Evolving the Theory of Everything for Health Decision-Making

EbM-Kongress 2025 – 26. Jahrestagung des EbM-Netzwerks

Holger Schünemann, MD, MSc, PhD, MGIN, FRCPC

Clinical Epidemiology and Research Center
Humanitas University, Milan, Italy

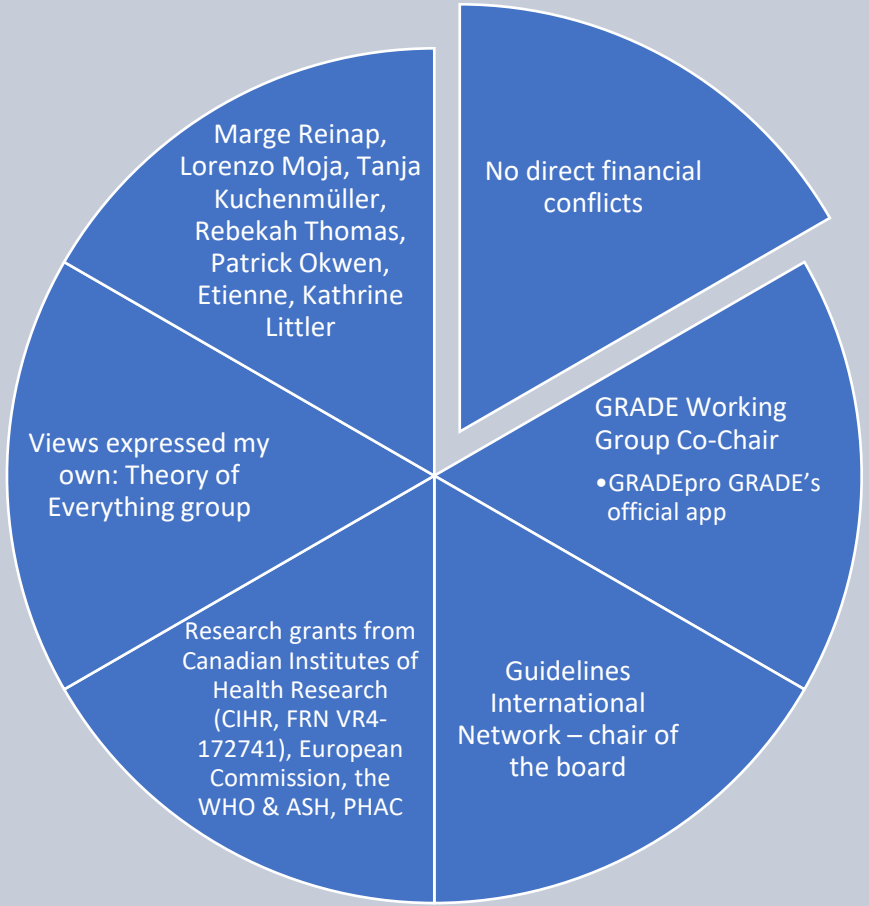
holger.schunemann@hunimed.eu

HUMANITAS UNIVERSITY

WHO Collaborating Center for Evidence-based Decision-making in Health

GRADE working group

Disclosures



GRADE working group

"25 Jahre"

GIN
Guidelines International Network

WHO Collaborating Center for Evidence-based Decision-making in Health

Why This Matters

Health decisions shape equity, impact lives, and allocate billions.

But decisions are often fragmented, politicized, or ad hoc.

What if we had a map for the whole decision ecosystem, even globally?

That's what the Theory of Everything (ToE) in health decision-making aims to offer.

Presentation Roadmap

1. From Theory to Action
2. Filling the Gaps: Equity, Power, Ethics
3. Mapping the Decision Genome
4. WHO based Coalition to connect to key actors
5. Future-Proofing Decisions and setting research priorities

What is the Theory of Everything (ToE)?

A conceptual map of the health decision-making ecosystem, its actors and approaches.

Includes framework of decision-criteria, evidence, research, values, actors, systems, and outcomes.

Our theory of everything in health decision-making

Goals

Lay out overarching concepts of decision-making and then create bridges between actors & disciplines & move to action

- go beyond the partnerships that are already established (HTA, guidelines & systematic reviews)

GRADE working group

Health Policy

The ecosystem of health decision making: from fragmentation to synergy

Holger J Schünemann, Marge Reinap, Thomas Piggott, Erki Laidmäe, Kristina Köhler, Mariliis Põld, Brendalynn Ens, Alar Irs, Elie A Akl, Carlos A Cuello, Maicon Falavigna, Michelle Gibbens, Luciana Neamtiu, Elena Parmelli, Mouna Jameleddine, Lisa Pyke, Ilse Verstijnen, Pablo Alonso-Coello, Peter Tugwell, Yuan Zhang, Zuleika Saz-Parkinson, Tanja Kuchenmüller, Lorenzo Moja

Clinicians, patients, policy makers, funders, programme managers, regulators, and science communities invest considerable amounts of time and energy in influencing or making decisions at various levels, using systematic reviews, health technology assessments, guideline recommendations, coverage decisions, selection of essential medicines and diagnostics, quality assurance and improvement schemes, and policy and evidence briefs. The criteria and methods that these actors use in their work differ (eg, the role economic analysis has in decision making), but these methods frequently overlap and exist together. Under the aegis of WHO, we have brought together representatives of different areas to reconcile how the evidence that influences decisions is used across multiple health system decision levels. We describe the overlap and differences in decision-making criteria between different actors in the health sector to provide bridging opportunities through a unifying broad framework that we call theory of everything. Although decision-making activities respond to system needs, processes are often poorly coordinated, both globally and on a country level. A decision made in isolation from other decisions on the same topic could cause misleading, unnecessary, or conflicted inputs to the health system and, therefore, confusion and resource waste.

