at the level of the systematic review
  - Quality rating at the level of the important outcomes affected by the policy option

Quality rating at the level of the systematic review

- Several instruments have been designed to appraise individual studies that are being included in reviews or how certain steps should be conducted (e.g. Cochrane risk of bias tool)
- AMSTAR (A MeaSurement Tool to Assess systematic Reviews), first published in 2007, and revised in 2017 (AMSTAR 2), is one of the most widely used instrument
  - Practical critical appraisal tool for use by policy makers for rapid and reproducible assessments of the quality of the conduct of reviews

AMSTAR 2

16-item instrument assessing all aspects of the review using “Partially yes”, “Yes” and “No” responses: https://amstar.ca/docs/AMSTAR-2.pdf

1. Did the research questions and inclusion criteria for the review **include the components of PICO?**

2. Did the report of the review contain an explicit statement that **the review methods were established prior to the conduct of the review** and did the report justify any significant deviations from the protocol?

3. Did the review authors explain their **selection of the study designs for inclusion** in the review?

4. Did the review authors use a **comprehensive literature search strategy**?

AMSTAR 2

5. Did the review authors perform **study selection in duplicate**?

6. Did the review authors perform **data extraction in duplicate**?

7. Did the review authors provide a **list of excluded studies** and justify the exclusions?

8. Did the review authors **describe the included studies in adequate detail**?

9. Did the review authors use a **satisfactory technique for assessing the risk of bias** (RoB) in individual studies that were included in the review?

10. Did the review authors report on the **sources of funding for the studies** included in the review?

11. If meta-analysis was performed did the review authors use **appropriate methods for statistical combination of results**?

AMSTAR 2

12. If meta-analysis was performed, did the review authors assess the **potential impact of RoB in individual studies on the results of the meta-analysis** or other evidence synthesis?

13. Did the review authors **account for RoB in individual studies** when interpreting/ discussing the results of the review?

14. Did the review authors provide a satisfactory explanation for, and discussion of, **any heterogeneity observed in the results** of the review?

15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of **publication bias** (small study bias) and discuss its likely impact on the results of the review?

16. Did the review authors report any **potential sources of conflict of interest**, including any funding they received for conducting the review?

AMSTAR 2 critical domains

- Protocol registered before commencement of the review (item 2)
- Adequacy of the literature search (item 4)
- Justification for excluding individual studies (item 7)
- Risk of bias from individual studies being included in the review (item 9)
- Appropriateness of meta-analytical methods (item 11)
- Consideration of risk of bias when interpreting the results of the review (item 13)
- Assessment of presence and likely impact of publication bias (item 15)

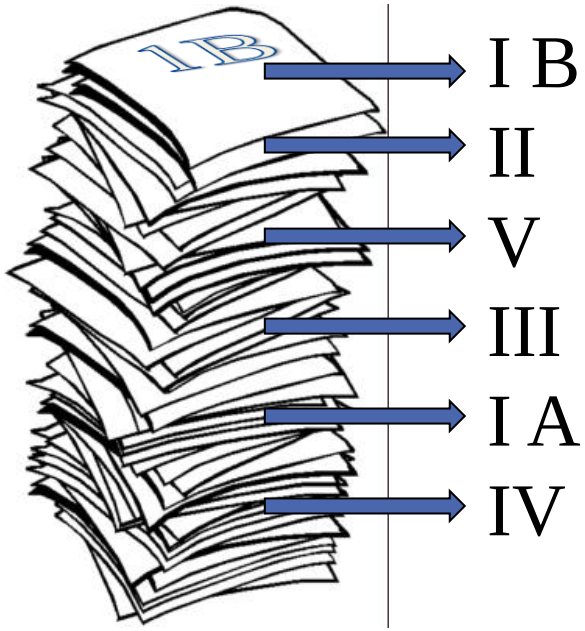
AMSTAR 2 – rating overall confidence in the findings of the review

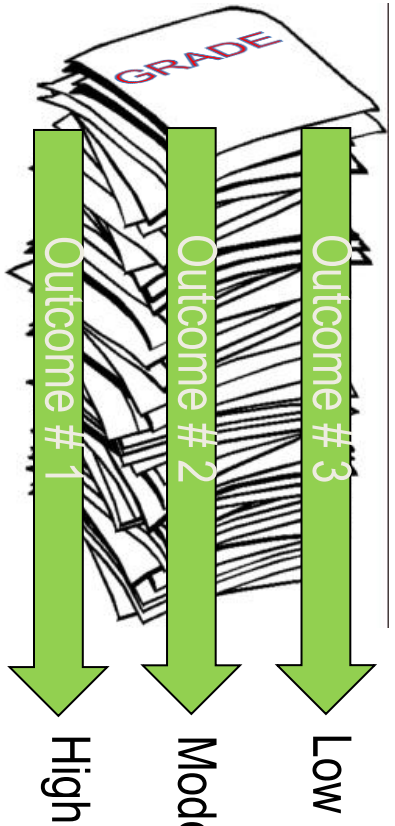
High	No or one non-critical weakness: the systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest
Moderate	More than one non-critical weakness: the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review
Low	One critical flaw with or without non-critical weaknesses: the review has a critical flaw and may not provide an accurate and comprehensive summary of the available studies that address the question of interest
Critically low	More than one critical flaw with or without non-critical weaknesses: the review has more than one critical flaw and should not be relied on to provide an accurate and comprehensive summary of the available studies

