APRIL 10-11, 2026 • FRENCH EMBASSY, WASHINGTON D.C.

KIDNEY DISEASE CLINICAL TRIALISTS

The processes of developing KDIGO recommendations

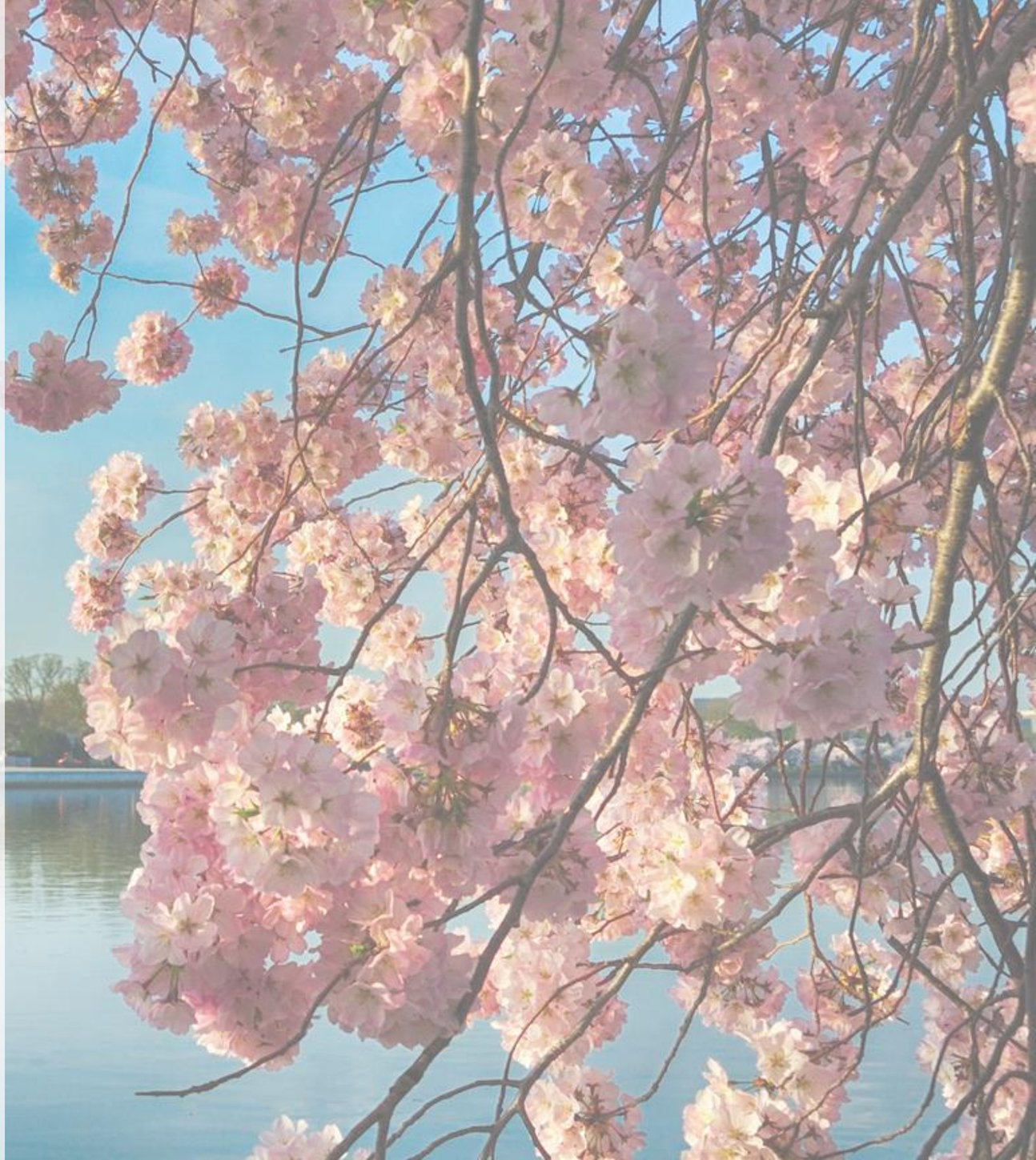
Session:7B Saturday April 11

Reem A. Mustafa, MD, MPH, PhD

Professor of Medicine/Nephrology

Chair of the KDIGO Methods committee

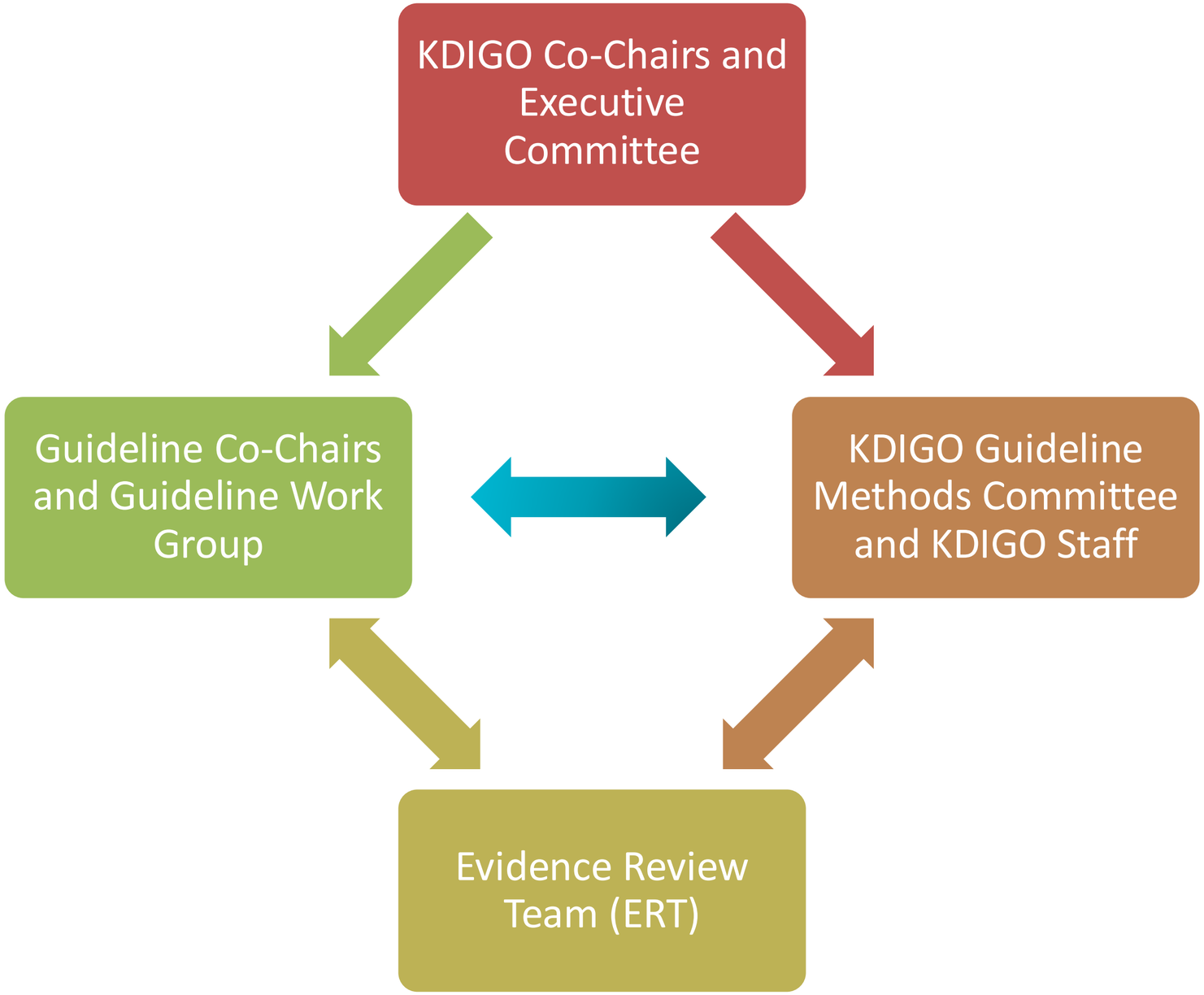
University of Kansas Medical Center



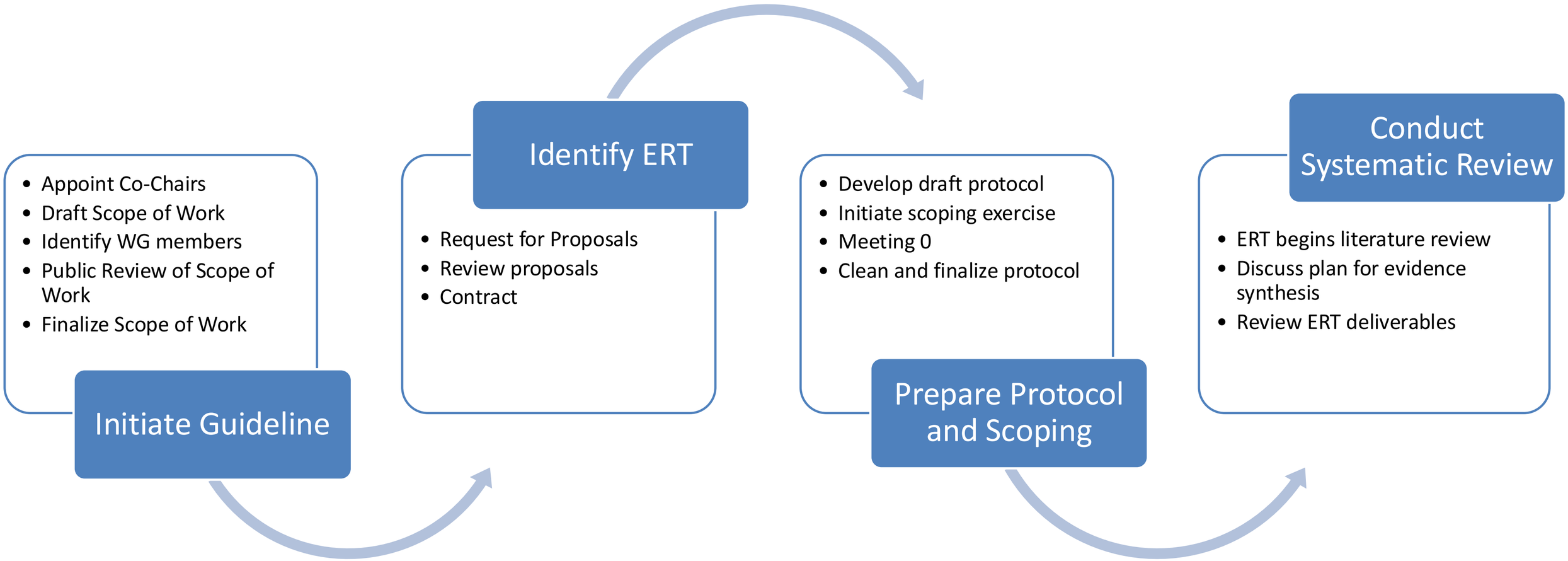
Disclosures

- Co-founder of the US GRADE network and the Evidence Foundation
- GRADE Guidance Group (GRADE executive committee)
- Chair-Methods committee and Executive Committee : KDIGO guidelines
- Chair and methodologist: WHO, ASH, AGA, ACR, KSA MoH, IDSA, AASM, AAP…
- Clinical Practice Guideline Committee: Canadian Society of Nephrology
- Cochrane Collaboration: member of the Cochrane GRADEing methods group
- Past-Chair of the Midwest Comparative Effectiveness Public Advisory Council (Midwest CEPAC) convened by the Institute for Clinical and Economic Review (ICER)
- President-Elect, Women In Nephrology (WIN)
- PCORI, CTSA, NIDDK, AHRQ, PKDOC funding
- No speaker bureau, no consulting arrangement with for profit entities
- Site PI of EMPA-Kidney study- revenue KUMC- ended 9/2022