Lancet Public Health 2022; 7: e378–90

Department of Health Product Policy and Standards (L Moja MD) and Department of Research for Health (T Kuchenmüller MA), WHO, Geneva, Switzerland; Michael G. DeGroote Cochrane Canada Centre and McMaster GRADE Centres, Department of Health Research Methods, Evidence, and Impact, Department of Medicine, and WHO Collaborating Center for Infectious Diseases, Research Methods and Recommendations (Prof H J Schünemann MD, T Piggott MD, C A Cuello MD, Y Zhang PhD), McMaster University, Hamilton, ON, Canada; Dipartimento di Scienze Biomediche, Humanitas University, Milan, Italy (Prof H J Schünemann); Division of Country Health Policies and Systems, WHO Regional Office for Europe, Copenhagen, Denmark (M Reinap MA); Department of Medicines and Medical Devices, Estonian Health Insurance

Introduction

Many actors influence or make decisions at various levels using systematic reviews, health technology assessments (HTAs), guideline recommendations, coverage decisions, selection of essential medicines or diagnostics, quality improvement, and policy or evidence briefs.[1–7] We provide a brief overview of the different actors who influence or make health decisions (panel 1; appendix p 2).

The extent to which these different actors use structured processes and transparent criteria differs, and these processes and criteria are often poorly coordinated, despite a general increased interest in the intersectionality of using key evidence and how closely partnership of systematic review authors and guideline developers.

Under the aegis of WHO (Regional Office for Europe, the Country Office in Estonia, and the headquarter WHO Department of Health Product Policy and Standards), and in collaboration with the Estonian Health Insurance Fund, we brought together representatives of different areas to discuss similarities in the criteria and processes encompassing evidence evaluation across different health system decision levels. This process, beginning in 2019, was inclusive, involving key stakeholders, such as systematic reviewers, guideline developers and panelists, regulators, policy makers, and payers, whose

Background to the work

Decade of working with WHO country office in Estonia on guideline methods and capacity building in evidence-based decision-making

Estonia national guideline making conditional recommendation for DOACs in atrial fibrillation – cost too high for strong recommendation based on systematic review and HTA

Submitted application for inclusion of new oral anticoagulants (direct oral anticoagulants/DOACs) in WHO EML 2015 – rejected, no WHO guideline

Price negotiations with Estonian Health Insurance Fund – manufacturer lowering price → strong recommendation

2019 WHO EML approval of DOACs

Evidence to Decision Criteria and Framework

GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction

Pablo Alonso-Coello,[1,2] Holger J Schünemann,[2,3] Jenny Moberg,[4] Romina Brignardello-Petersen,[2,5] Elie A Akl,[2,6] Marina Davoli,[7] Shaun Treweek,[8] Reem A Mustafa,[2,9] Gabriel Rada,[10,11,12] Sarah Rosenbaum,[4] Angela Morelli,[4] Gordon H Guyatt,[2,3] Andrew D Oxman[4] the GRADE Working Group