Quality rating at the level of systematic review outcomes

Grades of **R**ecommendation **A**ssessment, **D**evelopment and **E**valuation

Quality of evidence across studies





GRADE: determinants of quality of evidence

- RCTs start high → observational studies start low
- 5 factors lower the quality of evidence
  - Limitations of the study design
  - Inconsistency
  - Indirectness
  - Imprecision
  - Reporting bias
- 3 factors raise the quality of the evidence
  - large magnitude can upgrade one level
  - dose response relation
  - Residual confounding unlikely to be responsible for observed effect

Quality of the evidence, four categories

High	Further research is very unlikely to change our confidence in the estimate of effect	⨁⨁⨁⨁
Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.	⨁⨁⨁◯
Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate	⨁⨁◯◯
Very low	Any estimate of effect is very uncertain	⨁◯◯◯

Scenario – 1

You work in the Ministry of Health and are preparing a document regarding options to address the high rates of maternal deaths from postpartum bleeding in a your country situated in a tropical climate region. Three systematic reviews of the effects of options that can be scaled to prevent postpartum bleeding have been identified and you have been asked to make an assessment of how much confidence can be placed in each review.

Review #1 shows that compared to usual care, "Option A" reduces the likelihood of postpartum bleeding by 20% and review authors assigned "moderate" overall AMSTAR 2 rating

Review #2 shows that compared to usual care, "Option B" probably reduces the likelihood of postpartum bleeding by 35% and review authors assigned "low" overall AMSTAR 2 rating

Review #3 shows that compared to usual care, "Option C" probably reduces the likelihood of postpartum bleeding by 15% and review authors assigned "high" overall AMSTAR 2 rating

Based on the level of confidence that can be placed on these findings, which options would you recommend for the Ministry to scale up, and in what order, in a situation where other contextual factors are not considered?

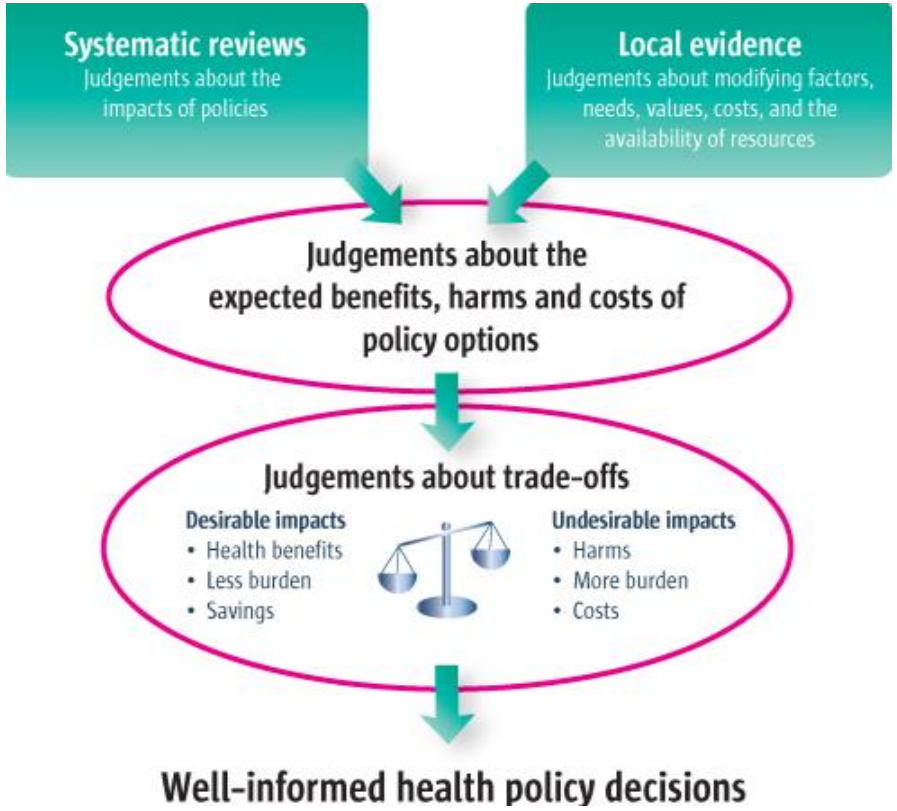
Consolidating systematic review findings with local evidence