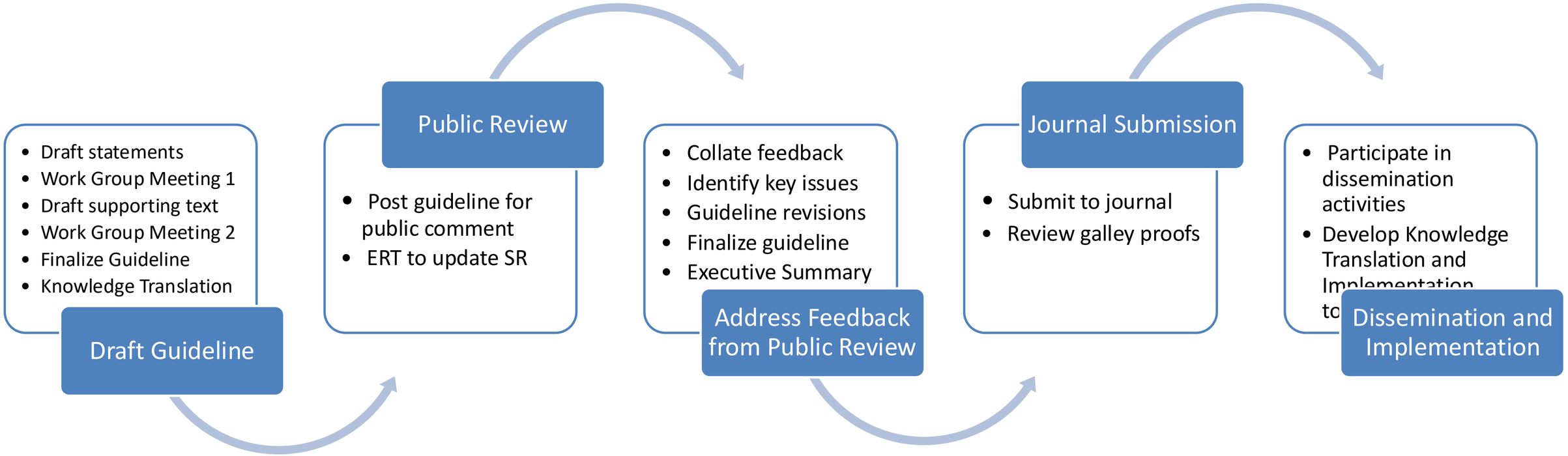
KEY PLAYERS IN GUIDELINE DEVELOPMENT



Guideline Development Process



Guideline Development Process

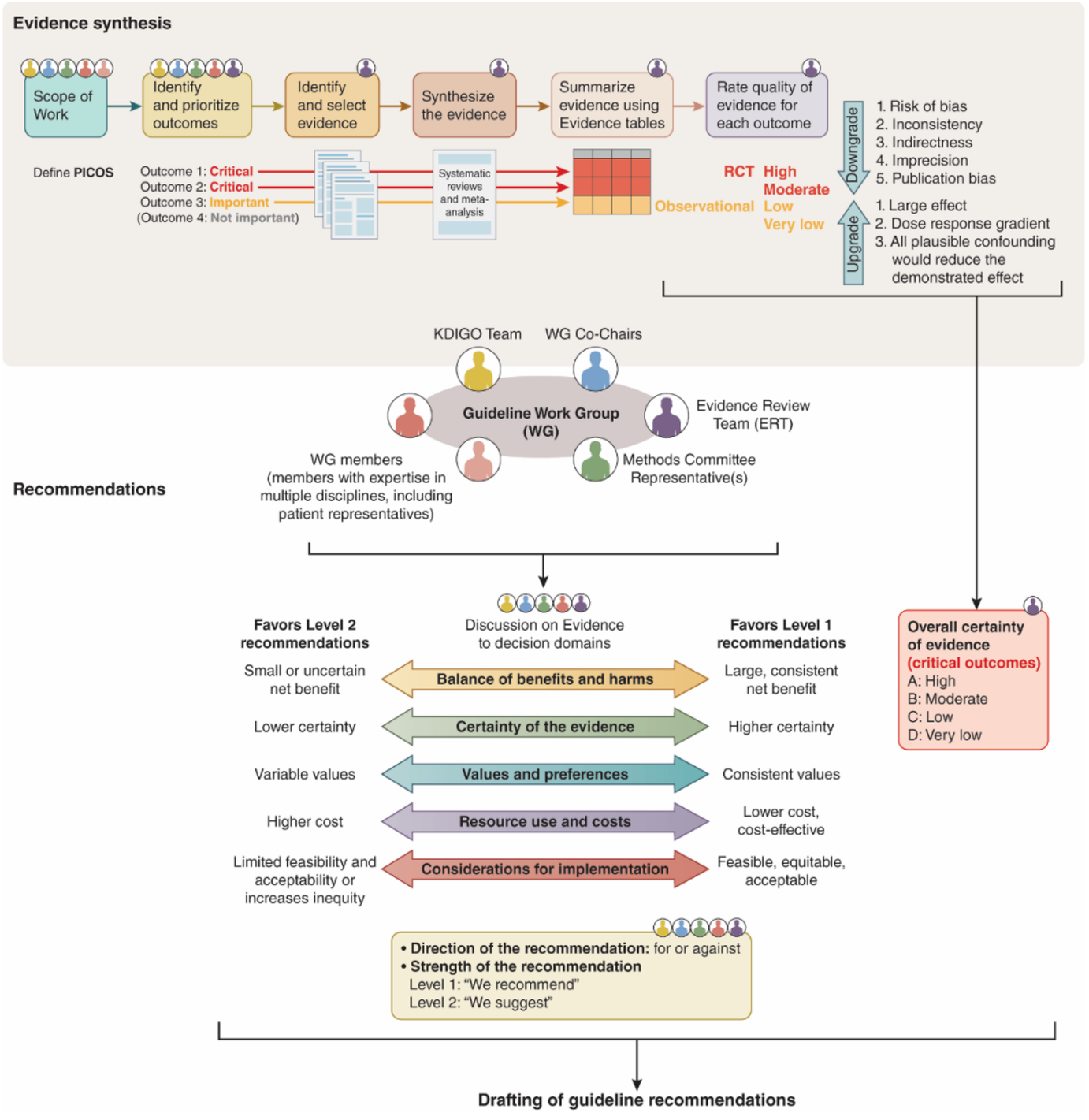


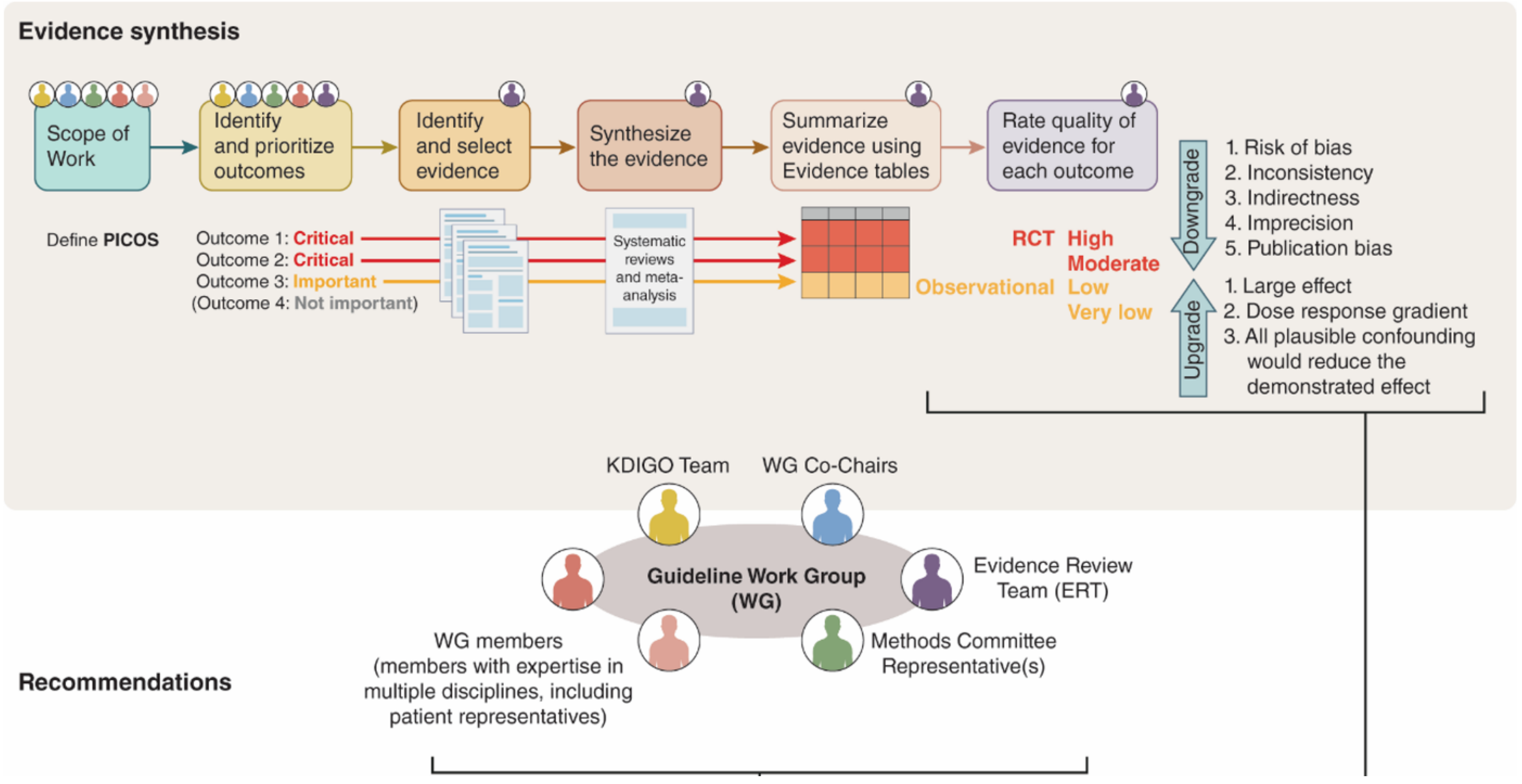
Guideline Process

- The Work Group has the final responsibility for the content of the guideline.
- Decisions are reached through consensus.
- Members are asked to fully disclose any potential conflicts before participation.
- Members may be asked to recuse themselves from any guideline development decisions on issues or topics that are related to compensated relationships.
- Members keep all discussions, debates, or expressed opinions confidential
- All Work Group decisions are made based solely on improving outcomes for patients without undue influence from other organizations or individuals.
- KDIGO guidelines and guideline updates are developed based on science and evidence without obligation to regulatory, financial, or country-specific resources and practices.

Guideline Process

- Work Groups use the GRADE System in rating the quality of evidence and strength of recommendations when evaluating findings and determinations from the independent Evidence Review Team (ERT).
- KDIGO is accountable to patients, and its mission is to improve the outcomes of their care; as such, there are times when guidance is critical to patients despite the paucity of high-quality evidence. To address this unmet need, the Work Group may decide to issue practice points rather than graded recommendations to bridge the evidence gap and provide optimal patient care.
- Members are listed as authors on all guideline publications.


Evidence synthesis
Scope of Work
Identify and prioritize outcomes
Identify and select evidence
Synthesize the evidence
Summarize evidence using Evidence tables
Rate quality of evidence for each outcome
Define PICOS
Outcome 1: Critical
Outcome 2: Critical
Outcome 3: Important
(Outcome 4: Not important)
Systematic reviews and meta-analysis
RCT High
Moderate
Observational Low
Very low
Downgrade
1. Risk of bias
2. Inconsistency
3. Indirectness
4. Imprecision
5. Publication bias
Upgrade
1. Large effect