Starting point

Horizon project, many partners, including WHO, NICE and other actors

GRADE working group

Table 2 | Detailed judgments in Evidence to Decision (EtD) frameworks

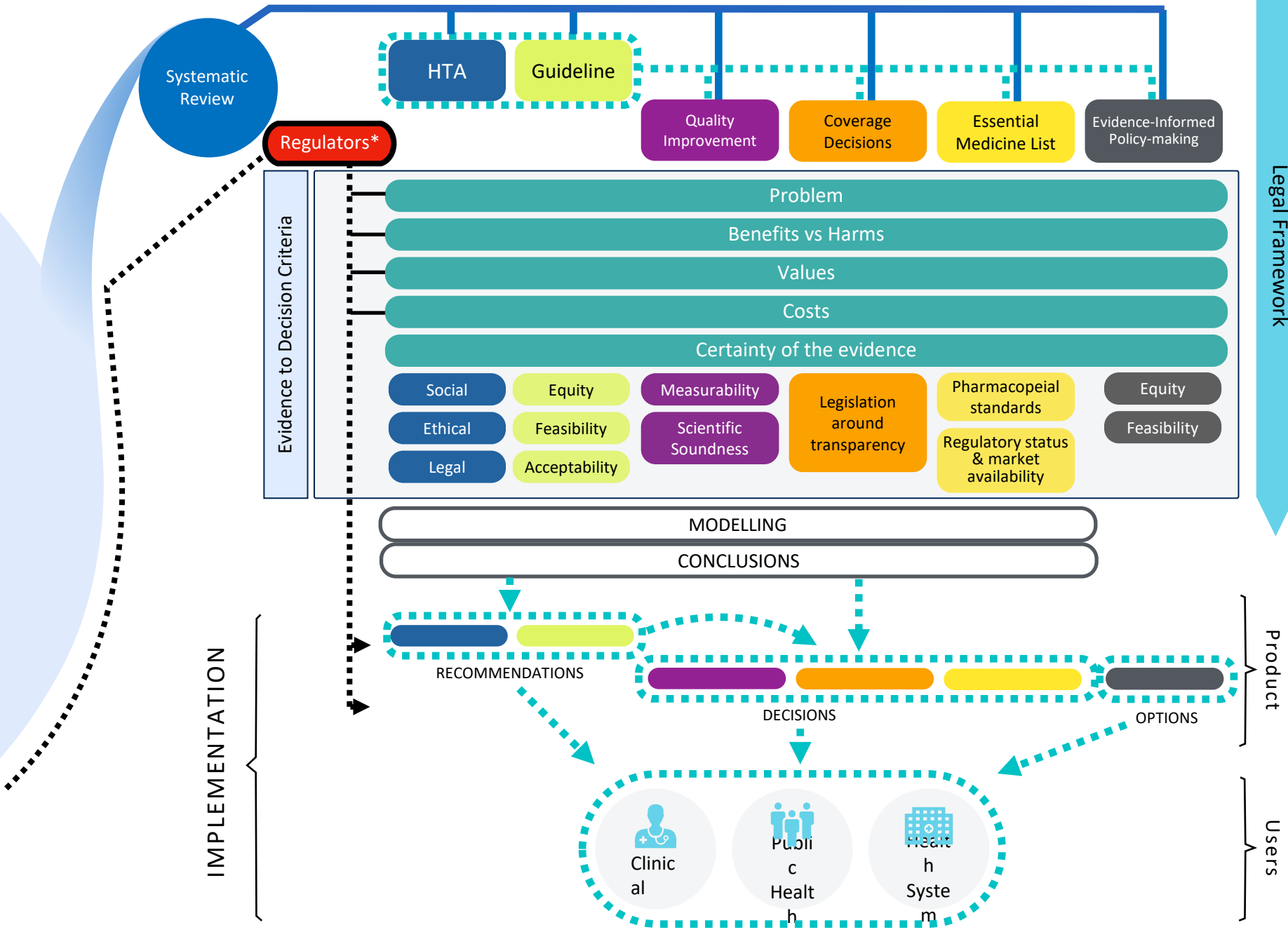
Criterion	Detailed judgments
Is the problem a priority?*	• Are the consequences of the problem serious (that is, severe or important in terms of the potential benefits or savings)? • Is the problem urgent? [Not relevant for coverage decisions] • Is it a recognised priority (such as based on a political or policy decision)? [Not relevant when an individual patient perspective is taken]
How substantial are the desirable anticipated effects?	• Judgments for each outcome for which there is a desirable effect
How substantial are the undesirable anticipated effects?	• Judgments for each outcome for which there is an undesirable effect
What is the overall certainty of the evidence of effects?	• See GRADE guidance regarding detailed judgments about the quality of evidence or certainty in estimates of effects[30,31]
Is there important uncertainty about or variability in how much people value the main outcomes?	• Is there important uncertainty about how much people value each of the main outcomes? • Is there important variability in how much people value each of the main outcomes? [Not relevant for coverage decisions]
Do the desirable effects outweigh the undesirable effects?	• Judgments regarding each of the four preceding criteria • To what extent do the following considerations influence the balance between the desirable and undesirable effects: - How much less people value outcomes that are in the future compared to outcomes that occur now (their discount rates)? - People's attitudes towards undesirable effects (how risk averse they are)? - People's attitudes towards desirable effects (how risk seeking they are)?
How large are the resource requirements?†	• How large is the difference in each item of resource use for which fewer resources are required? • How large is the difference in each item of resource use for which more resources are required?
What is the certainty of the evidence of resource requirements?†	• Have all important items of resource use that may differ between the options being considered been identified? • How certain is the evidence of differences in resource use between the options being considered (see GRADE guidance regarding detailed judgments about the quality of evidence or certainty in estimates)? • How certain is the cost of the items of resource use that differ between the options being considered? • Is there important variability in the cost of the items of resource use that differ between the options being considered?
Are the net benefits worth the incremental cost?*	• Judgments regarding each of the six preceding criteria • Is the cost effectiveness ratio sensitive to one-way sensitivity analyses? • Is the cost effectiveness ratio sensitive to multivariable sensitivity analysis? • Is the economic evaluation on which the cost effectiveness estimate is based reliable? • Is the economic evaluation on which the cost effectiveness estimate is based applicable to the setting(s) of interest?
What would be the impact on health equity?*†	• Are there groups or settings that might be disadvantaged in relation to the problem or interventions (options) that are considered? • Are there plausible reasons for anticipating differences in the relative effectiveness of the intervention (option) for disadvantaged groups or settings? • Are there different baseline conditions across groups or settings that affect the absolute effectiveness of the intervention or the importance of the problem for disadvantaged groups or settings? • Are there important considerations that should be made when implementing the intervention (option) in order to ensure that inequities are reduced, if possible, and that they are not increased?
Is the intervention/option acceptable to key stakeholders?*	• Are there key stakeholders who would not accept the distribution of the benefits, harms and costs? • Are there key stakeholders who would not accept the costs or undesirable effects in the short term for desirable effects (benefits) in the future? • Are there key stakeholders who would not agree with the importance (value) attached to the desirable or undesirable effects (because of how they might be affected personally or because of their perceptions of the relative importance of the effects for others)? • Would the intervention adversely affect people's autonomy? • Are there key stakeholders who would disapprove of the intervention morally, for reasons other than its effects on people's autonomy (such as in regard to ethical principles such as no maleficence, beneficence, or justice)?
Is the intervention feasible to implement?*	For decisions other than coverage decisions: • Is the intervention or option sustainable? • Are there important barriers that are likely to limit the feasibility of implementing the intervention (option) or require consideration when implementing it?[30,31] For coverage decisions: • Is coverage of the intervention sustainable? • Is it feasible to ensure appropriate use for approved indications? • Is inappropriate use (indications that are not approved) an important concern? • Is access to the intervention an important concern? • Are there important legal or bureaucratic or legal constraints that that make it difficult or impossible to cover the intervention?