- Systematic reviews identified often do not reflect the local setting in the country for which the policy brief is being developed
- Critical to understand to what extent the findings of a review is applicable to the local context
  - Were the studies included in a systematic review conducted in the same setting or were the findings consistent across settings or time periods?
  - Are there important differences in on-the-ground realities and constraints that might substantially alter the feasibility and acceptability of a policy or programme option?
  - Are there important differences in health system arrangements that may mean an option could not work in the same way?
  - Are there important differences in the baseline conditions that might yield different absolute effects even if the relative effectiveness were the same?
  - What insights can be drawn about scaling up, implementation, and monitoring and evaluation (M&E)?

Source: Lavis, Oxman, et al., 2009.

Consolidating systematic review findings with local evidence

- Complex issues often require consideration of a variety of multiple types of evidence together to support or refute the identified policy options
- Local evidence, data or research, if available, should always be systematically identified and considered alongside systematic review evidence
  - Degree of need
  - Values
  - Feasibility
  - Acceptability
  - Resource use, cost, and cost-effectiveness



Source: Oxman et al. 2009

Scenario – 2

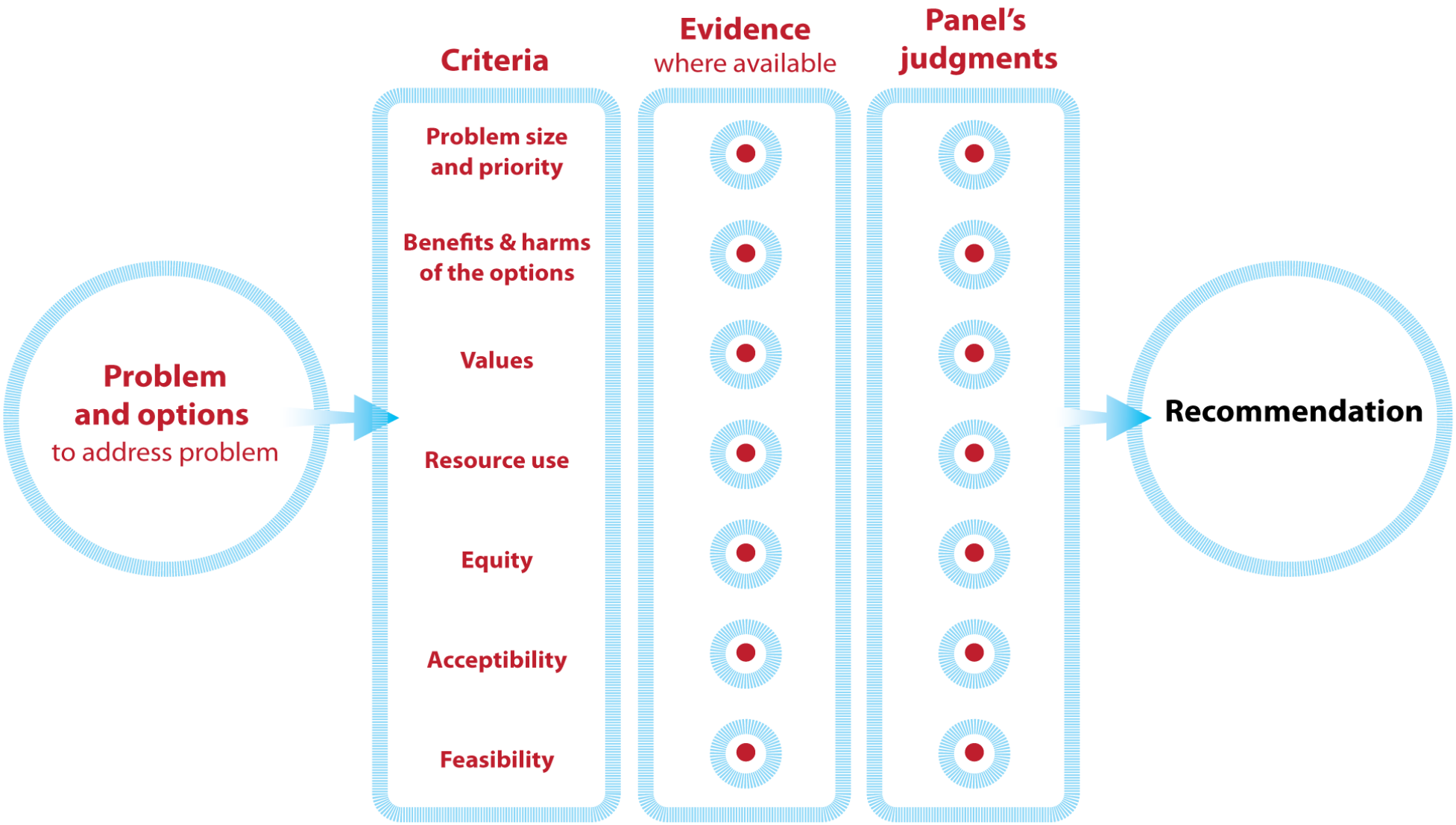
The three systematic reviews for Options A, B, and C included intervention studies that were largely conducted in Europe. ***Option A*** *can only be delivered by injections provided skilled health care providers, and remains potent when exposed to room temperature for a prolonged period of time;* ***Option B*** *can be administered as oral tablets by all cadres of health care providers, including community health workers, and remains potent when exposed to room temperature for a prolonged period of time;* ***Option C*** *can only be delivered by injections provided skilled health care providers, but gets degraded when exposed to room temperature for a prolonged period of time and therefore needs to be refrigerated to remain potent.*