2. Dose response gradient
3. All plausible confounding would reduce the demonstrated effect
Recommendations
KDIGO Team
WG Co-Chairs
Guideline Work Group (WG)
Evidence Review Team (ERT)
WG members (members with expertise in multiple disciplines, including patient representatives)
Methods Committee Representative(s)
Discussion on Evidence to decision domains
Favors Level 2 recommendations
Favors Level 1 recommendations
Small or uncertain net benefit
Balance of benefits and harms
Large, consistent net benefit
Lower certainty
Certainty of the evidence
Higher certainty
Variable values
Values and preferences
Consistent values
Higher cost
Resource use and costs
Lower cost, cost-effective
Limited feasibility and acceptability or increases inequity
Considerations for implementation
Feasible, equitable, acceptable
Overall certainty of evidence (critical outcomes)
A: High
B: Moderate
C: Low
D: Very low
• Direction of the recommendation: for or against
• Strength of the recommendation
Level 1: "We recommend"
Level 2: "We suggest"
Drafting of guideline recommendations


Evidence synthesis
Scope of Work
Identify and prioritize outcomes
Identify and select evidence
Synthesize the evidence
Summarize evidence using Evidence tables
Rate quality of evidence for each outcome
Define PICOS
Outcome 1: Critical
Outcome 2: Critical
Outcome 3: Important
(Outcome 4: Not important)
Systematic reviews and meta-analysis
RCT High
Moderate
Observational Low
Very low
Downgrade
1. Risk of bias
2. Inconsistency
3. Indirectness
4. Imprecision
5. Publication bias
Upgrade
1. Large effect
2. Dose response gradient
3. All plausible confounding would reduce the demonstrated effect
KDIGO Team
WG Co-Chairs
Guideline Work Group (WG)
Evidence Review Team (ERT)
WG members (members with expertise in multiple disciplines, including patient representatives)
Methods Committee Representative(s)
Recommendations

Assessing Certainty of Evidence (quality of evidence) by Outcome

Table: GRADE's approach to rating quality of evidence (aka certainty in effect estimates)

For each outcome based on a systematic review and across outcomes (lowest quality across the outcomes critical for decision making)