*The certainty of the evidence could be considered as a detailed judgement for these criteria.
†These criteria are not included when an individual patient perspective is taken.

about the strength of recommendation or type of decision; for example, a strong or weak (sometimes called conditional, discretionary, or qualified) recommendation for or against an intervention or option. In addition, the panel states the recommendation or decision in a concise, clear and actionable manner,[18] and provides the justification for their recommendation or decision. The conclusions also include relevant considerations about subgroups, implementation, monitoring and evaluation, and research priorities (see box 3 for the conclusions reached in the bedaquiline example).

Guideline panels may be reluctant to make a recommendation for or against an intervention or option.

Evidence from Primary Research



Evidence ecosystem of health decision-making

The actors

Are fragmented…..

Slovenia – country case study

Population of 2,118,697 (2024 mid year)

GDP in Slovenia reached USD 32,164 per capita.

Ranked 87 of the major economies

Source: Worlddata

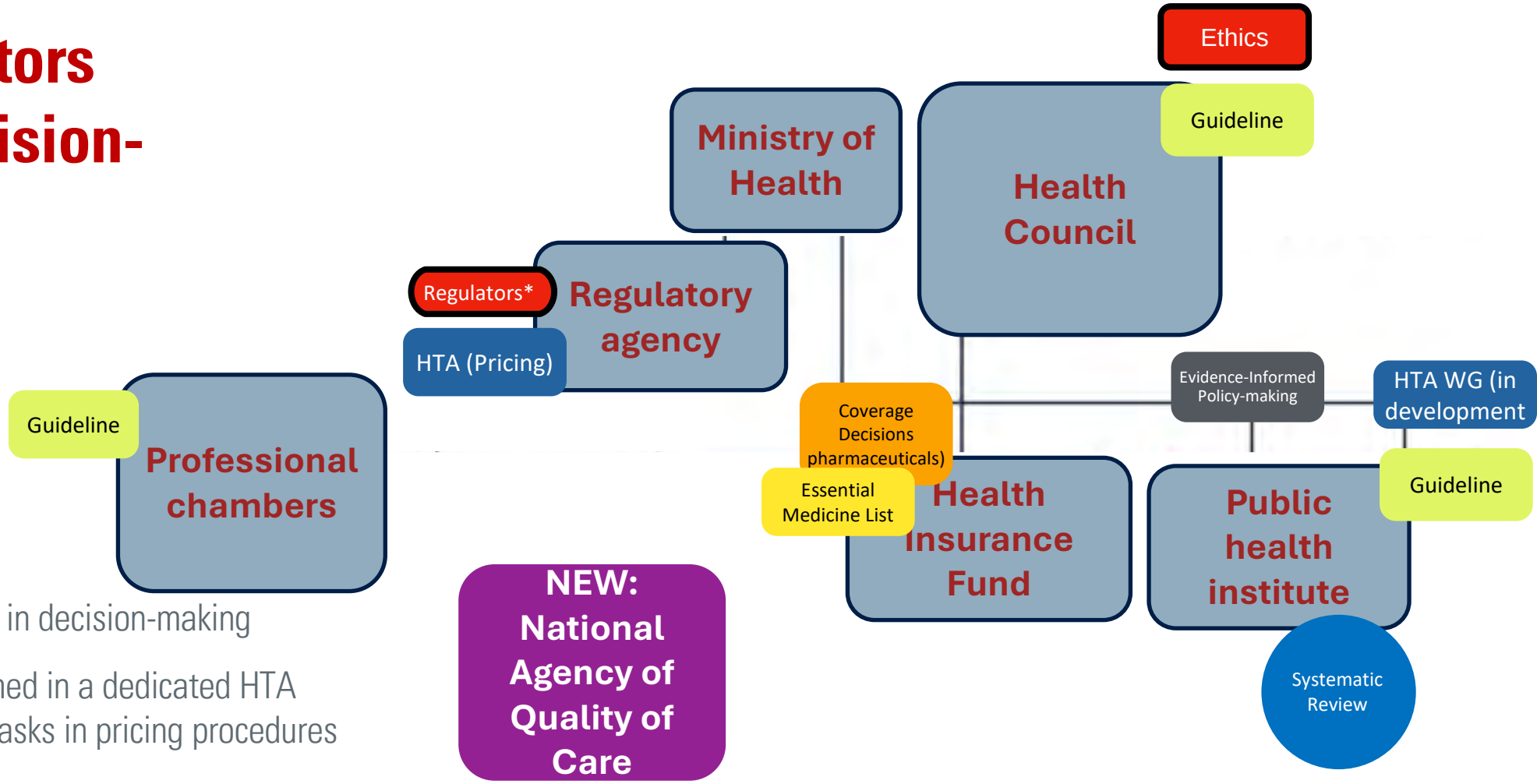
A social health insurance system financed by a single payer provides centralized hospital care and primary care devolved to the municipalities.