To decide on which policy option(s) to recommend for your country, what types of additional questions would want to address using locally available data, and why are these important?

Consolidating systematic review findings with local evidence

- No single, agreed framework for synthesizing diverse forms of evidence
- Combined evidence types include:
  - Interpretive (narrative and qualitative)
  - Integrative (quantitative and Bayesian meta-/decision analysis)
- Most common approach is narrative synthesis of different systematic reviews, qualitative research, local data and research

Leveraging effectiveness evidence for identification of policy options (evidence to recommendation)



Decision-making on policy options

The extent to which we are confident that the desirable effects of a policy options outweigh the undesirable effects across the range of people for whom the policy options are intended

Desirable effects
- health benefits
- less burden
- savings

Undesirable effects
- harms
- more burden
- costs

Decision-making on policy options

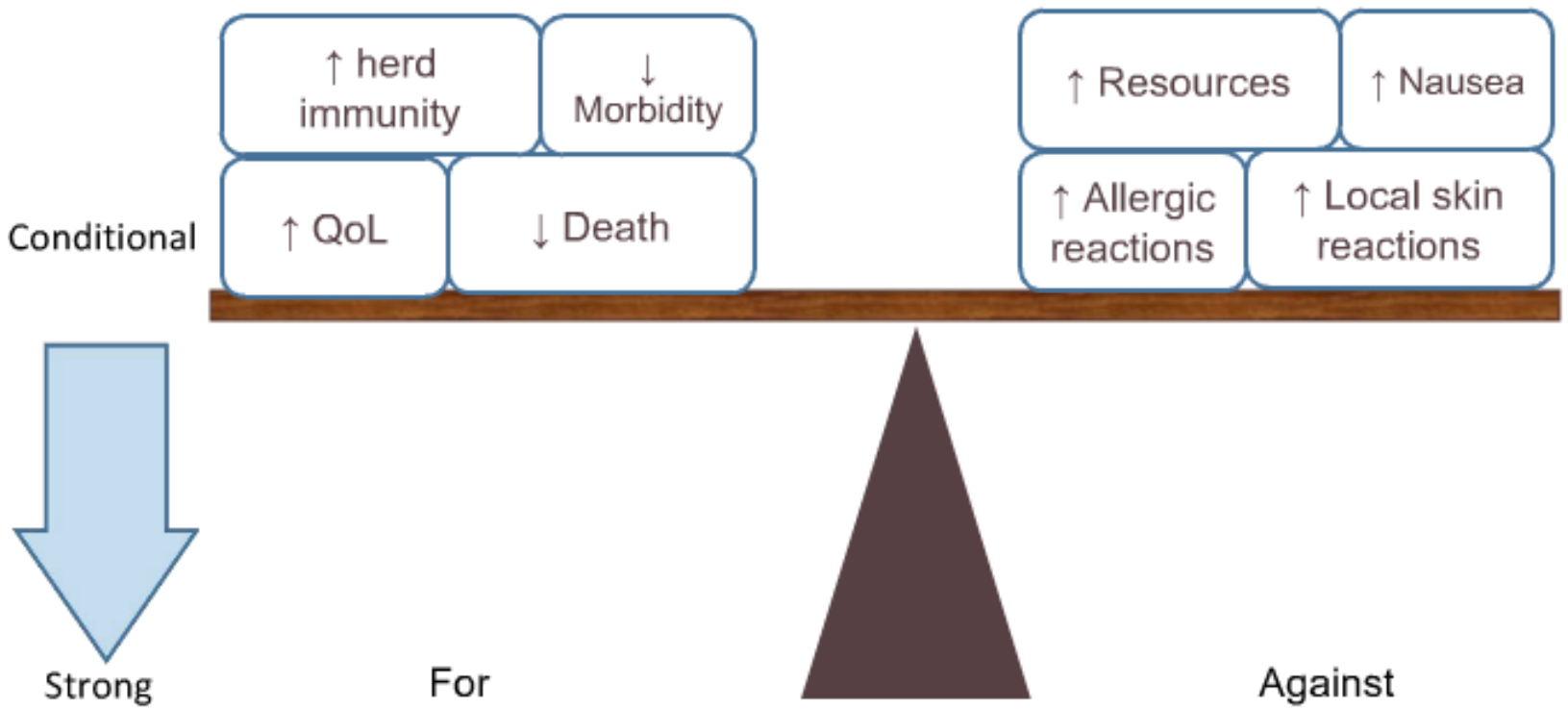
- The choice of policy recommendation will be determined by assessing each policy option on the basis of:
  - (i) **quality** of available evidence
  - (ii) **balance** of desirable versus undesirable consequences ("trade-offs")
  - (iii) **values** and **preferences** related to policy option in local settings/populations;
  - (iv) **resource use (costs)** for policy options in local setting
- By default, a moderate or high quality evidence often lead to a decision to recommend a policy option (for or against) while a low or very low quality evidence often lead to provisional or contextual recommendation of the policy option …**but this is not the rule!**