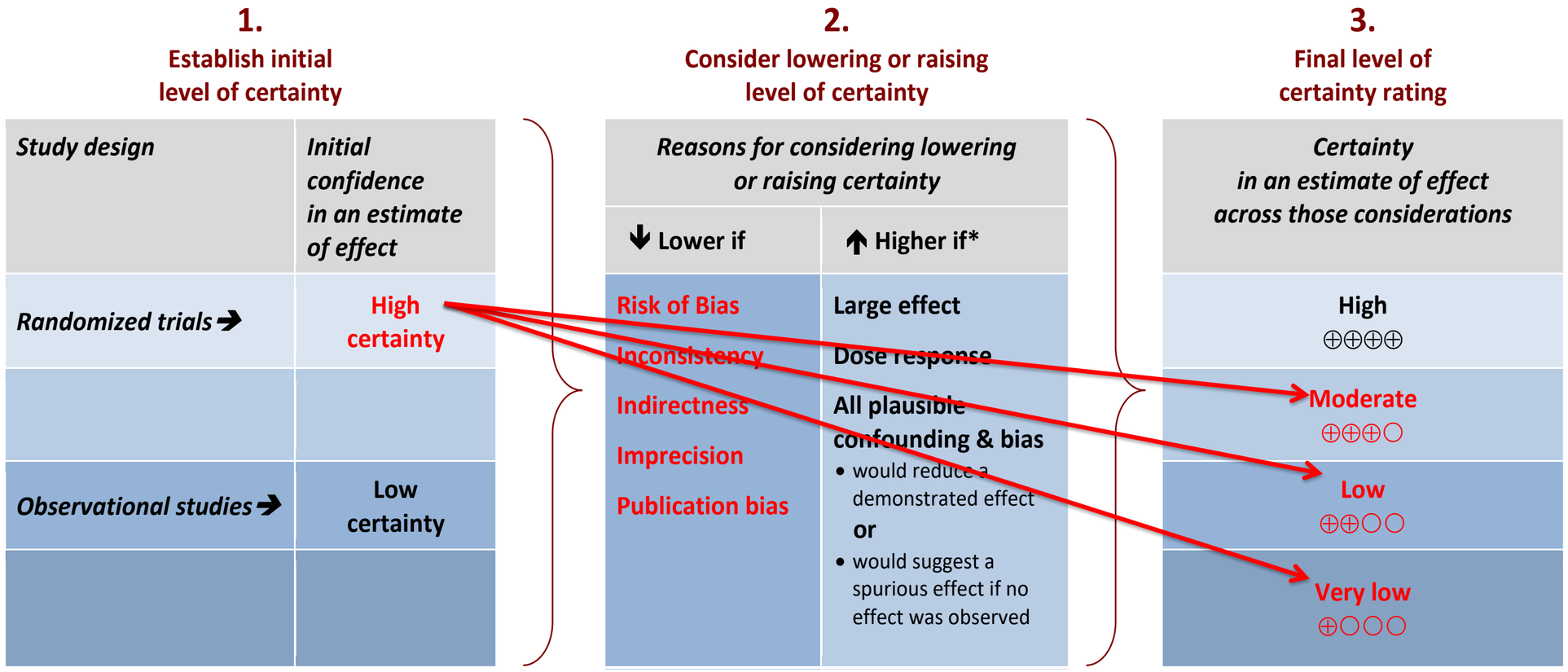
1. Establish initial level of certainty		**2. Consider lowering or raising level of certainty**		**3. Final level of certainty rating**
Study design	***Initial certainty in an estimate of effect***	***Reasons for considering lowering or raising certainty***		***Certainty in an estimate of effect across those considerations***
		⬇ Lower if	**⬆ Higher if***	
Randomized trials ➔	**High certainty**	**Risk of Bias** **Inconsistency** **Indirectness** **Imprecision** **Publication bias**	**Large effect** **Dose response** **All plausible confounding & bias** • would reduce a demonstrated effect **or** • would suggest a spurious effect if no effect was observed	**High** ⨁⨁⨁⨁
				Moderate ⨁⨁⨁◯
Observational studies ➔	**Low certainty**			**Low** ⨁⨁◯◯
				Very low ⨁◯◯◯

*upgrading criteria are usually applicable to observational studies only.

Lowering certainty in RCTs

Table: GRADE's approach to rating quality of evidence (aka certainty in effect estimates)

For each outcome based on a systematic review and across outcomes (lowest quality across the outcomes critical for decision making)