Work with M. Reinap, WHO EURO

WHO Collaborating Center for Evidence-based Decision-making in Health

„However, many elements that could improve efficiency, such as a clear methodology for budget allocation based on population health needs, strategic purchasing, or the formal use of health technology assessment (HTA) to support coverage decisions, are still missing.“

Slovenian actors in health decision-making



- Health council advising MoH in decision-making
- „HTA is currently not performed in a dedicated HTA institution, but HTA-related tasks in pricing procedures are performed“
- Guideline development individually driven by medical societies and no proper oversight
- Perception that public health GL≠Clinical GL
- ToE helped uncover gaps in coordination and transparency.

WHO Collaborating Center for Evidence-based Decision-making in Health

With M. Reinap, WHO EURO

Country case studies

Brazil (Luciane Lopes)

Australia – under review (Zachary Munn)

GIN Guidelines International Network

GIN Africa

Health Decision-Making (HDM) Ecosystem ToE: Cameroon, Nigeria, & South Africa

- All 3 countries demonstrate varying levels of HDM dimensions with EML being the commonest approach used for guidance.
- Only South Africa demonstrated a clear use of HTA and government involvement in funding while Nigeria and Cameroon mostly depend on external funders for ecosystems.
- Nigeria and Cameroon prioritize some dimensions (Regulation and EML) over others
- Several sub-ecosystems exist within the HDM ecosystem
- Missing dimensions: Local data (eg from DHIS2) and including **patients/people with lived experience explicitly**

Slides from Etienne Ngeh

Theory of Everything

New elements: clarifying legitimacy, inclusion, epistemic justice.

Participatory priority setting, actor to value mapping.

What is “value”?

What is the role of ethics?

Equity, Politics & Power

ToE must confront structural inequities and decision hierarchies.

Health information systems and what is data is a core upgrade in ToE.

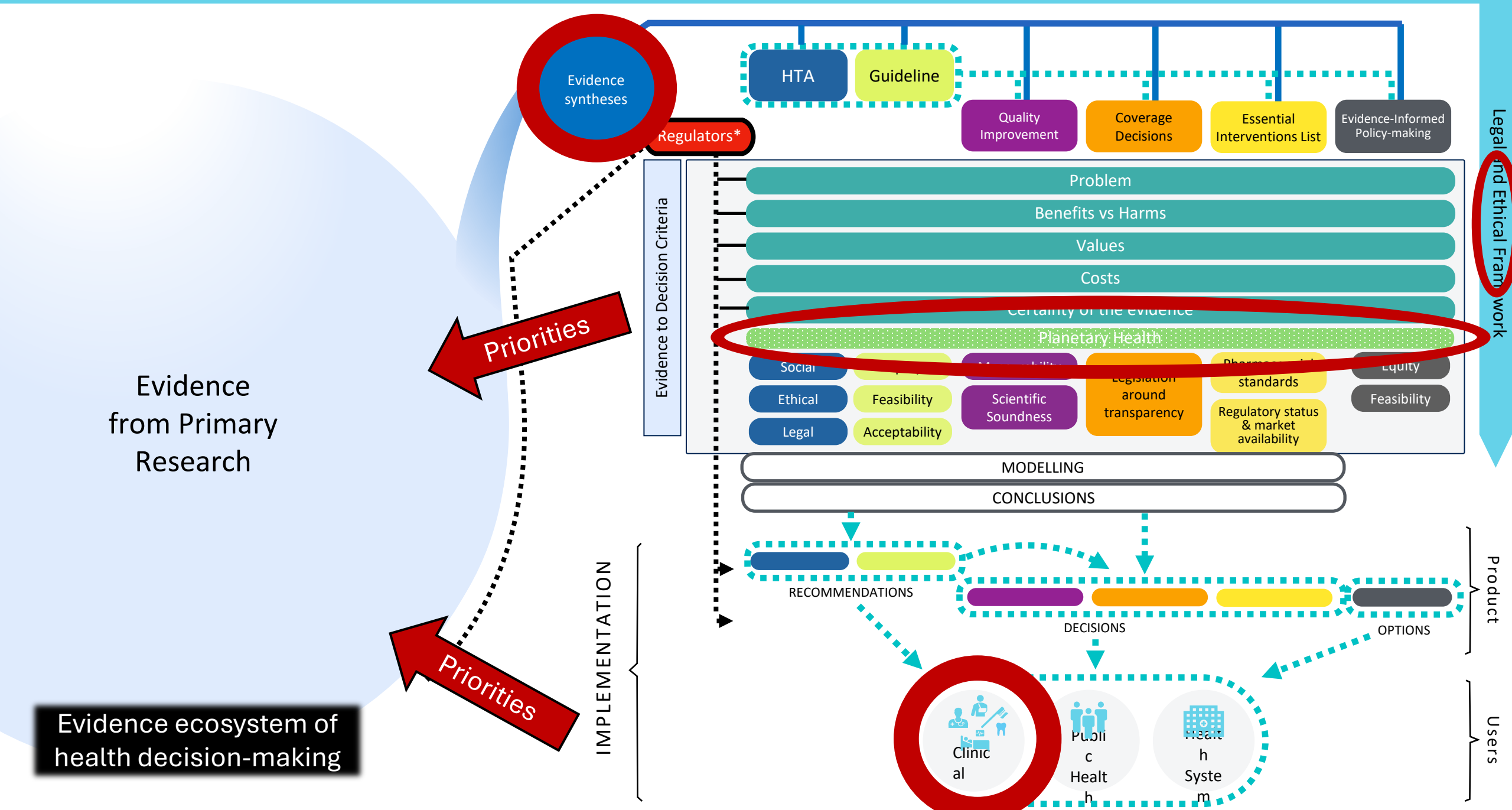
Who defines what counts as evidence? Who is excluded – if any – and who is included?