Decision-making on policy options

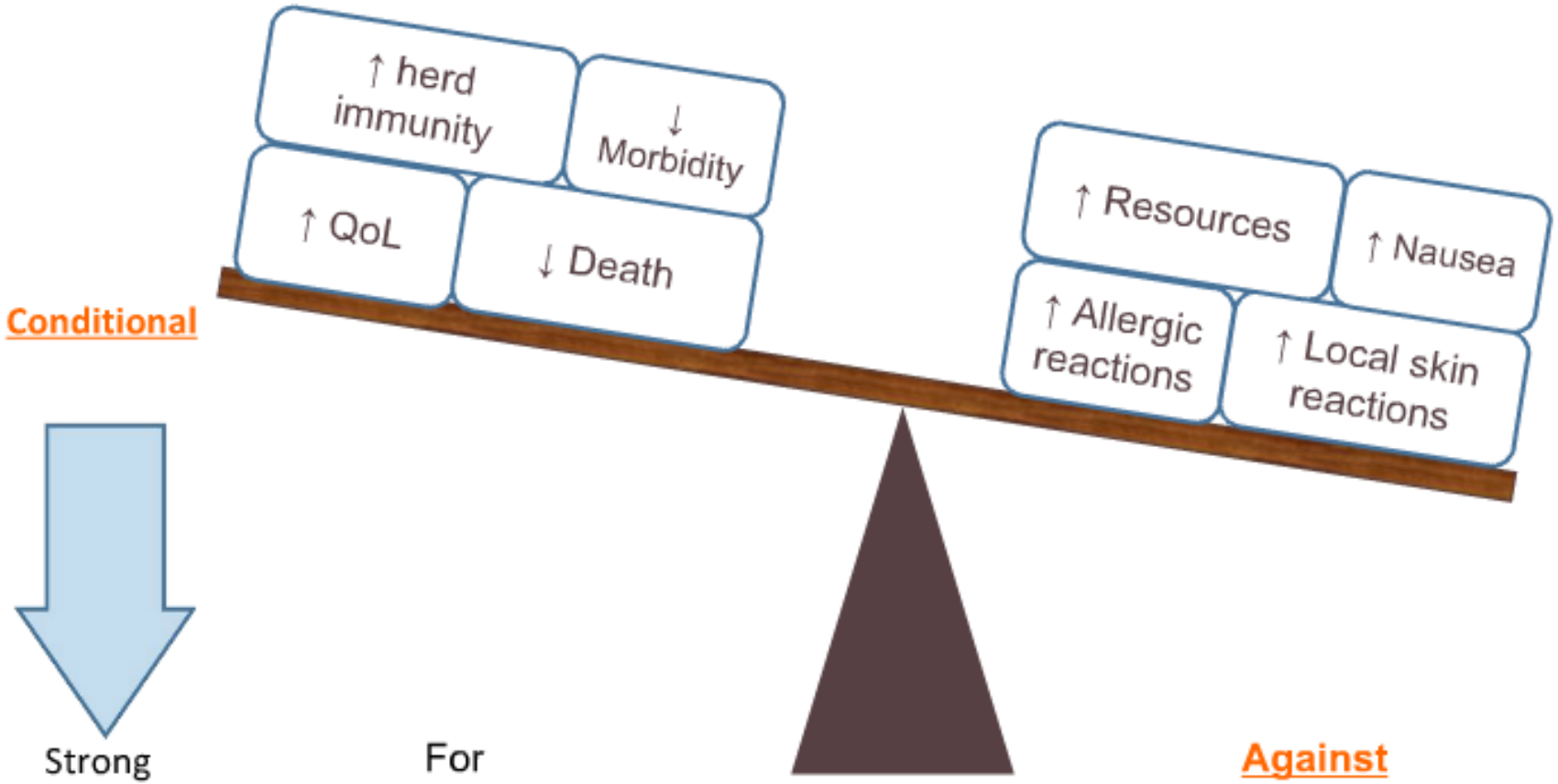
Domains	Comments
Overall **quality of evidence** (confidence in the magnitude of estimates of effect the intervention on critical outcomes)	The higher the quality of evidence, the more likely a policy option is warranted
Balance between desirable and undesirable outcomes (**trade-offs**)	The larger the differences between the desirable and undesirable consequences, the more likely a policy option is warranted
Confidence in **values and preferences**	The greater the variability in values and preferences, or uncertainty in values and preferences, the more likely a provisional policy option is warranted
Resource use	The higher the costs of an intervention (the more the resources consumed), the less likely a policy option is warranted