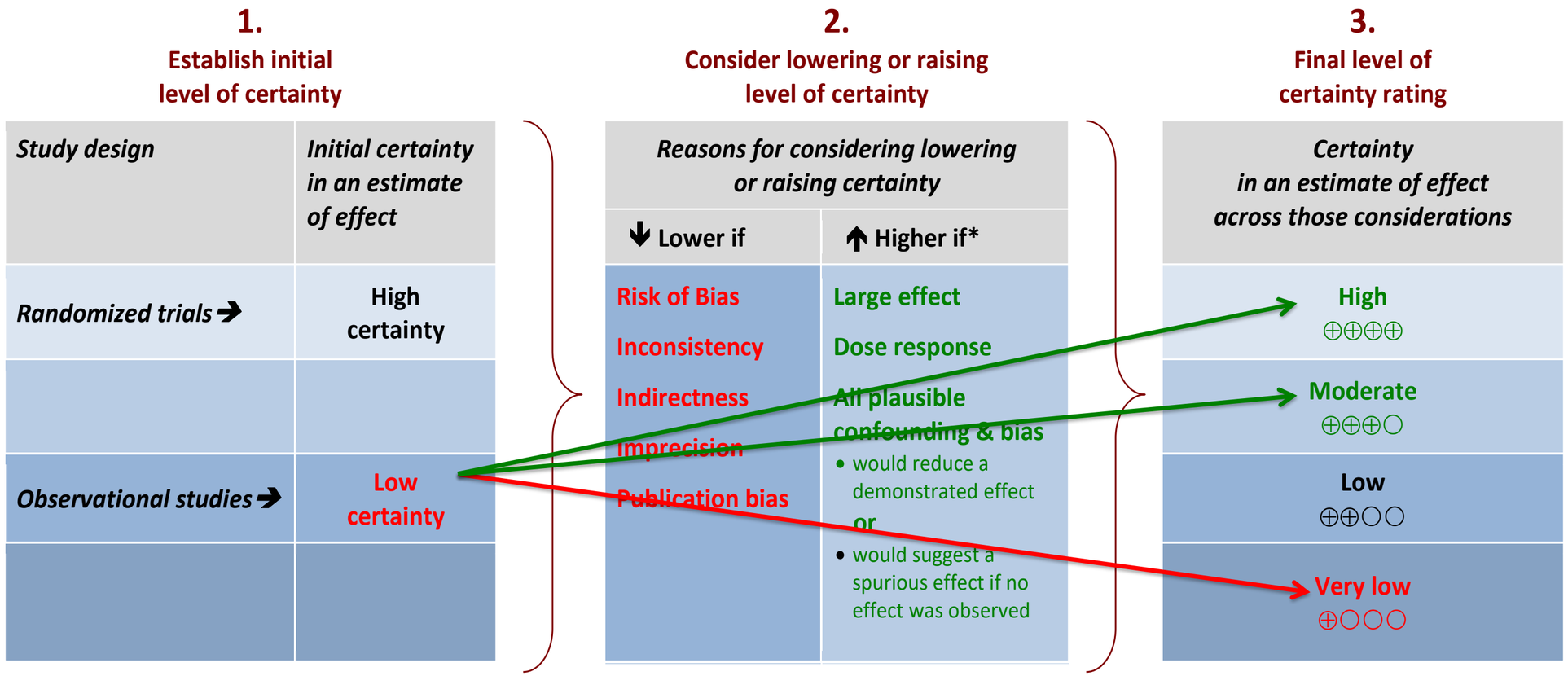


*upgrading criteria are usually applicable to observational studies only.

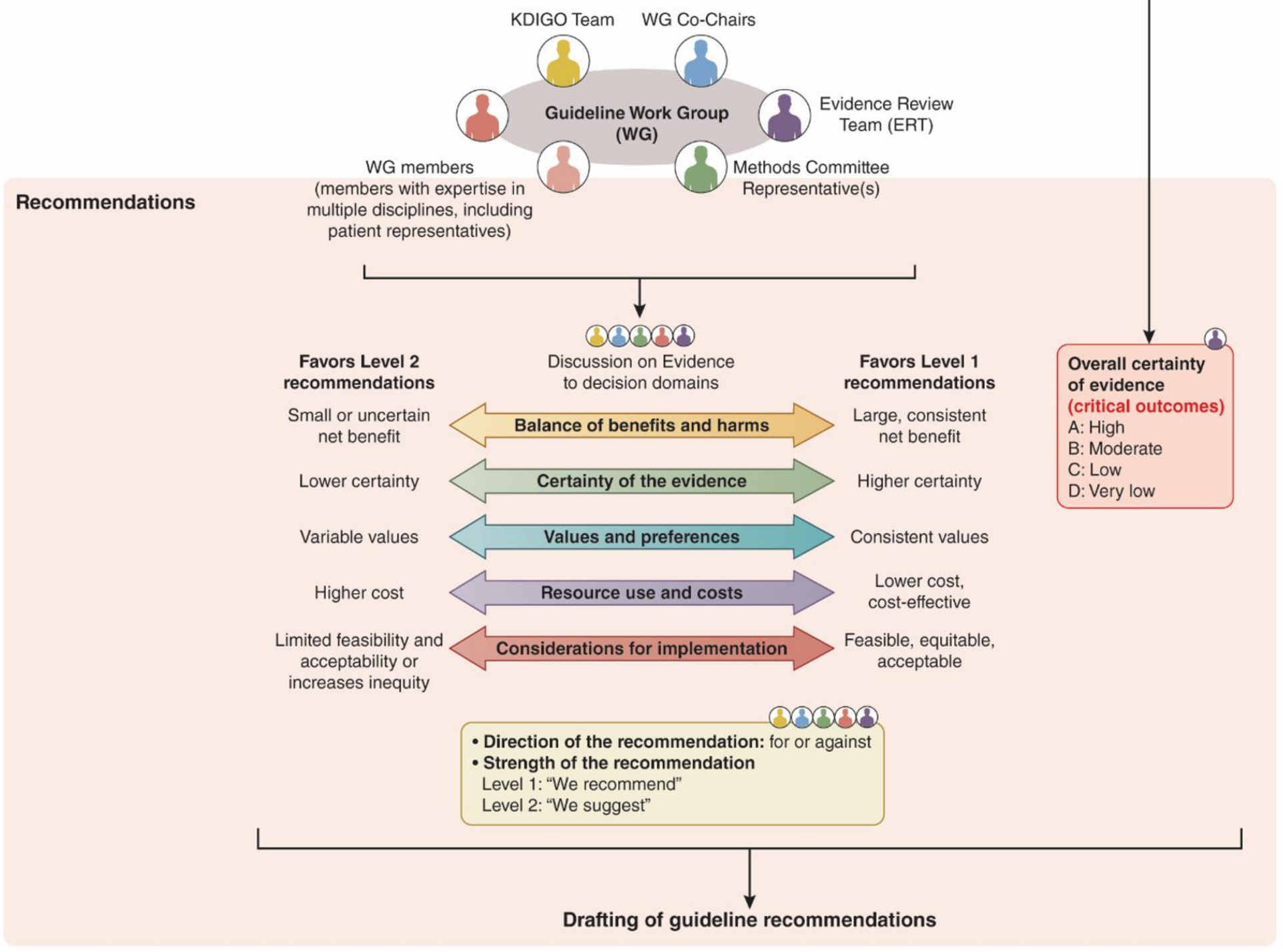
Altering certainty in observational studies

Table: GRADE's approach to rating quality of evidence (aka certainty in effect estimates)

For each outcome based on a systematic review and across outcomes (lowest quality across the outcomes critical for decision making)



*upgrading criteria are usually applicable to observational studies only.


KDIGO Team
WG Co-Chairs
Guideline Work Group (WG)
Evidence Review Team (ERT)
WG members (members with expertise in multiple disciplines, including patient representatives)
Methods Committee Representative(s)
Recommendations
Discussion on Evidence to decision domains
Favors Level 2 recommendations
Favors Level 1 recommendations
Small or uncertain net benefit
Balance of benefits and harms
Large, consistent net benefit
Lower certainty
Certainty of the evidence
Higher certainty
Variable values
Values and preferences
Consistent values
Higher cost
Resource use and costs
Lower cost, cost-effective
Limited feasibility and acceptability or increases inequity
Considerations for implementation
Feasible, equitable, acceptable
Overall certainty of evidence (critical outcomes)
A: High
B: Moderate
C: Low
D: Very low
• Direction of the recommendation: for or against
• Strength of the recommendation
Level 1: "We recommend"
Level 2: "We suggest"
Drafting of guideline recommendations

Implications of the Level of Recommendation

	Patients	Clinicians	Policy
Level 1, "We recommend"	**Most** people in your situation would want the recommended course of action, and only a small proportion would not.	**Most** patients should receive the recommended course of action.	The recommendation can be evaluated as a candidate for developing a policy or a performance measure.
Level 2, "We suggest"	The **majority** of people in your situation would want the recommended course of action, but **many** would not.	**Different choices will be appropriate for different patients.** Each patient needs help to arrive at a management decision consistent with her or his values and preferences.	The recommendation is likely to require substantial debate and involvement of stakeholders before policy can be determined.