Ethics and Evidence – What We're Missing

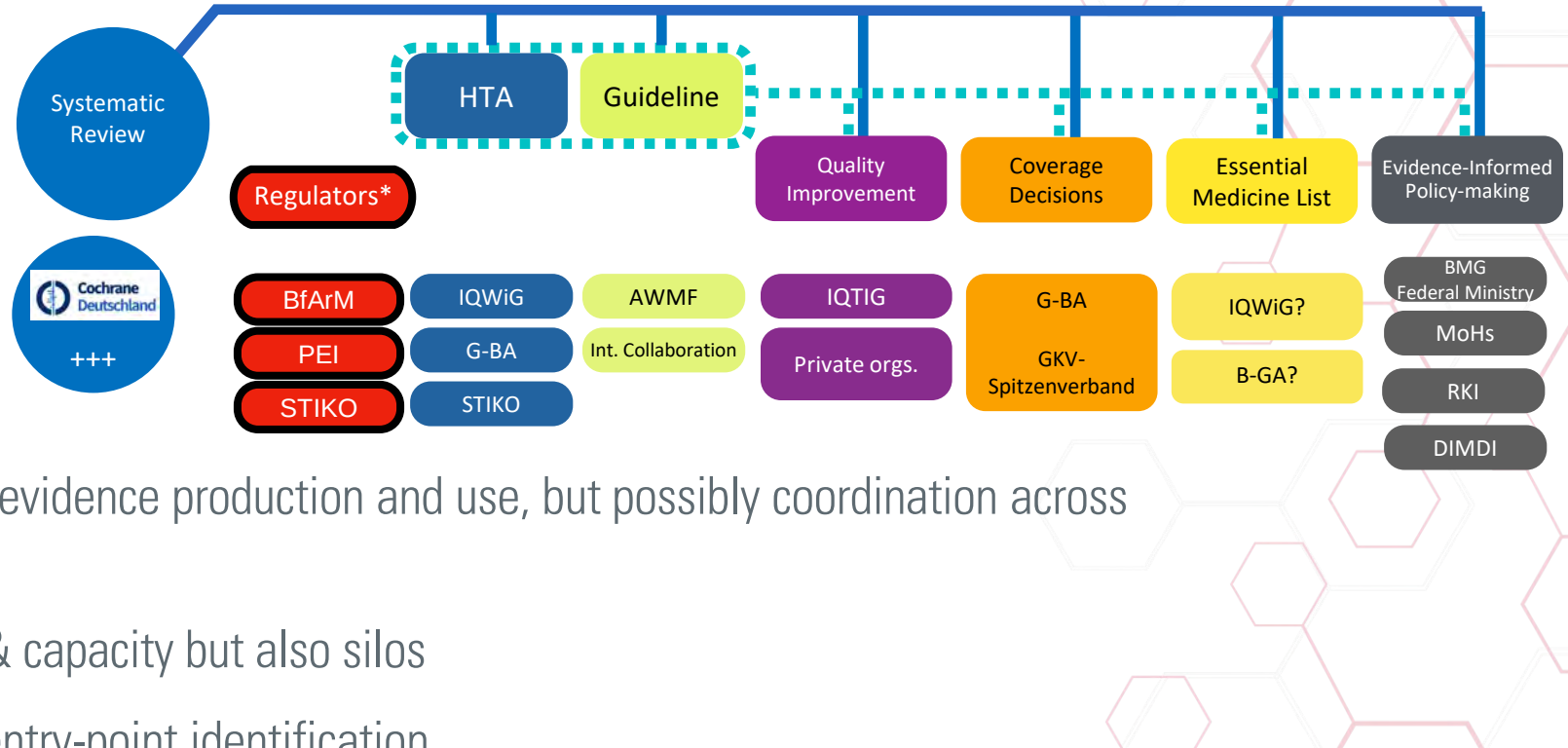
Ethics often treated as background, not structure.

Kathrin Littler: 'We need ethical reflection at every decision layer.'

Proposal: ethics guidance to be aligned with ToE logic model.