Source: J Andrew et al. JCE 2013

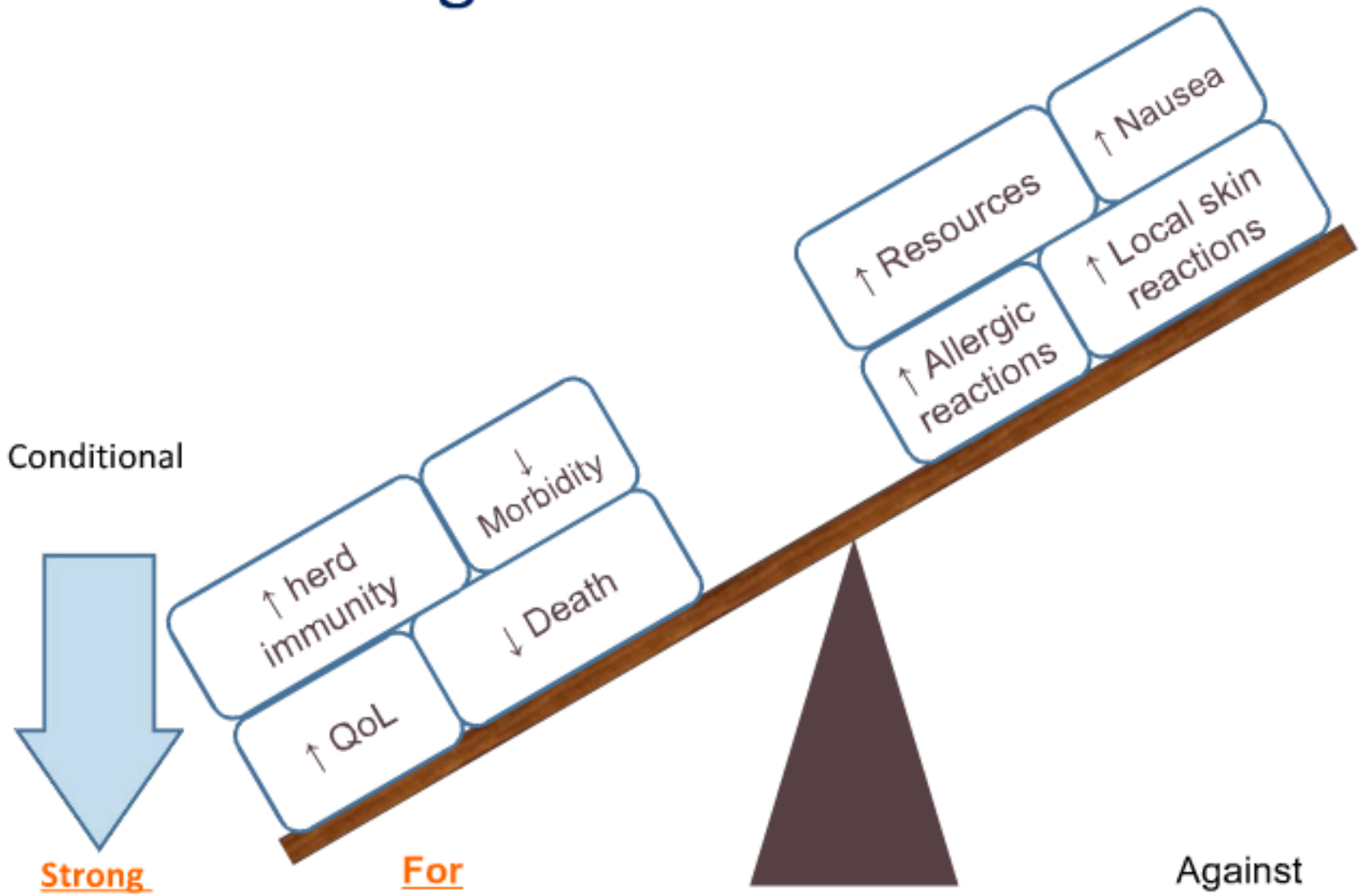
Balancing benefits and downsides



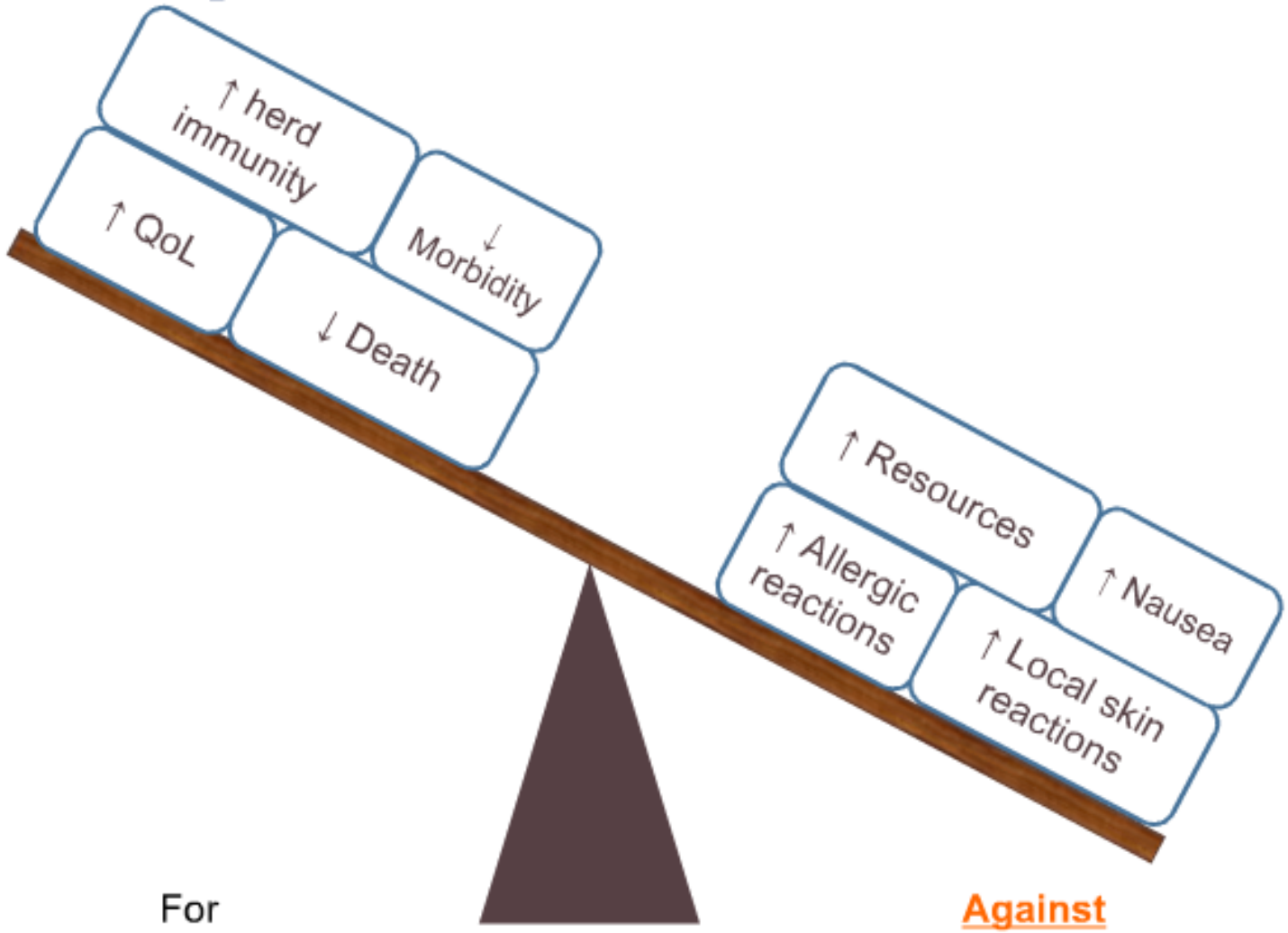
Balancing benefits and downsides



Balancing benefits and downsides



Balancing benefits and downsides



Conditional

Strong

For

Against

Example of how to put it altogether

Desirable effects	- Don't know	- Varies		- Trivial	- Small	✓ Moderate	- Large
Undesirable effects	- Don't know	- Varies		- Large	- Moderate	- Small	✓ Trivial
Certainty of the evidence	- No included studies			- Very low	✓ Low	- Moderate	- High
Values				- Important uncertainty or variability	- Possibly important uncertainty or variability	✓ Probably no important uncertainty or variability	- No important uncertainty or variability
Balance of effects	- Don't know	- Varies	- Favours the comparison	- Probably favours the comparison	- Does not favour either the intervention or the comparison	✓ Probably favours the intervention	- Favours the intervention
Resources required	- Don't know	- Varies	- Large costs	- Moderate costs	- Negligible costs or savings	✓ Moderate savings	- Large savings
Certainty of evidence of required resources	- No included studies			- Very low	✓ Low	- Moderate	- High
Cost-effectiveness	- Don't know	- Varies	- Favours the comparison	- Probably favours the comparison	- Does not favour either the intervention or the comparison	✓ Probably favours the intervention	- Favours the intervention
Equity	- Don't know	- Varies	- Reduced	- Probably reduced	- Probably no impact	✓ Probably increased	- Increased
Acceptability	- Don't know	- Varies		- No	- Probably No	✓ Probably Yes	- Yes
Feasibility	- Don't know	- Varies		- No	- Probably No	✓ Probably Yes	- Yes

☐ Strong recommendation against the option	☐ Conditional recommendation against the option	☐ Conditional recommendation for either the option or the comparison	☐ Conditional recommendation for the option	☒ Strong recommendation for the option

Summarizing key findings from systematic reviews relevant to a policy option

CATEGORY OF FINDING	KEY FINDINGS