Factors Influencing the Strength of Recommendation

Factors influencing the strength of recommendation	Comment
Balance of benefits and harms	The larger the difference between desirable and undesirable effects, the more likely a Level 1 recommendation is warranted. The narrower the gradient, the more likely a Level 2 recommendation is warranted.
Certainty of the evidence	The higher the certainty of evidence, the more likely a Level 1 recommendation is warranted. However, there are exceptions for which low or very low certainty of the evidence will warrant a Level 1 recommendation.
Values and preferences	The more variability or the more uncertainty in values and preferences, the more likely a Level 2 recommendation is warranted. Values and preferences were obtained from the literature, when possible, or were assessed in the judgment of the Work Group when robust evidence was not identified.
Resource use and costs	The higher the cost of an intervention-that is, the more resources consumed- the less likely a Level 1 recommendation is warranted.
Considerations for implementation	Other factors that could be considered when determining the strength of the recommendation might include feasibility of global implementation, equity, sex differences etc. These are factors that may inform the uptake of the recommendation.

Developing a Recommendation

Strength	Level 1 for the intervention	Level 2 for the intervention	Neutral balance (Level 2 for either the intervention or the alternative)	Level 2 for the alternative	Level 1 for the alternative
Balance of EtD factors	Desirable effects **clearly** outweigh undesirable effects	Desirable effects **probably** outweigh undesirable effects	**Trade-offs** equally balanced or uncertain	Undesirable effects **probably** outweigh desirable effects	Undesirable effects **clearly** outweigh desirable effects

Guideline Terminology

- The entire document is a guideline and is divided into chapters.
- Each chapter contains:
  - a set of recommendations with supporting text,
  - practice points for healthcare providers presented as text and/or tables, figures, or algorithms, and
  - reference to evidence tables outlining the evidence upon which the recommendation is based.
- Finally, each chapter will contain research recommendation outlining unmet needs and areas of future research for a given chapter.

Practice Points

- Practice points are consensus-based statements representing the expert judgment of the WG and are not graded.

- They are issued when a clinical question was not supported by a systematic review, often to help readers implement the guidance from graded recommendation.
    - Specific examples include: frequency of monitoring, provision of standard care (such as regular clinic visits), referral to specialist care, etc.

- Practice points may be formatted as text, a table, a figure, or an algorithm.

- Because practice points add length to the guideline document (and consume the Work Group's time), they should be used judiciously.

Situations When Practice Points Can/Should Be Used

1.
Not Testable PICO Question

- There is no reasonable alternative
- Rarely needed

2.
Implementation Guidance

- Commonly used as figures, tables and explanatory text
- Variable as needed

3.
Testable PICO Question

- Not included in a systematic review due to resource limitations
- Exceedingly rare

Recommendations vs Practice points

Recommendations	**Practice points**
Based on systematic review of the evidence.	Not based on systematic reviews.
Graded for both the strength of recommendation and certainty of the evidence.	Not graded for either strength of recommendation or certainty of evidence.
Presented as text (Level 1 “We recommend”; Level 2 “We suggest”)	Presented in various formats, including text, tables, figures or algorithms.
Guidance is always actionable.	Guidance is discretionary for the healthcare providers.

Challenges and Opportunities

- Rapidly changing landscape in Nephrology 🎉🎉🎉
  - Commentaries vs guideline updates
  - Living guidelines- methods under review and discussion

Commentary vs Guideline update

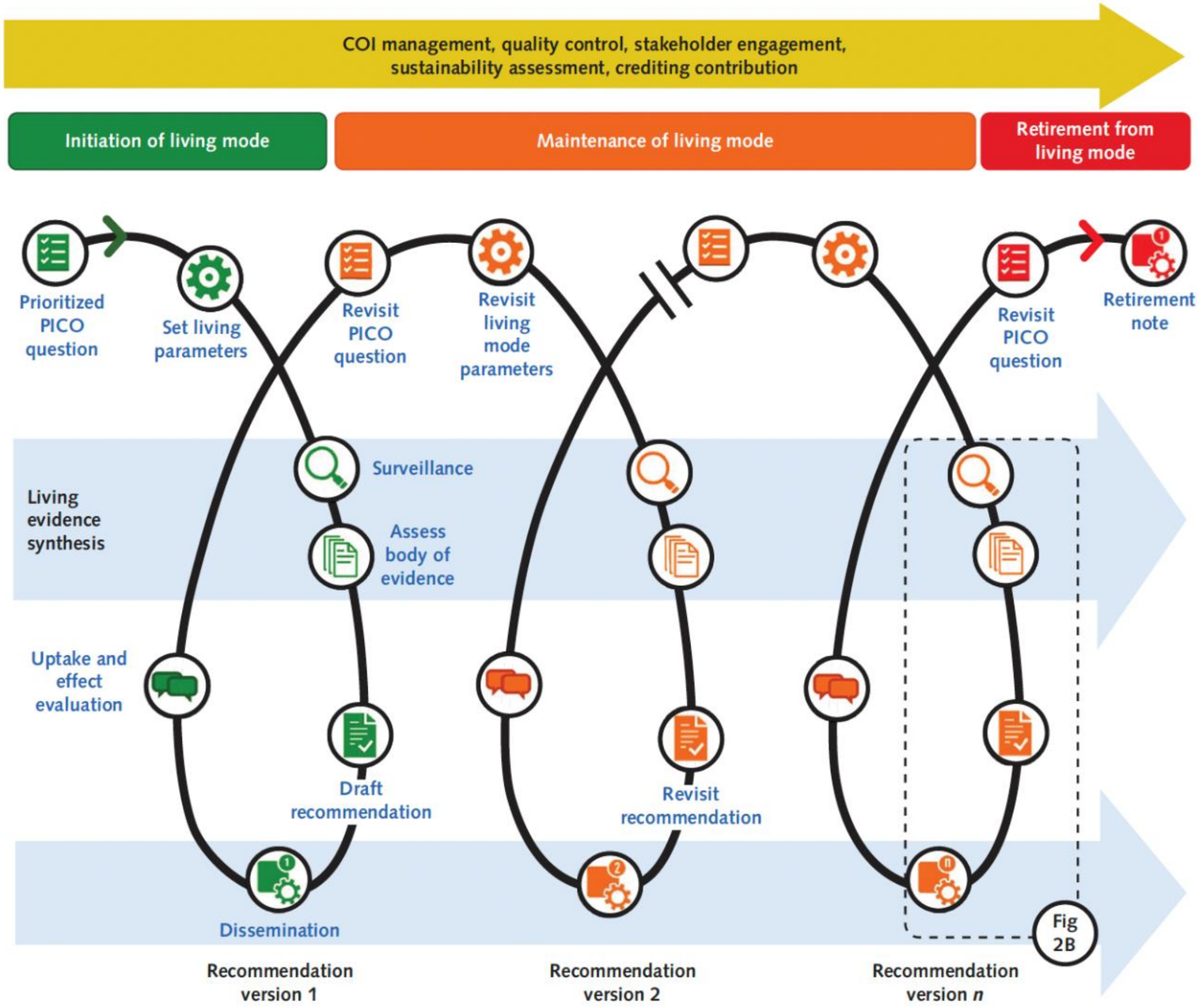
Criteria	Commentary	Guideline update [a]
Impact of new evidence (consider strength, direction, novelty)	Important, but may not be practice-changing	Practice-changing
Long-term data for patient-important outcomes, including safety	Not required	Yes
New studies are anticipated during the update timeline	Not required	No
New evidence repeatedly changing the direction of the recommendation for or against an intervention (i.e., Yo-yo effect)	Possible	Unlikely