Germany – ToE Actor Mapping



Germany has strong institutions for evidence production and use, but possibly coordination across them may be improved

Parallel structures = specialization & capacity but also silos

ToE supports system mapping and entry-point identification

What is the role of the EBM network?

What is the **role and connection to primary research**: Network of University Medicine to fill gap

WHO Global Coalition for Evidence (GC4E)

WHO-led initiative for collaboration, coordination, consolidation.

Four working groups, launched at GES 2025

ToE is being applied in working group 2 to integrate different evidence streams.

Collaboration, Coordination, Consolidation

System transformation through practical tool integration.

GRADE working group

Mapping the Decision Genome against ToE

Not all decisions are equal: technical, political, procedural.

Decision genome clarifies types, criteria, and actor roles in the ToE

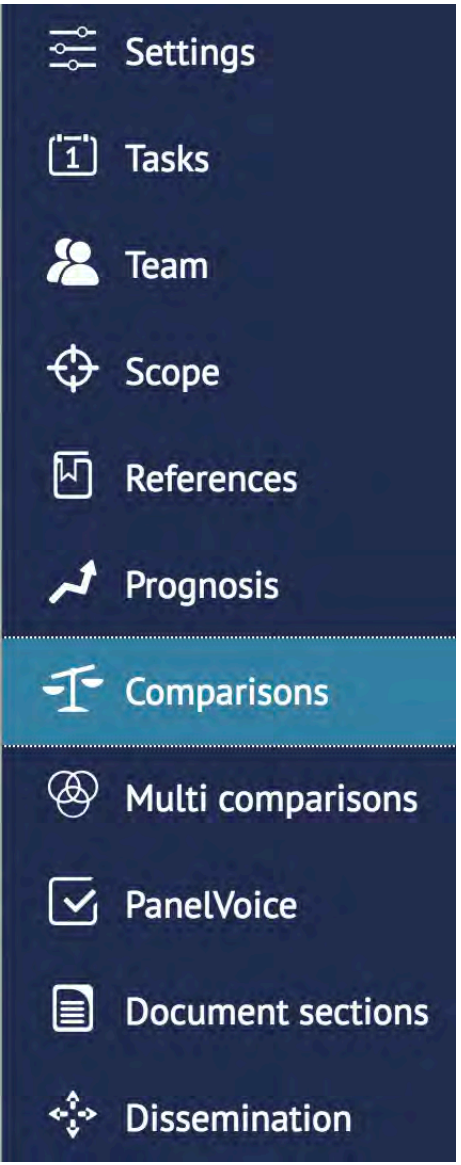
Types of science and research (social, economics) needed

Using GRADE to help identify where research is needed to fill gaps by mapping against many elements of the ToE framework.

GRADE approach, method or system

- Certainty assessment
- Evidence to Decision Frameworks

Evidence to Decision Frameworks



Question

- Details – PICO Subgroups
- Background and conflicts of interest

Assessment

- Criteria (Certainty of evidence)
- Judgements
- Research evidence (HTA and Systematic Reviews)
- Additional considerations

Conclusions

- Type of decision - recommendation
- Justification
- Implementation considerations - monitoring and evaluation
- Research considerations

GRADEpro

Perspectives

- Clinical – individual
- Clinical – population
- Health Systems & Public Health
- Health Systems & Policy

Type of decisions

- Recommendation
- Policy
- Coverage

Use

- Group decision making
- In person/online

BMJ, JCE, IJHTA, HARPS, 2016-18

Implications for research: Evidence to Decision Frameworks

Problems with the process

Research priorities

- Phase 3 clinical trial(s) of safety and efficacy of bedaquiline, with particular attention to mortality (including causes of death), in the treatment of MDR-TB should be accelerated
- Development of a reliable test for bedaquiline resistance.
- Pharmacokinetics, safety, and efficacy studies in specific populations (children, HIV patients, alcohol and drug misusers, elderly, pregnant women, diabetics, and people with extrapulmonary TB).
- Well designed safety studies events (short and long term), including type, frequency, and severity of adverse events.
- Drug-drug interactions, including with existing and other newly developed anti-TB drugs and antiretroviral drugs.
- Mortality (including cause of death).
- Acquisition of resistance to bedaquiline and to other anti-TB drugs.
- Duration and dosing of treatment.
- Patient acceptability.
- Further research on the validity of culture conversion as a surrogate marker of treatment outcome.

*Adapted from a WHO guideline.[22] This should not be considered as a WHO recommendation. An interactive version of this framework which includes subgroup information can be found at http://ietd.epistemonikos.org/#/frameworks/54992ce9352a502d58179c5c/question and at http://dbep.gradepro.org/profile/3879A46D-7E19-FEBA-9B96-BC2B3F996EB1

Alonso-Coello, et al. BMJ 2016, WHO Bedaquiline guidelines 2013

GRADE for research

Objective of this project group

- Guidance for the identification of research gaps and how to fill them
- Higher certainty evidence for a body of evidence that informs the evidence to decision framework
- Across the EtD: Question, background, decision criteria, conclusions
- So far not specific
- Having ToE and decision genome in mind

Certainty of evidence

The GRADE domains that are influenced by primary research:

Risk of bias, imprecision, indirectness, inconsistency and dissemination (publication) bias

Upgrading domains

- less susceptible to being the suggested focus of additional research.