Benefits	[Insert one or more bulleted key messages about the benefits that have been found for each component of the option, ensuring that findings are presented with reference to the recency, quality, local applicability, prioritized group applicability and issue applicability]
Potential harms	[Insert one or more bulleted key messages about the harms that have been found for each component of the option, ensuring that findings are presented with reference to the recency, quality, local applicability, prioritized group applicability and issue applicability]
Resource use, costs and/or cost–effectiveness	[Insert one or more bulleted key messages about the resource use, costs and/or cost–effectiveness that have been found for each component of the option]
Uncertainty regarding benefits and potential harms (so M&E could be warranted if the option was to be pursued)	Uncertainty because no systematic reviews were identified [Insert a brief description of option components for which no reviews were identified] Uncertainty because no studies were identified despite an exhaustive search as part of a systematic review [Insert a brief description of option components for which "empty" reviews were identified] No clear message from studies included in a systematic review [Insert a brief description of option components for which there is insufficient evidence]
Key elements of the policy option if it was tried elsewhere	[Insert one or more bulleted key messages about the key elements of the policy option, ensuring that findings are presented with reference to the recency, quality, local applicability, prioritized group applicability and issue applicability]
Stakeholders' views and experiences	[Insert one or more bulleted key messages about stakeholders' views and experiences, ensuring that findings are presented with reference to the recency, quality, local applicability, prioritized group applicability and issue applicability]

Source: Evidence Briefs for Policy. Using the Integrated Knowledge Translation Approach. Guiding Manual. Copenhagen: WHO Regional Office for Europe; 2020

Scenario – 3

Category of finding	Option A	Option B	Option C