All criteria will be met for a guideline update, whereas meeting any criteria for a commentary is sufficient

GUIDELINE UPDATES

Living framework consist of 3 phases

1) Initiation of the living mode
2) Maintenance of living mode
3) Retirement from living mode



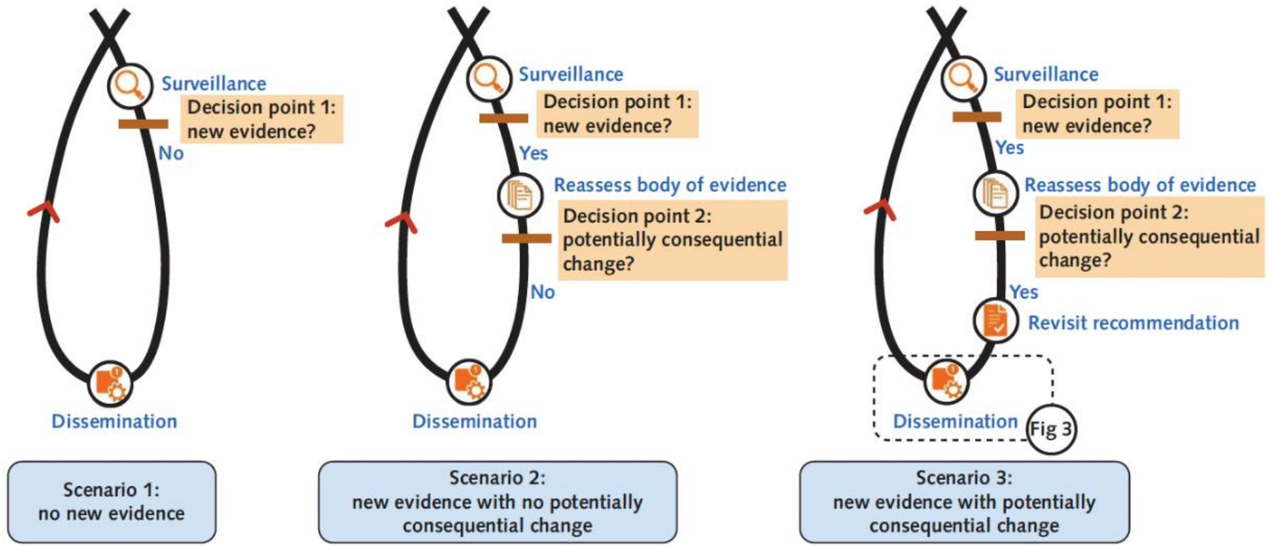
Mikati et al. A Framework for the Development of Living Practice Guidelines in Health Care in guidelines. Annals of Internal Medicine, 2022 Volume 175, Number 8

Guideline Updates

- The panel (including KDIGO Chairs, WG Chairs, KDIGO team, Methods Representative) should define when a short commentary is adequate, when a single recommendation requires updating, when an entire chapter should be revisited, and when a full guideline update is needed based on.
  - **Emergence of New Evidence:** Substantial new evidence, such as pivotal practice changing clinical trials that significantly impact existing recommendations unless additional new evidence is soon anticipated.
  - **Changes in Clinical Practice:** Advances in technology, diagnostic tools, or treatment modalities that necessitate revised recommendations.
  - **Feedback from Stakeholders:** Input from clinicians, patients, or policymakers highlighting areas where guidelines may be silent, outdated or unclear.
  - **Regulatory Changes:** New regulations, approvals, or safety advisories that directly affect guideline recommendations.

Guideline Updates

- Revisiting a recommendation can result in several outcomes:
  - No modification
  - Modification of elements such as scope, direction, strength, or certainty
  - Merging recommendation with another recommendation
  - Splitting of the recommendation into 2 or more recommendations
  - Retirement from the living mode: i.e., the question is still relevant and the recommendation is still valid but it would not be updated using a living mode.
  - Removal of recommendations



Challenges and Opportunities

- Rapidly changing landscape in Nephrology 🎉🎉🎉
  - Commentary vs GL update
  - Living guidelines
    - When to invoke an update? any trial, new outcomes, new interventions
    - How to maintain the rigor of the process and remain responsive to the community needs
    - Unit of the update
    - Publication platforms
    - Acknowledging the contribution of the WG
    - Avoid working group fatigue
- Lack of comparative effectiveness
  - NMA and IPDMA
- Lack of long-term safety data
  - When is it reasonable to issue recommendations on a new class with accelerated approval based on a reasonably likely surrogate?
  - HARMS?
  - What are the other options for treatment?

Challenges and Opportunities

- Lack of data on outcomes that matter to patients and on PROs
  - We have to include all patient important outcomes in trials
  - Need to develop validated and responsive PROs
- Lack of tools that allows appropriate elicitation of people's value and risk tolerance
  - Individual, generation, gender, ethnic…etc differences
- Mismatch between trial inclusion criteria and those who need the treatment
  - Real world evidence
- Lack of direct data in certain population (e.g. pediatrics vs adults)
  - Role of bridging biomarkers
  - HARM?
- Inconsistency in inclusion criteria
  - Reclassification and updating classification and how does that impact using the data- Strength of recommendations?
- Perfect should not be the enemy of the good.
  - UACR vs PACR
  - Consistency in guidance


APRIL 10-11, 2026 • FRENCH EMBASSY, WASHINGTON D.C.
KIDNEY
DISEASE
CLINICAL
TRIALISTS
Thank you