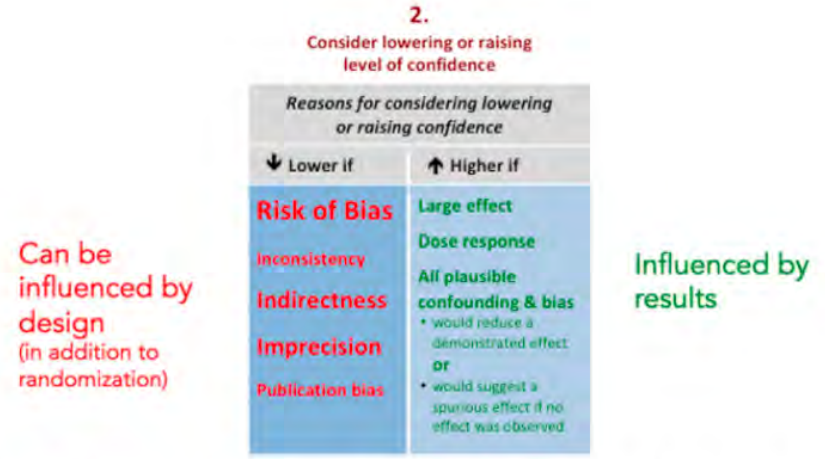


Figure 2 GRADE's criteria for the confidence in the effects. Legend: Generally, factors that lead to lowering the confidence are influenced by the design of studies (where the size of the font indicate how much influence can typically be exerted). Risk of bias and to a lesser degree indirectness (through broader PICO criteria) and imprecision (e.g., through planned larger sample sizes with higher event rates) and to an even smaller degree inconsistency (e.g., through appropriate a priori hypotheses to evaluate inconsistency) and publication bias (e.g. by ensuring that all studies are published) can be influenced through design features. The confidence can be raised by findings such as a large effect or dose response relations of issues related to plausible residual confounding and bias.

Methodological idiosyncracies, frameworks and challenges of non-pharmaceutical and non-technical treatment interventions

Methodische Besonderheiten, Rahmenbedingungen und Herausforderungen bei Studien mit nicht-medikamentösen und nicht-technischen Behandlungsverfahren

Holger J. Schünemann*

Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ) (2013) 107, 214–220

Table 2. Interpretation of the certainty of a body of evidence for research and practice, according to individual GRADE domains (Cochrane Handbook * Zhang et al, 2018).

By outcome	Implications for research	Examples
Risk of bias	Need for methodologically better designed and executed studies	All studies suffered from lack of blinding of outcome assessors. Trials blinding outcome assessors are required.
Inconsistency	Unexplained inconsistency: need for individual participant data meta-analysis (IPDMA) to explore subgroup effects; need for studies in relevant subgroups	Studies in patients with small cell lung cancer are needed to understand if the effects differ from those in patients with pancreatic cancer.
Indirectness	Need for studies that more directly address the PICO elements and question of interest	Studies in patients with early cancer are needed because the evidence is from studies with advanced cancer.
Imprecision	Need for more studies with more participants to reach optimal information size.	Studies with approximately 200 more events in the treatment and control group are required.
Publication bias	Need to investigate and identify unpublished data; large studies might help resolve this issue	Unpublished studies about flavonoids for hemorrhoids
Large effects	No implications	No implications
Dose effects	No implications	No implications
Opposing bias and confounding	Studies controlling for the residual bias and confounding may be needed to better estimate the effects.	Studies controlling for following possible confounders may be required smoking, degree of education.

Our approach: all sections of the EtD framework : Structured conclusion section

	Status quo	Research gaps Synthesis, primary, policy, implementation	Recommendations for research
Question	PICO elements missing	Missing subgroup Missing interventions	New drug regiments containing bedaquiline
Background	There is resistance to bedaquiline	Mechanims In whom	Primary research
EtD criterion	**Certainty of evidence (if available)**	**Research gaps** Synthesis, primary, policy, implementation	**Recommendations for research**
Problem			Who, what, how?
Values			Studies on the relative importance (utility values) for tuberculosis outcomes
Desirable and undesirable health effects			
Certainty of the evidence*	Very low for mortality (bedaquiline containing regiments)	Individual participant data-meta analysis Programmatic research	large trials (is always the answer)
Resource implications			
Cost effectiveness			
Equity			
Acceptability	gap of concepts (on to do list)		Research on how to overcome barriers
Feasibility	gap of concepts (on to do list)		

Future-Proofing Decisions

AI-enhanced EtD and research conclusions generated based on trained models (requires framework for drawing conclusions), system dashboards

Another project on target experiment approach to best possible evidence will inform

GRADE + ToE as operating system of decision-making and research priority setting.

Summary

1. Theory of everything as a roadmap for actors and factors

    Country studies to enhance ToE

2. Filling the Gaps in ToE: Equity, Power, Ethics
3. WHO based Coalition to connect to key actors
4. Mapping the Decision Genome:

    Clarifies types, criteria, and actor roles in the ToE

5. Use legitimate tools to set research priorities and make research happen based on what actors in health decision making ecosystem need

Thank You – to all the collaborators

Contact: holger.schunemann@hunimed.eu

Let's make the next breakthrough in health decisions.