Benefits	Probably reduces the likelihood of postpartum bleeding by 20% (Moderate confidence)	Possibly reduces the likelihood of postpartum bleeding by 35% (Low confidence)	Reduces the likelihood of postpartum bleeding by 15% (High confidence)
Potential harms	Nausea and occasional vomiting	Nausea, vomiting, shivering, fever, diarrhoea	Minor nausea. Generally not associated with side effects
Resource use, cost and cost-effectiveness	Supply cost is approximately 1.5 times more than that of Option C	Supply cost is approximately 0.5 times more than that of Option C	Supply cost is $1 for one woman's treatment
Uncertainty regarding balance between benefits and potential harms	Probably favours this Option.	Possibly favours this Option.	Favours this Option
Key elements of the options if tried elsewhere	Does not have any special storage requirements. Can only be administered via injections. Requires trained maternity staff	Does not have any special storage requirements. Tablets should be kept in tightly closed containers and protected from humidity. Trained lay health workers can also administer	Cold chain storage and transport; requires protection from light, and storage at 2–8 °C to prolong shelf life. Can only be administered via injections
Stakeholders views and experiences	Recently registered in national EML, not yet widely available. Health care providers keen to apply in local practice	Has been in use for decades in the country. Registered in national EML. Widely available. Evidence of substandard products in local markets. Health care providers at primary care level believes it is life-saving	Has been in use for decades in the country. Registered in national EML. Widely available but large quantity of substandard products in local markets. Health care providers have lost confidence in its potency – using double dose for routine care.

Based on the systematic review of effects of Options A, B, and C, and additional local evidence gathered as presented above, which policy option(s) would you suggest to address the priority problem (maternal deaths from postpartum bleeding) and why?